JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Exactly Solvable Many-Boson Model*

R. W. RICHARDSON
Physics Department, New York University, University Heights, New York

(Received 16 October 1967)

A many-boson model is formulated and expressions for its exact eigenstates and energies are obtained for both an arbitrary finite and an infinite number of bosons. The Hamiltonian of the model contains interactions between bosons whose momenta have equal magnitudes but opposite directions. The matrix elements of this interaction are taken to be a constant over a range of momenta surrounding $\mathbf{k} = 0$. The ground state of the $2N$-particle system is shown to be a product of N pair-creation operators acting on the vacuum state. Each of these pair-creation operators depends upon one of N parameters which are called pair energies. The N pair energies are shown to satisfy a coupled system of nonlinear algebraic equations. The energy of the state is the sum of the pair energies and the occupation probabilities of the single-particle levels are given as simple functions of the pair energies. Similar results are derived for the excited states of the system and for the states of an odd number of particles. These results are valid for both a repulsive and an attractive interaction, since they only depend upon the form of the interaction. The equations are solved algebraically for two model systems. The first of these is one whose single-particle kinetic energy takes on only one value. The equations for this system are solved for an arbitrary interaction strength and it is shown that the pair energies are proportional to the zeros of certain Laguerre polynomials. The second system is one in which the single-particle kinetic energy can take on two values. The equations for this system are solved in the strong repulsive-interaction limit and it is shown that the pair energies are proportional to the zeros of certain Jacobi polynomials. The excitation energies of this second system are shown to be proportional to $1/n$ and the occupations of the two single-particle levels in the ground state are shown to be proportional to n, where n is the total number of particles. For a repulsive interaction and an arbitrary single-particle spectrum, the algebraic equations for the pair energies are converted into an approximate integral equation for the density of roots which is accurate to order $1/n$. This integral equation is solved for a strong interaction which, in the context of this model, means an interaction whose strength is greater than a constant times $1/V^{\frac{1}{2}}$ in the limit of a large volume. From this solution, the following results are obtained: (1) the lowest two single-particle levels have occupations of order n; (2) the excitation spectrum is that of a set of noninteracting quasiparticles; and (3) the quasiparticle spectrum has two zeros corresponding to the lowest two single-particle levels. Apart from the presence of two zeros, the quasiparticle spectrum does not differ significantly from that of the noninteracting particles.

1. INTRODUCTION

The Hamiltonian of an interacting many-boson system may be written as[1]

$$H = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a_{\mathbf{k}}^{+} a_{\mathbf{k}} + \frac{1}{2V} \sum_{\mathbf{k}\mathbf{k}'\mathbf{q}} v(\mathbf{q}) a_{\mathbf{k}+\mathbf{q}}^{+} a_{\mathbf{k}'-\mathbf{q}}^{+} a_{\mathbf{k}'} a_{\mathbf{k}}, \quad (1.1)$$

where $\epsilon_{\mathbf{k}} = k^2/2m$ is the energy of a boson with momentum $\mathbf{k}$ in units such that $\hbar = 1$, V is the volume of the system, $v(\mathbf{q})$ is the Fourier transform of the two-body interaction, and $a_{\mathbf{k}}^{+}$ and $a_{\mathbf{k}}$ are boson creation and annihilation operators which satisfy the usual Bose commutation rules

$$[a_{\mathbf{k}}, a_{\mathbf{k}'}^{+}] = \delta_{\mathbf{k}\mathbf{k}'}. \quad (1.2)$$

One widely used approach to the eigenstates of (1.1) is to replace it by a model Hamiltonian which may be easily diagonalized and which, in some sense, is a good approximation to H. Two examples of this approach are the Bogoliubov approximation[1,2] and the pair Hamiltonian.[1,3]

In the Bogoliubov approximation, it is assumed that the occupation of the $\mathbf{k} = 0$ level will be a finite fraction of the total number of particles in the thermodynamic limit when the volume of the system and the total number of particles go to infinity in such a way that their ratio, the density, is kept constant. The interaction in (1.1) is then truncated by ignoring all terms which have fewer than two operators referring to this $\mathbf{k} = 0$ state. The terms of the interaction that are retained in this approximation are those that involve the operators $a_0^{+}a_0^{+}a_0a_0$, $a_0^{+}a_{\mathbf{k}}^{+}a_{\mathbf{k}}a_0$, $a_{\mathbf{k}}^{+}a_0^{+}a_{\mathbf{k}}a_0$, $a_{\mathbf{k}}^{+}a_{-\mathbf{k}}^{+}a_0a_0$, and $a_0^{+}a_0^{+}a_{-\mathbf{k}}a_{\mathbf{k}}$, for $\mathbf{k} \neq 0$. In order to

* This research was supported in part by the National Science Foundation.

[1] E. H. Lieb, *Lectures in Theoretical Physics* (The University of Colorado Press, Boulder, Colo., 1965), Vol. VIIc, p. 175.

[2] N. N. Bogoliubov, J. Phys. (USSR) **11**, 23 (1947); N. N. Bogoliubov and D. N. Zubarev, Sov. Phys.—JETP **1**, 83 (1955); N. M. Hugenholtz, Rept. Progr. Phys. **28**, 201 (1965).

[3] D. N. Zubarev and Iu. A. Tserkovnikov, Sov. Phys.—Dokl. **3**, 603 (1958); M. Girardeau and R. Arnowitt, Phys. Rev. **113**, 755 (1959); M. Girardeau, Phys. Rev. **115**, 1090 (1959); M. Girardeau, Phys. Rev. **127**, 1809 (1962); M. Girardeau, J. Math. Phys. **3**, 131 (1962); G. Wentzel, Phys. Rev. **120**, 1572 (1960); M. Luban, Phys. Rev. **120**, 965 (1962).

diagonalize the resulting Hamiltonian, the further approximation of replacing the operators $a_0^\dagger$ and a_0 $(n_0)^{\frac{1}{2}}$, where n_0 is the occupation of the $\mathbf{k} = 0$ state, is made. Thus, the commutator of $a_0^\dagger$ and a_0, which is one, is neglected relative to the expectation value of $a_0^\dagger a_0$ which, by assumption, is of order n, the total number of particles. The Hamiltonian resulting from these approximations is a bilinear form in the operators referring to the states with $\mathbf{k} \neq 0$ and may be diagonalized by a canonical transformation of these operators. The resulting excitation spectrum is given by the energies of a system of noninteracting quasiparticles whose energies are given by

$$e_k = [\epsilon_k(\epsilon_k + 2\rho v(k))]^{\frac{1}{2}}, \tag{1.3}$$

where the density of the system is $\rho = n/V$.

The pair Hamiltonian is also the result of a truncation of the interaction in (1.1). In this approximation, only those terms in the interaction which involve the operators $a_{\mathbf{k}}^\dagger a_{\mathbf{k}}^\dagger a_{\mathbf{k}'} a_{\mathbf{k}}$, $a_{\mathbf{k}}^\dagger a_{\mathbf{k}'}^\dagger a_{\mathbf{k}} a_{\mathbf{k}'}$, and $a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger a_{-\mathbf{k}'} a_{\mathbf{k}'}$ are kept. This includes the terms of the Bogoliubov approximation plus many more. The resulting Hamiltonian has been the subject of extensive studies,[3] where it is shown that, in the thermodynamic limit, its excitation spectrum is that of a system of noninteracting quasiparticles. However, the expression for the energies of the quasiparticles is much more complicated than (1.3).

It is important to note that, in both these approximations, the terms of the interaction that are retained are those that are quadratic in the number operators $a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$ or are products of the pair operators $a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger$ and $a_{-\mathbf{k}} a_{\mathbf{k}}$. In this paper, we consider a special Hamiltonian whose interaction contains terms of this second kind and show how it can be exactly diagonalized.

The Hamiltonian that we treat may be written as

$$H = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \frac{g}{2} \sum_{\mathbf{k},\mathbf{k}'} \theta(k)\theta(k') a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger a_{-\mathbf{k}'} a_{\mathbf{k}'}, \tag{1.4}$$

where

$$\begin{aligned} \theta(k) &= 1, \quad \text{for} \quad 0 < k < K, \\ &= 0, \quad \text{for} \quad k > K. \end{aligned} \tag{1.5}$$

This Hamiltonian is the boson analog of the reduced Hamiltonian of the Bardeen–Cooper–Schrieffer (BCS) theory of superconductivity.[4] As such, it has been studied by Valatin and Butler,[5] who transcribed the quasiparticle formulation of the BCS theory[6] into a form appropriate for bosons. In (1.4) and (1.5), g is an interaction strength and $1/K$ is the "range" of the separable nonlocal interaction. If we transform this interaction into configuration space, we have

$$\langle \mathbf{r}_1 \mathbf{r}_2 | \, v \, | \mathbf{r}_1' \mathbf{r}_2' \rangle = g\bar{\theta}(|\mathbf{r}_1 - \mathbf{r}_2|)\bar{\theta}(|\mathbf{r}_1' - \mathbf{r}_2'|), \tag{1.6}$$

where

$$\bar{\theta}(r) = \frac{1}{V} \sum e^{i\mathbf{k}\cdot\mathbf{r}} \theta(k) \xrightarrow[V\to\infty]{} \frac{K^3}{(2\pi)^2} \left[\frac{\sin Kr - Kr \cos Kr}{(Kr)^3} \right]. \tag{1.7}$$

The interaction strength of g may be taken to be proportional to $1/V$, as would be the case if (1.4) were the result of a truncation of (1.1). However, if this is done, then the interaction (1.6) is proportional to $1/V$ rather than being independent of the volume. Therefore, in Secs. 2–4, we use g as a parameter which may or may not be taken proportional to $1/V$. In Sec. 5, where we consider the infinite volume limit of a realistic system, we explicitly put g equal to G/V.

While it may well be argued that the form of the interaction in (1.4) is nonphysical, our point of view is that the exact solvability of the model more than makes up for this deficiency. Exactly solvable models are very useful testing grounds for the approximations of many-body theory and they provide a lot of insight into the properties of real many-body systems. Thus, the Hamiltonian (1.4) joins the small number of model Hamiltonians such as the model studied by Bassichis and Foldy[7] and the BCS reduced Hamiltonian[8] that may be analyzed exactly and in complete detail.

In Sec. 2, equations for all the eigenstates of (1.4) are derived. These equations are a set of coupled, nonlinear algebraic equations for a set of parameters which we call pair energies, which characterize a given eigenstate. The energy of a state is given in terms of the sum of these pair energies. In Sec. 3, an expression for the occupation probabilities of the single-particle levels in one of these states is derived. The evaluation of this expression is shown to require the solution of a system of algebraic equations whose coefficients depend upon the pair energies. Two model systems, for which the equations for the pair energies can be solved exactly, are treated in Sec. 4. The first of these is the system in which the single-particle kinetic energy takes on only one value and in the second it takes on two values. The equations for the first system are solved for an arbitrary interaction strength and those of the second are solved in the strong repulsive interaction limit. In Sec. 5, we solve our equations for a system with an arbitrary single-particle spectrum and a repulsive interaction

[4] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

[5] J. G. Valatin and D. Butler, Nuovo Cimento **10**, 37 (1958).

[6] J. G. Valatin, Nuovo Cimento **7**, 843 (1958); N. N. Bogoliubov, Sov. Phys.—JETP **7**, 41 (1958).

[7] W. H. Bassichis and L. L. Foldy, Phys. Rev. **133**, A935 (1964).

[8] R. W. Richardson, Phys. Letters **3**, 277 (1963); R. W. Richardson and N. Sherman, Nucl. Phys. **52**, 221 (1964).

in the limit of an infinite number of particles. This is done by first converting the algebraic equations for the pair energies into an integral equation for the density of pair energies. This equation is then solved and expressions are given for the energies and occupation probabilities of the states. The excitation energies of the system are shown to be sums of independent quasiparticle energies. However, in contrast to (1.3), the quasiparticle energy is zero for the two lowest values of $\mathbf{k}$ instead of just the $\mathbf{k} = 0$ level. Corresponding to these two zeros in the quasiparticle energy, the occupations of the lowest two single-particle levels are finite fractions of the total number of particles. Thus, the model exhibits a very special kind of generalized Bose condensation of the form discussed by Girardeau.[9]

2. EIGENSTATES OF THE MODEL

We consider the eigenstates of n bosons contained in a volume V with periodic boundary conditions imposed. The single-particle states of the system are labeled by their momentum $\mathbf{k}$. In terms of these states, the model Hamiltonian may be written as

$$H = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a_{\mathbf{k}}^{+} a_{\mathbf{k}} + \frac{g}{2} \sum_{\mathbf{k}\mathbf{k}'} a_{\mathbf{k}}^{+} a_{-\mathbf{k}}^{+} a_{-\mathbf{k}'} a_{\mathbf{k}'}, \qquad (2.1)$$

where $\epsilon_{\mathbf{k}} = k^2/2m$ (we choose units such that $\hbar = 1$), g is the interaction strength, and $a_{\mathbf{k}}^{+}$ and $a_{\mathbf{k}}$ are boson creation and annihilation operators satisfying the usual Bose commutation rules (1.2). In (2.1), all sums over $\mathbf{k}$ are restricted to the range $0 < k < K$, where $k = |\mathbf{k}|$, and, in what follows, we will assume that all vectors $\mathbf{k}$ lie in this range. The neglected part of the Hamiltonian,

$$\sum_{\mathbf{k}, k > K} a_{\mathbf{k}}^{+} a_{\mathbf{k}},$$

only plays a role if we consider excitations with momenta greater than K and, since these states can be easily expressed in terms of the states of (2.1), we shall not discuss them in this paper.

In this section we will develop a set of equations for the eigenvalues and eigenstates of (2.1). In order to have a specific system in mind, we have written (2.1) in terms of plane-wave single-particle states. However, it should be pointed out that this is in no way essential for the analysis that follows. All we need is some set of single-particle states $\mathbf{k}$ for which $-\mathbf{k}$ is uniquely related to $\mathbf{k}$, e.g., by time reversal, and which satisfies $\epsilon_{\mathbf{k}} = \epsilon_{-\mathbf{k}}$. Thus, the analysis is equally valid for one-, two-, or three-dimensional systems with any single-particle spectrum that is invariant under time reversal.

[9] M. Girardeau, Phys. Fluids **5**, 1468 (1962); J. Math. Phys. **6**, 1083 (1965).

The analysis is also independent of the value of g and is therefore valid for both repulsive and attractive interactions. Before turning to this analysis, we will consider two points. First, we will rewrite (2.1) so as to emphasize the arbitrary nature of the single-particle states $\mathbf{k}$ and then we will consider in detail the problem of labeling the eigenstates of (2.1).

It will turn out that the discrete nature of the single-particle spectrum will play an important role in the equations that we are about to derive and this remains true even in the infinite volume limit. We therefore order the possible values that ϵ_k can take on with the integers as $\epsilon_0, \epsilon_1, \cdots, \epsilon_m$ with $\epsilon_l < \epsilon_{l+1}$. We also define the degeneracy of the lth level, Ω_l, as being the number of different values of $\mathbf{k}$ for which $\epsilon_{\mathbf{k}} = \epsilon_l$. It is then useful to perform the sums in (2.1) over the states in each degenerate level and write it as

$$H = \Sigma_l \epsilon_l \hat{n}_l + \frac{g}{2} \Sigma_{ll'} A_l^{+} A_{l'}, \qquad (2.2)$$

where

$$\hat{n}_l = \sum_{(\epsilon_{\mathbf{k}} = \epsilon_l)} a_{\mathbf{k}}^{+} a_{\mathbf{k}}, \qquad (2.3)$$

and

$$A_l^{+} = \sum_{(\epsilon_{\mathbf{k}} = \epsilon_l)} a_{\mathbf{k}}^{+} a_{-\mathbf{k}}^{+}, \quad A_l = \sum_{(\epsilon_{\mathbf{k}} = \epsilon_l)} a_{-\mathbf{k}} a_{\mathbf{k}}. \qquad (2.4)$$

The sums in (2.3) and (2.4) may contain as few as one or two terms if the single-particle state only has time-reversal degeneracy, or they may contain many terms if there are other degeneracies. In general, the number of terms is Ω_l. In the form (2.2), the single-particle spectrum ϵ_l and the associated degeneracies Ω_l are a set of arbitrary numbers and the Hamiltonian is manifestly independent of the particular single-particle states. For future reference, it will be useful to complete this transcription of the Hamiltonian by listing the commutation rules of the operators defined in (2.3) and (2.4). These follow from the Bose commutation rules (1.2) of the operators $a_{\mathbf{k}}^{+}$ and $a_{\mathbf{k}}$, and are given by

$$[\hat{n}_l, A_{l'}^{+}] = 2\delta_{ll'} A_l^{+} \qquad (2.5)$$

and

$$[A_l, A_{l'}^{+}] = 2\delta_{ll'}(\Omega_l + 2\hat{n}_l). \qquad (2.6)$$

In labeling the n-particle states of this Hamiltonian, it is useful to introduce the concept of an unpaired particle. From (1.6) and (1.7), it is clear that the interaction is only effective between two particles with zero total linear momentum and which are coupled to zero angular momentum. Therefore, a particle that is not coupled to zero linear momentum and zero angular momentum with any other particle will not interact. We will call such a particle an unpaired particle. To be more explicit, let us first define a ν-particle state with

all the particles unpaired. This state is defined by the equations

$$\hat{n}\,|\varphi_\nu\rangle = \nu\,|\varphi_\nu\rangle,$$
$$A_l\,|\varphi_\nu\rangle = 0, \quad \text{for all } l, \tag{2.7}$$

where $\hat{n}$ is the total number of particles operator

$$\hat{n} = \Sigma_l \hat{n}_l. \tag{2.8}$$

These equations, of course, do not uniquely determine the state $|\varphi_\nu\rangle$. However, since the unpaired particles in this model do not interact, we may further require that the state $|\varphi_\nu\rangle$ be an eigenstate of the operators $\hat{n}_l$, i.e.,

$$\hat{n}_l\,|\varphi_\nu\rangle = \nu_l\,|\varphi_\nu\rangle. \tag{2.9}$$

Here ν_l is the number of unpaired particles occupying level l. Equations (2.7) and (2.9) still do not determine $|\varphi_\nu\rangle$ uniquely. However, as we shall see, the energies of the eigenstates of (2.2) will only depend upon the numbers ν_l and are therefore degenerate with respect to the quantum numbers of $|\varphi_\nu\rangle$ that we have not defined. The possible values of the quantum numbers ν_l are

$$\begin{aligned} \nu_l &= 0, 1, && \text{for} \quad \Omega_l = 1, \\ &= 0, 1, \cdots, n, && \text{for} \quad \Omega_l > 1, \end{aligned} \tag{2.10}$$

where n is the total number of particles in the state. Here, the first case is appropriate for a $\mathbf{k} = 0$ level and the second case for a $\mathbf{k} \neq 0$ level of a system. Note that with (2.7) and (2.9) we have

$$\nu = \Sigma \nu_l$$

and

$$H\,|\varphi_\nu\rangle = (\Sigma_l \nu_l \epsilon_l)\,|\varphi_\nu\rangle. \tag{2.11}$$

Such states may be constructed by putting particles in states $\mathbf{k}$ but not in states $-\mathbf{k}$, e.g., for $\nu = 2$,

$$|\varphi_2\rangle = (a_{\mathbf{k}}^+)^2\,|0\rangle,$$

where $|0\rangle$ is the vacuum state, or they may be made up of pairs of particles with nonzero total angular momentum, e.g.,

$$|\varphi_2\rangle = \sum_{(\epsilon_{\mathbf{k}}=\epsilon_l)} \Sigma_{\mathbf{k}} \alpha(\mathbf{k}) a_k^+ a_{-k}^+\,|0\rangle,$$

where $\Sigma_{\mathbf{k}} \alpha(\mathbf{k}) = 0$. An arbitrary state that has $n = 2N + \nu$ particles and ν unpaired particles may now be constructed as a linear combination of the states

$$A_{l_1}^+ \cdots A_{l_N}^+\,|\varphi_\nu\rangle. \tag{2.12}$$

We will refer to such states as having N pairs of particles and ν unpaired particles. The quantum numbers ν_l are good quantum numbers of such states and are therefore good quantum numbers of the eigenstates of H.

In order to distinguish between those states of n bosons that have the same quantum numbers ν_l, we must introduce labels that distinguish the various states formed from those of (2.12). We do this by introducing the set of labels $l_{10} \cdots l_{N0}$ defined as follows: if $|\psi\rangle$ is the eigenstate of H that is being labeled, then the set $l_{10} \cdots l_{N0}$ is defined by

$$\lim_{g \to 0} |\psi\rangle = A_{l_{10}}^+ \cdots A_{l_{N0}}^+\,|\varphi_\nu\rangle, \tag{2.13}$$

i.e., $l_{10} \cdots l_{N0}$ are the levels that are occupied by paired bosons in the limit of zero interaction strength. Note that many of the indices $l_{\alpha 0}$ may be equal. Thus, the ground state of an even number of particles is labeled by the quantum numbers

$$l_{\alpha 0} = 0, \quad \nu_l = 0, \quad \alpha = 1 \cdots n/2,$$

and that of an odd number of particles by

$$l_{\alpha 0} = 0, \quad \nu_0 = 1, \quad \nu_l = 0, \quad l > 0,$$
$$\alpha = 1 \cdots (n-1)/2.$$

With these notational preliminaries over, we will now show that an n-boson eigenstate of H, with N pairs of bosons and ν unpaired bosons, can be written as

$$|\psi\rangle = B_1^+ \cdots B_N^+\,|\varphi_\nu\rangle, \tag{2.14}$$

where we have disregarded a normalization constant. The pair-creation operators B_α^+ in (2.14) are given by

$$B_\alpha^+ = \sum_l u_\alpha(l) A_l^+, \quad \alpha = 1 \cdots N, \tag{2.15}$$

and in proving (2.14) we will determine the form of the amplitudes u_α as well as the energy of the state. In order to demonstrate that (2.14) is an eigenstate of H, we calculate

$$\begin{aligned} H\,|\psi\rangle &= H B_1^+ \cdots B_N^+\,|\varphi_\nu\rangle \\ &= \left(\sum_l \nu_l \epsilon_l\right) |\psi\rangle + [H, B_1^+ \cdots B_N^+]\,|\varphi_\nu\rangle, \end{aligned} \tag{2.16}$$

where we have used (2.11). We now do some commutator algebra on the second term of (2.16) to write it as

$$\begin{aligned} &[H, B_1^+ \cdots B_N^+]\,|\varphi_\nu\rangle \\ &= \left\{ \sum_{\alpha=1}^{N} \left(\prod_{\gamma \neq \alpha} B_\gamma^+ \right) [H, B_\alpha^+] \right. \\ &\qquad \left. + \frac{1}{2} \sum_{\alpha,\beta=1}^{N}{}' \left(\prod_{\gamma \neq \alpha,\beta} B_\gamma^+ \right) [[H, B_\alpha^+], B_\beta^+] \right\} |\varphi_\nu\rangle, \end{aligned} \tag{2.17}$$

where the prime on the sum on α and β indicates that it is over those values of α and β satisfying $\alpha \neq \beta$. The commutators in (2.17) may be evaluated using (2.5), (2.6), and (2.15) with the results

$$[H, B_\alpha^+] = \sum_l 2\epsilon_l u_\alpha(l) A_l^+ + g \sum_{ll'} A_l^+ u_\alpha(l')(\Omega_{l'} + 2\hat{n}_{l'}) \tag{2.18}$$

and

$$[[H, B_\alpha^+], B_\beta^+] = 4g \sum_{ll'} A_l^+ u_\alpha(l') u_\beta(l') A_{l'}^+ . \quad (2.19)$$

When these results are substituted into (2.16), we obtain a Schrödinger equation for the amplitude of the state $|\psi\rangle$. However, before doing so, we will discuss the form of the double commutator (2.19).

In order to proceed further, we assume that the product $u_\alpha u_\beta$ in the double commutator (2.19) can be written as

$$u_\alpha(l) u_\beta(l) = M_{\alpha\beta} u_\alpha(l) + M_{\beta\alpha} u_\beta(l), \quad (2.20)$$

for all l and all $\alpha \neq \beta$. This is to be regarded as a set of equations that must be satisfied by the u_α and it introduces the as yet undetermined matrix M. This assumption is made with the advantage of hindsight and we will show that it can be exactly satisfied. The skeptical reader may wish to treat (2.20) as an approximation and place a remainder term on its right-hand side. The matrix M would then be chosen so as to minimize the magnitude of this remainder. He would then proceed as we do and derive equations for the u_α that are identical to ours. However, these equations would be valid in the approximation in which the remainder term introduced into (2.20) is neglected. The solutions to these equations will be the same as ours given below and, as we will show, these solutions satisfy (2.20) exactly. Therefore, the remainder term is zero and the approximation is exact. We will therefore proceed under the assumption that (2.20) is exactly satisfied and we shall demonstrate that this is so after we determine the form of the amplitudes u_α. Using (2.20), the double commutator (2.19) becomes

$$[[H, B_\alpha^+], B_\beta^+] = 4g\left(\sum_l A_l^+\right)(M_{\alpha\beta} B_\alpha^+ + M_{\beta\alpha} B_\beta^+). \quad (2.21)$$

If we write the energy of the state $|\psi\rangle$ as

$$E = \sum_l \nu_l \epsilon_l + \sum_{\alpha=1}^{N} E_\alpha, \quad (2.22)$$

where the pair energies E_α are yet to be determined, then, substituting (2.18) and (2.21) into (2.17) and the result into (2.16), we obtain

$$(H - E)|\psi\rangle = \sum_{\alpha=1}^{N}\left(\prod_{\gamma\neq\alpha} B_\gamma^+\right)\sum_l A_l^+\Big[(2\epsilon_l - E_\alpha)u_\alpha(l) + g\sum_{l'}(\Omega_{l'} + 2\hat{n}_{l'})u_\alpha(l') + 4g\sum_\beta{}' M_{\beta\alpha}\Big]|\varphi_\nu\rangle. \quad (2.23)$$

Since the operators $\hat{n}_{l'}$ in this expression operate on the state $|\varphi_\nu\rangle$, they may be replaced by the quantum numbers $\nu_{l'}$ of that state using (2.9). The expression (2.23) will now vanish if the u_α satisfy the system of equations

$$(2\epsilon_l - E_\alpha)u_\alpha(l) + g\sum_{l'}{}'(\Omega_{l'} + 2\nu_{l'})u_\alpha(l') + 4g\sum_\beta{}' M_{\beta\alpha} = 0, \quad \alpha = 1 \cdots N. \quad (2.24)$$

Equations (2.20) and (2.24) provide a complete set of equations for the amplitudes $u_\alpha(l)$, the matrix M, and the pair energies E_α.

The solution of Eqs. (2.20) and (2.24) starts with the observation that only the first term of (2.24) depends upon l. We can therefore immediately solve for the l dependence of the amplitudes u_α which is given by

$$u_\alpha(l) = \frac{-gC_\alpha}{2\epsilon_l - E_\alpha}, \quad (2.25)$$

where C_α is given by

$$C_\alpha = \sum_l{}'(\Omega_l + 2\nu_l)u_\alpha(l) + 4\sum_\beta{}' M_{\beta\alpha}. \quad (2.26)$$

Furthermore, using (2.25) in the product $u_\alpha u_\beta$, we have

$$u_\alpha(l)u_\beta(l) = \frac{g^2 C_\alpha C_\beta}{(2\epsilon_l - E_\alpha)(2\epsilon_l - E_\beta)} = \frac{g^2 C_\alpha C_\beta}{E_\alpha - E_\beta}\left[\frac{1}{2\epsilon_l - E_\alpha} - \frac{1}{2\epsilon_l - E_\beta}\right], \quad (2.27)$$

where we have assumed $E_\alpha \neq E_\beta$ in order to perform the partial-fraction expansion. The validity of this assumption will be discussed in the following paragraph, where we will show that it is always satisfied. Comparing (2.20), (2.25), and (2.27), we have

$$M_{\beta\alpha} = \frac{-gC_\alpha}{E_\beta - E_\alpha} \quad (2.28)$$

for the matrix M, and Eq. (2.20) is exactly satisfied. Substituting (2.25) and (2.28) into (2.26), we then have

$$C_\alpha = -g\left[\sum_l{}'\frac{\Omega_l + 2\nu_l}{2\epsilon_l - E_\alpha} + 4\sum_\beta{}'\frac{1}{E_\beta - E_\alpha}\right]C_\alpha, \quad \alpha = 1 \cdots N,$$

which, for nonvanishing C_α, yields the set of equations

$$1 + 4g\sum_\beta{}'\frac{1}{E_\beta - E_\alpha} + g\sum_l{}'\frac{\Omega_l + 2\nu_l}{2\epsilon_l - E_\alpha} = 0, \quad \alpha = 1 \cdots N, \quad (2.29)$$

for the pair energies E_α. Since the coefficients C_α are arbitrary and the over-all factor of $\Pi_\alpha(-gC_\alpha)$ can be absorbed in the normalization coefficient of the state, we may write the u_α as

$$u_\alpha(l) = \frac{1}{2\epsilon_l - E_\alpha}. \quad (2.30)$$

Therefore, Eqs. (2.14), (2.15), (2.30), (2.29), and (2.22)

are the equations for the eigenstates of H. There are two aspects of these equations that need further discussion. The first is the validity of the assumption $E_\alpha \neq E_\beta$ used in the derivation of (2.29) and the second is the question of whether all the solutions of (2.29) are in one-to-one correspondence with all the states of H. We now turn to a discussion of these two points.

The question of the validity of the assumption $E_\alpha \neq E_\beta$ may be investigated using the same methods that were used in discussing the same point in the corresponding many-fermion problem. We will briefly outline the application of these methods to this many-boson problem. The question to be answered is under what conditions is the assumption $E_\alpha \neq E_\beta$ for all $\alpha \neq \beta$ incompatible with the E_α being roots of (2.29). In the corresponding many-fermion problem, whose equations differ from (2.29) by presence of some minus signs instead of plus signs, we found that these two requirements on the E_α are indeed incompatible for a finite set of isolated values of g. For the boson problem, we will show that the signs in (2.29) are such that the two requirements on the E_α are always compatible. Clearly, in the limit $g \to 0$, the conditions $E_\alpha \neq E_\beta$ are not satisfied for any state that has more than one pair in a particular single-particle level in this limit. In particular, for the ground state, all the E_α are $2\epsilon_0$ in the limit $g \to 0$. However, one can show that

$$\lim_{g\to 0} \frac{1}{g}(E_\alpha - E_\beta) \neq 0, \tag{2.31}$$

and this is sufficient for the validity of (2.27). We defer the proof of (2.31) until Sec. 4, where we treat a one-level kinetic-energy model as an example of our equations. There we show that if the sum on l in (2.29) is restricted to one term, then its roots are g times the zeros of certain Laguerre polynomials. Since the zeros are distinct, (2.31) is satisfied for this system. However, in the limit $g \to 0$, the single-particle level spacing becomes very large compared to the interaction strength and each degenerate set of single-particle states can be treated as an isolated one-level kinetic-energy system. The result (2.31) is therefore proven for an arbitrary system. Furthermore, from the known properties of the zeros of the Laguerre polynomials, we know that the pair energies are all real in this limit. We have thus shown that our assumption is satisfied for an arbitrary state and system in a neighborhood about $g = 0$.

In order to study the validity of the assumption for $g \neq 0$, we first consider the conditions under which all N pair energies may become equal. For convenience, we adjust the energy scale so that the value that the pair energies have when they are equal is zero. We therefore seek conditions that must be satisfied if Eqs. (2.29) are to have the solution $E_\alpha = 0$, $\alpha = 1 \cdots N$, for $g \neq 0$. Since we have shown that Eqs. (2.30) have solutions that satisfy $E_\alpha \neq E_\beta$ for small g, we can think of the violation of this condition occurring at some value of g which we call g_0. We therefore have the situation that is described by

$$E_\alpha \neq E_\beta, \quad \text{for} \quad |g| < |g_0|,$$
$$E_\alpha = 0, \quad \text{for} \quad g = g_0 \quad \text{and} \quad \alpha = 1 \cdots N. \tag{2.32}$$

We now show that this is impossible by considering (2.29) in the limit $g \to g_0$ with $|g| < |g_0|$. If we multiply Eqs. (2.29) by E_α and then sum over α, we have

$$\sum_{\alpha=1}^{N} E_\alpha - 2gN(N-1) + g\sum_l \sum_{\alpha=1}^{N} \frac{E_\alpha(\Omega_l + 2\nu_l)}{2\epsilon_l - E_\alpha} = 0, \tag{2.33}$$

where ϵ_l is now the energy of the single-particle level l after we have readjusted the zero of energy as described above. We now take the limit of (2.33) as $g \to g_0$. This limit is given by

$$g_0 N(\omega + 2N - 2) = 0, \tag{2.34}$$

where $\omega = 0$ if $\epsilon_l \neq 0$ for all l and $\omega = \Omega_{l_0} + 2\nu_{l_0}$ if $\epsilon_{l_0} = 0$ for some value l_0 of l. However, since $\omega \geq 0$, (2.34) cannot be satisfied for $g_0 \neq 0$. Thus, we have shown that Eqs. (2.29) do not have a solution with all the E_α equal to each other.

The above argument may be easily generalized to the cases in which fewer than N pair energies become equal. Thus, we have shown that the assumption $E_\alpha \neq E_\beta$ is implied by Eqs. (2.29) and it is not an additional set of conditions that the roots must satisfy. Furthermore, since we have shown that the roots are real and distinct for small g and that no two become equal, we have shown that the pair energies are real and distinct for all values of $g \neq 0$. This is in contrast to the corresponding many-fermion problem,[10] in which there are isolated values of g, where the conditions $E_\alpha \neq E_\beta$ are violated. The fact that the pair energies are real for all values of g in the boson problem makes the analysis of their properties much simpler than the analysis of the pair energies that describe the eigenstates of the fermion pairing Hamiltonian.

In the light of the preceding discussion on the conditions $E_\alpha \neq E_\beta$, it is now easy to set up a one-to-one correspondence between the states of the interacting system, i.e., the roots of (2.29), and the states

[10] R. W. Richardson, J. Math. Phys. 6, 1034 (1965).

of the noninteracting system and therefore show that we have equations for all the states of the system. We have indicated that the states of the paired particles are labeled by the quantum numbers $l_{10}\cdots l_{N0}$, which indicate the levels that are occupied by pairs in the limit $g\to 0$. Transcribing this label into a property of pair energies, we have

$$\lim_{g\to 0} E_\alpha = 2\epsilon_{l_{\alpha 0}}, \quad \alpha = 1\cdots N, \tag{2.35}$$

and from our discussion of the one-level kinetic energy model we have indicated that such a limit exists for all possible choices of $l_{10}\cdots l_{N0}$. We can therefore single out the roots of Eqs. (2.29) which correspond to a state with any set of quantum numbers $l_{10}\cdots l_{N0}$. This concludes the proof of completeness. However, knowing that the pair energies E_α are real and distinct for all values of g, we can rewrite (2.35) in the more practical form

$$2\epsilon_{l_{\alpha 0}} < E_\alpha < 2\epsilon_{l_{\alpha 0+1}}, \quad \text{for} \quad g>0 \quad \text{and} \quad \alpha = 1\cdots N, \tag{2.36}$$

and

$$2\epsilon_{l_{\alpha 0-1}} < E_\alpha < 2\epsilon_{l_{\alpha 0}}, \quad \text{for} \quad g<0 \quad \text{and} \quad \alpha = 1\cdots N, \tag{2.37}$$

where in this last expression we have defined $\epsilon_{-1} = -\infty$. Thus, for a given value of g, a state is labeled by the way the roots E_α are distributed between the values of $2\epsilon_l$. The equivalence of these last two expressions with (2.35) follows from two properties of the pair energies. First, from the one-level kinetic-energy model mentioned above and treated in Sec. 4, we have $(\partial E_\alpha/\partial g)_{g=0} > 0$. And second, since the pair energies are real and distinct, the only time when $E_\alpha = 2\epsilon_l$, for some value of l, is when $g=0$. Therefore, for a repulsive (attractive) interaction, the pair energies are increasing (decreasing) functions of $|g|$ at $g=0$ and for larger values of $|g|$ are bounded by the values of $2\epsilon_l$ as given in (2.36) and (2.37).

3. OCCUPATION PROBABILITIES

The occupation probabilities for the single-particle levels

$$n_l = \langle\psi|\,\hat{n}_l\,|\psi\rangle \tag{3.1}$$

may be calculated using a general theorem of quantum mechanics that is applicable to Hamiltonians that depend linearly on a parameter. In this case, the parameter is the single-particle energy ϵ_l and the theorem states that

$$n_l = \frac{\partial E}{\partial \epsilon_l}. \tag{3.2}$$

This may be easily proven using the fact that the expectation value of H in one of its eigenstates $|\psi\rangle$ is stationary with respect to variations of that state. Using the expression (2.22) for the energy in (3.2), we have

$$n_l = \nu_l + \sum_{\alpha=1}^{N} \frac{\partial E_\alpha}{\partial \epsilon_l} \tag{3.3}$$

for the occupation probability.

Equations for the derivatives of the pair energies that appear in (3.3) may be obtained by differentiating (2.29) with respect to ϵ_l. This yields the set of equations

$$\left[C_\alpha + 4\sum_\beta{}' \frac{1}{(E_\alpha - E_\beta)^2}\right]\frac{\partial E_\alpha}{\partial \epsilon_l} - 4\sum_\beta{}' \frac{1}{(E_\alpha - E_\beta)^2}\frac{\partial E_\beta}{\partial \epsilon_l} = \frac{2(\Omega_l + 2\nu_l)}{(2\epsilon_l - E_\alpha)^2}, \quad \alpha = 1\cdots N, \tag{3.4}$$

where

$$C_\alpha = \sum_l \frac{(\Omega_l + 2\nu_l)}{(2\epsilon_l - E_\alpha)^2}. \tag{3.5}$$

Solving (3.4) for $\partial E_\alpha/\partial\epsilon_l$ and substituting the result into (3.3), after rearranging the terms we get

$$n_l = \nu_l + 2\sum_{\alpha=1}^{R} \frac{\Omega_l + 2\nu_l}{(2\epsilon_l - E_\alpha)^2} D_\alpha, \tag{3.6}$$

where the D_α satisfy the system of equations

$$\left[C_\alpha + 4\sum_\beta{}' \frac{1}{(E_\alpha - E_\beta)^2}\right]D_\alpha - 4\sum_\beta{}' \frac{1}{(E_\alpha - E_\beta)^2} D_\beta = 1, \quad \alpha = 1\cdots N. \tag{3.7}$$

We may readily verify that the total number of particles in the state is $n = \nu + 2N$. For, summing (3.6) on l, we have

$$n = \sum_l n_l = \nu + 2\sum_{\alpha=1}^{N} C_\alpha D_\alpha,$$

and summing (3.7) on α, we have

$$\sum_{\alpha=1} C_\alpha D_\alpha = N,$$

which proves our point. Equations (3.6) and (3.7) provide a starting point for the study of the dependence of n_l on the interaction strength g and the single-particle spectrum ϵ_l.

4. TWO EXAMPLES

In order to exhibit the structure of our equations, we will discuss two simple models in some detail. The first model is a one-level kinetic-energy model in which the index l takes on the single value $l=0$ and we adjust the energy scale so that $\epsilon_0 = 0$. This model is used to prove (2.31) and thereby fill in the remaining hole in the proofs of Sec. 2. Physically, this model would approximate the strong-coupling limit of an attractive interaction in which the splitting of the

single-particle levels ϵ_l is ignored or it can be used to obtain the first-order perturbation theory expressions for the E_α as mentioned in Sec. 2. The second model has a two-level kinetic energy with l taking on the two values $l = 0$ or 1 and $\epsilon_0 = 0$ and $\epsilon_1 = 1$ in the appropriate units. The equations for this model are solved in the limit of a strong repulsive interaction and they show some interesting features which, as we will show in Sec. 5, are true of systems with more realistic single-particle spectra.

A. One-Level Kinetic Energy Model

If the sum over l in (2.29) contains just the single term with $l = 0$, then we have the equations

$$1 + 4\sum_{\beta}{}' \frac{1}{E_\beta - E_\alpha} = \frac{\omega}{E_\alpha}, \quad \alpha = 1 \cdots N, \quad (4.1)$$

where we have set $\epsilon_0 = 0$ and chosen the units of energy so that $g = 1$. In this expression, we have set $\omega = \Omega + 2\nu$, where we have dropped the subscript $l = 0$, and $N = \frac{1}{2}(n - \nu)$. We will obtain an explicit expression for the energies of the states of this system and we will show that the pair energies are proportional to the zeros of certain Laguerre polynomials. The interest in this model is not in the energies and eigenstates of the Hamiltonian since they can be calculated by much simpler means. Rather, the interest lies in the structure of Eqs. (4.1) and, in particular, in the proof of (2.31), i.e., the proof that the pair energies are distinct in the limit $g \to 0$. Nevertheless, as an introduction to (4.1), we will first derive the energies of the states of the model before we consider the structure of the equations.

The energy of a state of this model is the sum of the pair energies and we may obtain an expression for it by multiplying Eqs. (4.1) by E_α and then summing on α. Using the result

$$\sum_{\alpha,\beta=1}^{N}{}' \frac{E_\alpha}{E_\beta - E_\alpha} = -\tfrac{1}{2}N(N-1), \quad (4.2)$$

we then have

$$E = N(\omega + 2N - 2)$$

or, labeling the states with n and ν,

$$E(n\nu) = \tfrac{1}{2}(n - \nu)(n + \Omega + \nu - 2). \quad (4.3)$$

Note that the ground-state energy is proportional to n^2 in units of g. This indicates a condensation for an attractive interaction that is independent of the volume of the system [see (1.6)]. If we assume that $g < 0$, then the state with $\nu = 0(1)$ is the ground state of an even (odd) number of particles and the excitation energies are given by

$$E(n, \nu) - E(n, 0) = -\tfrac{1}{2}(\Omega + \nu - 2)\nu,$$

for n and ν even, and

$$E(n, \nu) - E(n, 1) = -\tfrac{1}{2}(\Omega + \nu - 1)(\nu - 1),$$

for n and ν odd. Recall that these excitation energies are positive, since they are in units of g which is negative. For $g > 0$, the state with $\nu = n$ is the ground state. However, this is unrealistic, since Eq. (4.1) will only represent the strong coupling limit of a system with a normal single-particle spectrum if the pair energies are proportional to g as $|g| \to \infty$, and this is not possible due to the bounding of the pair energies given in Eq. (2.35).

In order to study the structure of Eqs. (4.1), we construct a polynomial whose roots are the N pair energies E_α. This is done by first considering the symmetric functions of the reciprocals of the pair energies defined by

$$S_m = \sum_{\alpha_1 \cdots \alpha_m}{}' \frac{1}{E_{\alpha_1} \cdots E_{\alpha_m}}, \quad m = 1 \cdots N, \quad (4.4)$$

where the primed sum is over all values of $\alpha_1 \cdots \alpha_m$, each one ranging from 1 to N, such that no two α's are equal. In terms of these functions, the pair energies are the N roots of the single equation

$$\sum_{m=0}^{N} \frac{1}{m!} S_m \left(-\frac{1}{x}\right)^{N-m} = \left(-\frac{1}{x}\right)^{N} \sum_{m=0}^{N} \frac{1}{m!} S_m (-x)^m = 0, \quad (4.5)$$

where $S_0 \equiv 1$. Using Eq. (4.1), we can obtain a recursion relation between S_m and S_{m-1}, which, when solved, shows that the left-hand side of (4.5) is proportional to the Laguerre polynomial $L_N^{(a)}(x/2)$, with $a = \frac{1}{2}\omega - 1$. The study of Eqs. (4.1) is thus reduced to the study of the zeros of the Laguerre polynomials, the properties of which are well known to mathematicians.[11] We now turn to the construction of the symmetric functions S_m from Eqs. (4.1).

We simplify our notation for S_m by letting the index i stand for α_i in sums such as (4.4). With this shorthand, (4.4) becomes

$$S_m = \sum_{1 \cdots m}{}' \frac{1}{E_1 \cdots E_m}. \quad (4.6)$$

The first symmetric function S_1 may be easily evaluated by summing (4.1) on α, with the result that

$$S_1 = \frac{N}{\omega}. \quad (4.7)$$

[11] G. Szego, *Orthogonal Polynomials* (American Mathematical Society, New York, 1939); Bateman Manuscript Project, *Higher Transcendental Functions* (McGraw-Hill Book Co., New York, 1953).

In order to evaluate S_m, for $m > 1$, we write (4.6) as

$$S_m = \left(\sum_{1\cdots m-1}' \frac{1}{E_1\cdots E_{m-1}}\right)\left(\sum_m \frac{1}{E_m}\right) - (m-1)\sum_{1\cdots m-1}' \frac{1}{E_1\cdots E_{m-2}E_{m-1}^2} = S_{m-1}S_1 - (m-1)\sum_{1\cdots m-1}' \frac{1}{E_1\cdots E_{m-2}E_{m-1}^2}. \tag{4.8}$$

The second term in (4.8) may be evaluated by using (4.1) for one of the factors $1/E_{m-1}$ in the summand, i.e.,

$$\sum_{1\cdots m-1}' \frac{1}{E_1\cdots E_{m-2}E_{m-1}^2} = \frac{1}{\omega}\sum_{1\cdots m-1}' \frac{1}{E_1\cdots E_{m-1}}\left(1 + 4\sum_{m\neq m-1} \frac{1}{E_m - E_{m-1}}\right) = \frac{1}{\omega}S_{m-1} + \frac{4}{\omega}\sum_{1\cdots m-1}' \sum_{m\neq m-1} \frac{1}{E_1\cdots E_{m-1}(E_m - E_{m-1})}. \tag{4.9}$$

The second term in this expression may be evaluated by separating it into two parts. The first part includes those terms in the double sum for which the index m, i.e., α_m, does not equal any one of the indices $1\cdots m-1$ and the second part is made up of the remaining terms in which m equals one of the indices $1\cdots m-2$. With this separation, we have

$$\sum_{1\cdots m-1}' \sum_{m\neq m-1} \frac{1}{E_1\cdots E_{m-1}(E_m - E_{m-1})} = \sum_{1\cdots m}' \frac{1}{E_1\cdots E_{m-1}(E_m - E_{m-1})} + (m-2)\sum_{1\cdots m-1}' \frac{1}{E_1\cdots E_{m-2}E_{m-1}(E_{m-2} - E_{m-1})}. \tag{4.10}$$

The second term in (4.10) vanishes because the summand is antisymmetric in the indices $m-2$ and $m-1$. The summand in the first term of (4.10) can be symmetrized in the summation indices $m-1$ and m, yielding

$$\sum_{1\cdots m}' \frac{1}{E_1\cdots E_{m-1}(E_m - E_{m-1})} = \frac{1}{2}\sum_{1\cdots m}' \frac{1}{E_1\cdots E_{m-2}(E_m - E_{m-1})}\left(\frac{1}{E_{m-1}} - \frac{1}{E_m}\right) = \tfrac{1}{2}S_m. \tag{4.11}$$

Substituting (4.9), (4.10), and (4.11) into (4.8), we have

$$S_m = \frac{\omega S_1 - m + 1}{\omega + 2m - 2}S_{m-1} = \frac{N - m + 1}{\omega + 2m - 2}S_{m-1}, \tag{4.12}$$

where we have used (4.7) for S_1. Iterating (4.12), we then obtain

$$S_m = (-\tfrac{1}{2})^m \frac{(-N)_m}{\left(\frac{\omega}{2}\right)_m}, \tag{4.13}$$

where $(a)_m$ is defined in terms of Γ functions as

$$(a)_m = \frac{\Gamma(a+m)}{\Gamma(a)} = a(a+1)\cdots(a+m-1). \tag{4.14}$$

We therefore have obtained an explicit expression for the symmetric functions of the reciprocals of the pair energies for this model.

Substituting (4.13) into (4.5), we obtain the polynomial equation for the pair energies,

$$\sum_{m=0}^{N} \frac{(-N)_m}{m!\left(\frac{\omega}{2}\right)_m}\left(\frac{x}{2}\right)^m = {}_1F_1\left(-N;\frac{\omega}{2};\frac{x}{2}\right) = 0, \tag{4.15}$$

where we have written the sum as a confluent hypergeometric function. However, the confluent hypergeometric function in (4.15) is proportional to the Laguerre polynomial $L_N^{(a)}(x/2)$, with $a = (\omega/2) - 1$. Thus, the pair energies are proportional to the zeros of $L_N^{(a)}$ and as such they satisfy two conditions[11] that are of interest in the present context. The first is that the pair energies are real, positive, and distinct. This proves Eq. (2.31) and fills in the gap in our proof that the pair energies are real and distinct for an arbitrary single-particle spectrum and for all values of g. The second condition establishes upper and lower bounds on the values of the pair energies. For if the pair energies are ordered so that $E_\alpha < E_{\alpha+1}$, for $\alpha = 1\cdots N-1$, then they are bounded by

$$\frac{2j_\alpha^2}{4N+\omega} < E_\alpha < \frac{4\alpha+\omega}{4N+\omega} \times \{4\alpha + \omega + [(4\alpha+\omega)^2 + 1 - (\omega-2)^2]^{\frac{1}{2}}\}, \tag{4.16}$$

where $j_1\cdots j_N$ are the first N positive zeros of the Bessel function $J_{\frac{1}{2}\omega-1}(x)$. For large α, we have $j_\alpha \simeq (\pi\alpha)^2$, and (4.16) becomes

$$\frac{2\pi^2\alpha^2}{4N+\omega} < E_\alpha < \frac{32\alpha^2}{4N+\omega}, \quad \alpha \gg 1. \tag{4.17}$$

In the particular cases $\omega = 1$ or 3, the zeros of the Bessel function $J_{\frac{1}{2}\omega-1}$ are those of the trigonometric functions and (4.16) becomes

$$\frac{2\pi^2(\alpha - \frac{1}{2})^2}{4N+1} < E_\alpha < \frac{2(4\alpha+1)^2}{4N+1}, \qquad (4.18)$$

for $\omega = 1$, and

$$\frac{2\pi^2\alpha^2}{4N+3} < E_\alpha < \frac{2(4\alpha+3)^2}{4N+3}, \qquad (4.19)$$

for $\omega = 3$. These two special cases are important since they correspond to an even or an odd number of particles in a nondegenerate level such as the $\mathbf{k} = 0$ level. Equations (4.19) and (4.20) thus give bounds on the first-order perturbation theory expressions for the pair energies, in units of g, for the ground state of an even or odd number of bosons. They also indicate that this perturbation theory result is accurate as long as $2N|g|$, the upper bound on the largest pair energy $|E_m|$, is small compared to twice the single-particle level spacing, $2(\epsilon_1 - \epsilon_0)$.

B. Two-Level Kinetic Energy Model

Another special case of Eqs. (2.29) that can be reduced to the study of the zeros of a classical polynomial is the strong repulsive-interaction limit of the idealized system, in which the single-particle kinetic energy takes on only two values. The polynomials in this case will be shown to be the Jacobi polynomials and the methods that we will use to show this are similar to those used in the preceding subsection on the one-level kinetic-energy model. That is, from the equation for the pair energies, we will derive a recursion relation for the symmetric functions S_m defined in (4.4). When the solution of this recursion relation is substituted into (4.5), we will show that this equation becomes an equation for the zeros of a Jacobi polynomial. However, before turning to the construction of the functions S_m, we will first derive explicit expressions for the energies of the states of this system and the occupation probabilities of the two single-particle levels in one of these states. We will show that the excitation energies of the low-lying states have the interesting property that they are proportional to $1/n$ for large n, where n is the total number of particles. The occupation probabilities, which are obtained from the energies by using (3.2), show that, for large n, the particles are evenly distributed over the two levels. In the next section, we will show that these features are present in a system with an arbitrary single-particle spectrum.

For a system with a two-level kinetic energy, Eqs. (2.29) may be written as

$$\frac{1}{g} + 4\sum_\beta{}' \frac{1}{E_\beta - E_\alpha} + \frac{\omega_0}{2\epsilon_0 - E_\alpha} + \frac{\omega_1}{2\epsilon_1 - E_\alpha} = 0, \quad \alpha = 1 \cdots N, \qquad (4.20)$$

where $\omega_0 = \Omega_0 + 2\nu_0$, $\omega_1 = \Omega_1 + 2\nu_1$, and $N = \frac{1}{2}(n - \nu)$. In order to keep the model as physically reasonable as possible, we will assume that $\Omega_0 = 1$ and that $\Omega_1 = \Omega$ is arbitrary, although this specialization is not at all necessary for the analysis of the model. We are going to consider the solutions of (4.20) in the limit $g \to +\infty$. These solutions fall into two classes. In the first class the states are labeled with the quantum numbers $I_{\alpha 0} = 0$, $\alpha = 1 \cdots N$. According to (2.36), the pair energies for these states lie in a bounded interval and satisfy $2\epsilon_0 < E_\alpha < 2\epsilon_1$, $\alpha = 1 \cdots N$, for all positive values of g. For the states in the second class, we have $I_{\alpha 0} = 1$ for some values of α. The corresponding pair energies lie in the unbounded interval $2\epsilon_1 < E_\alpha < \infty$ and it is easily seen that they are proportional to g in the limit $g \to \infty$. We shall only consider the states of the first class here since they are the only states with a finite excitation energy in this limit. The states of the second class could be treated by a combination of the methods used in this and the preceding subsection. However, we shall not do that here. Since the pair energies all approach finite limits as $g \to \infty$ for the states that we are considering, we may neglect the term $1/g$ in (4.20). In order further to simplify the equations, we introduce the dimensionless quantities x_α defined by

$$E_\alpha = 2\epsilon_0 + (\epsilon_1 - \epsilon_0)x_\alpha, \quad \alpha = 1 \cdots N. \qquad (4.21)$$

In terms of these new quantities, we have the equations

$$4\sum_\beta{}' \frac{1}{x_\beta - x_\alpha} + \frac{\omega_0}{-x_\alpha} + \frac{\omega_1}{2 - x_\alpha} = 0, \quad \alpha = 1 \cdots N, \qquad (4.22)$$

and we seek roots of these equations on the interval $0 < x_\alpha < 2$. We will first calculate the energies and occupation probabilities for the states described by (4.22) as functions of n, ν_0, and ν_1, after which we will show that the x_α are given in terms of the zeros of certain Jacobi polynomials.

The energies of the states may be obtained from (4.22) using techniques similar to those used in the preceding subsection. As a first step, we sum Eqs. (4.22) over the index α. The first term, being antisymmetric in α and β, does not contribute to the sum and we therefore have the relation

$$\omega_0 \sum_{\alpha=1}^{N} \frac{1}{-x_\alpha} + \omega_1 \sum_{\alpha=1}^{N} \frac{1}{2 - x_\alpha} = 0. \qquad (4.23)$$

For a second relation, we first multiply (4.22) by x_α and then sum over the index α. The contribution of the first term of (4.22) was evaluated in (4.2). The contribution of the second term is $-N\omega_0$ and the contribution of the third term is

$$\omega_1 \sum_{\alpha=1}^{N} \frac{x_\alpha}{2 - x_\alpha} = \omega_1 \sum_{\alpha=1}^{N} \left(-1 + \frac{2}{2 - x_\alpha}\right) = -N\omega_1 + 2\omega_1 \sum_{\alpha=1}^{N} \frac{1}{2 - x_\alpha}. \quad (4.24)$$

Combining these results and using (4.23), we have

$$\omega_0 \sum_{\alpha=1}^{N} \frac{1}{x_\alpha} = \omega_1 \sum_{\alpha=1}^{N} \frac{1}{2 - x_\alpha} = \tfrac{1}{2}N(2N + \omega_0 + \omega_1 - 2). \quad (4.25)$$

We obtain a third and final relation by multiplying (4.22) by x_α^2 and then summing on α. The contribution of the first term of (4.22) to this sum is

$$4 \sum_{\alpha,\beta}{}' \frac{x_\alpha^2}{x_\beta - x_\alpha} = 2 \sum_{\alpha,\beta}{}' \frac{x_\alpha^2 - x_\beta^2}{x_\beta - x_\alpha} = 4(N - 1)x, \quad (4.26)$$

where we have introduced

$$x = \sum_{\alpha=1}^{N} x_\alpha, \quad (4.27)$$

which is the energy of the paired bosons in the units of (4.21). The contribution of the second term of (4.22) to this sum is $-\omega_0 x$ and the third term yields

$$\omega_1 \sum_{\alpha=1}^{N} \frac{x_\alpha^2}{2 - x_\alpha} = \omega_1 \sum_{\alpha=1}^{N} \left(-x_\alpha - 2 + \frac{2}{2 - x_\alpha}\right) = -\omega_1 x - 2N\omega_1 + N(2N + \omega_0 + \omega_1 - 2), \quad (4.28)$$

where we have used (4.25). Combining these results and solving for x, we have

$$x = \frac{2N(2N + \omega_0 - 2)}{4N + \omega_0 + \omega_1 - 4} \quad (4.29)$$

for the energy of the paired bosons or, returning to the original units of energy, we have

$$\sum_{\alpha=1}^{N} E_\alpha = 2N\epsilon_0 + \frac{2N(2N + \omega_0 - 2)}{4N + \omega_0 + \omega_1 - 4}(\epsilon_1 - \epsilon_0). \quad (4.30)$$

Adding to this the energy of the unpaired bosons, $\nu_0\epsilon_0 + \nu_1\epsilon_1$, and expressing the result in terms of n, ν_0, and ν_1, we then have

$$E(n, \nu_0, \nu_1) = n\epsilon_0 + \left[\nu_1 + \frac{(n - \nu_0 - \nu_1)(n + \nu_0 - \nu_1 - 1)}{2n + \Omega - 3}\right](\epsilon_1 - \epsilon_0) \quad (4.31)$$

for the energies of the states of the system, where we have used $\Omega_0 = 1$ and $\Omega_1 = \Omega$. However, since $\Omega_0 = 1$, ν_0 can only take on the values 0 or 1 [see (2.10)] and this expression is independent of ν_0. Therefore, we may write the energies of these states as

$$E(n, \nu_1) \equiv E(n, \nu_0, \nu_1) = n\epsilon_0 + \left[\nu_1 + \frac{(n - \nu_1)(n - \nu_1 - 1)}{2n + \Omega - 3}\right](\epsilon_1 - \epsilon_0), \quad (4.32)$$

for $\nu_1 = 0, 1, \cdots, n$. If we denote the excitation energy of the state $n\nu_1$ by $e(n, \nu_1)$,

$$e(n, \nu_1) = E(n, \nu_1) - E(n, 0),$$

then we have

$$e(n, \nu_1) = \frac{(\Omega + \nu_1 - 2)\nu_1}{2n + \Omega - 3}, \quad \nu_1 = 1 \cdots n, \quad (4.33)$$

in units of $\epsilon_1 - \epsilon_0$. Thus, for fixed Ω and large n, the excitation energies are proportional to $1/n$ and there are roughly $(2n)^{\frac{1}{2}}$ states with excitation energies less than that of the first excited state of the noninteracting system. In the next section we will see that these properties are characteristic of the interaction and are independent of the single-particle spectrum.

The occupation probabilities of the two levels may be calculated using (3.2). Differentiating (4.32) with respect to ϵ_0 and ϵ_1, we then have

$$n_0 = \frac{(n - \nu_1)(n + \Omega + \nu_1 - 2)}{2n + \Omega - 3} \quad (4.34)$$

and

$$n_1 = \nu_1 + \frac{(n - \nu_1)(n - \nu_1 - 1)}{2n + \Omega - 3} \quad (4.35)$$

for the occupations of the two levels. Note that, for the low-lying levels of a large number of particles, both these occupations are of the order $n/2$. In the next section, we will show that this feature of the states is also independent of the single-particle spectrum.

A polynomial equation for the pair energies may be derived using the same methods that were used for the one-level kinetic-energy model. We first derive and solve a recursion relation for the symmetric functions S_m of (4.6) and then substitute the results into (4.5). The resulting equation indicates that the pair energies can be given in terms of the zeros of certain Jacobi polynomials which, for large N, can be given explicitly from the asymptotic forms of these polynomials.

We can initiate a recursion relation for S_l with S_1

which, using (4.25), is given by

$$S_1 = \frac{N(2N + \omega_0 + \omega_1 - 2)}{2\omega_0}, \tag{4.36}$$

where we are using the dimensionless pair energies x_α rather than the E_α. This recursion relation is obtained by using (4.22) to evaluate the second term in (4.8). For if we solve (4.22) for $1/x_{m-1}$ [recall that we use the notation of (4.6), where $m-1$ stands for α_{m-1}] and substitute the result for one of the two factors in the second term of (4.8), we have

$$\sum_{\cdots m-1}{}' \frac{1}{x_1 \cdots x_{m-2} x_{m-1}^2} = \frac{1}{\omega_0} \sum_{1 \cdots m-1}{}' \frac{1}{x_1 \cdots x_{m-1}} \times \left[\frac{\omega_1}{2 - x_{m-1}} + 4 \sum_{m \neq m-1} \frac{1}{x_m - x_{m-1}} \right]. \tag{4.37}$$

The first term of this expression can be evaluated by performing a partial-fraction expansion of the factor $1/x_{m-1}(2 - x_{m-1})$ with the result

$$\frac{\omega_1}{\omega_0} \sum_{1 \cdots m-1}{}' \frac{1}{x_1 \cdots x_{m-1}(2 - x_{m-1})} = \frac{\omega_1}{2\omega_0} \sum_{1 \cdots m-1}{}' \frac{1}{x_1 \cdots x_{m-2}} \left(\frac{1}{x_{m-1}} + \frac{1}{2 - x_{m-1}} \right) = \frac{\omega_1}{2\omega_0} S_{m-1} + \frac{\omega_1}{2\omega_0} \sum_{1 \cdots m-1}{}' \frac{1}{x_1 \cdots x_{m-2}(2 - x_{m-1})}. \tag{4.38}$$

Next, Eqs. (4.22) are solved for $\omega_1/(2 - x_{m-1})$ and the result substituted into the second term of (4.38), giving

$$\frac{\omega_1}{2\omega_0} \sum_{1 \cdots m-1}{}' \frac{1}{x_1 \cdots x_{m-2}(2 - x_{m-1})} = \frac{1}{2\omega_0} \sum_{1 \cdots m-1}{}' \frac{1}{x_1 \cdots x_{m-2}} \times \left[\frac{\omega_0}{x_{m-1}} - 4 \sum_{m \neq m-1} \frac{1}{x_m - x_{m-1}} \right] = \tfrac{1}{2} S_{m-1} - \frac{2}{\omega_0} \sum_{1 \cdots m-1}{}' \sum_{m \neq m-1} \frac{1}{x_1 \cdots x_{m-2}(x_m - x_{m-1})}. \tag{4.39}$$

We next evaluate the second term in this expression. This may be done by first splitting the sum into two parts—the first part coming from those terms in the sum in which the index m equals one of the indices $1 \cdots m-2$ and the second part coming from the remaining terms in the sum in which the indices $1 \cdots m$ are all distinct. That is, we may write the second term of (4.39) as

$$\frac{2}{\omega_0} \sum_{1 \cdots m-1}{}' \sum_{m \neq m-1} \frac{1}{x_1 \cdots x_{m-2}(x_m - x_{m-1})} = \frac{2(m-2)}{\omega_0} \sum_{1 \cdots m-1}{}' \frac{1}{x_1 \cdots x_{m-2}(x_{m-2} - x_{m-1})} + \frac{2}{\omega_0} \sum_{1 \cdots m}{}' \frac{1}{x_1 \cdots x_{m-2}(x_m - x_{m-1})}. \tag{4.40}$$

The second sum of (4.40) vanishes due to the antisymmetry of the summand in the indices m and $m-1$. The first sum may be evaluated in the same manner as (4.11) with the result

$$\frac{2}{\omega_0} \sum_{1 \cdots m-1}{}' \sum_{m \neq m-1} \frac{1}{x_1 \cdots x_{m-2}(x_m - x_{m-1})} = \frac{-(m-2)}{\omega_0} S_{m-1}. \tag{4.41}$$

Substitution of this result into (4.39) and that into (4.38) yields

$$\frac{\omega_1}{\omega_0} \sum_{1 \cdots m-1}{}' \frac{1}{x_1 \cdots x_{m-1}(2 - x_{m-1})} = \frac{\omega_0 + \omega_1 + 2m - 4}{2\omega_0} S_{m-1}, \tag{4.42}$$

which completes the evaluation of the first term of (4.37). The second term of (4.37) has been given in (4.11). Combining these two terms, we then have for (4.37)

$$\sum_{1 \cdots m-1}{}' \frac{1}{x_1 \cdots x_{m-2} x_{m-1}^2} = \frac{\omega_0 + \omega_1 + 2m - 4}{2\omega_0} S_{m-1} + \frac{2}{\omega_0} S_m, \tag{4.43}$$

and, substituting this into (4.8), we have

$$S_m = \frac{2\omega_0 S_1 - (m-1)(\omega_0 + \omega_1 + 2m - 4)}{2\omega_0(\omega_0 + 2m - 2)} S_{m-1} = \frac{(N - m + 1)(2N + \omega_0 + \omega_1 + 2m - 4)}{2(\omega_0 + 2m - 2)} S_{m-1}, \tag{4.44}$$

where we have used (4.36) for S_1. The solution of (4.44) is

$$S_m = (-\tfrac{1}{2})^m \frac{(-N)_m \left(N + \frac{\omega_0 + \omega_1}{2} - 1 \right)_m}{\left(\frac{\omega_0}{2} \right)_m}, \tag{4.45}$$

in the notation of (4.14). Substituting this result into (4.5), we then have the pair energies given as the roots

of the equation

$$\sum_{m=0}^{N} \frac{(-N)_m \left(N + \frac{\omega_0 + \omega_1}{2} - 1\right)_m}{m! \left(\frac{\omega_0}{2}\right)_m} \left(\frac{x}{2}\right)^m$$

$$= {}_2F_1\left(-N, N + \frac{\omega_0 + \omega_1}{2} - 1; \frac{\omega_0}{2}; \frac{x}{2}\right) = 0, \quad (4.46)$$

where we have written it in terms of a hypergeometric function. Equation (4.54) can be rewritten in terms of a Jacobi Polynomial as[11]

$$P_N^{(a,b)}(1 - x) = 0, \quad (4.47)$$

where

$$a = \tfrac{1}{2}\omega_0 - 1 = \nu_0 - \tfrac{1}{2} = \pm\tfrac{1}{2} \quad (4.48)$$

and

$$b = \tfrac{1}{2}\omega_1 - 1 = \tfrac{1}{2}\Omega + \nu_1 - 1$$
$$= \tfrac{1}{2}\Omega - 1, \cdots, \tfrac{1}{2}\Omega + n - 1. \quad (4.49)$$

Equation (4.47) is the polynomial equation for the pair energies that we set out to derive.

Equation (4.47) takes on a particularly simple form[11] if we artificially set $b = \pm\frac{1}{2}$ and, since the roots are monotonic functions of b, this sets bounds on the roots for $b = 0$, i.e., $\Omega = 2$ and $\nu_1 = 0$. If we write the roots of (4.47) as $x_\alpha(a, b)$ and order them so that $x_\alpha < x_{\alpha+1}$, then these bounds are given by

$$1 - \cos\frac{2\alpha - 1}{2N + 1}\pi < x_\alpha(-\tfrac{1}{2}, 0) < 1 - \cos\frac{2\alpha - 1}{2N}\pi$$

and

$$1 - \cos\frac{\alpha}{N + 1}\pi < x_\alpha(\tfrac{1}{2}, 0) < 1 - \cos\frac{2\alpha}{2N + 1}\pi, \quad (4.50)$$

for $\alpha = 1 \cdots N$. In the limit of large N, we then have the pair energies given by

$$x_\alpha = 1 - \cos\frac{2\alpha - 1}{2N}\pi. \quad (4.51)$$

The spacing between the pair energies is then given by

$$x_{\alpha+1} - x_\alpha = \frac{\pi}{N}\sin\frac{2\alpha - 1}{2N}\pi, \quad (4.52)$$

which is equal to the reciprocal of the density of the roots. If we approximate this distribution by a continuous one with the same density, we obtain

$$\Delta_N(x) = \frac{N}{\pi[1 - (1 - x)^2]^{\frac{1}{2}}}, \quad 0 < x < 2, \quad (4.53)$$

for the density of roots. In the next section, we will show that similar results may be obtained for an arbitrary single-particle spectrum.

For other values of b, we may use an asymptotic expansion of the Jacobi polynomial to obtain the pair energies in the limit of large N. From such an expansion,[11] we obtain the result

$$x_\alpha(a, b) = 1 - \cos\left(\frac{4\alpha + 2a + 1}{4N + 2a + 2b + 2}\right)\pi + O\left(\frac{1}{N}\right), \quad (4.54)$$

which is valid for those E_α that lie in the fixed interval $\epsilon < E_\alpha < 2 - \epsilon$. Comparing this with (4.51), we see that the qualitative features of the example with $b = 0$ are independent of b.

5. INFINITE SYSTEM WITH A REPULSIVE INTERACTION

We will now solve Eqs. (2.29) in the limit $N \to \infty$ for a repulsive interaction in a system with an arbitrary single-particle spectrum. This is done for an arbitrary value of the volume so that the solution can be evaluated in the thermodynamic limit in which N, $V \to \infty$ in such a way that the density N/V is fixed. The energies of the states are obtained to order $1/N$ and it is shown that, to this order, the excitations behave like a gas of noninteracting quasiparticles. However, as is suggested by the results of the preceding section, the quasiparticle energy associated with the first excited single-particle level is zero to this order and the ground state is barely stable. Corresponding to this zero in the quasiparticle spectrum, the occupations of the lowest two single-particle levels are of order N. Thus, the model exhibits a very special type of generalized Bose condensation[9] in which the particles condense into two single-particle levels. We will first consider the states of the system with the quantum numbers $l_{\alpha 0} = 0$, i.e., those states which correspond to the states of the noninteracting system in which all the paired bosons occupy the $l = 0$ level. These states can be characterized as having roots E_α of Eqs. (2.29) satisfying $2\epsilon_0 < E_\alpha < 2\epsilon_1$. We will then consider states with a finite number of pairs excited out of the $l = 0$ state.

Since the roots of (2.29) will be located between values of $2\epsilon_l$ for a repulsive interaction [see (2.35)], it will never be possible to replace sums over the single-particle states by integrals. Therefore, the analysis of the equations is facilitated by choosing a volume-dependent unit of energy that keeps ϵ_l fixed as the volume varies. We therefore choose the energy $\epsilon_1 - \epsilon_0$ as our unit of energy. For plane-wave states, this unit is $\frac{1}{2}(2\pi/L)^2$, where $V = L^3$ and we have set the mass of the bosons equal to one. We also readjust the zero of energy so that it lies midway between ϵ_0 and ϵ_1. We therefore introduce the dimensionless

quantities x_α and y_l, defined by

$$E_\alpha = (\epsilon_1 + \epsilon_0) + (\epsilon_1 - \epsilon_0)x_\alpha \tag{5.1}$$

and

$$2\epsilon_l = (\epsilon_1 + \epsilon_0) + (\epsilon_1 - \epsilon_0)y_l, \tag{5.2}$$

where, in particular, we have

$$y_0 = -1, \quad y_1 = 1. \tag{5.3}$$

In terms of these quantities, Eqs. (2.29) become

$$\gamma + 4\sum_\beta{}' \frac{1}{x_\beta - x_\alpha} + \sum_l \frac{\omega_l}{y_l - x_\alpha} = 0, \quad \alpha = 1 \cdots N, \tag{5.4}$$

where

$$\gamma = (\epsilon_1 - \epsilon_0)/g \tag{5.5}$$

and the sum on β is from 1 to N while the sum on l is from 0 to M, where $M + 1$ is the total number of single-particle levels included in the Hamiltonian, and

$$\omega_l = \Omega_l + 2v_l. \tag{5.6}$$

For the ground state and those states with $l_{\alpha 0} = 0$, for $\alpha = 1 \cdots N$, we seek roots of (5.4) which satisfy $-1 < x_\alpha < 1$ for a repulsive interaction, $\gamma > 0$. Equations (5.4) are solved for these states by converting it into an integral equation for the density of roots on this interval. This integral equation is accurate to order $1/N$ and can be solved explicitly. For the states which have a finite number of the $l_{\alpha 0}$ different from zero, Eqs. (5.4) can be written as an integral equation for the density of roots with $l_{\alpha 0} = 0$ in the sense of (2.34), which is coupled to a set of algebraic equations for the pair energies corresponding to $l_{\alpha 0} \neq 0$. These equations can also be solved explicitly.

Before turning to the solution of (5.4), we will discuss the range of values of the interaction strength for which our solution will be valid. The discussion will be given in the specific framework of plane-wave single-particle states and the thermodynamic limit will always be taken. Our solution is predicated upon the assumption of a smooth distribution of roots of (5.4) in the interval $(-1, 1)$. However, for $g = 0$, we know that the distribution is $N\delta(x + 1)$, i.e., all the pair energies satisfy $E_\alpha = 2\epsilon_0$. Therefore, we expect our solution to be a strong coupling solution with a distribution similar to (4.53). We therefore need to determine a lower bound on the interaction strength for the validity of our solution. In order to do this, let us consider the volume dependence of (5.4). The volume of the system appears in γ, which, setting $g = G/V$ since we are interested in a lower bound, becomes $\gamma = (\epsilon_1 - \epsilon_0)V/G$ with $\epsilon_1 - \epsilon_0 = \frac{1}{2}(2\pi/L)^2$. It also appears in the limits on the sum on l, since the sum is over all states with $|\mathbf{k}| < K$. We approximate the sum and its volume dependence by its first two terms, which restrict the roots to the interval $(-1, 1)$, plus a constant contribution from the remaining terms, i.e.,

$$\sum_l \frac{\omega_l}{y_l - x_\alpha} \cong \frac{\omega_0}{-1 - x_\alpha} + \frac{\omega_1}{1 - x_\alpha} + 2KL. \tag{5.7}$$

This approximation is valid because it is only the first two terms of the sum on l that depend strongly upon x_α. With this approximation, (5.4) becomes

$$\frac{(\epsilon_1 - \epsilon_0)V}{G} + 2KL + 4\sum_{\beta=1}^{N}{}' \frac{1}{x_\beta - x_\alpha} + \frac{\omega_0}{-1 - x_\alpha} + \frac{\omega_1}{1 - x_\alpha} = 0, \quad \alpha = 1 \cdots N, \tag{5.8}$$

which is (4.20) with a redefinition of g and the zero of energy. If we can neglect the first two terms in (5.8), then we have (4.22), whose solution we have discussed in detail in the preceding section. From this solution, we know that the roots of (5.8), still neglecting the first two terms, are spread over the interval $(-1, 1)$ with spacings of the order of $1/N$. Thus, the dominant terms in the sum on β will be of order N. This leads to the criterion

$$\frac{(\epsilon_1 - \epsilon_0)V}{G} + 2KL \ll 4N = 2n$$

for the validity of the assumption of a continuous distribution of roots. For fixed $\rho = n/V$ and K, this becomes

$$2\rho G \gg (\epsilon_1 - \epsilon_0) \tag{5.9}$$

in the limit $n, V \to \infty$. Since $\epsilon_1 - \epsilon_0 = \frac{1}{2}(2\pi/L)^2$, this is a very weak condition. However, it should be emphasized that our results are not valid for $G \to 0$.

Returning to the solution of (5.4) for the states with $l_{\alpha 0} = 0$, $\alpha = 1 \cdots N$, we introduce the density of roots $\Delta_N(x)$ defined by

$$\Delta_N(x) = \sum_{i=1}^{N} \delta(x - x_i), \tag{5.10}$$

which in the limit $N \to \infty$ we replace by the smooth function $\Delta(x)$, defined by

$$\int_a^b \Delta(x)\, dx = \lim_{N\to\infty} \int_a^b \Delta_N(x)\, dx, \tag{5.11}$$

for any a, b in the interval $(-1, 1)$. The second term in (5.4) can then be written as

$$\sum_{\beta=1}^{N}{}' \frac{1}{x_\beta - x_\alpha} = P\int_{-1}^{1} \frac{\Delta_N(x)\, dx}{x - x_\alpha} \xrightarrow[n\to\infty]{} P\int_{-1}^{1} \frac{\Delta(x)\, dx}{x - x_\alpha}, \tag{5.12}$$

where P indicates that the principal value of the integral is to be taken. Equation (5.12) is just the finite Hilbert transform of the density Δ, and Eq. (5.4) becomes an equation for this Hilbert transform,[12] i.e.,

$$4\pi D(x) = -\gamma - \sum \frac{\omega_l}{y_l - x}, \quad -1 < x < 1, \quad (5.13)$$

where $D(x)$ is defined by

$$D(x) = \frac{1}{\pi} P\int_{-1}^{1} \frac{\Delta(x')\,dx'}{x' - x}. \quad (5.14)$$

In going from the discrete variable x_α to the continuous variable x, we have made an error of order $1/N$, since this is the order of the spacing between the x_α's. This will be verified by comparing the solution of (5.13) with the exact results of Sec. 4. Thus, the problem is reduced to one of inverting this integral equation for Δ. Note that from its definition Δ must satisfy

$$\int_{-1}^{1} \Delta(x)\,dx = N, \quad (5.15)$$

in addition to (5.13).

The general form of (5.13) is known as the airfoil equation, and the inversion of such transforms has been studied by Tricomi,[13] who gives the solution

$$\Delta(x) = \frac{1}{\pi(1 - x^2)^{\frac{1}{2}}}\left[C - P\int_{-1}^{1} \frac{(1 - x'^2)^{\frac{1}{2}} D(x')\,dx'}{x' - x}\right], \quad (5.16)$$

where C is an arbitrary constant. The solution (5.16) is valid when the function $D(x)$ belongs to the class $L^p(-1, 1)$ with $p > \frac{4}{3}$. However, this last requirement is not satisfied by the terms in (5.13) with $l = 0$ and 1, which have poles at -1 and $+1$, respectively. In order to remove these poles from the interval $(-1, 1)$, we introduce small shifts in the corresponding single-particle energies that move the poles out of this interval. That is, instead of (5.3), we let

$$y_0 = -1 - \delta, \quad y_1 = 1 + \delta, \quad (5.17)$$

and, in all calculated quantities, we will take the limit $\delta \to 0$. The strong coupling assumption (5.9) is necessary for justifying the introduction of this shift. For it is certainly not possible to introduce such a shift at zero interaction strength when all the particles occupy the $l = 0$ level. However, when the roots are spread out over the interval $(-1, 1)$ and no root has the value y_0 or y_1, then the solution of (5.14) using (5.17) will be a good approximation to the equation using (5.3). Furthermore, when we calculate the energies of the states and their occupation probabilities, we may take the limit $\delta \to 0$. These results, when specialized to the two-level model of Sec. 4B, will be shown to agree with the exact results to order $1/n$. We therefore proceed under the assumption that the interaction is strong enough to justify the use of (5.17) rather than (5.3). With this method of removing the singularities, (5.13) may be substituted into (5.16) and all the integrations performed. The result of this calculation is

$$\Delta(x) = \frac{1}{4\pi(1 - x^2)^{\frac{1}{2}}} \times \left[4N - \gamma x + \sum_l \omega_l\left(1 - \frac{(y_l^2 - 1)^{\frac{1}{2}}}{y_l - x}\right)\right], \quad (5.18)$$

where we have used (5.15) to fix the value of the constant $C = N$.

The energies of the states are given by (2.22), which may be written as

$$E = \sum_l \nu_l \epsilon_l + N(\epsilon_0 + \epsilon_1) + (\epsilon_1 - \epsilon_0)\sum_{\alpha=1}^{N} x_\alpha = \sum_l \nu_l \epsilon_l + N(\epsilon_0 + \epsilon_1) + (\epsilon_1 - \epsilon_0)\int_{-1}^{1} x\Delta(x)\,dx, \quad (5.19)$$

where we have used (5.1) for the pair energies E_α. The integral over x may be evaluated using (5.18), with the result

$$\int_{-1}^{1} x\Delta(x)\,dx = -\frac{\gamma}{8} + \frac{1}{4}\sum_l \omega_l[(y_l^2 - 1)^{\frac{1}{2}} - y_l] \quad (5.20)$$

$$= -\frac{\gamma}{8} + \frac{1}{4}\sum_l (\Omega_l + 2\nu_l)[(y_l^2 - 1)^{\frac{1}{2}} - y_l], \quad (5.21)$$

where we have used (5.6) for ω_l. Substituting this into (5.19), we have

$$E = \sum_l \nu_l \tfrac{1}{2}(y_l^2 - 1)^{\frac{1}{2}}(\epsilon_1 - \epsilon_0) + \tfrac{1}{2}n(\epsilon_1 + \epsilon_0) + \frac{1}{4}\left\{-\frac{\gamma}{2} + \sum_l \Omega_l[(y_l^2 - 1)^{\frac{1}{2}} - y_l]\right\}(\epsilon_1 - \epsilon_0) \quad (5.22)$$

for the energy of the states, where we have used (5.2) for y_l and $n = 2N + \nu$. Note that this expression is a well behaved function of δ and we have therefore set δ equal to zero.

The ground state of an even (odd) number of particles has $\nu = 0$ ($\nu = 1$ with $\nu_0 = 1$). Therefore, the term

$$\sum_l \nu_l \tfrac{1}{2}(y_l^2 - 1)^{\frac{1}{2}}(\epsilon_1 - \epsilon_0) \quad (5.23)$$

[12] The author is indebted to Professor Jerome K. Percus for this observation.

[13] F. G. Tricomi, Quart. J. Math. (Oxford) (2) **2**, 199 (1951).

in (5.22) is the excitation energy of the state. This has the character of the states of a set of noninteracting quasiparticles whose energy is given by $\frac{1}{2}(y_l^2 - 1) \times (\epsilon_1 - \epsilon_0)$. If we set $\epsilon_0 = 0$, this expression may be written as

$$e_l = [\epsilon_l(\epsilon_l - \epsilon_1)]^{\frac{1}{2}}. \tag{5.24}$$

Therefore, the quasiparticle spectrum has two zeros corresponding to the lowest two single-particle levels. This result, which is valid up to terms of order $1/n$, is in agreement with the exact results of the two level model treated in Sec. 4B. For the other single-particle levels, the quasiparticle energies are not qualitatively different from those of the noninteracting particles.

The accuracy of (5.22) may be checked by comparing it with the exact results for the two-level model given in Sec. 4B. From (4.32), we have the exact ground state energy of this model given by

$$\begin{aligned} E &= \tfrac{1}{2}n(\epsilon_1 + \epsilon_0) - \frac{n(\Omega - 1)}{2(2n + \Omega - 3)}(\epsilon_1 - \epsilon_0) \\ &= \tfrac{1}{2}n(\epsilon_1 + \epsilon_0) - \frac{\Omega - 1}{4}(\epsilon_1 - \epsilon_0) + O(1/n). \end{aligned} \tag{5.25}$$

If we set $\gamma = 0$, $\Omega_0 = 1$, and $\Omega_1 = \Omega$ in (5.22), we have

$$E = \tfrac{1}{2}n(\epsilon_1 + \epsilon_0) + \tfrac{1}{4}\{1 - \Omega\}(\epsilon_1 - \epsilon_0), \tag{5.26}$$

which agrees with (5.25) to order $1/n$. This result is in accord with our estimate of the errors in (5.22) to be of order $1/n$. The excitation energies cannot be checked in this way since they are all of order $1/n$ in the two-level model.

The occupation probabilities for the single-particle levels may be calculated using (3.2). In order to use (3.2), we need to write the energy as an explicit function of the single-particle energies ϵ_l. This expression is obtained by substituting the definitions of the various quantities into (5.22) with the result

$$\begin{aligned} E = \frac{n}{2}(\epsilon_1 + \epsilon_0) - \frac{(\epsilon_1 - \epsilon_0)^2}{8g} - \left(\frac{\Omega_1 - \Omega_0}{4}\right)(\epsilon_1 - \epsilon_0) \\ + \frac{1}{4}\sum_{l>1}\{2(\Omega_l + 2\nu_l)[(\epsilon_l - \epsilon_1)(\epsilon_l - \epsilon_0)]^{\frac{1}{2}} \\ - \Omega_l(2\epsilon_l - \epsilon_0 - \epsilon_1)\}. \end{aligned} \tag{5.27}$$

Differentiating this expression with respect to the various single-particle energies according to (3.2), we obtain

$$\begin{aligned} n_0 = \tfrac{1}{2}n + \tfrac{1}{4}\gamma + \tfrac{1}{4}(\Omega_1 - \Omega_0) \\ + \frac{1}{4}\sum_{l>1}\left[\Omega_l - (\Omega_l + 2\nu_l)\left(\frac{\epsilon_l - \epsilon_1}{\epsilon_l}\right)^{\frac{1}{2}}\right], \end{aligned} \tag{5.28}$$

$$\begin{aligned} n_1 = \tfrac{1}{2}n - \tfrac{1}{4}\gamma - \tfrac{1}{4}(\Omega_1 - \Omega_0) \\ + \frac{1}{4}\sum_{l>1}\left[\Omega_l - (\Omega_l + 2\nu_l)\left(\frac{\epsilon_l}{\epsilon_l - \epsilon_1}\right)^{\frac{1}{2}}\right], \end{aligned} \tag{5.29}$$

$$n_l = \frac{1}{2}\left\{\frac{(\Omega_l + 2\nu_l)(\epsilon_l - \tfrac{1}{2}\epsilon_1)}{[\epsilon_l(\epsilon_l - \epsilon_1)]^{\frac{1}{2}}} - \Omega_l\right\}, \quad \text{for} \quad l > 1, \tag{5.30}$$

where we have set $\epsilon_0 = 0$. For plane-wave states in the limit $n, V \to \infty$, these expressions may be simplified to

$$n_{\mathbf{k}} = \frac{1}{2}\left[n + \left(\frac{2\pi^2}{G} + K\right)L\right], \quad \mathbf{k} = 0, \tag{5.31}$$

$$= \frac{1}{12}\left[n - \left(\frac{2\pi^2}{G} + K\right)L\right], \quad k^2 = \left(\frac{2\pi}{L}\right)^2, \tag{5.32}$$

$$= \frac{1}{2}\left\{\frac{m^2 - \frac{1}{2}}{[m^2(m^2 - 1)]^{\frac{1}{2}}} - 1\right\}, \quad k^2 = \left(\frac{2\pi}{L}\right)^2 m^2, \quad m^2 > 1, \tag{5.33}$$

where the interaction strength is G, the range of the interaction in momentum space is K, $V = L^3$, and we have assumed that $v = 0$. These expressions are the occupations of the individual single-particle states since we have divided the previous expressions by the degeneracies Ω_l. Thus, the system exhibits a Bose condensation into the lowest two single-particle levels. Note that the strong coupling requirement (5.9) requires that G be large enough so that the terms $\pm[(2\pi^2/G) + K]L$ in (5.31) and (5.32) be small corrections to the occupations of the lowest two levels. This condition is satisfied for any nonvanishing G.

To show that the quasiparticle interpretation of the excitation energies holds even for states that have pairs excited out of the $l = 0$ level, we calculate the energies of the states with one pair excited out of this level, i.e., those states with $l_{1,0} = \cdots = l_{N-1,0} = 0$ and $l_{N,0} \neq 0$. The method can be generalized to treat any state with a finite number of pairs excited out of the $l = 0$ state. Equations (5.4) for a state with one pair excited out of the $l = 0$ level, may be written as

$$\gamma + 4\sum_{\beta=1}^{N-1}{}' \frac{1}{x_\beta - x_\alpha} + \frac{4}{x_N - x_\alpha} + \sum_l \frac{\omega_l}{y_l - x_\alpha} = 0, \quad \alpha = 1 \cdots N - 1, \tag{5.34}$$

and

$$\gamma + 4\sum_{\beta=1}^{N-1} \frac{1}{x_\beta - x_N} + \sum_l \frac{\omega_l}{y_l - x_N} = 0, \tag{5.35}$$

where $-1 < x_i < 1$, for $i = 1 \cdots N - 1$, and $x_N > 1$. Introducing the density of roots Δ which is defined on the interval $-1 < x < 1$, we may write these

equations as

$$4\pi D(X) = -\gamma - \frac{4}{x_N - x} - \sum_l \frac{\omega_l}{y_l - x}, \quad -1 < x < 1, \quad (5.36)$$

and

$$\gamma + 4\int_{-1}^{1} \frac{\Delta(x)\,dx}{x - x_N} + \sum_l \frac{\omega_l}{y_l - x_N} = 0, \quad x_N > 1, \quad (5.37)$$

where $D(x)$ is again the finite Hilbert transform of $\Delta(x)$ given by (5.14). We may now use (5.16) to solve (5.36) for $\Delta(x)$ subject to the normalization condition

$$\int_{-1}^{1} \Delta(x)\,dx = N - 1. \quad (5.38)$$

The result of this calculation is

$$\Delta(x) = \frac{1}{4\pi(1-x^2)^{\frac{1}{2}}}\left\{4N - \gamma x - \frac{4(x_N^2-1)^{\frac{1}{2}}}{x_N - x} + \sum_l \omega_l\left[1 - \frac{(y_l^2-1)^{\frac{1}{2}}}{y_l - x}\right]\right\}. \quad (5.39)$$

This result may now be used in (5.37) to obtain an equation for x_N. Using (5.39), the second term of (5.37) becomes

$$\begin{aligned} 4\int_{-1}^{1} \frac{\Delta(x)\,dx}{x - x_N} &= -\frac{4x_N}{x_N^2 - 1} + \frac{1}{(x_N^2-1)^{\frac{1}{2}}} \\ &\quad \times \left\{-4N + \gamma x_N + \sum_l \omega_l\left[\frac{(y_l^2-1)^{\frac{1}{2}}}{y_l - x_N} - 1\right]\right\} \\ &\quad - \gamma - \sum_l \frac{\omega_l}{y_l - x_N}, \end{aligned} \quad (5.40)$$

which, when substituted into (5.37), yields the equation

$$-\frac{4x_N}{(x_N^2-1)^{\frac{1}{2}}} + \gamma x_N - 4N + \sum_l \omega_l\left[\frac{(y_l^2-1)^{\frac{1}{2}}}{y_l - x_N} - 1\right] = 0 \quad (5.41)$$

for x_N. The solutions of this equation, to order $1/n$, are just the values

$$x_N = y_{l_0}, \quad l_0 > 1. \quad (5.42)$$

The energies of these states are given by

$$\begin{aligned} E &= \sum_l \nu_l \epsilon_l + N(\epsilon_1 + \epsilon_0) + (\epsilon_1 - \epsilon_0) \\ &\quad \times \left[x_N + \int_{-1}^{1} x\Delta(x)\,dx\right] \\ &= \left[(x_N^2-1)^{\frac{1}{2}} + \sum_l \nu_l \tfrac{1}{2}(y_l^2-1)^{\frac{1}{2}}\right] \\ &\quad \times (\epsilon_1 - \epsilon_0) + \tfrac{1}{2}n(\epsilon_1 + \epsilon_0) \\ &\quad + \frac{1}{4}\left\{-\frac{\gamma}{2} + \sum_l \Omega_l[(y_l^2-1)^{\frac{1}{2}} - y_l]\right\}(\epsilon_1 - \epsilon_0), \end{aligned} \quad (5.43)$$

where we have used (5.39) to perform the integration over x. In view of (5.42), the contribution of the paired particles to the energy of the state is just that of two quasiparticles in the level l_0. Thus, the quasiparticle interpretation of the excited states holds true even for those states with pairs excited out of the $l = 0$ level.

It is interesting to compare the above results with those of the Bogoliubov approximation.[1,2] In this approximation, we neglect all terms in the interaction that contain fewer than two operators associated with the $l = 0$ level. Furthermore, a_0 and $a_0^\dagger$ are replaced by $n_0^{\frac{1}{2}}$, where

$$n_0 = n - \sum_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}, \quad (5.44)$$

in the notation of (2.1). Then, keeping only the leading terms in powers of n and setting $g = G/V$, (2.1) becomes

$$H_B = \tfrac{1}{2}n\rho G + \sum_{\mathbf{k}\neq 0}[(\epsilon_k - \rho G)a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \tfrac{1}{2}\rho G(a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger + a_{-\mathbf{k}}a_{\mathbf{k}})], \quad (5.45)$$

where we have set $\epsilon_0 = 0$ and $\rho = n/V$.

This is a bilinear form in the operators $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^\dagger$ and can be diagonalized by a Bogoliubov uv transformation, with the result that the ground-state energy is given by

$$E_B = \tfrac{1}{2}n\rho G + \frac{1}{2}\sum_{\mathbf{k}\neq 0}\{[\epsilon_k(\epsilon_k - 2\rho G)]^{\frac{1}{2}} - (\epsilon_k - \rho G)\} \quad (5.46)$$

and the excitation energies are those of a system of independent quasiparticles whose energies are given by

$$e_k = [\epsilon_k(\epsilon_k - 2\rho G)]^{\frac{1}{2}}. \quad (5.47)$$

These expressions become complex for $2\rho G > \epsilon_1$, indicating a failure of the approximation. We can therefore say that the Bogoliubov approximation fails unless

$$2\rho G < \epsilon_1, \quad (5.48)$$

a limit that vanishes as the volume increases. It is interesting to note that this upper bound on the strength of the interaction for the validity of the Bogoliubov approximation is precisely the same as the lower bound on G (5.9) for the validity of our calculation. The two calculations are therefore complementary, with one appropriate for weak coupling and the other for strong coupling. The dividing line between these two domains of interaction strengths is essentially at zero interaction, i.e., $G \cong \epsilon_1/2\rho \sim V^{-\frac{2}{3}}$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Isotropic Solutions of the Einstein–Liouville Equations

J. EHLERS,* P. GEREN, AND R. K. SACHS*
The University of Texas, Austin, Texas

(Received 11 July 1966)

The gravitational field generated by a gas whose one-particle distribution function obeys the Liouville equation is examined under the following assumptions: First, the distribution is locally isotropic in momentum space with respect to some world-velocity field; second, if the particles have rest-mass zero, the gas is irrotational. It is shown that the model is then either stationary or a Robertson–Walker model. The time dependence of the radius in the Robertson–Walker models is given in terms of integrals containing the distribution function.

1. INTRODUCTION

In galactic dynamics it is useful to relate the velocity dependence of the stellar distribution function to the spatial configuration of the galaxy and to the galaxy's gravitational field. In this paper we give some analogous general-relativistic results for the very simple case of a locally isotropic distribution function. We have in mind applications to cosmology.

Einstein's gravitational field equation

$$G_{ab} = -T_{ab} \tag{1.1}$$

relates the metric of space–time to the stress-energy–momentum distribution of matter. It is necessary to supplement (1.1) by assumptions about the structure of matter. We must specify the dependence of T_{ab} on the basic matter (or field) variables, and state the nongravitational equations of motion, constitutive equations, etc., which these additional variables are supposed to obey.

The model of matter used in this paper is that of kinetic theory. We imagine space–time contains a system of particles all having the same[1] proper mass m (≥ 0). We think of the metric g_{ab} in (1.1) as the macroscopic gravitational potential generated collectively by all the particles, and we assume that each particle moves as a test particle in this average field except during point collisions. Moreover, we restrict ourselves to two cases: either collisions are completely neglected—Case A; or there is collisional equilibrium (detailed balancing)—Case B.

Let $f(x, p)$ be the one-particle distribution function, defined on the seven-dimensional manifold of pairs (x, p), where x is a space–time point and p a tangent vector at x with $p^2 = -m^2$. [We use the signature $(+++-)$ for g_{ab}.] The function f determines the energy–momentum tensor via the equation

$$T_{ab}(x) = \int_{P_m(x)} p_a p_b f(x, p)\, dP_m; \tag{1.2}$$

here $P_m(x)$ denotes the mass hyperboloid $p^2 = -m^2$ in the tangent space of space–time at x, and dP_m is the Lorentz-invariant measure on $P_m(x)$.

Either Case A or B above implies that f satisfies the Liouville condition[2]

$$\begin{aligned} f[x(s), p(s)] = {} & \text{const along each timelike (if } m > 0) \\ & \text{or lightlike (if } m = 0) \\ & \text{geodesic } \{x(s), p(s)\}. \end{aligned} \tag{1.3}$$

The system of equations (1.1)–(1.3) is the general-relativistic analog of the basic equations of stellar dynamics; (1.1) corresponds to Poisson's equation and (1.3) corresponds to the collisionless Boltzmann equation with gravitational forces.

Equations (1.1)–(1.3) are not independent; either (1.1) or the pair (1.2) and (1.3) imply[2]

$$T^{ab}{}_{;b} = 0. \tag{1.4}$$

Real systems for which Case A above seems to be a reasonable model are the system of galaxies now[3] and the galaxies themselves, considered as systems of stars.[4] Case B, with $m = 0$, may be applicable to the early state of the universe in a big-bang model. In the latter case, pertaining to epochs earlier than 10^3 years, we may think of a mixture of photons, perhaps neutrinos and even gravitons, and some electrons and nucleons, with most of the energy due to rest-mass zero or to ultrarelativistic particles. For photons the

* Research supported by Aerospace Research Labs., OAR, AF-33 (615) 1029.

[1] The assumption of equal masses could easily be relaxed; it is made here for simplicity and because of the special role played by a rest-mass-zero gas.

[2] G. E. Tauber and J. W. Weinberg, Phys. Rev. **122**, 1342 (1961).

[3] It is difficult to estimate reliably the relaxation time, but if one uses the usual Newtonian formulas (cf., e.g., Ref. 4) with a cutoff distance $\sim 10^{10}$ light years, one obtains relaxation times which are at least not short compared to the Hubble time.

[4] S. Chandrasekhar, *Principles of Stellar Dynamics* (Dover Publ. Inc., New York, 1960), especially Chap. II; see also the article by L. Woltjer in *Lectures in Applied Mathematics*, J. Ehlers, Ed. (American Mathematical Society, Providence, R.I., 1967), Vol. 9, especially Appendix I.

collisional equilibrium could be catalyzed by the electrons via scattering and free-free transitions; the average time a photon takes to Thomson-scatter at $t = 10^3$ years, assuming a temperature $T \approx 10^5$ °K and a mass density $\rho \approx 10^{-16}$ g/cm^3 (see Ref. 5), is of order 10^{-2} years, and this average collision time decreases rapidly if we consider still earlier epochs.

In this paper we consider those solutions of Eqs. (1.1), (1.2), and (1.3) in which the distribution is everywhere isotropic: There exists a timelike unit-vector field $u^a(x)$ such that $f(x, p)$ is, at any event x, invariant with respect to all those restricted homogeneous Lorentz transformations in the tangent space which leave u^a unchanged. In physical terms, this property means that there exists a preferred state of motion at each event x in the universe, with respect to which the peculiar motions of the particles near x are isotropically distributed. Analytically this means that f has the form $f(x, p) = h(x, -u(x) \cdot p)$. In Case B this isotropy follows from the assumed collisional equilibrium[6]; in Case A it is, of course, an independent assumption.

We show that this assumption (and, in the case $m = 0$, the additional assumption that either the acceleration or the rotation of the mean flow vanishes) leads, without any *a priori* assumptions about the symmetry of space-time, to a Robertson-Walker metric or to stationary space-times. In general-relativistic cosmology (we now have in mind Case A, $m > 0$) the cosmological principle and the Weyl postulate (see, e.g., Ref. 7) can, therefore, both be considered as consequences of the apparently weaker postulate of an isotropic distribution of peculiar velocities. The dependence of the scale factor $a(t)$ of the universe on the distribution function is given [Eq. (4.7)]; this corresponds to the dependence of $a(t)$ on the "equation of state" in hydrodynamical models.

Our result and the method of proof are extensions of the work of Tauber and Weinberg on general relativistic gases (Ref. 2). These authors have determined the restrictions imposed on the metric and the mean flow by the Liouville equation and the condition of isotropy; they did not consider the further restrictions imposed by the Einstein field equation. Because we want to point out the special role of rest-mass zero gases, and also because we need a more detailed description of the case of irrotational flows with expansion than that given in the paper mentioned, we shall rederive some of the relevant results.

[5] R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson, Astrophys. J. **142**, 414 (1965).
[6] K. Bichteler, Z. Physik **182**, 521 (1965).
[7] H. Bondi, *Cosmology* (Cambridge University Press, Cambridge, England, 1961).

2. GEOMETRICAL AND KINEMATICAL PRELIMINARIES

In this section we describe a few properties of congruences of timelike curves in normal hyperbolic Riemannian spaces. We use these properties in the proof of our main theorem.

Let u^a be the normalized tangent vector to a congruence of timelike curves $u_a u^a = -1$. The vector u^a may be interpreted physically as the local average particle world velocity.

The quantities ω_{ab}, σ_{ab}, $\dot{u}_a$, and θ, defined by

$$u_{a;b} = \omega_{ab} + \sigma_{ab} - \dot{u}_a u_b + \tfrac{1}{3}\theta(g_{ab} + u_a u_b), \quad (2.1)$$

$$\omega_{(ab)} = \sigma_{[ab]} = \sigma^a{}_a = 0, \quad \omega_{ab}u^b = \sigma_{ab}u^b = 0, \quad (2.2)$$

are known, respectively, as the angular velocity (or vorticity tensor), the shear velocity, the acceleration, and the expansion velocity of the congruence (see, e.g., Refs. 8 and 9).

We use the brackets ( ) and [ ] for symmetrization and antisymmetrization, respectively, and use throughout the dot to indicate covariant differentiation in the u^a direction, e.g., $\dot{u}_a = u_{a;b}u^b$.

The definitions imply the following lemmas:

Lemma 1: A flow is irrotational, $\omega_{ab} = 0$, if and only if the streamlines are hypersurface-orthogonal, i.e., if and only if there exists a scalar t such that

$$\dot{t}u_a = -t_{,a} \neq 0. \quad (2.3)$$

Lemma 2: The property

$$(\dot{u}_{[a} - \tfrac{1}{3}\theta u_{[a})_{,b]} = 0 \quad (2.4)$$

is necessary and sufficient for the existence of a metric $\bar{g}_{ab}$ conformally related to g_{ab} such that the congruence is geodesic and expansion-free with respect to $\bar{g}_{ab}$; if (2.4) holds, we may put

$$\dot{u}_a - \tfrac{1}{3}\theta u_a = U_{,a}, \quad \bar{g}_{ab} = e^{-2U}g_{ab}. \quad (2.5)$$

The properties discussed in these two lemmas are conformally invariant, that is, they are preserved under transformations

$$\bar{g}_{ab} = W^2 g_{ab}, \quad \bar{u}^a = W^{-1}u^a, \quad (2.6)$$

where W is an arbitrary positive scalar field. The vanishing of shear, $\sigma_{ab} = 0$, is likewise conformally invariant.

[8] J. L. Synge, *Relativity: The General Theory* (North-Holland Publ. Co., Amsterdam, 1960).
[9] J. Ehlers, Akad. Wiss. Lit. (Mainz) Abhandl. Math.-Nat. Kl. **11**, 793 (1961).

By combining the preceding lemmas we obtain further:

Lemma 3: The curves of a congruence are the orbits of a one-dimensional (local) group of conformal mappings of space–time into itself if and only if the congruence is shearfree and satisfies (2.4); if these conditions are satisfied and U is defined by (2.5), $\xi^a = e^U u^a$ generates the group. If, in addition, $\theta = 0$, the mappings are isometries.

We shall now prove:

Lemma 4: If a congruence satisfies $\omega_{ab} = \sigma_{ab} = 0$ and (2.4), then the metric is conformally decomposable; that is, there exist coordinates $(x^a) = (x^\nu, t)$, $\nu = 1, 2, 3$, such that

$$G \overset{\text{DEF}}{=} g_{ab}\,dx^a\,dx^b = e^{2U}\{d\sigma^2 - dt^2\},$$
$$d\sigma^2 = \gamma_{\lambda\mu}(x^\nu)\,dx^\lambda\,dx^\mu, \quad u^a = e^{-U}\delta^a_4. \tag{2.7}$$

In fact, if $\omega_{ab} = \sigma_{ab} = 0$ and (2.4) holds, we find from Lemma 2 that, with respect to $\bar{g}_{ab}$, $\bar{\omega}_{ab} = \bar{\sigma}_{ab} = \dot{\bar{u}}_a = \bar{\theta} = 0$, $\bar{u}^a$ is then covariant-constant with respect to $\bar{g}_{ab}$ by Eq. (2.1), and consequently $\bar{g}_{ab}$ is locally the direct product of a 3-space and a line (see Ref. 10, p. 286), so that g_{ab} can be written as in Eq. (2.7).

Finally, we shall establish two properties of Ricci proper congruences defined by

$$u^a R_{a[b}u_{c]} = 0. \tag{2.8}$$

From the contracted Ricci identity $u^a{}_{;[ab]} = \frac{1}{2}R_{bc}u^c$ and Eq. (2.1), we compute

$$u^a R_{a[b}u_{c]} = \tfrac{2}{3}\theta_{,[b}u_{c]} + \text{terms containing } \omega_{ab} \text{ or } \sigma_{ab}.$$

Hence:

Lemma 5: If a Ricci proper congruence satisfies $\omega_{ab} = \sigma_{ab} = 0$, then its expansion velocity θ is constant on each hypersurface orthogonal to the streamlines, so that

$$\theta = \theta(t) \tag{2.9}$$

with t as in Eq. (2.3).

If we specialize further by combining Lemmas 4 and 5 taking into account that, for the case (2.7), $\theta = 3e^{-U}(\partial U/\partial t)$, we get:

Lemma 6: If an irrotational, shearfree Ricci proper congruence satisfies Eq. (2.4), then coordinates exist such that (2.7) holds with

$$e^{-U} = X(t) + Y(x^\nu). \tag{2.10}$$

[10] J. A. Schouten, *Ricci Calculus* (Springer-Verlag, Berlin, 1954).

3. ISOTROPIC SOLUTIONS OF LIOUVILLE'S EQUATION[11]

We now proceed to analyze Liouville's equation (1.3), ignoring the field equation (1.1) for the moment. We have to find $g_{ab}(x)$, $u_a(x)$, and $h(x, E)$ such that, for a given mass $m \geq 0$, the distribution function

$$f(x, p) = h[x, -u(x)\cdot p]$$

is constant on each geodesic $\{x^a(s), p^a(s)\}$ with $p^a = dx^a/ds$, $p^2 = g_{ab}p^a p^b = -m^2$. Here E is an auxiliary real variable ($E \geq m$) to be interpreted as the energy of a particle with respect to that local frame (with time axis u^a) with respect to which f is isotropic in momentum space.

Since $h(x, E) > 0$ and $h(x, E) \to 0$ as $E \to \infty$ on physical grounds, we know that $h' = \partial h/\partial E \neq 0$ for some open E interval. For E in this interval let us put $h(x, E) = F$ and, for the solution with respect to E, write $E = g(x, F)$. Then Liouville's equation is equivalent to the statement that

$$\frac{dE}{ds} = -\frac{d}{ds}(u_a p^a) = -u_{a;b}p^a p^b = p^a g_{,a} \tag{3.1}$$

on each geodesic, where we define $g_{,a} = \partial g/\partial x^a$ with F fixed. If we split the 4-momentum in the form

$$p^a = Eu^a + (E^2 - m^2)^{\frac{1}{2}}e^a,$$
$$u_a e^a = 0, \quad e_a e^a = 1, \tag{3.2}$$

and insert Eqs. (3.2) and (2.1) into Eq. (3.1), we obtain

$$g\dot{g} + \frac{\theta}{3}(g^2 - m^2) + (g^2 - m^2)^{\frac{1}{2}}(g\dot{u}_a + g_{,a})e^a + (g^2 - m^2)\sigma_{ab}e^a e^b = 0.$$

This equation has to hold identically in the seven independent variables x^a, F, e^a; e^a may be considered as a point on a Euclidean, two-dimensional unit sphere. Hence, since spherical harmonics of different degrees are linearly independent,

$$\sigma_{ab} = 0, \quad \dot{u}_a + (\log g)_{,a} = \alpha u_a, \quad \frac{\theta}{3} = \frac{-g\dot{g}}{g^2 - m^2}. \tag{3.3}$$

The last two of these equations can be replaced by the single relation

$$\dot{u}_a - \tfrac{1}{3}\theta u_a = -(\log g)_{,a} - \frac{m^2\theta}{3g^2}u_a. \tag{3.4}$$

Differentiating this equation with respect to F and inserting the resulting expression for u_a into Eq. (3.4), we obtain

$$\dot{u}_a - \tfrac{1}{3}\theta u_a = -\tfrac{1}{2}(\log gg')_{,a} = U_{,a}, \tag{3.5}$$

[11] For this whole section, compare Ref. 2, Sec. III.

where $g' = \partial g/\partial F$ and $U(x)$ is defined by Eq. (3.5) up to an additive constant.

According to Lemma 3 of Sec. 2, *the congruence associated with an isotropic distribution is conformal* (Tauber and Weinberg, 1962).

The function $g(x, F)$ is related to $U(x)$ by $dU = -\frac{1}{2}d(\log gg')$, d referring to the variables x^a only with F treated as a parameter. Integrating gives

$$e^{2U}(g^2 - m^2) = l(F) - k(x) \tag{3.6}$$

with some functions l and k. But from (3.3) and (3.5)

$$\dot{U} = \frac{\theta}{3}\frac{-(g^2 - m^2)^{\cdot}}{2(g^2 - m^2)};$$

consequently, differentiation of Eq. (3.6) in the u^a direction gives

$$k = 0; \tag{3.7}$$

thus k is constant on each streamline.

Combining Eqs. (3.4) and (3.5), we get a further condition:

$$\frac{m^2\theta}{3}u_a = m^2\dot{U}u_a = -g^2(U + \log g)_{,a}. \tag{3.8}$$

To summarize: Characterizing properties of an isotropic solution of Liouville's equation are Eqs. (3.6), (3.7), (3.8), and the conformal character of the congruence generated by u^a.

According to Eq. (3.8), two possibilities exist:

A. $m\theta = 0$: In this case (3.8) requires that g^2e^{2U} is a function of F only; then the distribution function has the form

$$f(x, p) = j(\xi_a(x)p^a), \tag{3.9}$$

where $\xi^a = e^U u^a$ generates a conformal group and j is some function. If $\theta = 0$, which is necessarily so if $m \neq 0$, the group is an isometry group.

It is well known that Eq. (3.9) gives first integrals for the equations of geodesics; the remarkable fact is that these are the only ones of the form $h(x^a, -u_b(x)\dot{x}^b)$.

The case $\theta = 0$ is not of interest in cosmology, and we shall not consider it in detail.

B. $m\theta \neq 0$: In this case, Eq. (3.8) and Lemma 1 of Sec. 2 show that the congruence must be irrotational; consequently, Lemma 4 applies. Moreover, Eqs. (2.3) and (3.8) show that the preferred time variable t must be related to g and U by

$$ig^2d(U + \log g) = m^2\dot{U}\,dt.$$

Hence, $e^{2U}g^2$ must depend functionally on t and F; this fact, together with Eqs. (3.6) and (3.7), restricts the functional relation to the form

$$e^{2U}g^2 = \bar{l}(F) - q(t) \tag{3.10}$$

with some functions $\bar{l}$, q. The distribution function is therefore

$$f(x, p) = j([\xi_a(x)p^a]^2 + q(t)). \tag{3.11}$$

Using the preferred coordinates of Eq. (2.7), we have, then, the result

$$G = \frac{k(x^\nu) - q(t)}{m^2}[d\sigma^2 - dt^2], \tag{3.12}$$

$$f(x, p) = j\left(\frac{k(x^\nu) - q(t)}{m^2}E^2 + q(t)\right), \tag{3.13}$$

E being the energy of p^a with respect to u^a.

When $m\theta \neq 0$, the irrotationality of the flow follows, as we have seen, from the Liouville equation and the isotropy condition. ω_{ab} might be different from zero if $m\theta = 0$, at least so long as no field equations are imposed. It is, however, of interest to note that, if the flow is geodesic and has expansion, $\dot{u}_a = 0 \neq \theta$, Eq. (3.5) and Lemma 1 show that $\omega_{ab} = 0$. For $m = 0$ and $\theta \neq 0$, we therefore have the subcases A_1: $\dot{u}_a = \omega_{ab} = 0$ and A_2: $\dot{u}_a \neq 0$. In the former, Lemma 4 applies again, and the metric can be written in the form (2.7).

4. SOLUTIONS OF THE FIELD EQUATION FOR ISOTROPIC DISTRIBUTIONS

We now ask which restrictions are imposed on the solutions $\{g_{ab}, f\}$ of Liouville's equation by the field equation (1.1) with the source (1.2). The isotropy of f with respect to u^a implies that

$$T^{ab} = (\mu + p)u^a u^b + pg^{ab}, \tag{4.1}$$

where the mean energy density μ and pressure p can be expressed in terms of f (see below). From (4.1) and (1.1) it is obvious that u^a is an eigenvector of the Ricci tensor, i.e., Eq. (2.8) holds. In Case B of the preceding section and also in Case A, if either $\dot{u}_a = 0$ or $\omega_{ab} = 0$ is assumed, we can apply Lemma 6 of Sec. 1; we then obtain the metric

$$[X(t) + Y(x^\nu)]^{-2}[d\sigma^2 - dt^2]. \tag{4.2}$$

In Case B, comparison of this expression with Eq. (3.12) shows that the conformal factor can depend only on t or on x^ν, but not on both variables. Since $\theta \neq 0$, we conclude that $k = \text{const}$, $Y = \text{const}$; hence, without loss of generality, $Y = 0$ in (4.2). The resulting metric satisfies the field equation with (4.1) only if it is a Robertson–Walker metric (see Ref. 12, p. 107)

$$a^2(t)\,d\sigma^2 - dt^2 \tag{4.3}$$

(t is a new time coordinate), where $d\sigma^2$ has constant curvature $\epsilon = \pm 1, 0$. From (3.13) the distribution

[12] P. Jordan, *Schwerkraft und Weltall* (Vieweg and Sohn, Braunschweig, Germany, 1955).

function is then of the form

$$f(x, p) = \frac{1}{4\pi} g(a^2(t)\mathbf{p}^2), \tag{4.4}$$

where

$$\mathbf{p}^2 = (g_{ab} + u_a u_b)p^a p^b$$

is the squared 3-momentum of a particle relative to the preferred local frame defined by u^a, and g is some positive function of a real variable.

From (1.2) and (4.4), introducing $x = a\,|\mathbf{p}|$, we get

$$\mu = a^{-4}\int_0^\infty x^2 g(x^2)(a^2m^2 + x^2)^{\frac{1}{2}}\,dx,$$
$$p = \tfrac{1}{3}a^{-4}\int_0^\infty x^4 g(x^2)(a^2m^2 + x^2)^{-\frac{1}{2}}\,dx. \tag{4.5}$$

These relations imply, as is well known,[2] energy conservation, $(\mu a^3)\cdot + p(a^3)\cdot = 0$, and therefore the only remaining field equation is

$$3a^2(\dot{a}^2 + \epsilon) = \int_0^\infty x^2 g(x^2)(m^2a^2 + x^2)^{\frac{1}{2}}\,dx. \tag{4.6}$$

Since all these universes have, according to Raychaudhuri's theorem, a singular state $a = 0$ which we may take as the t origin, the time development of a generalized Friedmann model is determined by the function g, the distribution, through

$$t = \frac{\sqrt{3}}{2}\int_0^{a^2}\left[-3\epsilon u + \int_0^\infty x^2 g(x^2)(m^2u + x^2)^{\frac{1}{2}}dx\right]^{-\frac{1}{2}} du. \tag{4.7}$$

Equations (4.3), (4.4), (4.5), and (4.7) determine completely the model universe in Case B.

We now return to Case A and restrict attention to the subcase $\theta \neq 0$ so that $m = 0$. Since, in this case, $T^a{}_a = 0$ from (1.2), in Eq. (4.1) we have

$$p = \tfrac{1}{3}\mu. \tag{4.8}$$

Independently of kinetic theory, it follows that, for an energy-momentum tensor (4.1) together with (4.8), the conservation law $T^{ab}{}_{;b} = 0$ is equivalent to the relation

$$\dot{u}_a - \tfrac{1}{3}\theta u_a = -\tfrac{1}{4}(\log\mu)_{,a}, \tag{4.9}$$

which implies the conservation law $\dot{\mu} + \frac{4}{3}\mu\theta = 0$. Its geometrical meaning is described in Lemma 2 of Sec. 2. (The quantity whose density is $\mu^{\frac{3}{4}}$ is conserved during the motion. For thermal radiation, this conserved quantity is the entropy.)

Combining (4.9) with the arguments which led to the metric (4.2) [cf. Eqs. (2.5) and (2.10)], we see that in the case $m = 0$ the source quantity μ is related to the conformal factor by

$$\mu = [X(t) + Y(x^\nu)]^4. \tag{4.10}$$

Now we use the "4,4 component" of the field equation (1.1):

$$G_{ab}u^a u^b = -\mu,$$

where the left-hand side can easily be computed from (4.2) by means of the equations for conformal transformations,[13] and the right-hand side is given by Eq. (4.10). We obtain

$$6\left(\frac{dx}{dt}\right)^2 - 2(X + Y)^4 - \bar{R}(X + Y)^2 - 4\Delta Y(X + Y) + 6DY = 0. \tag{4.11}$$

Here $\bar{R}$ is the Ricci scalar of $d\sigma^2$, Δ is the Laplace operator of $d\sigma^2$, and $DY = \gamma^{\lambda\mu}Y_{,\lambda}Y_{,\mu}$. Since $\theta \neq 0$ implies $dX/dt \neq 0$, we can introduce $t' = X(t)$ as a new time variable and write $(dX/dt)^2 = F(t')$. Then (4.11) becomes

$$6F(t') = 2(t' + Y)^4 + \bar{R}(t' + Y)^2 + 4\Delta Y(t' + Y) + 6DY.$$

This equation holds identically in t' and x^ν; the left-hand side is independent of x^ν; therefore, the right-hand side (in particular, the coefficient $8Y$ of t'^3) is independent of x^ν; then $Y = \text{const}$. We absorb Y into $X(t)$ so that $Y = 0$. The further analysis is identical with the one in Case A, following Eq. (4.3), with the specialization $m = 0$ in Eqs. (4.5) to (4.7). Then the models are precisely the Tolman models.

We have proven the following:

Theorem 1: The most general solution of the Einstein–Liouville equations (1.1), (1.2), and (1.3) with an isotropic distribution function for particles with nonvanishing mass is either stationary or a generalized Friedmann model {(4.3), (4.4), (4.7)}; for particles with vanishing mass, the solution is either stationary, or a Tolman model, or nonstationary with $\dot{u}_a \neq 0 \neq \omega_{ab}$.[14]

If one looks at the proof, one recognizes that a result can also be formulated which is independent of kinetic-theory assumptions.

Theorem 2: The only solution of the Einstein field equation (1.1) with a "perfect-radiation" source

$$T^{ab} = \frac{\mu}{3}(4u^a u^b + g^{ab})$$

[13] L. P. Eisenhart, *Riemannian Geometry* (Princeton University Press, Princeton, N.J., 1956); P. Jordan, J. Ehlers, and W. Kundt, Akad. Wiss. Lit. (Mainz) Abhandl. Math.-Nat. Kl. No. 2, 23 (1960).

[14] Whether the last case actually admits solutions is not known at present. Some perturbation calculations suggest this case is empty. Of course stationary solutions are known: see O. Klein, Arkiv Mat. Astr. Fys. 34A, Paper 19 (1947).

in shearfree, irrotational motion is the Tolman universe.

We also note the following:

Corollary: The gravitational field generated by a spherically symmetric "perfect-radiation" source in shearfree motion is either static or the Tolman universe.

In fact, a timelike vector field u^a, invariant under the group O_3 (acting on spacelike spheres), is automatically hypersurface-orthogonal; the gas is then irrotational, and the corollary follows from Theorem 2.

We end this section with a few additional remarks:

(1) Equation (4.5) can be considered as a parameter representation of an "equation of state" $\mu = \varphi(p)$ determined by the distribution g. If $m = 0$, $\mu = 3p$ for all g's.

(2) The original Friedmann universes, i.e., the dust models ($p = 0$), are contained in {(4.3), (4.4), (4.7)} as the limiting case in which

$$g(x^2) = \frac{4M}{m}\frac{\delta(x^2)}{x}, \quad \mu a^3 = M = \text{const};$$

they are the only models without any random particle motions.

(3) For $t \to 0$, and hence $a \to 0$, all the models (except the dust model) behave, according to Eqs. (4.5) and (4.7), asymptotically like a Tolman radiation universe; if a model expands indefinitely, it behaves for $t \to \infty$ and $a \to \infty$ asymptotically like a dust model; more precisely, one has $\mu \sim a^{-3}$ and $p/\mu \sim a^{-2}$.

(4) A Planck distribution

$$f(x, p) = \frac{2}{h^3}\left[\exp\left(\frac{-u_a p^a}{kT}\right) - 1\right]^{-1}$$

is rigorously compatible with (4.4) if $m = 0$ and $T \sim a^{-1}$; an equilibrium distribution for $m > 0$, however, is incompatible with an isotropically expanding universe.[9]

According to Eq. (3.13), the general solution of Liouville's equation in a Robertson–Walker universe has the form $f(x, p) = j(a^2(t)\mathbf{p}^2)$; hence if at $t = t_0$ we have, say, a (relativistic) Boltzmann distribution

$$c \exp\left(\frac{-E}{kT_0}\right) = c \exp\left(\frac{-(m^2 + \mathbf{p}^2)^{\frac{1}{2}}}{kT_0}\right),$$

then we obtain later

$$f(x, p) = c \exp\left(\frac{-1}{kT_0}\left\{m^2 + \left[\frac{a(t)}{a(t_0)}\right]^2 \mathbf{p}^2\right\}^{\frac{1}{2}}\right),$$

which is *not* an exact equilibrium distribution. For $(a(t)/a(t_0))^2\mathbf{p}^2 \ll m^2$ we have, however, approximately

$$f(x, p) \approx c' \exp(\mathbf{p}^2/2mkT),$$

with

$$T = T_0(a(t_0)/a(t))^2,$$

which is a (nonrelativistic) Boltzmann distribution with a temperature $T \sim a^{-2}$ (compare Ref. 5).

5. DISCUSSION

Unfortunately, the result presented cannot be taken to mean that the universe in its earliest stages was necessarily a Friedmann model with detailed balance established by rapid collisions of a gas whose particles have zero or negligible rest mass. There are various difficulties. First, nothing is known as yet about the case where a rest-mass zero gas rotates, not even if time-dependent detailed-balance rotational solutions exist. Second, it is known that in a Friedmann model there are particle horizons.[15] For example, with the parameters mentioned in the Introduction a given particle has had time at $t = 10^3$ years to communicate with only about 10^{14} solar masses of matter. There must be particle horizons in more general models as well; we can hardly suppose that portions of the gas which have not had time to communicate have been able to establish detailed balance. More generally, our equilibrium considerations do not indicate how quickly detailed balance is established, if at all.

[15] W. Rindler, Monthly Notices Roy. Astron. Soc. **116**, 662, 1956.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Asymptotic Behavior of Stieltjes Transforms. II

W. S. WOOLCOCK*
Department of Physics, University College London, London, England

(Received 28 July 1967)

The results of a previous paper concerning the asymptotic behavior of Stieltjes transforms for large $|z|$ are extended to prove theorems which hold uniformly for all directions in the complex plane. Special additional assumptions, which hold for all sufficiently large values of the argument of the function whose transform is taken, are required to obtain these extended results.

I. INTRODUCTION

In an earlier paper[1] we proved several theorems concerning the asymptotic behavior of Stieltjes transforms for large $|z|$, which were shown to hold uniformly in a sector of the complex z plane which does not include the cut in the transform. It was pointed out that such asymptotic properties are important in considering the exact form of dispersion relations for particle-scattering processes. For such considerations it is in fact necessary to have results which hold uniformly for all directions in the complex plane and apply, in particular, to principal-value integrals. The purpose of this paper is to extend the results of Ref. 1 in just this way.

The first task is to *define* the Stieltjes transform on the upper and lower sides of the cut. This leads naturally to the imposition of a Lipschitz condition on the original function at each point. The boundary value of the transform can then be defined in terms of a Cauchy principal-value integral, and the transform thus defined in the cut plane has a continuity property as the cut is approached from above and from below. These ideas will be made precise in Sec. II.

Section III will prove, under two inequivalent sets of conditions, that the transform $F(z)$ approaches zero as $|z|$ approaches infinity in any direction; indeed uniform convergence for all directions will be demonstrated. As one would expect, the conditions on the original function are more restrictive than in Ref. 1. Section IV establishes that, under stricter conditions on the original function, $z^{\alpha}F(z) \to 0$ as $|z| \to \infty$, where $0 < \alpha < 1$. Finally, in Sec. V a bound on $|F(z)|$ is obtained from a bound on the original.

A number of results about the asymptotic behavior of principal-value integrals were given by Hamilton and Woolcock.[2] The results in Secs. III and IV of this paper considerably extend many of the results given in that paper. The original inspiration for these results derives from the late E. C. Titchmarsh. The result of Sec. V is similar to one of Lanz and Prosperi,[3] who derive a bound on $|F(z)|$ by putting conditions on the derivative of the original function.

II. PRELIMINARIES

We take $g(x)$ to be a real-valued function, defined for all $x > 0$, which belongs to $L([a, b])$ for any choice of a, b with $0 < a < b$ and for which the limit $\int_{\to 0} g(x)\,dx$ exists.[4] These conditions will be assumed to hold for every theorem and corollary in the paper and will not be explicitly stated.

If the limit $\int^{\to \infty} g(x)\,dx/x$ exists,[4] the function $F(z)$ may be defined by

$$F(z) = \int_{\to 0}^{\to \infty} \frac{g(t)\,dt}{t - z},$$

for z not belonging to $[0, \infty)$. The Stieltjes transform as usually defined is then $F(-z)$. The function $F(z)$ is an analytic function, regular in the whole z-plane cut along the nonnegative real axis.

We first remove the complication at the lower end by noting that, if

$$\phi(z) = \int_{\to 0}^{a} \frac{g(t)\,dt}{t - z}$$

for z not belonging to $[0, a]$, then

$$z\phi(z) \to -\int_{\to 0}^{a} g(t)\,dt,$$

as $|z| \to \infty$ uniformly for all directions. For, writing

$$\psi(x) = \int_{\to 0}^{x} g(t)\,dt \quad (0 < x \leq a), \quad \psi(0) = 0,$$

we have

$$z\phi(z) + \int_{\to 0}^{a} g(t)\,dt = \frac{a\psi(a)}{a - z} + z\int_0^a \frac{\psi(t)\,dt}{(t - z)^2}.$$

* Present address: Department of Theoretical Physics, Institute of Advanced Studies, Australian National University, GPO Box 4, Canberra, ACT 2600, Australia.

[1] W. S. Woolcock, J. Math. Phys. **8**, 1270 (1967).

[2] J. Hamilton and W. S. Woolcock, Rev. Mod. Phys. **35**, 737 (1963).

[3] L. Lanz and G. M. Prosperi, Nuovo Cimento **33**, 201 (1964).

[4] The use of an arrow to denote that a limit is taken is due to E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Oxford University Press, New York, 1948), 2nd ed., p. 9.

For any z, with $|z| = r > a$,

$$\left| z\int_0^a \frac{\psi(t)\,dt}{(t-z)^2} \right| \leq \frac{r}{(r-a)^2}\int_0^a |\psi(t)|\,dt,$$

and the result follows.

This simple result will mean that results stated for $F(z)$ later in the paper need be proved only for the function

$$f(z) = \int_a^{\to\infty} \frac{g(t)\,dt}{t-z},$$

which is defined when z does not belong to $[a, \infty)$. We put the natural extension of the definition of $f(z)$ to the upper and lower sides of the cut $[a, \infty)$ in the form of a theorem.

Theorem A: Suppose that there exist constants $K > 0$, μ $(1 \geq \mu > 0)$, h $(x_0 - a \geq h > 0)$ such that

$$|g(x) - g(x_0)| \leq K\,|x - x_0|^\mu$$

for all x for which $|x - x_0| \leq h$. Then

(i) $$P\int_a^{\to\infty} \frac{g(t)\,dt}{t-x_0} \quad \text{exists;}$$

(ii) $$\int_a^{\to\infty} \frac{g(t)\,dt}{t-x_0-iy} \to P\int_a^{\to\infty} \frac{g(t)\,dt}{t-x_0} + i\pi g(x_0) \quad \text{as} \quad y \downarrow 0.$$

The proof uses straightforward analysis but lies outside the purpose of this paper.[5] Since $g(x)$ is real-valued,

$$\int_a^{\to\infty} \frac{g(t)\,dt}{t-x_0-iy} \to P\int_a^{\to\infty} \frac{g(t)\,dt}{t-x_0} - i\pi g(x_0) \quad \text{as} \quad y \uparrow 0.$$

If the Lipschitz condition of Theorem A holds uniformly, the conclusions of the theorem can be strengthened. This leads to:

Theorem B: Suppose that for $x_0 > a$ there is an interval $I = [x_0 - h, x_0 + h]$, where $0 < h \leq (x_0 - a)$, such that, for any points x_1, x_2 belonging to I,

$$|g(x_1) - g(x_2)| \leq K\,|x_1 - x_2|^\mu \quad (0 < \mu \leq 1).$$

Then

(i) $$\phi(x) = P\int_a^{\to\infty} \frac{g(t)\,dt}{t-x}$$

[which exists for $(x_0 - h) < x < (x_0 + h)$ by Theorem A] satisfies

$$|\phi(x_1) - \phi(x_2)| \leq K'\,|x_1 - x_2|^{\mu'} \quad (0 < \mu' < \mu),$$

for any two points x_1, x_2 belonging to the interval $[x_0 - h', x_0 + h']$, where $0 < h' < h$[6];

(ii) given $\epsilon > 0$ there exists $\delta > 0$ such that

$$|f(x + iy) - \phi(x_0) - i\pi g(x_0)| < \epsilon$$

for all x, y satisfying $0 < |x - x_0| < \delta$, $0 < y < \delta$.

In the theorems of Secs. III–V it will be possible to find $a > 0$ such that the condition on $g(x)$ in Theorem B holds for *each* $x > a$; the correct definition of $f(z)$ for the upper and lower sides of the cut is therefore clear. Write $z = r\exp(i\theta)$ and distinguish $\theta = 0$ and $\theta = 2\pi$ as two separate possibilities. Thus from now on f is taken as a function of two real variables r, θ which is defined as follows:

$$f(r, \theta) = \int_a^{\to\infty} \frac{g(t)\,dt}{t - re^{i\theta}},$$

for $z = re^{i\theta}$ not belonging to $[a, \infty)$;

$$f(r, 0) = P\int_a^{\to\infty} \frac{g(t)\,dt}{t-r} + i\pi g(r) \quad \text{for} \quad r > a;$$

$$f(r, 2\pi) = P\int_a^{\to\infty} \frac{g(t)\,dt}{t-r} - i\pi g(r) \quad \text{for} \quad r > a.$$

The complete transform $F(r, \theta) = \phi(r, \theta) + f(r, \theta)$ is defined provided $z = re^{i\theta}$ does not belong to $[0, a]$ and satisfies $F(r, \theta) = F^*(r, 2\pi - \theta)$.

We shall prove results concerning the behavior of $F(r, \theta)$ as $r \to \infty$ which hold uniformly for $0 \leq \theta \leq 2\pi$. In view of the reflection property above, only the interval $0 \leq \theta \leq \pi$ need be considered. The corresponding theorems of Ref. 1 will show in each case that the desired asymptotic property holds uniformly for $\beta \leq \theta \leq \pi$, where $0 < \beta < \pi$. Thus we need consider only the interval $0 \leq \theta \leq \beta$, which we denote by S_β. For reasons which will appear later, β is taken to be a fixed angle satisfying $0 < \beta \leq \tan^{-1}\frac{1}{2}$.

III. CONDITIONS UNDER WHICH THE TRANSFORM TENDS TO ZERO

Theorem 1: Suppose that $\int^{\to\infty} g(x)\,dx/x$ exists and that $g(x)$ satisfies *one* of the two following conditions:

(a) There exist constants $K > 0$, $h > 0$, $b > 0$, μ $(1 \geq \mu > 0)$ such that

$$|g(x_2) - g(x_1)| \leq K\,|x_2 - x_1|^\mu$$

for all x_1, x_2 satisfying $x_1, x_2 \geq b$, $|x_2 - x_1| \leq h$.

Further, $g(x) \ln x \to 0$ as $x \to \infty$.

[5] The techniques of the proof will be found, for example, in N. I. Muskheliskvili, *Singular Integral Equations* (P. Noordhoff Ltd., Groningen, The Netherlands, 1953), Chap. 2.

[6] Any positive $\mu' < \mu$ will do; K' will depend on the values of μ', h' chosen.

(b) Given $\epsilon > 0$, there exists $X > 0$ (depending on ϵ) such that

$$\frac{|g(x_2) - g(x_1)|}{x_2 - x_1} < \frac{\epsilon}{x_1}$$

for all x_1, x_2 satisfying $x_2 > x_1 \geq X$.

Then $F(r, \theta) \to 0$ as $r \to \infty$, uniformly for $0 \leq \theta \leq 2\pi$.

Proof: Choose $a > 0$ as in Sec. II and consider $f(r, \theta)$ for $\theta \in S_\beta$. From now on we take $x > a$, where $x = r\cos\theta$, $y = r\sin\theta$. Let

$$f(r, \theta) = (f_1 + f_2 + f_3 + f_4 + f_5 + f_6)(r, \theta),$$

where the range of integration has been divided into the intervals $[a, \Delta]$, $[\Delta, \frac{1}{2}x]$, $[\frac{1}{2}x, x - \eta]$, $[x - \eta, x + \eta]$, $[x + \eta, 2x]$, $[2x, \infty)$, respectively.

(a) Consider first f_2 and f_6. We have

$$\operatorname{Re} f_2(r, \theta) = -\int_\Delta^{x/2} dt\, \frac{g(t)}{t}\, \frac{t(x - t)}{(x - t)^2 + y^2},$$

$$\operatorname{Im} f_2(r, \theta) = y\int_\Delta^{x/2} dt\, \frac{g(t)}{t}\, \frac{t}{(x - t)^2 + y^2},$$

$$\operatorname{Re} f_6(r, \theta) = \int_{2x}^{\to\infty} dt\, \frac{g(t)}{t}\, \frac{t(t - x)}{(t - x)^2 + y^2},$$

$$\operatorname{Im} f_6(r, \theta) = y\int_{2x}^{\to\infty} dt\, \frac{g(t)}{t}\, \frac{t}{(t - x)^2 + y^2}.$$

Now the functions $t(x - t)/[(x - t)^2 + y^2]$ and $t/[(x - t)^2 + y^2]$ are monotonically increasing in $[\Delta, x/2]$, while the functions $t(t - x)/[(t - x)^2 + y^2]$ and $t/[(t - x)^2 + y^2]$ are monotonically decreasing in $[2x, \infty)$, provided $\tan\theta = y/x \leq 1/\sqrt{3}$. Applying the second mean-value theorem for integrals,[7]

$$\operatorname{Re} f_2(r, \theta) = \frac{-(x/2)^2}{(x/2)^2 + y^2}\int_{\xi_1}^{x/2} dt\, \frac{g(t)}{t},$$

$$\operatorname{Im} f_2(r, \theta) = \frac{(x/2)y}{(x/2)^2 + y^2}\int_{\xi_2}^{x/2} dt\, \frac{g(t)}{t},$$

$$\operatorname{Re} f_6(r, \theta) = \frac{2x^2}{x^2 + y^2}\int_{2x}^{\xi_3} dt\, \frac{g(t)}{t},$$

$$\operatorname{Im} f_6(r, \theta) = \frac{2xy}{x^2 + y^2}\int_{2x}^{\xi_4} dt\, \frac{g(t)}{t},$$

where $\Delta < \xi_1, \xi_2 < x/2$, $2x < \xi_3, \xi_4$. Now, since $\int_a^{\to\infty} g(t)\, dt/t$ exists, given $\epsilon > 0$, we can choose Δ so that

$$|f_2(r, \theta)| + |f_6(r, \theta)| < \epsilon/4$$

for all r, θ satisfying $r\cos\theta > 2\Delta$, $\theta \in S_\beta$. *Henceforth* Δ *is fixed*; it is taken to be always > 1. Next,

$$|f_1(r, \theta)| \leq \int_a^\Delta \frac{|g(t)|\, dt}{|t - z|} \leq \frac{1}{(x - \Delta)}\int_a^\Delta |g(t)|\, dt \leq \frac{2}{x}\int_a^\Delta |g(t)|\, dt,$$

since $x \geq 2\Delta$.

Hence $|f_1(r, \theta)| < \epsilon/4$, provided

$$r\cos\theta > \frac{8}{\epsilon}\int_a^\Delta |g(t)|\, dt.$$

Consider now

$$\operatorname{Re} f_4(r, \theta) = P\int_{-\eta}^{\eta} du\, \frac{u(g(x + u) - g(x))}{u^2 + y^2};$$

the principal value need be taken only when $y = 0$. If $\eta \leq h$, we have

$$|\operatorname{Re} f_4(r, \theta)| \leq 2K\int_{-0}^{\eta} du\, \frac{u^{1+\mu}}{u^2 + y^2} \leq \frac{2K\eta^\mu}{\mu}.$$

Therefore by choosing η sufficiently small we can make $|\operatorname{Re} f_4(r, \theta)| < \epsilon/4$, independently of r and θ. *Henceforth* η *is fixed*; it is taken to be always < 1.

Finally, choose X so that

$$|g(t)| \ln t < \frac{\epsilon \ln \Delta}{4(2\ln\Delta + 2\ln\eta^{-1} + \pi)} (= \epsilon', \text{ say})$$

for all $t \geq X$, and suppose that $x > 2X$. Then

$$|f_3(r, \theta)| < \epsilon'\int_{\frac{1}{2}x}^{x-\eta} \frac{dt}{\ln t\, |t - z|} \leq \epsilon'\int_{\frac{1}{2}x}^{x-\eta} \frac{dt}{(x - t)\ln t} < \epsilon'(\ln x/2)^{-1}(\ln x/2 + \ln\eta^{-1}),$$

and similarly

$$|f_5(r, \theta)| < \epsilon'(\ln x)^{-1}(\ln x + \ln\eta^{-1}).$$

Since, for $y \neq 0$,

$$\operatorname{Im} f_4(r, \theta) = y\int_{x-\eta}^{x+\eta} \frac{g(t)\, dt}{(t - x)^2 + y^2},$$

we have

$$|\operatorname{Im} f_4(r, \theta)| < \frac{y\epsilon'}{\ln x/2}\int_{x-\eta}^{x+\eta} \frac{dt}{(t - x)^2 + y^2} = \frac{2\epsilon'}{\ln x/2}\tan^{-1}\frac{\eta}{y} < \frac{\pi\epsilon'}{\ln x/2}.$$

But $\operatorname{Im} f_4(r, 0) = \pi g(r)$, so this inequality extends to $\theta = 0$. Thus, provided $r\cos\theta > 2X$, we have

$$|f_3(r, \theta)| + |\operatorname{Im} f_4(r, \theta)| + |f_5(r, \theta)| < \epsilon/4,$$

since $x > 2\Delta$.

Combining the above estimates, we conclude that $|f(r, \theta)| < \epsilon$ for all (r, θ) such that $\theta \in S_\beta$, $r \geq \sec\beta \max\{2\Delta, 2X, 8/\epsilon\int_a^\Delta |g(t)|\, dt\}$.

[7] See, for example, E. W. Hobson, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series* (Cambridge University Press, New York, 1927), Vol. I, 3rd ed., Sec. 422.

(b) We prove first that $g(x) \to 0$ as $x \to \infty$. Given $\epsilon > 0$, we have, for $x \geq X(\epsilon)$,

$$\left| \int_x^{2x} \frac{g(t) - g(x)}{t}\, dt \right| < \epsilon \int_x^{2x} \frac{(t-x)\, dt}{tx} = \epsilon(1 - \ln 2),$$

and so

$$\int_x^{2x} \frac{g(t) - g(x)}{t}\, dt \to 0 \quad \text{as} \quad x \to \infty.$$

But

$$\int_x^{2x} \frac{g(t)}{t}\, dt \to 0, \quad \text{as} \quad x \to \infty,$$

and

$$\int_x^{2x} \frac{g(t)}{t}\, dt = g(x) \ln 2 + \int_x^{2x} \frac{g(t) - g(x)}{t}\, dt,$$

which establishes the required result.

The estimates of f_2, f_6, and f_1 given in (A) may be taken over to this case, thus fixing Δ. It remains to consider

$$\begin{aligned}(f_3 + f_4 + f_5)(r, \theta) &= g(x) \int_{x/2}^{2x} \frac{dt}{t - re^{i\theta}} + \int_{x/2}^{2x} \frac{g(t) - g(x)}{t - re^{i\theta}}\, dt \\ &= \phi_1(r, \theta) + \phi_2(r, \theta).\end{aligned}$$

The definition of ϕ_1, ϕ_2 for $\theta = 0$ is clear from Sec. II. Now

$$\phi_1(r, \theta) = g(x)\left[\tfrac{1}{2} \ln \left(\frac{x^2 + y^2}{(x/2)^2 + y^2}\right) + i\left(\pi - \tan^{-1} \frac{y}{x} - \tan^{-1} \frac{2y}{x}\right)\right],$$

and so $|\phi_1(r, \theta)| \leq |g(x)|\,(\ln 2 + \pi)$. Since $g(x) \to 0$ as $x \to \infty$, there exists X_1 such that $|\phi_1(r, \theta)| < \epsilon/4$ for all (r, θ) satisfying $x \geq X_1$, $y \geq 0$.

Finally, choose X_2 so that

$$\frac{g(x_2) - g(x_1)}{x_2 - x_1} < \frac{\epsilon}{8x_1} \quad \text{for} \quad x_2 > x_1 \geq X_2.$$

Then, for $x \geq 2X_2$, $y \geq 0$, we have

$$\begin{aligned}&|\phi_2(r, \theta)| \\ &\leq \lim_{\xi \to 0} \left(\int_{x/2}^{x-\xi} \frac{|g(x) - g(t)|}{x - t}\, dt + \int_{x+\xi}^{2x} \frac{|g(t) - g(x)|}{t - x}\, dt \right) \\ &< \frac{\epsilon}{8}\left(\int_{x/2}^{x} \frac{dt}{t} + \frac{1}{x} \int_x^{2x} dt \right) = \frac{\epsilon}{8}(1 + \ln 2) < \frac{\epsilon}{4}.\end{aligned}$$

Hence $|f(r, \theta)| < \epsilon$ for all (r, θ) satisfying $\theta \in S_\beta$,

$$r \geq \sec \beta \max \left\{ 2\Delta, X_1, 2X_2, \frac{8}{\epsilon} \int_a^{\Delta} |g(t)|\, dt \right\}.$$

The theorem is therefore established under both the conditions (a) and (b).

The conditions (a) (with $\mu = 1$, of course) are not equivalent to that of (b). Two simple examples show this. Note first that, if $g(x)$ is differentiable for all sufficiently large x, condition (b) becomes $xg'(x) \to 0$ as $x \to \infty$. Now consider $g(x) = \sin x(\ln x)^{-2}$. Then $\int^{\to\infty} g(x)\, dx/x$ exists and $g(x) \ln x \to 0$ as $x \to \infty$. Further, $g'(x) \to 0$ as $x \to \infty$ and so, if

$$M = \sup \{g'(x) : x \geq 2\},$$

then for $y > x \geq 2$ we have $|g(y) - g(x)|/(y - x) \leq M$. This means that all the conditions in (a) are satisfied, with $\mu = 1$. However, since $xg'(x)$ does not approach 0 as $x \to \infty$, condition (b) is not satisfied. On the other hand, if

$$g(x) = \sin (\ln x)(\ln x)^{-1},$$

$\int^{\to\infty} g(x)\, dx/x$ exists and $xg'(x) \to 0$ as $x \to \infty$ so that (b) is satisfied. But $g(x) \ln x$ does not approach 0 as $x \to \infty$ and so one of the conditions in (a) fails.

We turn now to two corollaries.

Corollary 1: Suppose that $xg(x)$ satisfies the conditions on $g(x)$ in Theorem 1. Then

$$re^{i\theta} F(r, \theta) \to -\int_{\to 0}^{\to\infty} g(t)\, dt \text{ as } r \to \infty$$

uniformly for $0 \leq 0 \leq 2\pi$.

This is an immediate consequence of Theorem 1.

Corollary 2: Let $xg(x) = A + h(x)$, where $h(x)$ satisfies the conditions on $g(x)$ in Theorem 1. Then

$$re^{i\theta} F(r, \theta) = -A(\ln r + i(\theta - \pi)) + \Phi(r, \theta),$$

where

$$\Phi(r, \theta) \to -\int_{\to 0}^{1} g(t)\, dt - \int_1^{\to\infty} \frac{h(t)\, dt}{t}$$

as $r \to \infty$, uniformly for $0 \leq \theta \leq 2\pi$.

Proof: When $z = r \exp(i\theta)$ does not belong to $[a, \infty)$,

$$\begin{aligned}&re^{i\theta} f(r, \theta) \\ &= A \int_a^{\to\infty} dt \left(\frac{1}{t - re^{i\theta}} - \frac{1}{t} \right) + re^{i\theta} \int_a^{\to\infty} \frac{h(t)\, dt}{t(t - re^{i\theta})}.\end{aligned}$$

The extension to $r > a$, $\theta = 0, 2\pi$ is clear. By Corollary 1, the second integral on the right side

$$\to -\int_a^{\to\infty} \frac{h(t)\, dt}{t}$$

as $r \to \infty$, uniformly for $0 \leq \theta \leq 2\pi$. The first integral on the right side is

$$-A[\ln (a^2 + r^2 - 2ar \cos \theta)^{\frac{1}{2}} + i(\theta' - \pi) - \ln a],$$

where $\tan\theta' = r\sin\theta/(r\cos\theta - a)$. For $r > a$, θ' increases monotonically from 0 to 2π as θ increases from 0 to 2π. Now

$$\left|\ln\frac{(a^2 + r^2 - 2ar\cos\theta)^{\frac{1}{2}}}{r}\right| \leq -\ln\left(1 - \frac{a}{r}\right) \to 0$$

as $r \to \infty$, uniformly for $0 \leq \theta \leq 2\pi$, while

$$|\tan(\theta' - \theta)| = a|\sin\theta|/(r - a\cos\theta) \leq a/(r - a) \to 0$$

as $r \to \infty$, uniformly for $0 \leq \theta \leq 2\pi$.

From Sec. II,

$$re^{i\theta}\phi(r, \theta) \to -\int_{\to 0}^{a} g(t)\,dt$$

as $r \to \infty$, uniformly for $0 \leq \theta \leq 2\pi$. Thus

$$re^{i\theta}F(r, \theta) = -A(\ln r + i(\theta - \pi)) + \Phi(r, \theta),$$

where

$$\Phi(r, \theta) \to A\ln a - \int_{\to 0}^{a} g(t)\,dt - \int_{a}^{\to\infty}\frac{h(t)\,dt}{t} = -\int_{\to 0}^{1} g(t)\,dt - \int_{1}^{\to\infty}\frac{h(t)\,dt}{t}$$

as $r \to \infty$, uniformly for $0 \leq \theta \leq 2\pi$.

IV. ANOTHER THEOREM

In this section we show that strengthening the conditions on $g(x)$ leads to the behavior $r^\alpha F(r, \theta) \to 0$ as $r \to \infty$, with $0 < \alpha < 1$.

Theorem 2: Suppose that $g(x) = h(x)/x^\alpha$, where $0 < \alpha < 1$ and $h(x)$ satisfies *one* of the following three conditions.

(a), (b). These are identical with the conditions (a) and (b) on $g(x)$ given in Theorem 1, together with the condition that $\int^{\to\infty} h(x)\,dx/x$ exists.

(c) $h(x) = 1/p(x)$, where:

(i) $p(x) > 0$ for all $x \geq a$ (> 0),

(ii) $p(x)$ is concave[8] in $[a, \infty)$,

(iii) $p(x) \to \infty$ as $x \to \infty$.

Then $r^\alpha F(r, \theta) \to 0$ as $r \to \infty$, uniformly for $0 \leq \theta \leq 2\pi$.

Proof: Choose $a > 0$ as in Sec. II. Since $\alpha < 1$ we need prove only that $r^\alpha f(r, \theta) \to 0$ as $r \to \infty$, uniformly for $\theta \in S_\beta$. Again we always take $x > a$.

[8] This means that for any x_1, x_2 such that $a \leq x_1 < x_2$ and any λ for which $0 \leq \lambda \leq 1$ we have

$$p(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda p(x_1) + (1 - \lambda)p(x_2).$$

See, for example, G. Choquet, *Topology* (Academic Press Inc., New York, 1966), Sec. 16.

(a) The proof is an adaption of that of Theorem 1(a). To deal with f_2 and f_6, note that if $0 < \alpha < 1$ and $0 \leq \theta \leq \tan^{-1}\frac{1}{2}$, the functions

$$\frac{t^{1-\alpha}(x - t)}{(x - t)^2 + y^2} \quad \text{and} \quad \frac{t^{1-\alpha}}{(x - t)^2 + y^2}$$

are monotonically increasing in $[\Delta, x/2]$, while

$$\frac{t^{1-\alpha}(t - x)}{(t - x)^2 + y^2} \quad \text{and} \quad \frac{t^{1-\alpha}}{(t - x)^2 + y^2}$$

are monotonically decreasing in $[2x, \infty)$. Using the second mean-value theorem as before, it follows from the existence of $\int_a^{\to\infty} h(t)\,dt/t$ that, given $\epsilon > 0$, we can choose Δ such that

$$r^\alpha|f_2(r, \theta)| + r^\alpha|f_6(r, \theta)| < \epsilon/4$$

for all (r, θ) satisfying $r\cos\theta > 2\Delta$, $\theta \in S_\beta$ $(0 < \beta \leq \tan^{-1}\frac{1}{2})$. Henceforth Δ is fixed (and is chosen > 1). Next,

$$|f_1(r, \theta)| \leq 2/x\int_a^{\Delta}|g(t)|\,dt$$

and so

$$r^\alpha|f_1(r, \theta)| \leq 2x^{-1+\alpha}(\sec\beta)^\alpha\int_a^{\Delta}|g(t)|\,dt.$$

Thus $r^\alpha|f_1(r, \theta)| < \epsilon/4$ provided $\theta \in S_\beta$ and $r\cos\theta > (8/\epsilon\,(\sec\beta)^\alpha\int_a^{\Delta}|g(t)|\,dt)^{(1-\alpha)^{-1}}$.

Special care is needed in estimating $\mathrm{Re}\,f_4(r, \theta)$. For $x \geq b + h$, $|u| \leq h$,

$$\begin{aligned}|x^\alpha(g(x + u) - g(x))| &\leq |((x + u)^\alpha - x^\alpha)g(x + u)| \\ &\quad + |(x + u)^\alpha g(x + u) - x^\alpha g(x)| \\ &< |g(x + u)|\alpha|u|(x - |u|)^{\alpha-1} + K|u|^\mu.\end{aligned}$$

But $g(x)$ is bounded for $x \geq b$ and so a constant K' can be found so that

$$|x^\alpha(g(x + u) - g(x))| \leq K'|u|^\mu$$

for $x \geq b + h$, $|u| \leq h$. Then estimating as before,

$$r^\alpha|\mathrm{Re}\,f_4(r, \theta)| \leq \frac{2(\sec\beta)^\alpha K'\eta^\mu}{\mu}$$

for $\eta \leq h$, $\theta \in S_\beta$. Thus by choosing η sufficiently small we can make $r^\alpha|\mathrm{Re}\,f_4(r, \theta)| < \epsilon/4$ for $\theta \in S_\beta$. Henceforth η is fixed (and is chosen < 1).

Finally, choose X so that

$$|h(t)\ln t| < \frac{\epsilon\ln\Delta}{4(2\ln\Delta + 2\ln\eta^{-1} + \pi)(\sec\beta)^\alpha}$$

for all $t \geq X$. Then for $r\cos\theta > 2X$, $\theta \in S_\beta$, we find that

$$r^\alpha(|f_3(r, \theta)| + |\mathrm{Im}\,f_4(r, \theta)| + |f_5(r, \theta)|) < \epsilon/4.$$

The proof is completed on combining the above estimates.

(b) The estimates of f_2, f_6, and f_1 in (a) may be used, thus fixing Δ. As in Theorem 1(b), $h(x) \to 0$ as $x \to \infty$. Now consider

$$r^\alpha(f_3 + f_4 + f_5)(r, \theta) = r^\alpha h(x)\int_{x/2}^{2x} \frac{dt}{t^\alpha(t - re^{i\theta})} + r^\alpha \int_{x/2}^{2x} \frac{h(t) - h(x)}{t^\alpha(t - re^{i\theta})}\, dt = \phi_1(r, \theta) + \phi_2(r, \theta).$$

The definition of ϕ_1, ϕ_2 for $\theta = 0$ is clear. Now

$$\operatorname{Re} \phi_1(r, \theta) = r^\alpha h(x)\int_{x/2}^{2x} \frac{(t - x)\, dt}{t^\alpha[(t - x)^2 + y^2]} = \left(\frac{r}{x}\right)^\alpha h(x)\int_{-\frac{1}{2}}^{1} \frac{u\, du}{(u + 1)^\alpha(u^2 + \tan^2\theta)}.$$

For fixed θ, the integral on the right side decreases monotonically as α increases from 0 to 1. For $\alpha = 0$, this integral $\leq \ln 2$ and, for $\alpha = 1$, $\geq -\ln 2$. Thus, for $0 < \alpha < 1$ and $\theta \in S_\beta$,

$$|\operatorname{Re} \phi_1(r, \theta)| < (\sec\beta)^\alpha(\ln 2)\, |h(x)|.$$

Also,

$$|\operatorname{Im} \phi_1(r, \theta)| = \left(\frac{r}{x}\right)^\alpha |h(x)| \tan\theta \int_{-\frac{1}{2}}^{1} \frac{du}{(1 + u)^\alpha(u^2 + \tan^2\theta)} < (\sec\beta)^\alpha \frac{3\pi}{2}\, |h(x)|$$

for $0 < \alpha < 1$ and $\theta \in S_\beta$. Since $h(x) \to 0$ as $x \to \infty$, there exists X_1 such that $|\phi_1(r, \theta)| < \epsilon/4$ for all (r, θ) satisfying $r\cos\theta \geq X_1$, $\theta \in S_\beta$.

Finally,

$$|\phi_2(r, \theta)| \leq \left(\frac{2r}{x}\right)^\alpha \lim_{\xi \to 0} \left(\int_{x/2}^{x-\xi} \frac{|h(x) - h(t)|\, dt}{x - t} + \int_{x+\xi}^{2x} \frac{|h(t) - h(x)|}{t - x}\, dt\right).$$

Now choose X_2 such that

$$\frac{|g(x_2) - g(x_1)|}{x_2 - x_1} < \frac{\epsilon}{8x_1(2\sec\beta)^\alpha},$$

for all x_1, x_2 satisfying $x_2 > x_1 \geq X_2$. Then we have $|\phi_2(r, \theta)| < \epsilon/4$ for $x \geq X_2$, $\theta \in S_\beta$. Combining the above estimates completes the proof.

(c) Since $p(x) > 0$ for all $x \geq a$, it follows that $p(x)$ is a monotonic increasing function of x for $x \geq a$. For, if there exist x_1, x_2 $(x_2 > x_1)$ such that $p(x_2) < p(x_1)$, it follows from the concavity of $p(x)$ that, for $x > x_2$,

$$p(x) \leq p(x_2) + [p(x_2) - p(x_1)](x - x_2)/(x_2 - x_1).$$

But the right side of this inequality becomes negative for $x > [x_2p(x_1) - x_1p(x_2)]/[p(x_1) - p(x_2)]$.

Since $1/p(t)$ is monotonically decreasing for $t \geq a$, the second mean-value theorem gives

$$\operatorname{Re} f_2(r, \theta) = -\int_\Delta^{x/2} \frac{dt(x - t)}{t^\alpha p(t)[(x - t)^2 + y^2]} = -\frac{1}{p(\Delta)x^\alpha}\int_{\Delta/x}^{\xi_1/x} \frac{du(1 - u)}{u^\alpha[(1 - u)^2 + \tan^2\theta]},$$

where $\Delta < \xi_1 < x/2$. Hence for $\theta \in S_\beta$

$$r^\alpha\, |\operatorname{Re} f_2(r, \theta)| < \frac{(\sec\beta)^\alpha}{p(\Delta)}\int_0^{\frac{1}{2}} \frac{du}{u^\alpha(1 - u)}.$$

Similarly we find that for $\theta \in S_\beta$

$$r^\alpha\, |\operatorname{Im} f_2(r, \theta)| < \frac{(\sec\beta)^\alpha \tan\beta}{p(\Delta)}\int_0^{\frac{1}{2}} \frac{du}{u^\alpha(1 - u)^2},$$

$$r^\alpha\, |\operatorname{Re} f_6(r, \theta)| < \frac{(\sec\beta)^\alpha}{p(2x)}\int_2^\infty \frac{du}{u^\alpha(u - 1)},$$

$$r^\alpha\, |\operatorname{Im} f_6(r, \theta)| < \frac{(\sec\beta)^\alpha \tan\beta}{p(2x)}\int_2^\infty \frac{du}{u^\alpha(u - 1)^2}.$$

Since $1/p(x) \to 0$ as $x \to \infty$, we may choose Δ so that

$$r^\alpha\, |f_2(r, \theta)| + r^\alpha\, |f_6(r, \theta)| < \epsilon/4$$

for all r, θ satisfying $r\cos\theta \geq 2\Delta$, $\theta \in S_\beta$. Henceforth Δ is fixed.

Exactly as in (a), $r^\alpha\, |f_1(r, \theta)| < \epsilon/4$ provided $\theta \in S_\beta$ and

$$r\cos\theta > \left(\frac{8}{\epsilon}(\sec\beta)^\alpha \int_a^\Delta \frac{dt}{t^\alpha p(t)}\right)^{(1-\alpha)^{-1}}.$$

Finally, consider

$$r^\alpha(f_3 + f_4 + f_5)(r, \theta) = r^\alpha \int_{x/2}^{2x} \frac{dt}{t^\alpha p(t)(t - re^{i\theta})} = \psi(r, \theta).$$

As in (b), for $0 < \alpha < 1$ and $\theta \in S_\beta$,

$$|\operatorname{Im} \psi(r, \theta)| < (\sec\beta)^\alpha \frac{3\pi}{2}\, \frac{1}{p(x/2)}.$$

Now write

$$\operatorname{Re} \psi(r, \theta) = \frac{r^\alpha}{p(x)}\int_{x/2}^{2x} \frac{dt(t - x)}{t^\alpha[(t - x)^2 + y^2]} - \frac{r^\alpha}{p(x)}\int_{x/2}^{2x} \frac{dt}{t^\alpha}\, \frac{(t - x)^2}{(t - x)^2 + y^2} \times \frac{1}{p(t)}\, \frac{p(t) - p(x)}{t - x}.$$

The first term on the right side is less in absolute value than $(\sec\beta)^\alpha(\ln 2)/p(x)$ for $\theta \in S_\beta$, by the same argument as in (b). To estimate the second term, note that,

since $p(x)$ is concave in $[a, \infty)$, the function $(p(t) - p(x))/(t - x)$, for fixed $x\ (> a)$, is a monotonic decreasing function of t for $a \le t < x$ and for $t > x$ (its left and right limits at x need not be equal, but of course the left limit $\ge$ the right limit). Thus

$$\frac{r^\alpha}{p(x)} \int_{x/2}^{2x} \frac{dt}{t^\alpha} \frac{(t-x)^2}{(t-x)^2 + y^2} \frac{1}{p(t)} \frac{p(t) - p(x)}{t - x}$$

$$\le \frac{r^\alpha}{p(x)p(x/2)} \frac{p(x) - p(x/2)}{x/2} \int_{x/2}^{2x} \frac{dt}{t^\alpha}$$

$$< 3\left(\frac{r}{x}\right)^\alpha \frac{p(x) - p(x/2)}{p(x)p(x/2)} < 3\left(\frac{r}{x}\right)^\alpha \frac{1}{p(x/2)}.$$

Combining the above inequalities,

$$|\psi(r, \theta)| < \left(\frac{r}{x}\right)^\alpha \left(\frac{3\pi}{2} + \ln 2 + 3\right) \frac{1}{p(x/2)},$$

so that we may choose X such that $|\psi(r, \theta)| < \epsilon/2$ for all r, θ satisfying $r \cos\theta \ge X$, $\theta \in S_\beta$. Combining the above estimates completes the proof of (c).

Clearly the conclusion of Theorem 2 holds if $g(x)$ can be written as the sum of a finite number of terms, each of which satisfies one of the conditions (a), (b), or (c). The simplest examples of functions $p(x)$ satisfying the condition (c) are $\ln x$ (with $a > 1$) and $\ln\ln x$ (with $a > e$). This means that we could have $h(x) \ln x \to a$ nonzero constant as $x \to \infty$ in (a).

Corollary: Suppose that $x^\alpha g(x) = A + h(x)$, where A, α are constants, $0 < \alpha < 1$, and $h(x)$ is the sum of a finite number of terms, each of which satisfies one of the conditions (a), (b), or (c) on $h(x)$ in Theorem 2. Then

$$r^\alpha e^{i\alpha\theta} F(r, \theta) \to A\pi(\cot \pi\alpha + i),$$

as $r \to \infty$, uniformly for $0 \le \theta \le 2\pi$.

Proof: For $0 < \theta < 2\pi$,

$$r^\alpha e^{i\alpha\theta} F(r, \theta)$$

$$= r^\alpha e^{i\alpha\theta} A \int_{\to 0}^{\to \infty} \frac{dt}{t^\alpha (t - re^{i\theta})}$$

$$- r^\alpha e^{i\alpha\theta} A \int_{\to 0}^{a} \frac{dt}{t^\alpha (t - re^{i\theta})}$$

$$+ r^\alpha e^{i\alpha\theta} \int_{\to 0}^{a} \frac{g(t)\, dt}{t - re^{i\theta}} + r^\alpha e^{i\alpha\theta} \int_a^{\to \infty} \frac{dt\, h(t)}{t^\alpha (t - re^{i\theta})}.$$

The extension to $\theta = 0, 2\pi$ is clear. The first term on the right side is $\pi A(\cot \pi\alpha + i)$. The result proved early in Sec. II applies to the second and third terms and Theorem 2 to the fourth term.

V. A BOUND ON THE TRANSFORM

The theorem to be proved in this section is similar to one of Lanz and Prosperi.[3]

Theorem 3: Suppose that $g(x) = h(x)/x^\alpha$ $(0 < \alpha \le 1)$, where $h(x)$ satisfies the conditions:

(i) $|h(x)| < A$, a constant, for $x \ge a \ge 1$, and

(ii) there exist constants $K > 0$, $h > 0$, μ $(1 \ge \mu > 0)$ such that

$$|h(x_2) - h(x_1)| \le K |x_2 - x_1|^\mu$$

for all x_1, x_2 satisfying $x_1, x_2 \ge a$, $|x_2 - x_1| \le h$. Then there are constants C, $R\ (> a)$ such that

$$|F(r, \theta)| < Cr^{-\alpha} \ln r$$

for all $r \ge R$, $\theta \in [0, 2\pi]$.

Proof: By Theorem 4 of Ref. 1 and the result of Sec. II it is sufficient to prove that there exist constants C', $R'\ (> a)$ such that

$$|f(r, \theta)| < C' r^{-\alpha} \ln r$$

for all $r \ge R'$, $\theta \in S_\beta$.

Take $x = r\cos\theta \ge (a + h)$ and divide the range of integration into three intervals $[a, x - h]$, $[x - h, x + h]$, and $[x + h, \infty)$, with

$$f(r, \theta) = (f_1 + f_2 + f_3)(r, \theta).$$

As in the proof of Theorem 2(a) there is a constant K' such that

$$|x^\alpha (g(x + u) - g(x))| \le K' |u|^\mu$$

for $x \ge (a + h)$, $|u| \le h$, and so

$$|\mathrm{Re}\, f_2(r, \theta)| \le \frac{2K'h^\mu}{\mu x^\alpha}.$$

For the other estimates we have

$$|\mathrm{Im}\, f_2(r, \theta)| < \pi A/(x - h)^\alpha,$$

$$|f_1(r, \theta)| < A \int_a^{x-h} \frac{dt}{t^\alpha (x - t)} = \frac{A}{x^\alpha} \int_{h/x}^{1-a/x} \frac{du}{u(1-u)^\alpha}$$

$$\le \frac{A}{x^\alpha} \int_{h/x}^{1-a/x} \frac{du}{u(1-u)} < \frac{A}{x^\alpha}\left(\ln\frac{x}{a} + \ln\frac{x}{h}\right),$$

$$|f_3(r, \theta)| < A \int_{x+h}^{\infty} \frac{dt}{t^\alpha (t - x)} = \frac{A}{x^\alpha} \int_{h/x}^{\infty} \frac{du}{u(1+u)^\alpha}$$

$$< \frac{A}{x^\alpha}\left(\int_{h/x}^{1} \frac{du}{u} + \int_1^\infty \frac{du}{u^{1+\mu}}\right) = \frac{A}{x^\alpha}\left(\ln\frac{x}{h} + \alpha^{-1}\right).$$

Thus, for $\theta \in S_\beta$, $r \ge (a + h)\sec\beta$,

$$r^\alpha |f(r, \theta)|/\ln r < (\sec\beta)^\alpha (3A + B/\ln r),$$

where B is a constant. This proves the result.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

One- and Two-Center Expansions of the Breit–Pauli Hamiltonian*

PETER R. FONTANA†
Physics Department, University of Michigan, Ann Arbor, Michigan and Theoretical Chemistry Institute, University of Wisconsin, Madison, Wisconsin

AND

WILLIAM J. MEATH‡
Theoretical Chemistry Institute, University of Wisconsin, Madison, Wisconsin

(Received 17 January 1966)

The orbit–orbit, spin–spin, and spin–orbit Hamiltonians of the Breit–Pauli approximation are expressed in terms of irreducible tensors. One- and two-center expansions are given in a form in which the coordinate variables of the interacting particles are separated. In the one-center expansions of the orbit–orbit and spin–orbit Hamiltonians the use of the gradient formula reduces some of the infinite sums to finite ones. Two-center expansions are discussed in detail for the case of nonoverlapping charge distributions. The angular parts of the matrix elements of these Hamiltonians are evaluated for product wavefunctions.

1. INTRODUCTION

Relativistic effects cause energy splittings and energy shifts in atoms and molecules. They are responsible for certain "forbidden transitions," which are often significant in spectroscopy. These effects also modify the interaction between atoms and molecules at large separations.

The lowest-order relativistic corrections to the energy of a system can be calculated by using the Breit–Pauli Hamiltonian. Corrections of order higher than α^2 (where α is the fine-structure constant) cannot be obtained consistently in this approximation. This Hamiltonian is limited to systems containing nuclei with $Z \ll 137$. However, this does not seem to be a practical limitation for many problems since the valence electrons are shielded by the inner-shell electrons and thus are not appreciably affected by the bare nuclear charges. In long-range force calculations the Breit–Pauli approximation is valid for intermolecular separations less than the wavelength of the characteristic transition in the molecules.[1,2] At larger separations retardation effects become more important and quantum electrodynamics must be used to calculate the higher-order corrections.[3]

In this paper one- and two-center expansions for the orbit–orbit, spin–spin, and spin–orbit Hamiltonians are derived using the algebra of irreducible tensors.[4,5] This technique makes it possible to separate the coordinate variables of the interacting particles. If product wavefunctions are used, then the matrix elements can be evaluated in a straightforward manner.

In the one-center expansions the coefficient involving the radial variables contains an infinite sum. In the case of the orbit–orbit and spin–orbit Hamiltonians, the use of the gradient formula results in a finite sum. This technique has also been used by Blume and Watson[6] for the spin–orbit Hamiltonian.

In the two-center expansions only the expressions for nonoverlapping charge distributions are discussed in detail. The general case, however, can be treated using the same techniques.

For other expansions and integrations of the spin–spin Hamiltonian, see Ref. 7.

2. THE BREIT–PAULI HAMILTONIAN

The following Breit–Pauli Hamiltonian[8] describes the interactions of electrons moving in a nuclear Coulomb field. The operators for the spin and linear momentum of the jth electron are denoted by $\mathbf{s}_j$ and $\mathbf{p}_j = (1/i)\nabla_j$, respectively. All the results are in atomic units (energy in e^2/a_0 units, length in a_0 units where a_0 is the Bohr radius). The vector going from electron k to electron j is $\mathbf{r}_{jk} = \mathbf{r}_j - \mathbf{r}_k$. We use Greek

* This research was supported by the National Aeronautics and Space Administration, Grant NsG-275-62.

† Present address: Physics Department, Oregon State University, Corvallis, Oregon.

‡ Present address: Chemistry Department, University of Western Ontario, London, Ontario, Canada.

[1] E. A. Power and S. Zienau, J. Franklin Inst. **263**, 331 (1957).

[2] W. J. Meath and J. O. Hirschfelder, J. Chem. Phys. **44**, 3197, 3210 (1966).

[3] H. B. L. Casimir and D. Polder, Phys. Rev. **73**, 360 (1948); M. R. Aub, E. A. Power, and S. Zienau, Phil. Mag. **2**, 571 (1957); E. A. Power and S. Zienau, Nuovo Cimento **6**, 7 (1957); I. E. Dzialoshinskii, J. Exptl. Theoret. Phys. **3**, 977 (1957); C. Mavroyannis and M. J. Stephen, Mol. Phys. **5**, 629 (1962).

[4] M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1957).

[5] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (John Wiley & Sons, Inc., New York, 1957).

[6] M. Blume and R. E. Watson, Proc. Roy. Soc. (London) **A270**, 127 (1962).

[7] R. M. Pitzer, C. W. Kern, and W. N. Lipscomb, J. Chem. Phys. **37**, 267 (1962); M. Geller and R. W. Griffith, J. Chem. Phys. **40**, 2309 (1964); D. M. Schrader, J. Chem. Phys. **41**, 3266 (1964).

[8] The starting point for this Hamiltonian is the Breit–Hamiltonian: G. Breit, Phys. Rev. **34**, 553 (1929); **36**, 383 (1930); **39**, 616 (1932).

indices to designate nuclei and Roman indices to represent electrons.

The derivation of the Breit–Pauli Hamiltonian is discussed for a 2-electron atom by Bethe and Salpeter.[9,10] The generalization to a molecular system is given by Hirschfelder, Curtiss, and Bird.[11] The grouping of the terms is similar to the one used by Bethe and Salpeter[9]:

$$H = H_e + \alpha^2 H_{\rm rel}, \tag{2.1}$$

where $\alpha = e^2/\hbar c$ is the fine structure constant,

$$H_e = -\tfrac{1}{2}\sum_j \nabla_j^2 - \sum_{j,\beta}\frac{Z_\beta}{r_{j\beta}} + \sum_{k>j}\frac{1}{r_{jk}} + \sum_{\beta>\alpha}\frac{Z_\alpha Z_\beta}{r_{\alpha\beta}}, \tag{2.2}$$

and

$$H_{\rm rel} = H_{LL} + H_{SS} + H_{SL} + H_P + H_D, \tag{2.3}$$

with

$$H_{LL} = -\frac{1}{2}\sum_{k>j}\frac{1}{r_{jk}^3}[r_{jk}^2\mathbf{p}_j\cdot\mathbf{p}_k + \mathbf{r}_{jk}\cdot(\mathbf{r}_{jk}\cdot\mathbf{p}_j)\mathbf{p}_k], \tag{2.4}$$

$$H_{SS} = \sum_{k>j}\left\{-\frac{8\pi}{3}(\mathbf{s}_j\cdot\mathbf{s}_k)\delta^{(3)}(\mathbf{r}_{jk}) + \frac{1}{r_{jk}^5}[r_{jk}^2\mathbf{s}_j\cdot\mathbf{s}_k - 3(\mathbf{s}_j\cdot\mathbf{r}_{jk})(\mathbf{s}_k\cdot\mathbf{r}_{jk})]\right\}, \tag{2.5}$$

$$H_{SL} = \frac{1}{2}\sum_{\beta,j}\frac{Z_\beta}{r_{j\beta}^3}(\mathbf{r}_{j\beta}\times\mathbf{p}_j)\cdot\mathbf{s}_j - \frac{1}{2}\sum_{k\neq j}\frac{1}{r_{jk}^3}[(\mathbf{r}_{jk}\times\mathbf{p}_j)\cdot\mathbf{s}_j - 2(\mathbf{r}_{jk}\times\mathbf{p}_k)\cdot\mathbf{s}_j], \tag{2.6}$$

$$H_P = -\tfrac{1}{8}\sum_j p_j^4, \tag{2.7}$$

$$H_D = \frac{\pi}{2}\left[\sum_{\beta,j} Z_\beta\delta^{(3)}(\mathbf{r}_{j\beta}) - 2\sum_{k>j}\delta^{(3)}(\mathbf{r}_{jk})\right]. \tag{2.8}$$

Equation (2.2) is the usual nonrelativistic Hamiltonian for the system. Z_α is the nuclear charge of the αth nucleus.

The first term in the relativistic Hamiltonian $H_{\rm rel}$ gives the orbit–orbit interaction corresponding to the classical electromagnetic coupling of the electrons.

The coupling of the spin-magnetic moments is given by H_{SS}. The Fermi contact term involving the delta function gives the behavior of this Hamiltonian when $\mathbf{r}_{jk} = 0$. The second term is only applicable when $\mathbf{r}_{jk} \neq 0$.

H_{SL} represents the spin–orbit magnetic coupling between electrons.

H_P is the relativistic correction due to the variation of mass with velocity.

H_D is a term characteristic of the Dirac theory, which has no simple interpretation.

In the above equations the nuclei are considered fixed (Born–Oppenheimer approximation) and we assume no external electric or magnetic fields.

In order to derive the one- and two-center expansions of the Breit–Pauli Hamiltonian, it is convenient to use the algebra of irreducible spherical tensors.[4,5] This method allows the separation of the variables into product form and permits the application of the Wigner–Eckart theorem[12] in the calculation of matrix elements. The first step in this procedure is to write the various terms in the Breit–Pauli Hamiltonian as contractions of irreducible tensors. To illustrate the method of contraction, we consider H_{LL} specifically, and then state the results for the other relativistic Hamiltonians without derivation.

In the first term of H_{LL} one has to contract $\mathbf{p}_j\cdot\mathbf{p}_k$. This can be done by introducing the following spherical tensor of the arbitrary vector $\mathbf{A}$:

$$T_1^{\pm1}(\mathbf{A}) = \mp\frac{1}{\sqrt{2}}(A_x \pm iA_y); \quad T_1^0 = A_z. \tag{2.9}$$

Then

$$\mathbf{p}_j\cdot\mathbf{p}_k = \sum_{\omega=-1}^{1}(-1)^\omega T_1^\omega(\mathbf{p}_j)T_1^{-\omega}(\mathbf{p}_k). \tag{2.10}$$

The second term of H_{LL} can be written as a double contraction. The first contraction is as follows:

$$\mathbf{r}_{jk}\cdot\mathbf{p}_j = \left(\frac{4\pi}{3}\right)^{\frac{1}{2}}\sum_\omega(-1)^\omega\mathcal{Y}_1^\omega(\mathbf{r}_{jk})T_1^{-\omega}(\mathbf{p}_j), \tag{2.11}$$

where $\mathcal{Y}_1^\omega(\mathbf{r}_{jk})$ is a solid spherical harmonic which in general is defined as[13]

$$\mathcal{Y}_l^m(\mathbf{r}) = r^l Y_l^m(\theta,\varphi). \tag{2.12}$$

Then

$$\mathbf{r}_{jk}\cdot(\mathbf{r}_{jk}\cdot\mathbf{p}_j)\mathbf{p}_k = \frac{4\pi}{3}\sum_{\omega=-1}^{1}\sum_{\eta=-1}^{1}(-1)^{\omega+\eta}\mathcal{Y}_1^\omega(\mathbf{r}_{jk})\mathcal{Y}_1^\eta(\mathbf{r}_{jk})T_1^{-\omega}(\mathbf{p}_j)T_1^{-\eta}(\mathbf{p}_k). \tag{2.13}$$

[9] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Academic Press Inc., New York, 1957), p. 170.

[10] This Hamiltonian has recently been derived using quantum electrodynamics by T. Itoh, Rev. Mod. Phys. **37**, 159 (1965).

[11] J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *The Molecular Theory of Gases and Liquids* (John Wiley & Sons, Inc., New York, 1954), p. 1044.

[12] E. P. Wigner, Z. Phys. **43**, 624 (1927); C. Eckart, Rev. Mod. Phys. **2**, 305 (1930).

[13] The phase convention we use for the $Y_l^m(\theta,\varphi)$ is the same as that used, for example, in E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra* (Cambridge University Press, London, 1935), and in Refs. 4 and 5.

The two solid spherical harmonics can now be coupled together[14]:

$$\mathcal{Y}_1^{\omega}(\mathbf{r}_{jk})\mathcal{Y}_1^{\eta}(\mathbf{r}_{jk}) = r_{jk}^2 \sum_l \frac{3}{[4\pi(2l+1)]^{\frac{1}{2}}} C(11l;\omega,\eta) \times C(11l;00) Y_l^{\omega+\eta}(\theta_{jk},\varphi_{jk}). \quad (2.14)$$

The Clebsch–Gordan coefficient $C(11l;00)$ vanishes unless $(1+1+l)$ is even and l is in the range 0 to 2.

Using Eqs. (2.10)–(2.14), one obtains

$$H_{LL} = \sum_{k>j} \frac{1}{r_{jk}} \left[\begin{array}{l} -\frac{2}{3} \sum_{\omega} (-1)^{\omega} T_1^{\omega}(\mathbf{p}_j) T_1^{-\omega}(\mathbf{p}_k) \\ -\left(\frac{2\pi}{15}\right)^{\frac{1}{2}} \sum_{\omega,\eta} (-1)^{\omega+\eta} C(112;\omega,\eta) \\ \quad \times Y_2^{\omega+\eta}(\theta_{jk},\varphi_{jk}) T_1^{-\omega}(\mathbf{p}_j) T_1^{-\eta}(\mathbf{p}_k) \end{array} \right]. \quad (2.15)$$

The first part of Eq. (2.15) contains a contribution from the $l=0$ term of Eq. (2.14). The Clebsch–Gordan coefficient in Eq. (2.15) can be given in closed form[15]:

$$C(112;\omega\eta) = \left[\frac{(2+\omega+\eta)!\,(2-\omega-\eta)!}{6(1+\omega)!\,(1-\omega)!\,(1+\eta)!\,(1-\eta)!} \right]^{\frac{1}{2}}. \quad (2.16)$$

In a similar fashion the spin–spin Hamiltonian can be contracted to yield[16]

$$H_{SS} = \sum_{k>j} \left[\begin{array}{l} -\frac{8\pi}{3} \delta^{(3)}(\mathbf{r}_{jk}) \sum_{\omega} (-1)^{\omega} T_1^{\omega}(\mathbf{s}_j) T_1^{-\omega}(\mathbf{s}_k) \\ -\left(\frac{24\pi}{5}\right)^{\frac{1}{2}} \frac{1}{r_{jk}^3} \sum_{\omega,\eta} (-1)^{\omega+\eta} C(112;\omega\eta) \\ \quad \times Y_2^{\omega+\eta}(\theta_{jk},\phi_{jk}) T_1^{-\omega}(\mathbf{s}_j) T_1^{-\eta}(\mathbf{s}_k) \end{array} \right]. \quad (2.17)$$

It is sometimes convenient to couple the spins together to form a total spin tensor defined by[4,5]

$$T_l^{-\nu}(\mathbf{s}_j,\mathbf{s}_k) = \sum_{\kappa} C(11l;-\kappa,-\nu+\kappa) T_1^{-\kappa}(\mathbf{s}_j) T_1^{-\nu+\kappa}(\mathbf{s}_k). \quad (2.18)$$

Then the Fermi contact term contracts to a scalar and in the spin-dipole–dipole term the spin transforms like a second-rank tensor:

$$H_{SS} = \sum_{k>j} \left[\begin{array}{l} \frac{8\pi}{\sqrt{3}} \delta^{(3)}(\mathbf{r}_{jk}) T_0^0(\mathbf{s}_j,\mathbf{s}_k) \\ -\left(\frac{24\pi}{5}\right)^{\frac{1}{2}} \frac{1}{r_{jk}^3} \sum_{\nu} (-1)^{\nu} \\ \quad \times Y_2^{\nu}(\theta_{jk},\varphi_{jk}) T_2^{-\nu}(\mathbf{s}_j,\mathbf{s}_k) \end{array} \right]. \quad (2.19)$$

In the spin–orbit Hamiltonian H_{SL} one can first introduce a tensor $T_1^{\omega}(\mathbf{r}\times\mathbf{p})$ to give

$$H_{SL} = \frac{1}{2} \sum_{\beta} \sum_{j} \frac{Z_{\beta}}{r_{j\beta}^3} \sum_{\omega} (-1)^{\omega} T_1^{\omega}(\mathbf{r}_{j\beta}\times\mathbf{p}_j) T_1^{-\omega}(\mathbf{s}_j) - \frac{1}{2} \sum_{k\neq j} \frac{1}{r_{jk}^3} \sum_{\omega} (-1)^{\omega} \times [T_1^{\omega}(\mathbf{r}_{jk}\times\mathbf{p}_j) - 2T_1^{\omega}(\mathbf{r}_{jk}\times\mathbf{p}_k)] T_1^{-\omega}(s_j). \quad (2.20)$$

In the first term of Eq. (2.20), $(\mathbf{r}_{j\beta}\times\mathbf{p}_j)$, is the orbital angular-momentum operator of electron j with respect to nucleus β. The vectors $(\mathbf{r}_{jk}\times\mathbf{p}_j)$ and $(\mathbf{r}_{jk}\times\mathbf{p}_k)$, however, are not angular momentum operators about a fixed center. Here it is convenient to write them as a contraction which separates the position variables from the momentum operator:

$$T_1^{\omega}(\mathbf{r}_{jk}\times\mathbf{p}_j) = \frac{1}{i}\sqrt{\frac{8\pi}{3}} \sum_{\eta=-1}^{1} C(111;\eta,\omega-\eta) \mathcal{Y}_1^{\eta}(\mathbf{r}_{jk}) T_1^{\omega-\eta}(\mathbf{p}_j). \quad (2.21)$$

The Clebsch–Gordan coefficient in Eq. (2.21) is given by

$$C(111;\eta,\omega-\eta) = \frac{1}{\sqrt{8}} \left[\begin{array}{l} \left[\frac{(1+\eta)!\,(1-\omega+\eta)!\,(1+\omega)!\,(1-\omega)!}{(1-\eta)!\,(1+\omega-\eta)!\,[\eta!\,(\eta-\omega)!]^2} \right]^{\frac{1}{2}} \\ -\left[\frac{(1-\eta)!\,(1+\omega-\eta)!\,(1+\omega)!\,(1-\omega)!}{(1+\eta)!\,(1-\omega+\eta)!\,[(1-\eta)!\,(\omega-\eta)!]^2} \right]^{\frac{1}{2}} \end{array} \right]. \quad (2.22)$$

The remaining relativistic Hamiltonians H_P and H_D only involve scalars and thus need not be considered further.

3. ONE-CENTER EXPANSIONS

In general, the origin of the coordinate system is arbitrary. The vectors $\mathbf{r}_{\beta}$ and $\mathbf{r}_j$ denote the position of a nucleus and an electron, respectively.

The derivation of the one-center expansion for H_{LL}, H_{SS}, and H_{SL}, respectively, consists of three steps. First one has to express the $Y_l^{\mu}(\theta_{jk},\varphi_{jk})$ as a sum of products in the spherical harmonics of (θ_j,φ_j) and (θ_k,φ_k). Then $(1/r_{jk})^n$ is expanded in a similar

[14] See, for example, Ref. 4, p. 61.

[15] Closed form expressions for these coefficients are available (see Refs. 4 and 5), and they are tabulated in Ref. 13. The 3-j symbols, which are closely related to the Clebsch–Gordan coefficients, have been tabulated in detail by M. Rotenberg, R. Bivins, N. Metropolis, and J. K. Wooten, Jr., *The 3-j and 6-j Symbols* (The Technology Press, Cambridge, Mass., 1959).

[16] P. R. Fontana, Phys. Rev. **125**, 220 (1962).

manner. Finally the two expansions are coupled together.

The general addition theorem for the solid spherical harmonics is given by Rose[17]:

$$\mathcal{Y}_N^{\mu}(\mathbf{r}_{jk}) = [4\pi(2N+1)!]^{\frac{1}{2}} \times \sum_{L=0}^{N}\sum_{\kappa=-L}^{L}(-1)^L \frac{C(L, N-L, N; \kappa, \mu-\kappa)}{[(2L+1)!\,(2N-2L+1)!]^{\frac{1}{2}}} \times \mathcal{Y}_{N-L}^{\mu-\kappa}(\mathbf{r}_j)\mathcal{Y}_L^{\kappa}(\mathbf{r}_k). \tag{3.1}$$

The one-center expansion for $(1/r_{jk})^n$ can always be written in the form[18]

$$\frac{1}{r_{jk}^n} = 4\pi \sum_{l=0}^{\infty}\sum_{\nu=-l}^{l} \frac{R(-n, l)}{(2l+1)}(-1)^{\nu} Y_l^{\nu}(\theta_j, \varphi_j) Y_l^{-\nu}(\theta_k, \varphi_k), \tag{3.2}$$

where $R(-n, l)$ is a function of r_j and r_k. Only the radial coefficients for $n = 1$, 3, and 5 are required. In the limiting case $r_j = r_k$ the functions $R(-n, l)$ diverge for $n \geq 3$, and one has to introduce a special cutoff in the integrations. Letting $r_j = r_k(1-\epsilon)$ at the limit avoids these difficulties. After integrating and adding up the sums, ϵ can be set equal to zero. If one uses the Laplace expansion, then[18]

$$R(-1, l) = \frac{r_<^l}{r_>^{l+1}}, \tag{3.3}$$

$$R(-3, l) = (2l+1)\sum_{n=0}^{\infty} \frac{r_<^{2n+l}}{r_>^{2n+l+3}}, \tag{3.4}$$

$$R(-5, l) = \frac{(2l+1)}{3}\sum_{n=0}^{\infty}(n+1)(2l+2n+3)\frac{r_<^{2n+l}}{r_>^{2n+l+5}}, \tag{3.5}$$

where $r_>$ and $r_<$ stand for the greater or lesser of r_j and r_k. The coefficients $R(-n, l)$ can also be written symmetrically with respect to r_j and r_k. There are two such expansions; they involve powers of $(r_j^2 + r_k^2)^{\frac{1}{2}}$ and $(r_j + r_k)$, respectively[18,19]:

$$R(-1, l) = (2l+1)\sum_n \frac{(2n-1)!!\, r_j^n r_k^n}{(n+l+1)!!(n-l)!!\, r^{2n+1}}, \tag{3.6}$$

$$R(-3, l) = (2l+1)\sum_n \frac{(2n+1)!!\, r_j^n r_k^n}{(n+l+1)!!(n-l)!!\, r^{2n+3}}, \tag{3.7}$$

$$R(-5, l) = \frac{(2l+1)}{3}\sum_n \frac{(2n+3)!!\, r_j^n r_k^n}{(n+l+1)!!(n-l)!!\, r^{2n+5}}, \tag{3.8}$$

where

$$r = (r_j^2 + r_k^2)^{\frac{1}{2}},\quad n = l, l+2, l+4, \cdots,$$

$$(2k)!! = 2\cdot 4\cdots 2k,$$

and

$$(2k+1)!! = 1\cdot 3\cdots(2k+1);$$

$$R(-1, l) = 2(2l+1) \times \sum_{n=0}^{\infty} \frac{(2l+2n-1)!\,(l+n)(r_j r_k)^{l+n}}{(2l+n+1)!\, n!\, (r_j + r_k)^{2l+2n+1}}, \tag{3.9}$$

$$R(-3, l) = (2l+1) \times \sum_{n=0}^{\infty} \frac{(2l+2n+1)!\,(r_j r_k)^{l+n}}{(2l+n+1)!\, n!\, (r_j + r_k)^{2l+2n+3}}, \tag{3.10}$$

$$R(-5, l) = \frac{(2l+1)}{6} \times \sum_{n=0}^{\infty} \frac{(2l+2n+3)!\,(r_j r_k)^{l+n}}{(l+n+1)(2l+n+1)!\, n!\, (r_j + r_k)^{2l+2n+5}}. \tag{3.11}$$

Finally, Eqs. (3.1) and (3.2) are combined using the coupling theorem for spherical harmonics[14]:

$$\frac{1}{r_{jk}^n}\mathcal{Y}_N^{\mu}(r_{jk}) = [4\pi(2N+1)!]^{\frac{1}{2}} \sum_{l,\nu}\sum_{L,\kappa}\sum_{q,t} \chi_N^{\mu}(l, \nu; L, \kappa; q; t) \times r_j^{N-L} r_k^L R(-n, l) Y_q^{\mu+\nu-\kappa}(\theta_j, \varphi_j) Y_t^{\kappa-\nu}(\theta_k, \varphi_k), \tag{3.12}$$

where

$$\chi_N^{\mu}(l, \nu; L, \kappa; q; t) = (-1)^{L+\nu} \times \frac{C(L, N-L, N; \kappa, \mu-\kappa) C(N-L, l, q; \mu-\kappa, \nu)}{[(2q+1)(2t+1)(2L)!\,(2N-2L)!]^{\frac{1}{2}}} \times C(N-L, l, q; 00) C(L, l, t; \kappa, -\nu) C(L, l, t; 00). \tag{3.13}$$

Here the sums over q and t are controlled by the Clebsch–Gordan coefficients.

One can now apply Eq. (3.12) to the tensorial representation of the relativistic Hamiltonians [Eqs. (2.15), (2.17), and (2.20)]. The resulting one-center

[17] M. E. Rose, J. Math. & Phys. **37**, 215 (1958).
[18] R. A. Sack, J. Math. Phys. **5**, 245 (1964); **5**, 252 (1964).
[19] P. R. Fontana, J. Math. Phys. **2**, 825 (1961); Y. N. Chiu, J. Math. Phys. **5**, 283 (1964).

expansions are

$$H_{LL} = -8\pi \sum_{k>j} \Bigg[+ \tfrac{1}{3} \sum_{l=0}^{\infty} \sum_{\nu=-l}^{l} \sum_{\omega=-1}^{1} (-1)^{\nu+\omega} \times \frac{R(-1, l)}{(2l+1)} Y_l^{\nu}(\theta_j, \varphi_j) \times Y_l^{-\nu}(\theta_k, \varphi_k) T_1^{\omega}(\mathbf{p}_j) T_1^{-\omega}(\mathbf{p}_k) + \sum_{l=0}^{\infty} \sum_{\nu=-l}^{l} \sum_{L=0}^{2} \sum_{\kappa=-L}^{L} \sum_{q,t} \sum_{\omega,\eta=-1}^{1} \times (-1)^{\omega+\eta} C(112; \omega, \eta) \times \chi_2^{\omega+\eta}(l, \nu; L, \kappa; q; t) r_j^{2-L} r_k^{L} \times R(-3, l) Y_q^{\omega+\eta+\nu-\kappa}(\theta_j, \varphi_j) \times Y_t^{\kappa-\nu}(\theta_k, \varphi_k) T_1^{-\omega}(\mathbf{p}_j) T_1^{-\eta}(\mathbf{p}_k) \Bigg], \tag{3.14}$$

$$H_{SS} = -8\pi \sum_{k>j} \Bigg[+ \tfrac{1}{3} \delta^{(3)}(\mathbf{r}_{jk}) \sum_{\omega=-1}^{1} T_1^{\omega}(\mathbf{s}_j) T_1^{-\omega}(\mathbf{s}_k) + 6 \sum_{l=0}^{\infty} \sum_{\nu=-l}^{l} \sum_{L=0}^{2} \sum_{\kappa=-L}^{L} \sum_{q,t} \sum_{\omega,\eta=-1}^{1} \times (-1)^{\omega+\eta} C(112; \omega, \eta) \times \chi_2^{\omega+\eta}(l, \nu; L, \kappa; q; t) \times r_j^{2-L} r_k^{L} R(-5, l) \times Y_q^{\omega+\eta+\nu-\kappa}(\theta_j, \varphi_j) Y_t^{\kappa-\nu}(\theta_k, \varphi_k) \times T_1^{-\omega}(\mathbf{s}_j) T_1^{-\eta}(\mathbf{s}_k) \Bigg], \tag{3.15}$$

$$H_{SL} = \frac{1}{2} \sum_{\beta} \sum_{j} \frac{Z_\beta}{r_{j\beta}^3} \sum_{\omega=-1}^{1} (-1)^{\omega} T_1^{\omega}(\mathbf{r}_{j\beta} \times \mathbf{p}_j) T_1^{-\omega}(\mathbf{s}_j) - \frac{4\pi}{i} \sum_{k \neq j} \sum_{l=0}^{\infty} \sum_{\nu=-l}^{l} \sum_{L=0}^{1} \sum_{\kappa=-L}^{L} \sum_{q,t} \sum_{\omega,\eta=-1}^{1} (-1)^{\omega} \times C(111; \eta, \omega - \eta) \times \chi_1^{\eta}(l, \nu; L, \kappa; q; t) r_j^{1-L} r_k^{L} R(-3, l) \times Y_q^{\eta+\nu-\kappa}(\theta_j, \varphi_j) Y_t^{\kappa-\nu}(\theta_k, \varphi_k) \times [T_1^{\omega-\eta}(\mathbf{p}_j) - 2T_1^{\omega-\eta}(\mathbf{p}_k)] T_1^{-\omega}(\mathbf{s}_j). \tag{3.16}$$

There is a striking similarity between the second terms of H_{LL} and H_{SS}, the difference being the radial coefficient R and the appearance of linear-momentum operators in H_{LL} and spin in H_{SS}. It is interesting to note that the angular-momentum operators do not appear in H_{LL}. It is indeed possible to rewrite this Hamiltonian in such a way that it contains angular-momentum terms, but the transformed Hamiltonian does not simplify appreciably.[20] In all these expansions, the variables associated with electron j and k are now separated. In this form the angular part of the matrix elements of these Hamiltonians can be carried out in a straightforward manner (see Sec. 5). A difficulty arises in the radial integrations since the coefficients $R(-n, l)$ for $n > 1$ involve infinite sums.

In the case of H_{LL} and H_{SL} these infinite sums can be transformed into finite ones by applying the gradient formula to Eqs. (2.4) and (2.6), respectively. The procedure makes use of the fact that $\mathbf{r}_{jk}/r_{jk}^3$ appears in these two Hamiltonians. By making use of the relationship

$$\frac{\mathbf{r}_{jk}}{r_{jk}^3} = -\nabla_j \frac{1}{r_{jk}} \tag{3.17}$$

and the gradient formula[21]

$$T_1^{\mu}(\nabla)\Phi(r) Y_l^{\nu}(\theta, \varphi) = \left(\frac{l+1}{2l+3}\right)^{\frac{1}{2}} C(l, 1, l+1; \nu, \mu) Y_{l+1}^{\nu+\mu}(\theta, \varphi) \times \left(\frac{d\Phi}{dr} - \frac{l}{r}\Phi\right) - \left(\frac{l}{2l-1}\right)^{\frac{1}{2}} C(l, 1, l-1; \nu, \mu) \times Y_{l-1}^{\nu+\mu}(\theta, \varphi) \left(\frac{d\Phi}{dr} + \frac{l+1}{r}\Phi\right), \tag{3.18}$$

the μth component of $\mathbf{r}_{jk}/r_{jk}^3$ can be written as

$$T_1^{\mu}\left(-\nabla_j \frac{1}{r_{jk}}\right) = +4\pi \sum_{l,n,\nu} (-1)^{\nu} C(l1n; 00) \times C(l1n; \nu, -\mu) \times Y_n^{\mu-\nu}(\theta_j, \varphi_j) Y_l^{\nu}(\theta_k, \varphi_k) A_{l,n}, \tag{3.19}$$

where

$$A_{l,l+1} = \left(\frac{2l+1}{2l+3}\right)^{\frac{1}{2}} \frac{r_k^l}{r_j^{l+2}} \epsilon(r_j - r_k),$$

$$A_{l,l-1} = -\left(\frac{2l+1}{2l-1}\right)^{\frac{1}{2}} \frac{r_j^{l-1}}{r_k^{l+1}} \epsilon(r_k - r_j), \tag{3.20}$$

and $\epsilon(x - y) = 1$ for $x > y$, $\epsilon = 0$ for $y > x$. With these equations one can rewrite H_{LL} and H_{SL} in the following way:

$$H_{LL} = -\frac{(4\pi)^2}{6} \sum_{k>j} \sum_{l=0}^{\infty} \sum_{\nu=-l}^{l} \sum_{\omega=-1}^{1} (-1)^{\nu+\omega} \frac{R(-1, l)}{(2l+1)} \times Y_l^{\nu}(\theta_j, \varphi_j) Y_l^{-\nu}(\theta_k, \varphi_k) T_1^{\omega}(\mathbf{p}_j) T_1^{-\omega}(\mathbf{p}_k) - \frac{(4\pi)^2\sqrt{2}}{6} \sum_{k>j} \sum_{l=0}^{\infty} \sum_{\nu=-l}^{l} \sum_{L=0}^{1} \sum_{\kappa=-L}^{L} \sum_{n,q,t} \sum_{\omega,\eta=-1}^{1} G \times r_j^{1-L} r_k^{L} A_{l,n} Y_q^{\eta+\omega-\nu-\kappa}(\theta_j, \varphi_j) Y_t^{\nu+\kappa}(\theta_k, \varphi_k) \times T_1^{-\omega}(\mathbf{p}_j) T_1^{-\eta}(\mathbf{p}_k), \tag{3.21}$$

[20] The transformed Hamiltonian has the following form

$$H_{LL} = -\frac{1}{2} \sum_{k>j} \frac{1}{r_{jk}^3} [2r_{jk}^2(\mathbf{p}_j \cdot \mathbf{p}_k) - (\mathbf{r}_j \times \mathbf{p}_k)(\mathbf{p}_j \times \mathbf{p}_k) - (\mathbf{r}_{jk} \times \mathbf{p}_k) \cdot \mathbf{l}_j - (\mathbf{r}_{kj} \times \mathbf{p}_j) \cdot \mathbf{l}_k].$$

The terms of the form $(\mathbf{r}_{jk} \times \mathbf{p}_k) \cdot \mathbf{l}_j$ represent the coupling of the angular momentum of electron k relative to electron j with the angular momentum of electron j.

[21] M. E. Rose, *Multipole Fields* (John Wiley & Sons, Inc., New York, 1955), p. 28.

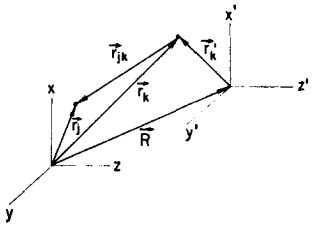


FIG. 1. Coordinate system for two-center expansions.

where

$$G = (-1)^{L+\nu+\omega+\eta}\left[\frac{(2n+1)(2l+1)}{(2L)!\,(2-2L)!\,(2q+1)(2t+1)}\right]^{\frac{1}{2}} \times C(L, 1-L, 1; \kappa, \omega-\kappa)C(l1n; 00) \times C(l1n; \nu, -\eta)C(n, 1-L, q; \eta-\nu, \omega-\kappa) \times C(n, 1-L, q; 00)C(lLt; \nu, \kappa)C(lLt; 00), \tag{3.22}$$

$$H_{SL} = \frac{1}{2}\sum_{\beta}\sum_{j}\frac{Z_\beta}{r_{j\beta}^3}\sum_{\omega=-1}^{1}(-1)^\omega T_1^\omega(\mathbf{r}_{j\beta}\times\mathbf{p}_j)T_1^{-\omega}(\mathbf{s}_j) - \frac{2\pi\sqrt{2}}{i}\sum_{k\neq j}\sum_{l=0}^{\infty}\sum_{\nu=-l}^{l}\sum_{n}\sum_{\omega,\eta=-1}^{1}C(111;\eta,\omega-\eta) \times C(l1n;\nu,-\eta)C(l1n;00) \times Y_n^{\eta-\nu}(\theta_j,\varphi_j)Y_l^\nu(\theta_k,\varphi_k) \times A_{l,n}[T_1^{\omega-\eta}(\mathbf{p}_j) - 2T_1^{\omega-\eta}(\mathbf{p}_k)]T_1^{-\omega}(\mathbf{s}_j). \tag{3.23}$$

In a calculation of matrix elements of H_{LL} the angular integration restricts the ranges of g and t and then $C(l, L, t; 00)$, say, limits the sum over l. For H_{SL} the angular integration directly limits the sum over l.

4. TWO-CENTER EXPANSION

One has to distinguish several regions in two-center expansions.[22] In long-range force calculations the distance R between the two centers is larger than the size of the charge distributions of the interacting molecules. In this case one can expand the Breit–Pauli Hamiltonian in a series in inverse powers of R. To obtain two-center expansions for the regions where the charge distributions overlap, one begins with the one-center result, transforms to the second center, and re-expands the result.

The coordinate system used in the two-center expansion is given in Fig. 1. The x, y, and z axes of the two coordinate systems are parallel. In general, $\mathbf{R}$ is not along the z axes. However, in most applications $\mathbf{R}$ is chosen to lie along the z axes. The vector $\mathbf{r}_j$ specifies the position of electron j with respect to center A, and $\mathbf{r}_k'$ the position of electron k referred to center B. The position of the nuclei are designated by $\mathbf{r}_\alpha$ and $\mathbf{r}_\beta'$, respectively.

The quantities $Y_l^\mu(\theta_{jk}, \varphi_{jk})$ and r_{jk}^{-n} which occur in the tensor forms of the Breit–Pauli Hamiltonian must be expressed in terms of the variables of the two coordinate systems. To generalize Eq. (3.1) to two centers, one makes use of the relations $\mathbf{r}_{jk} = \mathbf{r}_j - \mathbf{r}_k = \mathbf{r}_j - \mathbf{r}_k' - \mathbf{R}$. Then

$$\mathcal{Y}_N^\mu(\mathbf{r}_{jk}) = [4\pi(2N+1)!]^{\frac{1}{2}}\sum_{L=0}^{N}\sum_{\kappa=-L}^{L}(-1)^{N+L} \times \frac{C(L, N-L, N; \kappa, \mu-\kappa)}{[(2L+1)!\,(2N-2L+1)!]^{\frac{1}{2}}} \times \mathcal{Y}_{N-L}^{\mu-\kappa}(\mathbf{r}_k)\mathcal{Y}_L^\kappa(\mathbf{r}_j), \tag{4.1}$$

where on the right-hand side we have permuted $\mathbf{r}_k$ and $\mathbf{r}_j$, which introduces the phase factor $(-1)^N$. Since $\mathbf{r}_k = \mathbf{r}_k' + \mathbf{R}$, $\mathcal{Y}_{N-L}^{\mu-\kappa}(\mathbf{r}_k)$ can be expanded using Eq. (3.1) to give

$$\mathcal{Y}_N^\mu(\mathbf{r}_{jk}) = 4\pi\sum_{L=0}^{N}\sum_{\kappa=-L}^{L}\sum_{J=0}^{N-L}\sum_{\omega=-J}^{J}(-1)^{N+L}[(2N+1)!]^{\frac{1}{2}} \times C(L, N-L, N; \kappa, \mu-\kappa) \times \frac{C(J, N-L-J, N-L; \omega, \mu-\kappa-\omega)}{[(2L+1)!\,(2J+1)!\,(2N-2L-2J+1)!]^{\frac{1}{2}}} \times \mathcal{Y}_L^\kappa(\mathbf{r}_j)\mathcal{Y}_J^\omega(\mathbf{r}_k')\mathcal{Y}_{N-L-J}^{\mu-\kappa-\omega}(\mathbf{R}). \tag{4.2}$$

If $\mathbf{R}$ lies along the z axes, then[23]

$$\mathcal{Y}_{N-L-J}^{\mu-\kappa-\omega}(R_z) = R^{N-L-J}\left[\frac{2N-2L-2J+1}{4\pi}\right]^{\frac{1}{2}}\delta_{\mu-\kappa-\omega,0} \tag{4.3}$$

and

$$\mathcal{Y}_N^\mu(\mathbf{r}_{jk}) = r_{jk}^n Y_N^\mu(\theta_{jk}, \varphi_{jk}) = [4\pi(2N+1)]^{\frac{1}{2}}\sum_{L=0}^{N}\sum_{J=0}^{N-L}\sum_{\kappa=-L}^{L}\frac{(-1)^{N+L}}{(N-L-J)!} \times \left[\frac{(N+\mu)!\,(N-\mu)!}{(2L+1)(2J+1)(L+\kappa)!\,(L-\kappa)!\,(J-\mu+\kappa)!\,(J+\mu-\kappa)!}\right]^{\frac{1}{2}} \times r_j^L r_k'^J R^{N-L-J}Y_L^\kappa(\theta_j, \varphi_j)Y_J^{-\kappa}(\theta_k'\varphi_k'). \tag{4.4}$$

The two-center expansion of r_{jk}^{-n} for overlapping charge distributions is in general very complicated. For $n = 1$ the expansions have been done for the overlap regions.[22] A method[19] has been developed that can be used for the general expansion of r_{jk}^{-n}. For the nonoverlapping region a useful expansion has recently been derived by Sack.[24] In this case

[22] R. J. Beuhler and J. O. Hirschfelder, Phys. Rev. **83**, 628 (1951); **85**, 149 (1952).

[23] P. R. Fontana, Phys. Rev. **123**, 1865 (1961).

[24] R. A. Sack, J. Math. Phys. **5**, 260 (1964).

electrons j and k are associated with centers A and B, respectively. The following result is valid for R along the z axes[25]:

$$\frac{1}{r_{jk}^n} = 4\pi \sum_{\substack{l_1,l_2,l_3 \\ l_1+l_2+l_3=\text{even}}} \sum_{\nu} \sum_{q,t} \frac{(-1)^{l_1}(-1)^{\frac{1}{2}(l_1+l_2+l_3)}}{R^{n+l_1+l_2+2q+2t}} K(l_1, l_2, l_3; \nu) \times G(n; l_1, l_2, l_3; q, t; r_j, r_k') \times Y_{l_1}^{\nu}(\theta_j, \varphi_j) Y_{l_2}^{-\nu}(\theta_k', \varphi_k'), \tag{4.5}$$

where

$$K(l_1, l_2, l_3; \nu) = \frac{[\frac{1}{2}(l_1 + l_2 + l_3)]!}{[\frac{1}{2}(l_1 + l_2 - l_3)]! [\frac{1}{2}(l_1 + l_3 - l_2)]! [\frac{1}{2}(l_2 + l_3 - l_1)]!} \times \left[\frac{(2l_1 + 1)(2l_2 + 1)(2l_3 + 1)(l_1 + l_2 - l_3)! \times (l_1 + l_3 - l_2)! (l_2 + l_3 - l_1)!}{(l_1 + l_2 + l_3 + 1)!}\right]^{\frac{1}{2}} \times C(l_1 l_2 l_3; -\nu, \nu) \tag{4.6}$$

and

$$G(n; l_1, l_2, l_3; q, t; r_j, r_k') = \frac{2^{l_1+l_2+q+t}\Gamma[\frac{1}{2}(n + l_1 + l_2 + l_3) + q + t]}{\Gamma\left(\frac{n}{2}\right)\Gamma[\frac{1}{2}(n - 1)](2l_1 + 2q + 1)!! (2l_2 + 2t + 1)!! q! t!} \times \Gamma[\tfrac{1}{2}(n + l_1 + l_2 - l_3 - 1) + q + t] r_j^{2q+l_1} (r_k')^{2t+l_2}. \tag{4.7}$$

In Eq. (4.7) $\Gamma(x)$ is the Gamma function and the double factorials are defined in Sec. 3. The two center expansion for the orbit–orbit, spin–spin, and spin–orbit Hamiltonians can now be obtained by substituting Eqs. (4.4) and (4.5) in Eqs. (2.15), (2.17), and (2.20), respectively, and coupling the various spherical harmonics. In the resulting equations[26] the variables associated with centers A and B are separated.

The Wigner–Eckart theorem[12] when applied to the angular parts of the matrix elements of H_{LL}, H_{SS}, and H_{SL} yields selection rules for these Hamiltonians.

5. MATRIX ELEMENTS

The one- and two-center expansions of the Breit–Pauli Hamiltonians H_{LL}, H_{SS}, and H_{SL} are of the general form

$$H_{LL} \sim \sum \{\ \} Y_q^{\nu}(\theta_j, \varphi_j) Y_t^{\kappa}(\theta_k, \varphi_k) T_1^{-\omega}(\mathbf{p}_j) T_1^{-\eta}(\mathbf{p}_k), \tag{5.1}$$

$$H_{SS} \sim \sum \{\ \} Y_q^{\nu}(\theta_j, \varphi_j) Y_t^{\kappa}(\theta_k, \varphi_k) T_1^{-\omega}(\mathbf{s}_j) T_1^{-\eta}(\mathbf{s}_k), \tag{5.2}$$

$$H_{SL} \sim \sum \{\ \} Y_q^{\nu}(\theta_j, \varphi_j) Y_t^{\kappa}(\theta_k, \varphi_k) T_1^{-\omega}(\mathbf{p}_j) T_1^{-\eta}(\mathbf{s}_k). \tag{5.3}$$

If the wavefunction Ψ is of the type

$$\Psi = \sum_n c_n \prod_i \{\psi_n(\mathbf{r}_i, \mathbf{s}_i)\}, \tag{5.4}$$

where

$$\psi(\mathbf{r}_i, \mathbf{s}_i) = \Phi(r_i) Y_{l_i}^{m_i}(\theta_i, \varphi_i) \eta_{\mu_i}(\mathbf{s}_i), \tag{5.5}$$

then the matrix elements of the Hamiltonians can be calculated in a straightforward manner. In Eq. (5.5) $\eta_{\mu_i}(\mathbf{s}_i)$ is a two-component spinor ($\mu_i = \pm\frac{1}{2}$).

In H_{LL} and H_{SL} one first has to operate with $T_1^{-\omega}(\mathbf{p}_j)$ on the wavefunction. Application of the gradient formula [Eq. (3.18)] yields

$$T_1^{-\omega}(\mathbf{p}_j)\Phi(r_j) Y_l^m(\theta_j, \varphi_j) = \frac{1}{i} \sum_u C(l1u; m, -\omega) C(l1u; 00) \times Y_u^{m-\omega}(\theta_j, \varphi_j) A_{l,u}(r_j), \tag{5.6}$$

where

$$A_{l,l+1} = \left(\frac{2l+1}{2l+3}\right)^{\frac{1}{2}} \left(\frac{d\Phi}{dr_j} - \frac{l}{r_j}\Phi\right),$$
$$A_{l,l-1} = \left(\frac{2l+1}{2l-1}\right)^{\frac{1}{2}} \left(\frac{d\Phi}{dr_j} + \frac{l+1}{r_j}\Phi\right), \tag{5.7}$$

with all the other A's vanishing because of the triangular condition in $C(l, 1, u; 00)$. The angular integrations in H_{LL}, H_{SS}, and H_{SL} are now all of the same form, namely[27]

$$\langle Y_{l'}^{m'}(\theta_j, \varphi_j) |\, Y_q^{\nu}(\theta_j, \varphi_j) \,| Y_l^m(\theta_j, \varphi_j)\rangle = \left[\frac{(2l+1)(2q+1)}{4\pi(2l'+1)}\right]^{\frac{1}{2}} C(lql'; m, \nu, m') C(lql'; 00). \tag{5.8}$$

The selection rules for this angular integration can be directly obtained from the Clebsch–Gordan coefficients. The integral vanishes unless $l + l' \leq q \leq |l - l'|$ and the sum $l + l' + q$ is even. Also $m' = m + \nu$.

[25] The result for $n = 1$ agrees with the previous work of R. C. Carlson and L. S. Rushbrooke, Proc. Cambridge Phil. Soc. **46**, 626 (1950) and Refs. 17 and 22.

[26] See appendices 1.A–1.C of W. J. Meath, The University of Wisconsin Theoretical Chemistry Institute Technical Report WIS-TCI-75, April, 1965. For explicit expressions through $(1/R^3)$ see W. J. Meath and J. O. Hirschfelder, J. Chem. Phys. **44**, 3197 (1966).

[27] See, for example, Ref. 4, p. 62.

The integration over the spin variables is given by the expression[28]

$$\langle \eta_{\mu'}(\mathbf{s}_j)|\, T_1^{-\omega}(\mathbf{s}_j)\, |\eta_{\mu}(s_j)\rangle$$

$$= \frac{\sqrt{3}}{2}\, C(\tfrac{1}{2}1\tfrac{1}{2};\mu, -\omega, \mu')$$

$$= (-1)^{\frac{1}{2}-\mu}\tfrac{1}{2}[(1-\mu+\mu')!\,(1+\mu-\mu')!]^{\frac{1}{2}}. \quad (5.9)$$

[28] See, for example, Ref. 4, p. 89.

The remaining radial integrals depend on the particular choice of $\Phi(r_i)$ and cannot be done in a general manner.

ACKNOWLEDGMENTS

The authors are indebted to Professor J. O. Hirschfelder for his helpful remarks and critical comments. One of us (P. R. F.) gratefully acknowledges the hospitality extended to him by the members of the Theoretical Chemistry Institute at the University of Wisconsin.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Some Variational Principles for Integral Equations

P. D. ROBINSON AND A. M. ARTHURS
Department of Mathematics, University of York, England

(Received 19 January 1968)

Complementary variational principles are developed for the solution of Fredholm integral equations with symmetric positive-definite kernels. In particular, the theory is applied to linear equations of the type

$$\varphi(\mathbf{r}) = f(\mathbf{r}) + \lambda \int \mathcal{K}(\mathbf{r}, \mathbf{s})\varphi(\mathbf{s})\, d\mathbf{s},$$

and bounds are obtained for $\int f\varphi\, d\mathbf{r}$. When λ is negative, the bounds are complementary upper and lower ones. When λ is positive, the bounds are one-sided, but an improvement is made on a result of Strieder and Prager [J. Math. Phys. **8**, 514 (1967)]. A condition given by these authors for the existence of bounds does not seem to be strictly necessary, and alternative conditions are derived. Systematic improvement of bounds by iterative and scaling procedures is discussed.

1. INTRODUCTION

Recently Noble[1] and Rall[2] have developed complementary variational principles which are relevant in physical situations described by a pair of simultaneous equations

$$T\Phi = \frac{\partial W}{\partial U}, \quad (1a)$$

$$T^{\dagger} U = \frac{\partial W}{\partial \Phi}, \quad (1b)$$

T being a linear operator and $T^{\dagger}$ its adjoint. Applications of the theory have been made to ordinary differential equations[1] and also to partial differential equations of diffusion[3,4] and Poisson[5] type.

In this paper we show how the theory can be applied to integral equations with symmetric positive-definite kernels, and, in particular, to nonhomogeneous linear integral equations with parameter λ. When λ is negative, we obtain complementary upper and lower bounds; and when λ is positive, we obtain one-sided bounds which are an improvement on a result of Strieder and Prager.[6] A condition given by these authors for the existence of bounds does not seem to be necessary, and alternative conditions are derived. Systematic improvement of bounds by iterative and scaling procedures is also discussed.

2. THEORY

A. Complementary Variational Principles

In the simplest form of the theory (which suffices for the present paper), Φ and U are real functions of the position vector $\mathbf{r}$ and W is also real, depending on

[1] B. Noble, Univ. Wisconsin Math. Res. Center Rept. No. 473 (1964).
[2] L. B. Rall, J. Math. Anal. Appl. **14**, 174 (1966).
[3] A. M. Arthurs, Proc. Roy. Soc. (London) **A298**, 97 (1967).
[4] A. M. Arthurs and P. D. Robinson, Proc. Roy. Soc. (London) **A303**, 497 (1968).
[5] A. M. Arthurs and P. D. Robinson, Proc. Cambridge Phil. Soc. (to be published).
[6] W. Strieder and S. Prager, J. Math. Phys. **8**, 514 (1967).

The integration over the spin variables is given by the expression[28]

$$\langle \eta_{\mu'}(\mathbf{s}_j)|\, T_1^{-\omega}(\mathbf{s}_j)\, |\eta_{\mu}(s_j)\rangle$$

$$= \frac{\sqrt{3}}{2}\, C(\tfrac{1}{2}1\tfrac{1}{2};\mu, -\omega, \mu')$$

$$= (-1)^{\frac{1}{2}-\mu}\tfrac{1}{2}[(1-\mu+\mu')!\,(1+\mu-\mu')!]^{\frac{1}{2}}. \quad (5.9)$$

[28] See, for example, Ref. 4, p. 89.

The remaining radial integrals depend on the particular choice of $\Phi(r_i)$ and cannot be done in a general manner.

ACKNOWLEDGMENTS

The authors are indebted to Professor J. O. Hirschfelder for his helpful remarks and critical comments. One of us (P. R. F.) gratefully acknowledges the hospitality extended to him by the members of the Theoretical Chemistry Institute at the University of Wisconsin.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Some Variational Principles for Integral Equations

P. D. ROBINSON AND A. M. ARTHURS
Department of Mathematics, University of York, England

(Received 19 January 1968)

Complementary variational principles are developed for the solution of Fredholm integral equations with symmetric positive-definite kernels. In particular, the theory is applied to linear equations of the type

$$\varphi(\mathbf{r}) = f(\mathbf{r}) + \lambda \int \mathcal{K}(\mathbf{r}, \mathbf{s})\varphi(\mathbf{s})\, d\mathbf{s},$$

and bounds are obtained for $\int f\varphi\, d\mathbf{r}$. When λ is negative, the bounds are complementary upper and lower ones. When λ is positive, the bounds are one-sided, but an improvement is made on a result of Strieder and Prager [J. Math. Phys. **8**, 514 (1967)]. A condition given by these authors for the existence of bounds does not seem to be strictly necessary, and alternative conditions are derived. Systematic improvement of bounds by iterative and scaling procedures is discussed.

1. INTRODUCTION

Recently Noble[1] and Rall[2] have developed complementary variational principles which are relevant in physical situations described by a pair of simultaneous equations

$$T\Phi = \frac{\partial W}{\partial U}, \quad (1a)$$

$$T^{\dagger} U = \frac{\partial W}{\partial \Phi}, \quad (1b)$$

T being a linear operator and $T^{\dagger}$ its adjoint. Applications of the theory have been made to ordinary differential equations[1] and also to partial differential equations of diffusion[3,4] and Poisson[5] type.

In this paper we show how the theory can be applied to integral equations with symmetric positive-definite kernels, and, in particular, to nonhomogeneous linear integral equations with parameter λ. When λ is negative, we obtain complementary upper and lower bounds; and when λ is positive, we obtain one-sided bounds which are an improvement on a result of Strieder and Prager.[6] A condition given by these authors for the existence of bounds does not seem to be necessary, and alternative conditions are derived. Systematic improvement of bounds by iterative and scaling procedures is also discussed.

2. THEORY

A. Complementary Variational Principles

In the simplest form of the theory (which suffices for the present paper), Φ and U are real functions of the position vector $\mathbf{r}$ and W is also real, depending on

[1] B. Noble, Univ. Wisconsin Math. Res. Center Rept. No. 473 (1964).
[2] L. B. Rall, J. Math. Anal. Appl. **14**, 174 (1966).
[3] A. M. Arthurs, Proc. Roy. Soc. (London) **A298**, 97 (1967).
[4] A. M. Arthurs and P. D. Robinson, Proc. Roy. Soc. (London) **A303**, 497 (1968).
[5] A. M. Arthurs and P. D. Robinson, Proc. Cambridge Phil. Soc. (to be published).
[6] W. Strieder and S. Prager, J. Math. Phys. **8**, 514 (1967).

Φ, U, and $\mathbf{r}$. $T^\dagger$ is the adjoint of T in the sense that

$$\int U(T\Phi)\,d\mathbf{r} = \int (T^\dagger U)\Phi\,d\mathbf{r}, \tag{2}$$

the integrations being taken over the whole configuration space. If we define the functional

$$I(\Phi, U) = \int W(\mathbf{r}, \Phi, U)\,d\mathbf{r} - \int U(T\Phi)\,d\mathbf{r} \tag{3a}$$

$$= \int W(\mathbf{r}, \Phi, U)\,d\mathbf{r} - \int (T^\dagger U)\Phi\,d\mathbf{r}, \tag{3b}$$

then the following results can be derived[1]:

Stationary property: $I(\Phi, U)$ is stationary at (φ, u) if Eqs. (1a) and (1b) hold simultaneously at (φ, u).

First variational principle: Choose a trial Φ close to φ, and determine $U(\Phi)$ so that (1a) is satisfied identically.

Then, if (1b) holds at (φ, u), we have

$$G(\Phi) \equiv I(\Phi, U(\Phi)) = I(\varphi, u) + \tfrac{1}{2}\Delta_1(\Phi) + O(\Phi - \varphi)^3, \tag{4}$$

where

$$\Delta_1(\Phi) = \int \left\{ (\Phi - \varphi)^2 \left[\frac{\partial^2 W}{\partial \Phi^2}\right]_{\varphi,u} - [U(\Phi) - u]^2 \left[\frac{\partial^2 W}{\partial U^2}\right]_{\varphi,u} \right\} d\mathbf{r}. \tag{5}$$

Second variational principle: Choose a trial U close to u, and determine $\Phi(U)$ so that (1b) is satisfied identically.

Then, if (1a) holds at (φ, u), it follows that

$$J(U) \equiv I(\Phi(U), U) = I(\varphi, u) + \tfrac{1}{2}\Delta_2(U) + O(U - u)^3, \tag{6}$$

where

$$\Delta_2(U) = -\int \left\{ [\Phi(U) - \varphi]^2 \left[\frac{\partial^2 W}{\partial \Phi^2}\right]_{\varphi,u} - (U - u)^2 \left[\frac{\partial^2 W}{\partial U^2}\right]_{\varphi,u} \right\} d\mathbf{r}. \tag{7}$$

If terms of higher order than the second are neglected (or vanish), we see that the functionals $G(\Phi)$ and $J(U)$ furnish complementary upper and lower bounds for $I(\varphi, u)$ when Δ_1 and Δ_2 do not have the same sign. If Δ_1 and Δ_2 do have the same sign, then $G(\Phi)$ and $J(U)$ become different one-sided bounds.

B. Integral Equations with Symmetric, Positive-Definite Kernels

These variational principles are immediately applicable to integral equations of the type

$$m(\Phi) = \int \mathcal{K}(\mathbf{r}, \mathbf{s})\Phi(\mathbf{s})\,d\mathbf{s}, \tag{8}$$

whenever

$$\mathcal{K}(\mathbf{r}, \mathbf{s}) = \int \mathcal{C}(\mathbf{r}, \mathbf{t})\mathcal{C}(\mathbf{s}, \mathbf{t})\,d\mathbf{t} \tag{9}$$

and it is assumed that the functional m has an inverse. We may write (8) and (9) together as

$$m(\Phi) = K\Phi = T^\dagger T\Phi, \tag{10}$$

where K, T, and $T^\dagger$ are now integral operators. [If T corresponds to the real kernel $\mathcal{C}(\mathbf{r}, \mathbf{s})$, then the adjoint $T^\dagger$ corresponds to the kernel $\mathcal{C}(\mathbf{s}, \mathbf{r})$.] Condition (9) clearly implies that $\mathcal{K}(\mathbf{r}, \mathbf{s})$ is a symmetric kernel (or, equivalently, that K is a self-adjoint operator, assuming always that orders of integration can be changed). It also implies that $\mathcal{K}(\mathbf{r}, \mathbf{s})$ is a positive-definite kernel, since, for an arbitrary real function $\psi(\mathbf{r})$, we have

$$\int \psi K\psi\,d\mathbf{r} = \int \psi T^\dagger T\psi\,d\mathbf{r} = \int (T\psi)^2\,d\mathbf{r} \geq 0. \tag{11}$$

Further discussion of the kernel $\mathcal{K}$ is given in Sec. 5.

If we take

$$W = \tfrac{1}{2}U^2 + M(\Phi), \tag{12}$$

where

$$dM/d\Phi = m(\Phi), \tag{13}$$

then it is easy to see that Eq. (10) is equivalent to the pair of simultaneous equations

$$T\Phi = U = \frac{\partial W}{\partial U}, \tag{14a}$$

$$T^\dagger U = M(\Phi) = \frac{\partial W}{\partial \Phi}. \tag{14b}$$

The function φ is to be the solution of Eq. (8) for Φ, and u is $T\varphi$. On substituting from (12) and (13) into the various formulas of Sec. 2A, we obtain the expressions

$$G(\Phi) = \int \{-\tfrac{1}{2}(T\Phi)^2 + M(\Phi)\}\,d\mathbf{r} \tag{15}$$

$$= \int \{-\tfrac{1}{2}\Phi K\Phi + M(\Phi)\}\,d\mathbf{r}, \tag{16}$$

$$J(U) = \int \{\tfrac{1}{2}U^2 + M(m^{-1}(T^\dagger U)) - (T^\dagger U)m^{-1}(T^\dagger U)\}\,d\mathbf{r}, \tag{17}$$

and

$$I(\varphi, u) = \int \{-\tfrac{1}{2}\varphi m(\varphi) + M(\varphi)\}\,d\mathbf{r}. \tag{18}$$

It is interesting to note that expression (15) for $G(\Phi)$ does not involve the T operator explicitly. A similar expression for J is obtained if we restrict ourselves to trial functions U of the form

$$U = T\Theta, \tag{19}$$

when (17) becomes

$$J(T\Theta) = \int \{\tfrac{1}{2}\Theta K\Theta + M(m^{-1}(K\Theta)) - (K\Theta)m^{-1}(K\Theta)\}\, d\mathbf{r}. \tag{20}$$

From (5), (7), and (12) it is evident that, if

$$\left[\frac{\partial^2 W}{\partial \Phi^2}\right]_{\varphi,u} = \left[\frac{dm}{d\Phi}\right]_{\varphi} \leq 0, \tag{21}$$

then $G(\Phi)$ is a lower bound to $I(\varphi, u)$ and either $J(U)$ or $J(T\Theta)$ are upper bounds which are complementary to $G(\Phi)$.

3. NONHOMOGENEOUS LINEAR INTEGRAL EQUATIONS

A. Basic Formulas

The nonhomogeneous linear equation

$$\Phi(\mathbf{r}) = f(\mathbf{r}) + \lambda \int \mathcal{K}(\mathbf{r}, \mathbf{s})\Phi(\mathbf{s})\, d\mathbf{s} \tag{22}$$

provides an example of the foregoing theory with

$$m(\Phi) = \lambda^{-1}(\Phi - f) \tag{23}$$

and

$$W = \tfrac{1}{2}U^2 + \lambda^{-1}(\tfrac{1}{2}\Phi^2 - f\Phi). \tag{24}$$

The various functionals become

$$G(\Phi) = \int \{-\tfrac{1}{2}\Phi K\Phi + \lambda^{-1}(\tfrac{1}{2}\Phi^2 - f\Phi)\}\, d\mathbf{r}, \tag{25}$$

$$J(U) = \int \{\tfrac{1}{2}U^2 - \tfrac{1}{2}\lambda^{-1}(f + \lambda T^{\dagger}U)^2\}\, d\mathbf{r}, \tag{26}$$

$$J(T\Theta) = \int \{\tfrac{1}{2}\Theta K\Theta - \tfrac{1}{2}\lambda^{-1}(f + \lambda K\Theta)^2\}\, d\mathbf{r}, \tag{27}$$

and

$$I(\varphi, u) = -\tfrac{1}{2}\lambda^{-1}\int f\varphi\, d\mathbf{r}. \tag{28}$$

It should be noted that W is merely a quadratic functional of Φ and U, so that in this case there are no third- or higher-order terms in Eqs. (4) and (6). Thus the condition that Φ and U should be close to the exact φ and u can be dropped.

B. Complementary Bounds for Negative λ

Whenever λ is negative, condition (21) holds and we obtain complementary lower and upper bounds

$$G \leq I \leq J. \tag{29}$$

The quantity I acts as a measure of the accuracy of the solution of the integral equation (22). In situations where $\int f\varphi\, d\mathbf{r}$ is of physical interest, this technique is particularly advantageous.

C. Bounds for Positive λ

If λ is positive, a closer examination of the functionals $\Delta_1(\Phi)$ and $\Delta_2(U)$ is required. From (5), (14), and (24) we have

$$\Delta_1(\Phi) = \int \{\lambda^{-1}(\Phi - \varphi)^2 - [T(\Phi - \varphi)]^2\}\, d\mathbf{r} = \int (\Phi - \varphi)(\lambda^{-1} - K)(\Phi - \varphi)\, d\mathbf{r} \tag{30}$$

and

$$\Delta_2(U) = -\int \{\lambda[T^{\dagger}(U - u)]^2 - (U - u)^2\}\, d\mathbf{r} = +\lambda \int (U - u)(\lambda^{-1} - K)(U - u)\, d\mathbf{r}. \tag{31}$$

Thus the

$$\text{positive-definiteness of } (\lambda^{-1} - K) \tag{32}$$

is a necessary and sufficient condition for both $\Delta_1(\Phi)$ and $\Delta_2(U)$ to be nonnegative.

In this situation $G(\Phi)$ and $J(U)$ are each *upper* bounds to $I(\varphi, u)$, and the question arises as to which is the less. In general, we cannot say, but in the particular case when

$$U = T\Phi, \tag{33}$$

it is not difficult to see that J is less than G. From (4) and (6) we have

$$\begin{aligned} G(\Phi) - J(T\Phi) &= \tfrac{1}{2}\Delta_1(\Phi) - \tfrac{1}{2}\Delta_2(T\Phi) \\ &= \tfrac{1}{2}\lambda^{-1}\int (\Phi - \varphi)(1 - 2\lambda K + \lambda^2 K^2)(\Phi - \varphi)\, d\mathbf{r} \\ &= \tfrac{1}{2}\lambda^{-1}\int \{(1 - \lambda K)(\Phi - \varphi)\}^2\, d\mathbf{r} \geq 0. \end{aligned} \tag{34}$$

The evaluation of $J(T\Phi)$ requires only a single application of the K operator, as does $G(\Phi)$; in principle we have a better bound.

When K satisfies the Hilbert–Schmidt condition

$$\iint \mathcal{K}^2(\mathbf{r}, \mathbf{s})\, d\mathbf{r}\, d\mathbf{s} < \infty, \tag{35}$$

we note that (32) is equivalent to the condition

$$0 < \lambda < \lambda_1, \tag{36}$$

where λ_1 is the smallest eigenvalue of K (these eigenvalues are all positive, since K is positive-definite). This follows from standard integral-equation theory, which tells us that, as ψ varies, the quantity $\int \psi K\psi\, d\mathbf{r}$ has greatest value $\lambda_1^{-1}\int \psi^2\, d\mathbf{r}$.

D. Systematic Improvement of Bounds by Iteration

When λ is positive, condition (36) is precisely the one required in order that the Neumann series for the solution of the integral equation (22) by iteration should converge. Thus, if Φ_n is an approximation to φ, we expect Φ_{n+1}, defined by

$$\Phi_{n+1} = f + \lambda K\Phi_n, \tag{37}$$

to be a better approximation. From (34) it follows that

$$G(\Phi_n) \geq J(T\Phi_n). \tag{38}$$

Moreover, since

$$\begin{aligned}\Delta_1(\Phi_{n+1}) &= \int (f + \lambda K\Phi_n - \varphi)(\lambda^{-1} - K)(f + \lambda K\Phi_n - \varphi)\, d\mathbf{r} \\ &= \int (\Phi_n - \varphi)\lambda K(\lambda^{-1} - K)\lambda K(\Phi_n - \varphi)\, d\mathbf{r},\end{aligned} \tag{39}$$

we have

$$\begin{aligned}&J(T\Phi_n) - G(\Phi_{n+1}) \\ &= \tfrac{1}{2}\Delta_2(T\Phi_n) - \tfrac{1}{2}\Delta_1(\Phi_{n+1}) \\ &= \tfrac{1}{2}\int (\Phi_n - \varphi)(K - 2\lambda K^2 + \lambda^2 K^3)(\Phi_n - \varphi)\, d\mathbf{r} \\ &= \tfrac{1}{2}\int \{(1 - \lambda K)(\Phi_n - \varphi)\}K\{(1 - \lambda K)(\Phi_n - \varphi)\}\, d\mathbf{r} \\ &\geq 0,\end{aligned} \tag{40}$$

from (11). Thus, starting with any initial function Φ_1, it follows from (38) and (40) that, when condition (36) is satisfied,

$$\begin{aligned}G(\Phi_1) \geq J(T\Phi_1) \geq G(\Phi_2) \geq J(T\Phi_2) \\ \geq G(\Phi_3) \geq \cdots \geq I(\varphi, u).\end{aligned} \tag{41}$$

If λ is negative and satisfies

$$|\lambda| < \lambda_1, \tag{42}$$

then the result corresponding to (41) is

$$\begin{aligned}G(\Phi_1) \leq G(\Phi_2) \leq \cdots \leq I(\varphi, u) \\ \leq \cdots \leq J(T\Phi_2) \leq J(T\Phi_1).\end{aligned} \tag{43}$$

E. Improvement of Bounds by Scaling and Ritz Procedures

Because of the quadratic nature of the functionals G and J, they can readily be improved by scaling procedures. For example, if α is a parameter, it follows from (25) that

$$G(\alpha\Phi) = \tfrac{1}{2}\alpha^2 \int \Phi(\lambda^{-1} - K)\Phi\, d\mathbf{r} - \alpha\lambda^{-1}\int f\Phi\, d\mathbf{r}. \tag{44}$$

Let α_0 be the value of α which optimizes (44), chosen by setting $\partial G/\partial\alpha$ equal to zero. The optimum functional is

$$\tilde{G}(\Phi) = G(\alpha_0\Phi) = -\frac{\left\{\int f\Phi\, d\mathbf{r}\right\}^2}{2\lambda^2\int \Phi(\lambda^{-1} - K)\Phi\, d\mathbf{r}}. \tag{45}$$

For a given Φ, $\tilde{G}(\Phi)$ is a best lower bound when λ is negative and a best upper bound when $(\lambda^{-1} - K)$ is positive-definite. Similarly, from (26) and (27) we obtain

$$\tilde{J}(U) = -\tfrac{1}{2}\lambda^{-1}\int f^2\, d\mathbf{r} - \frac{\left\{\int fT^{\dagger}U\, d\mathbf{r}\right\}^2}{2\int U(1 - \lambda K)U\, d\mathbf{r}} \tag{46}$$

and

$$\tilde{J}(T\Theta) = -\tfrac{1}{2}\lambda^{-1}\int f^2\, d\mathbf{r} - \frac{\left\{\int fK\Theta\, d\mathbf{r}\right\}^2}{2\int (K\Theta)(1 - \lambda K)\Theta\, d\mathbf{r}}. \tag{47}$$

When $(\lambda^{-1} - K)$ is positive-definite, we can extend the result (34) to give

$$\tilde{G}(\Phi) \geq \tilde{J}(T\Phi). \tag{48}$$

This follows because

$$\tilde{G}(\Phi) = G(\alpha_0\Phi) \geq J(T\alpha_0\Phi) \geq \tilde{J}(T\Phi).$$

However, results (40), (41), and (43) do not necessarily hold for individually optimized G's and J's.

Instead of introducing a single scale factor we can adopt a Ritz procedure and set

$$\Phi = \alpha_1\psi_1 + \alpha_2\psi_2 + \cdots + \alpha_m\psi_m \tag{49}$$

in (25), where the ψ's are m linearly independent functions. The consequent optimized functional is

$$\tilde{G}_m(\Phi) = -\tfrac{1}{2}B'A^{-1}B = +\tfrac{1}{2}(\det A)^{-1}\begin{vmatrix} 0 & B_1 & B_2 & \cdots & B_m \\ B_1 & A_{11} & A_{12} & \cdots & A_{1m} \\ B_2 & A_{21} & A_{22} & \cdots & A_{2m} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ B_m & A_{m1} & A_{m2} & \cdots & A_{mm} \end{vmatrix}, \tag{50}$$

where

$$B_i = \lambda^{-1}\int f\psi_i\, d\mathbf{r}, \tag{51}$$

$$A_{ij} = A_{ji} = \int \psi_i(\lambda^{-1} - K)\psi_j\, d\mathbf{r}, \tag{52}$$

A is the square symmetric $m \times m$ matrix with elements A_{ij}, B is the column matrix with elements B_i,

and B' is the transpose of B. A similar result is obtainable for J. When $(\lambda^{-1} - K)$ is positive-definite, we have

$$\tilde{G}_m(\Phi) \geq \tilde{J}_m(T\Phi), \tag{53}$$

analogously to (48). Also it can be proved that[7]

$$\tilde{G}_1(\Phi) \geq \tilde{G}_2(\Phi) \geq \cdots \geq \tilde{G}_m(\Phi) \geq \cdots \geq I(\varphi, u) \tag{54}$$

and

$$\tilde{J}_1(U) \geq \tilde{J}_2(U) \geq \cdots \geq \tilde{J}_m(U) \geq \cdots \geq I(\varphi, u). \tag{55}$$

For negative λ the result is

$$\tilde{G}_1(\Phi) \leq \tilde{G}_2(\Phi) \leq \cdots \leq I(\varphi, u) \leq \cdots \leq \tilde{J}_2(U) \leq \tilde{J}_1(U). \tag{56}$$

4. COMPARISON WITH A RESULT OF STRIEDER AND PRAGER

In a paper on bounds for Knudsen flow rates, Strieder and Prager[6] discuss an integral equation like (22), subject to the conditions

$$\lambda > 0, \tag{57}$$

$$\mathcal{K}(\mathbf{r}, \mathbf{s}) = \mathcal{K}(\mathbf{s}, \mathbf{r}) \geq 0, \tag{58}$$

and

$$\lambda^{-1} - \int \mathcal{K}(\mathbf{r}, \mathbf{s})\, d\mathbf{r} \geq 0, \tag{59}$$

which arise naturally from physical considerations. They show in effect that $G(\Phi)$ and also $\tilde{G}(\Phi)$ are upper bounds to $I(\varphi, u)$ under conditions (57)–(59). (They actually take λ to be unity, but we need not do so.)

These conditions are in fact stronger than our condition (32) (which is both necessary and sufficient), as the following argument shows. Condition (32) is equivalent to stating that, for an arbitrary $\psi(\mathbf{r})$,

$$\lambda^{-1} \int \psi^2(\mathbf{s})\, d\mathbf{s} - \iint \mathcal{K}(\mathbf{r}, \mathbf{s})\psi(\mathbf{r})\psi(\mathbf{s})\, d\mathbf{r}\, d\mathbf{s} \geq 0. \tag{60}$$

Since $\mathcal{K}(\mathbf{r}, \mathbf{s})$ is a symmetric kernel, the left-hand side of (60) can be rearranged to give

$$\int \left\{\lambda^{-1} - \int \mathcal{K}(\mathbf{r}, \mathbf{s})\, d\mathbf{r}\right\} \psi^2(\mathbf{s})\, d\mathbf{s} + \tfrac{1}{2} \iint \mathcal{K}(\mathbf{r}, \mathbf{s})\{\psi(\mathbf{r}) - \psi(\mathbf{s})\}^2\, d\mathbf{r}\, d\mathbf{s}. \tag{61}$$

If the Strieder–Prager conditions (57)–(59) hold, expression (61) is never negative and thus our condition (60) [i.e., (32)] is satisfied. To show that (60) can be satisfied when (59) is not, it is enough to consider the simple example

$$\mathcal{K}(x, y) = 5xy/2\lambda, \quad \lambda > 0, \quad 0 \leq x, y \leq 1. \tag{62}$$

This kernel is positive-definite and has the single eigenvalue $(6\lambda/5)$; thus condition (60) holds [cf. (36)]. However, condition (59) is not satisfied when $0.8 < y \leq 1$.

Strieder and Prager do not obtain the bound $J(T\Phi)$, which is, in principle, superior to $G(\Phi)$; it appears to be new. Our methods do not reveal certain lower bounds derived by these authors when λ is positive.

5. DISCUSSION

The foregoing analysis has dealt with variational principles for certain integral equations. The derivation was based on the generalized canonical Euler equations (1a), (1b) to emphasize the complementary nature of the results, but it should be noted that the principles obtained in Secs. 3B and 3C are actually independent of the decomposition of the operator K into the form $T^\dagger T$. Thus, for $\lambda < 0$, the principles $G \leq I \leq J$ in (29) hold for any symmetric positive-definite kernel $\mathcal{K}(\mathbf{r}, \mathbf{s})$, while for $\lambda > 0$ the principles $I \leq G$, $I \leq J$ hold by (30) and (31) if, in addition, $(\lambda^{-1} - K)$ is positive-definite.

The decomposition $K = T^\dagger T$, although not strictly necessary, can give useful insight into the properties of K and it does permit a somewhat more flexible J bound. If $\mathcal{K}(\mathbf{r}, \mathbf{s})$ is a function of $(\mathbf{r} - \mathbf{s})$ only, the kernel $\mathcal{C}(\mathbf{r}, \mathbf{s})$ can be found in principle by Fourier-transform convolution techniques; Laplace transforms serve to this end when $\mathcal{K}(\mathbf{r}, \mathbf{s})$ depends only on $(r + s)$. More generally, if there is an expansion

$$\mathcal{K}(\mathbf{r}, \mathbf{s}) = \sum_n \alpha_n \theta_n(\mathbf{r})\theta_n(\mathbf{s}) \tag{63}$$

in terms of any orthonormal set $\{\theta_n(\mathbf{r})\}$, it follows that, formally,

$$\mathcal{C}(\mathbf{r}, \mathbf{s}) = \sum_n \alpha_n^{\frac{1}{2}} \theta_n(\mathbf{r})\theta_n(\mathbf{s}). \tag{64}$$

The case of negative λ has applications in potential theory,[8] in bound-state quantum-mechanical perturbation theory, and possibly also in time-series analysis.[9]

ACKNOWLEDGMENT

We are very grateful to Professor S. Prager for his valuable comments on this work.

[7] See, for example, I. M. Gel'fand and S. V. Fomin, *Calculus of Variations* (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963), Chap. 8.

[8] A. G. Webster, *Partial Differential Equations of Mathematical Physics* (Dover Publications, Inc., New York, 1955), Chap. 9.

[9] Professor D. G. Kendall (private communication).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Antisymmetric Projection in the Approximation of No Spin-Orbit Coupling

JOHN J. SULLIVAN
*Instituto da Fisica, Universidade Catolica, Rio de Janeiro, Brazil**

(Received 25 April 1967)

The explicit form of the projection operator for constructing antisymmetric wavefunctions for N fermions in the approximation of no spin–orbit coupling is developed. Projection is applied within the one particle approximation. It is shown that if the orbitals associated with the minority spin can be completely expanded in terms of the orbitals associated with the majority spin, then the projected Hartree–Fock scheme is completely equivalent to unprojected Hartree–Fock theory. In the unrestricted case, deviations from this condition are not expected to be large, and integral properties such as energies calculated in the projected scheme should not be significantly different from unprojected results. However, for such properties as spin density at the nucleus in atoms or ions with nominally closed s shell, there may be significant differences between projected and unprojected schemes.

INTRODUCTION

For a system of N interacting but indistinguishable fermions the Pauli principle requires that the system wavefunction transform antisymmetrically under permutations of the particle coordinates. This imposes additional correlations among the particles, which is evident by noting that the wavefunction must be identically zero in any region of phase space where the spatial and spin coordinates of any two are the same. This has important physical consequences, some of which are well known. Owing to the difficulties introduced by the many-body interactions, calculations and much of the theoretical development are inevitably carried out within some approximational scheme. Thus much of our theoretical understanding of the consequences of the Pauli principle is intrinsically linked to these approximating schemes. In the much utilized Hartree–Fock one-particle self-consistent field scheme, antisymmetry of the wavefunction gives rise to the so-called "exchange terms." If relativistic effects can be neglected, the one-particle functions are considered to be a product of separate spatial and spin parts. Exchange is effective only between particles with the same spin projection due to the orthonormality of the spin functions. This form of the exchange term has been the basis for postulating strong magnetic correlations via Coulombostatic coupling (as in the Heisenberg theory of magnetic interactions). More recently the exchange interaction has been the basis for developing the spin-polarized Hartree–Fock scheme (SPHF), which admits to different orbitals for different spin, in contrast to the usual (restricted) scheme in which a given orbital may be occupied twice corresponding to the two possibilities for the fermion spin. Contact hyperfine interactions in the iron-transition series and in somer are earths have been analyzed using SPHF calculations and theory.[1] In most of these cases surprisingly good results are obtained. However, for the lighter elements such as nitrogen and oxygen, SPHF calculations have not been reliable for predicting such electronic properties as the spin density at the nucleus.[2] The question then arises as to how much the SPHF results represent physical consequences of the Pauli principle and to what extent they are influenced by the *ad hoc* restrictions introduced in the approximation scheme. The question, of course, is appropriate to any of the various schemes for solving approximately the many-fermion problem.

In this article we treat two subproblems of this more general question. If relativistic effects can be ignored, then the intrinsic and extrinsic angular-momentum operators (S^2 and L^2) commute with the Hamiltonian and can be simultaneously diagonalized along with the energy. The Pauli principle requires that the permutation symmetry of the spatial parts of the wavefunction match the permutation symmetry of the spin parts of the wavefunction so that an antisymmetric linear combination of their products can be formed. The problem can be exactly formulated in group-theoretical terms using the permutation group of N items. This is done in Sec. I, and the explicit form for antisymmetric projection assuming no spin–orbit coupling is deduced.[3] The form of this operator allows significant reductions to be made

*Present address: Department of Physics and Astronomy, Southern Illinois University, Carbondale, Illinois.

[1] R. F. Watson and A. J. Freeman, *Hyperfine Interactions*, A. J. Freeman and R. B. Frankel, Eds. (Academic Press Inc., New York, 1967), p. 53.

[2] C. M. Moser, *Hyperfine Interactions*, A. J. Freeman and R. B. Frankel, Eds. (Academic Press, Inc., New York, 1967), p. 95.

[3] Since this work was completed, articles by W. A. Goddard [Phys. Rev. **157**, 73, 81, 93, (1967)] have appeared in which he also obtains the antisymmetric projection operator assuming no spin–orbit coupling. The method presented here is independent of his derivation, although our final result is the same. The form of our reduced matrix element (2) is equivalent to his Eq. (15). The other results and discussions of this paper are different from the work reported there. See also R. D. Poshusta and R. W. Kramling, Phys. Rev. **167**, 139 (1968).

when one considers the diagonal matrix element of an operator completely symmetric under the permutation of the particle coordinates. When the operator is also completely symmetric under permutation of the spatial coordinates alone (e.g., the system Hamiltonian), a further reduction can be made. These reduced matrix elements are given at the end of Sec. I.

Having solved for the explicit form of the antisymmetric projection operator assuming no spin–orbit coupling, we can ask how does a one-electron self-consistent field scheme incorporating this projection (often called projected or extended Hartree–Fock) compare to the SPHF scheme. Consideration of the projected one-electron approximation is made in Sec. II. It is shown that the projected scheme will be different from SPHF only when the intersection of the orbital spaces nominally associated with spin up and spin down does not exhaust either space. This is invariably the case when one allows different orbitals for different spins, as is strikingly illustrated in considering the singlet He ground state[4] or in using SPHF to consider spin density at the nucleus for atoms or ions with nominally closed s shells.[1]

Some consequences of using antisymmetric projection for no spin–orbit coupling are discussed in Sec. III.

I. ANTISYMMETRIC PROJECTION OPERATOR

In this section the antisymmetric projection operator assuming no spin–orbit coupling is developed. Reduced forms for the diagonal matrix elements of operators completely symmetric under the exchange of identical fermions are presented.

It is well known that the antisymmetric representation of the permutation group $\mathfrak{S}_N$ is contained only and only once in the Kronecker product of conjugate irreducible representations. It is easy to symbolically construct operators corresponding to a given spin value S that will project out of general spin and orbital spaces a spin space and a conjugate orbital space that are irreducible under permutations of $\mathfrak{S}_N$. The difficulty in this projection method is that the linear combination coefficients involved in the projection operators are, in general, unknown. However, out of all the possible linear combinations of the products of the two representations, one needs only that unique combination that is antisymmetric in $\mathfrak{S}_N$. This allows an explicit evaluation of the coefficients of interest and the desired antisymmetric projection operator.

The projection operator that will project out of any space a function that transforms as the rth row of the λ irreducible representation of $\mathfrak{S}_N$ with irreducible matrix elements $U^{\lambda}_{rs}(P)$ is defined by[5]

$$e^{\lambda}_{rs} \equiv \frac{n_\lambda}{N!}\sum_{P\subset\mathfrak{S}_N} PU^{\lambda}_{sr}(P^{-1}) = \frac{n_\lambda}{N!}\sum_{P\subset\mathfrak{S}_N} PU^{\lambda *}_{rs}(P) = e^{\lambda\dagger}_{sr}; \qquad r, s = 1, \cdots, n_\lambda.$$

These operators are the primitive elements of the group algebra and have the usual orthogonal multiplication rule

$$e^{\lambda}_{rs}e^{\nu}_{uv} = \delta_{\lambda\nu}\delta_{su}e^{\lambda}_{rv}.$$

In the N many-electron problem the irreducible representation associated with the spin angular momentum S can be uniquely designated by the bipartition of N, $[(N/2) + S, (N/2) - S]$. The conjugate representation for the orbital function is designated by the partition $[1^{2s}, 2^{(N/2)-S}]$. Although in this paper we are directly concerned with the many-electron problem, a number of the propositions made here are applicable or easily extended to Fermion problems involving other than bipartition representations, such as the nuclear-isospin problem.

We can symbolically write our desired antisymmetric projection operator as $\{A\}(e^{\tilde{S}}_{ij})[e^{S}_{kl}]$, where

$$\{A\} \equiv \frac{1}{N!}\sum_{P\subset\mathfrak{S}_N}(-1)^P\{P\}$$

is the antisymmetric projection operator operating in both orbital and spin space, $[e^{S}_{kl}]$ is the projection operator associated with the internal angular momentum $S\hbar$, and operates in spin space only, and $(e^{\tilde{S}}_{ij})$ is the projection operator for the irreducible representation conjugate to S defined by the relation

$$(-1)^P U^{S}_{ni}(P) = U^{\tilde{S}}_{ni}(P)$$

and operates in orbital space only. The derivation of the following identities for this projection operator is straightforward:

$$\{A\}(e^{\lambda}_{ij})[e^{S}_{kl}] = \delta^{\sim}_{\lambda S}\delta_{ik}\frac{1}{n_S}\sum_m (e^{\tilde{S}}_{mj})[e^{S}_{ml}] = \delta^{\sim}_{\lambda S}\delta_{ik}\{A\}[e^{S}_{jl}].$$

In general, the coefficients $U^{S}_{ij}(P)$ are unknown; however, as shown below, one can explicitly derive all the coefficients that enter into the final antisymmetric projection operator.

A well-known theorem states that, in an irreducible

[4] C. A. Coulson, *Quantum Theory of Atoms, Molecules and the Solid State*, P. O. Lowdin, Ed. (Academic Press Inc., New York, 1967), p. 97.

[5] D. F. Johnson, Rept. Progr. Phys. **23**, 66, (1960). See the general theory of Part I, and especially Sec. 4 of Part II that treats the N-fermion problem.

space transforming under the permutation group $\mathcal{S}_N$ as the $[\lambda_1, \lambda_2, \cdots]$ irreducible representation, there always can be found one and only one vector S_λ symmetric, that is, invariant under the permutations of the subgroup $\mathcal{S}_\lambda \equiv \mathcal{S}_{\lambda 1} \times \mathcal{S}_{\lambda 2} \times \cdots$. Or equivalently, there can always be found one and only one vector A_λ antisymmetric, that is, antisymmetric under the permutation of the subgroup $\mathcal{S}_\lambda$.[6] This S_λ symmetric vector is chosen to be the first member of the basis for the invariant spin space. Consequently the A_λ antisymmetric vector is the first member in the basis for the conjugate invariant orbital space. Let Q be a coset of the subgroup $\mathcal{S}_\lambda$ with respect to the full group $\mathcal{S}_N$:

$$\mathcal{S}_N = Q\mathcal{S}_\lambda = \mathcal{S}_\lambda Q.^7$$

(More exactly, Q is a set of coset representatives q of the subgroup $\mathcal{S}_\lambda$ with respect to the full group $\mathcal{S}_N$.) The spin projection operator can be written as

$$[e^S_{mk}] = \frac{n_S}{N!} \sum_{q \subset Q} \sum_n U^S_{mn}(q)[q] \sum_{h \subset \mathcal{S}_\lambda} U^S_{nk}(h)[h].$$

Let S_λ and A_λ be, respectively, the symmetrizer and antisymmetrizer on the subgroup $\mathcal{S}_\lambda$:

$$S_\lambda \equiv \sum_{h \subset \mathcal{S}_\lambda} h \quad \text{and} \quad A_\lambda \equiv \sum_{h \subset \mathcal{S}_\lambda} (-1)^h h.$$

Because there is one and only one S_λ symmetric vector in the spin basis and it is chosen to be first in order, the spin projection operator can be written as

$$[e^S_{m1}] = \frac{n_S \prod_i n_{\lambda_i}!}{N!} \sum_{q \subset Q} U^S_{m1}(q)[q][S_\lambda].$$

By an entirely equivalent argument the orbital projection operator can be written as

$$(\tilde{e}^S_{m1}) = \frac{n_S \prod_i n_{\lambda_i}!}{N!} \sum_{q \subset Q} (-1)^q U^S_{m1}(q)(q)(A_\lambda).$$

To obtain the desired antisymmetric projection operator, one takes the product of the above two operators and sums on m. Because all q are chosen to be self-inverse (see below) and the elements of the subgroup $\mathcal{S}_\lambda$ have the property

$$U^S_{n1}(h) = \delta_{n1} = U^S_{1n}(h)$$

by construction, the problem is completely resolved if the coefficients $U^S_{11}(q)$ of elements belonging to the coset Q are known.

The coset Q can be uniquely constructed in the following manner. The permutations of the coset Q correspond to all distinguishable ways of selecting j indices from the first $(N/2) + S$ indices and j indices from the last $(N/2) - S$ indices, both arranged in numerical order, and then interchanging the two sets. The elements of Q thus are jth-order products of mutually commuting transposes, i.e., they belong to the class $(1^{N-2j}, 2^j)$. The index j varies from 0 to $(N/2) - S$, and there are

$$\binom{\frac{N}{2}+S}{j}\binom{\frac{N}{2}-S}{j}$$

elements of Q with the same index j, where $\binom{\frac{N}{2}+S}{j}$ represents the binomial coefficient. It is convenient to define the operator $Q_j \equiv \sum q$ with the same index j.

The coset Q constructed in such a manner has the following useful properties:

(1) Q_j commutes with S_λ and A_λ

$$[Q_j, S_\lambda] = 0 = [Q_j, A_\lambda];$$

(2) Any product qQ contains one and only one element from each coset;

(3) The coefficient $U^S_{11}(q)$ depends only on the index j of q and is

$$U^S_{11}(j) = (-1)^j \Big/ \binom{\frac{N}{2}+S}{j}.$$

The proofs of these statements are given in Appendix A.

The desired antisymmetric projection operator in the approximation of no spin–orbit coupling is

$$\{A\}(\tilde{e}^S_{11})[e^S_{11}] = \mathcal{N} \sum_{Q,Q'} (-1)^q U^S_{11}(qq')(q)(A_\lambda)[q'][S_\lambda],$$

where $\mathcal{N}$ is an unimportant normalizing constant. Let Σ be a spin function that has maximum projected spin value along some axis $M_s = S$. The function is chosen to have spin up in the first $(N/2) + S$ positions and spin down in the remaining $(N/2) - S$ positions, and thus is already S_λ symmetric.

Consider the antisymmetric state projected from the product of this function with a general orbital function Φ. The expectation value with respect to this state of an operator symmetric under the permutation of the spin and orbital coordinates and which does not

[6] A. Messiah, *Quantum Mechanics* (John Wiley & Sons, New York, 1962), Vol. II, Appendix D, p. 1119.

[7] R. Gouavne, *Theory of Groups in Classical and Quantum Physics*, translated by T. Kahan (Oliver and Boyd, London, Edinburgh, 1965), Theorem 9, p. 287.

cause spin flip can be reduced to

$$\frac{\sum_{j,j'=0}^{(N/2)-S}\binom{\frac{N}{2}+S}{j}^{-1}\binom{\frac{N}{2}+S}{j'}^{-1}\langle(Q_{j'})(A_\lambda)\Phi|\langle\Sigma|O_p|\Sigma\rangle|(Q_j)(A_\lambda)\Phi\rangle}{\binom{N}{\frac{N}{2}-S}\frac{1}{n_S}\sum_{j=0}^{(N/2)-S}\binom{\frac{N}{2}+S}{j}^{-1}\langle(A_\lambda)\Phi\,|\,(Q_j)(A_\lambda)\Phi\rangle}, \tag{1}$$

where the group orthogonality relations and the relation

$$\langle[q']\Sigma|\,O_p\,|[q]\Sigma\rangle = \delta_{q'q}(q)\langle\Sigma|\,O_p\,|\Sigma\rangle(q)$$

have been used to accomplish the reduction. If the operator is purely orbital (e.g., the assumed Hamiltonian), one has the further reduction to

$$\frac{\sum_{j=0}^{(N/2)-S}\binom{\frac{N}{2}+S}{j}^{-1}\langle(A_\lambda)\Phi|\,O_{p\,\mathrm{orb}}\,|(Q_j)(A_\lambda)\Phi\rangle}{\sum_{j=0}^{(N/2)-S}\binom{\frac{N}{2}+S}{j}^{-1}\langle(A_\lambda)\Phi\,|\,(Q_j)(A_\lambda)\Phi\rangle}. \tag{2}$$

In the numerator and denominator of (1) and (2) above, one of the (A_λ) may be simply eliminated as this operator commutes with all the other operators. It is included here because in the next section $(A_\lambda)\Phi$ is treated as a unit.

II. THE ONE-PARTICLE APPROXIMATION

In this section the projection formulas developed above are applied within the one-particle approximation where the starting orbital function Φ is assumed to be a simple Nth-order product of one-particle functions. If the one-particle orbitals are determined by making the expectation value of the Hamiltonian an extremum, this is the projected Hartree–Fock scheme. We use the notation

$$(A_\lambda)\Phi \equiv d\varphi_{i_{n'},i_{n'}}\,d\varphi_{i_n,i_n},$$

where d signifies "the determinant of," i'_n is the numerically ordered set of the first $(N/2)+S$ indices, i_n is the numerically ordered set of the last $(N/2)-S$ indices, and φ_{i_n,i_n} is the matrix of one-electron orbitals $\varphi_i(\mathbf{r}_j)$ where the row indices denote the states and the column indices denote the coordinate variables (i.e., $d\varphi_{i_n,i_n}$ is a Slater determinant).

Because of the determinantal form, the orbitals within the same determinant can be assumed to be orthogonal.[8] Nothing is inferred as to the orthogonality of orbitals occurring in different determinants. Because of this lack of orthogonality, a special formalism for bookkeeping purposes is needed. Let $\{\theta_i\}$, $i = 1, \cdots, N$, be an orthogonal set that spans an N-dimensional space such that

$$\theta_i \equiv \varphi_i, \quad i = 1, \cdots, \frac{N}{2}+S,$$

$$\varphi_i \equiv \sum_{j=1}^{N} U_{ij}\theta_j, \quad i = \frac{N}{2}+S+1, \cdots, N.$$

Orthonormality requires

$$\sum_{j=1}^{N} U^*_{ij}U_{kj} = \delta_{ik}, \quad \frac{N}{2}+S < i, \quad \text{and} \quad k \leq N.$$

The second Slater determinant may be expanded in terms of the U_{ij}'s and the θ_i's by[9]

$$d\varphi_{i_{n'},i_n} = \sum_k dU_{i_{n'},k}\,d\theta_{k,i_n},$$

where k sums over all distinguishable sets of $(N/2)-S$ indices chosen from the N indices. The essential step for the analysis is noting that $(Q_j)(A_\lambda)\Phi$ corresponds to Laplace expansions of the two determinants by j columns and can be changed to a Laplace expansions of the two determinants by j rows. The sign factors involved in these expansions exactly compensate. Thus one can equivalently think of the permutations as acting on the coordinates or the state indices, and can write

$$(Q_j)(A_\lambda)\Phi = \sum_k dU_{i_n,k}\sum_{w,w'} d\theta_{i_{n'}+(w-w'),i_{n'}}\,d\theta_{k+(w'-w),i_n},$$

where w is summed over all distinguishable sets of j indices chosen from the set k of $(N/2)-S$ indices; w' is summed over all distinguishable sets of j indices chosen from the set i'_n; and the notation $(w-w')$ means the set w' is substituted by the set w and vice versa.

If we make the *ad hoc* restriction that all orbitals φ_i are completely expandable in the first $(N/2)+S$ orbitals, a nonzero result is obtained only when the set k and (thus) the set w also are contained within the set i'_n. For such choices the determinant $d\theta_{i_{n'}+(w-w'),i_{n'}}$ is nonzero only when the set w is identical with the set w'. Summing on w and w' multiplies the original determinants by an unimportant numerical factor which will cancel for normalized expectation values.

[8] In the variational calculation of the projected Hartree–Fock scheme, the "matrix" of Lagrange multipliers introduced by this orthonormality condition can be put in diagonal form, as can be shown by arguments identical to those used in the conventional Hartree–Fock scheme.

[9] G. Kowalewski, *Determinantentheorie* (Chelsea Publ. Co., London, 1948), Sec. 36.

We may conclude that if the orbitals associated with the minority spin $[\varphi_i;\ i = (N/2) + S + 1, \cdots, N]$ are restricted so as to be completely expandable in terms of the orbitals associated with the majority spin $[\varphi_i;\ i = 1, \cdots, (N/2) - S]$, then

$$(Q_j)(A_\lambda)\Phi = \mathcal{N}_j(A_\lambda)\Phi.$$

With this restriction a projected Hartree–Fock scheme is completely equivalent to an unprojected Hartree–Fock scheme.

While such an *ad hoc* restriction may not be serious for average properties such as one-electron energies, it would be untenable for investigating properties such as spin polarization, which depend strongly on the possibility that the minority spin oribtals cannot be completely expanded in terms of the majority spin orbitals.

Further reduction in the more general case can be accomplished by noting that the usual Hamiltonians contain, at most, two body operators. Only three cases for the set $i'_n + (w - w')$ are of interest:

Case 0: The set $i'_n + (w - w')$ is identical to the set i'_n;

Case 1: The set $i'_n + (w - w')$ differs from the set i'_n by one index $i'_n + (w - w') = i'_n + (\alpha - f)$, $f \in i'_n$ and $\kappa \in i_n$;

Case 2: The set $i'_n + (w - w')$ differs from the set i'_n by two indices $i'_n + (w - w') = i'_n + (\alpha + \beta - f - g)$, $g \in i'_n$ and $\beta \in i_n$.

Reduced forms for these cases are given in Appendix B.

III. DISCUSSION

While the forms we have presented for the antisymmetric projection operator assuming no spin–orbit coupling are complicated, they are, in principle, no more difficult than antisymmetrizing when spin–orbit coupling is significant. Correct projection seems desirable in order to distinguish results which follow due to the Pauli principle from those due to spin–orbit coupling. With modern computers the use of the forms presented here should not prove to be impractical.

The explicit factorization of the antisymmetrizer (A_λ) clearly shows that arguments based on the properties of this operator still are valid. Two particles associated with the same spin cannot occupy the same orbital position. The occurrence of the operators (Q_j) implies additional correlations in orbital space, but we can offer no simple interpretation for this at present. In the one-particle approximation the invariance of a Slater determinant to linear combination of its rows and columns has been used to argue that only the introduction of orbitals previously vacant can effect expectation values based on Slater determinants.[10] This is important in comparing an ionic situation with a covalent situation when the latter is described by linear combinations of the atomic orbitals (LCAO) used for describing the ionic configuration. Such qualitative arguments are seen to be unaffected by projection.

Because exchange is of secondary importance in determining the one-particle orbitals and their ionization energies in a variational scheme, one would expect that the orbitals associated with the minority spin could almost be expanded in terms of the orbitals associated with the majority spin. It follows from the results of Sec. II that SPHF or even restricted Hartree–Fock should not give results significantly different from projected Hartree–Fock for those properties that represent an average over all space, such as the one-electron ionization energies. However, in the SPHF theory a nonzero spin density at the nucleus in the transition elements is a direct consequence of the fact that the occupied spin-down s orbitals cannot be completely expanded in terms of the occupied spin-up s orbitals.[1] For a problem such as this, it is possible that the present scheme will give quantitative results significantly different from those of SPHF.[11]

Lastly, we wish to note that it seems desirable for qualitative reasons that a one-particle approximation scheme be such that the one-particle spatial symmetry group is the same as for the N-particle problem. Symmetry classification of the one-particle functions would significantly reduce the number of nonzero expansion coefficients that have to be considered in the expansions given in Sec. II.

ACKNOWLEDGMENTS

The author acknowledges a travel grant from the United States government under the Hays–Fulbright Act, some financial aid from the Conselho Nacional de Pesquisas do Brasil, and the hospitality of the Instituto da Física da Universidade Catolica do Rio de Janeiro, where this work was carried out.

APPENDIX A: PROPERTIES OF THE COSET Q

Let q_j be an element of Q with index j and h any element of the subgroup S_λ.

(a) The transformation $hq_jh^{-1} = q'_jh'$ is merely a substitution instruction for the indices appearing in q_j. The element q'_j must have the same index as q_j and h' must have even parity, because the transformation can neither change the index nor the parity of the permutation.

[10] R. E. Watson and A. J. Freeman, Phys. Rev. **134**, A1526 (1964).

[11] In Goddard's work (Ref. 3) his calculations of the lithium atom seem to support this conjecture, although the major part of the spin density is due to the single $2s$ orbital. The ^4S ground state of nitrogen would provide a more unambiguous test.

(b) By examining the possible results of multiplying q_j by a transpose contained in Q, one can conclude that

$$q_j'q_j = h'h^{-1} \quad \text{implies} \quad q_j' = q_j \quad \text{and} \quad h' = h.$$

(c) Because the group is closed we have

$$\mathcal{S}_N = q\mathcal{S}_N = qQ\mathcal{S}_\lambda = Q\mathcal{S}_\lambda.$$

From these statements one can conclude that the elements q_j do generate mutually exclusive right or left cosets, that the symmetrizer and antisymmetrizer of the subgroup commute with the operators Q_j

$$S_\lambda Q_j = Q_j S_\lambda \quad \text{and} \quad A_\lambda Q_j = Q_j A_\lambda,$$

and that qQ contains one and only one member from each coset.

One can deduce the coefficient $U_{11}^S(q_j)$ as follows:

(d) By statement (a) above we may infer that, for all q_j, q_j' with the same index j, there exists some element h of the subgroup such that

$$U_{11}^S(q_j') = \sum_{m,n} U_{1m}^S(h^{-1})U_{mn}^S(q_j)U_{n1}^S(h)$$
$$= U_{11}^S(q_j) = U_{11}^S(j).$$

This proves that the coefficient depends only on the index j.

(e) In our specific representation the basic group orthogonality relations take the form

$$\frac{N!}{n_s}\delta_{\lambda s}\delta_{iu}\delta_{jv} = \sum_{Q,m,n} U_{im}^S(q)U_{un}^\lambda(q)\sum_{\mathcal{S}_\lambda} U_{mj}^S(h)U_{nv}^\lambda(h),$$

where we have used the real and unitary property of the permutation group. For λ being the completely symmetric representation this becomes

$$\sum_{j=0}^{(N/2)-S}\binom{\frac{N}{2}-S}{j}\binom{\frac{N}{2}+S}{j}U_{11}^S(j) = 0.$$

For λ being the representation corresponding to S this becomes

$$\sum_{j=0}^{(N/2)-S}\binom{\frac{N}{2}-S}{j}\binom{\frac{N}{2}+S}{j}U_{11}^{S^2}(j) = \frac{\frac{N}{2}+S+1}{2S+1}.$$

Both these relations are satisfied by

$$U_{11}^S(j) = (-1)^j\binom{\frac{N}{2}+S}{j}^{-1}.$$

Although this cannot be strictly considered a proof, it can be shown by independent arguments that the formula is valid for $j = 0$, 1, and 2.

APPENDIX B: REDUCED FORMS IN THE ONE-ELECTRON APPROXIMATION

Let X_k be the number of indices of the set k that come from set i_n. Then we can write for

$$\sum_{j=0}^{(N/2)-S}\binom{\frac{N}{2}+S}{j}^{-1}(Q_j)(A_\lambda)\Phi$$

the following[12]:

Case 0:

$$\sum_k dU_{i_n,k}\left[1 - \frac{X_k}{2S+X_k+1}\right]d\theta_{i_{n'},i_{n'}}\,d\theta_{k,i_n},$$
$$0 \le X_k \le \frac{N}{2} - S.$$

Terms coming from the factor unity in the square bracket are entirely equivalent to unprojected Hartree–Fock. All other terms and those of Case 1 for one- and two-body operators and those of Case 2 for two-body operators are additional and particular to the projected Hartree–Fock scheme.

Case 1:

$$\sum_k dU_{i_n,k}\frac{2S+1}{(2S+X_k+1)(2S+X_k)} \times d\theta_{i_{n'}+(\alpha-f),i_{n'}}\,d\theta_{k+(f-\alpha),i_n}, \quad 1 \le X_k \le \frac{N}{2} - S,$$

where the sum on k is over the $\binom{N-2}{\frac{N}{2}-S-1}$ sets k that contain α and do not contain f.

Case 2:

$$\sum_k dU_{i_n,k}\frac{2(2S+1)}{(2S+X_k+1)(2S+X_k)(2S+X_k-1)} \times d\theta_{i_{n'}+(\alpha+\beta-f-g),i_{n'}}\,d\theta_{k+(f+g-\alpha-\beta),i_n},$$
$$2 \le X_k \le \frac{N}{2} - S,$$

where the sum on k is over the $\binom{N-4}{\frac{N}{2}-S-2}$ sets k that contain α and β and do not contain f and g.

[12] The sum formulas used can be derived from the identity

$$\frac{N}{n} = \sum_{j=0}^{N-n}\binom{j}{N-1}\binom{N-n}{j}$$

given in M. Boll, *Tables numeriques universelles* (Dunod et Cie., Paris, 1964), 3rd. ed., p. 539.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Irreducible Corepresentations of Groups Having a Compact Simple Lie Group as a Subgroup of Index 2

M. L. MEHTA AND P. K. SRIVASTAVA
Service de Physique Théorique, Centre d'Etudes Nucléaires de Saclay, Gif-sur-Yvette, France

(Received 29 November 1966)

For all possible extensions $\mathfrak{G}$ of the compact simple Lie groups G, such that G be a subgroup of $\mathfrak{G}$ of index 2, we determine those corepresentations of $\mathfrak{G}$ in which unitary operators correspond to the elements of the subgroup G, antiunitary operators to its coset. We find that any irreducible unitary representation of G can be extended to an irreducible corepresentation of $\mathfrak{G}$ in various ways summarized in Table IV.

1. INTRODUCTION

Consider a compact simple Lie group G and another group $\mathfrak{G}$ such that G is a subgroup of $\mathfrak{G}$ with index 2. The elements of G will be called the unitary operators and denoted by the letters $u, v, \cdots$. The elements of the coset of G in $\mathfrak{G}$ will be called the antiunitary operators and denoted by the letters $a, b, \cdots$. A set of unitary matrices $D(g)$, g an element of $\mathfrak{G}$, is said to form a corepresentation of $\mathfrak{G}$ if the following equations are satisfied:

$$D(u)D(v) = D(uv), \quad D(u)D(a) = D(ua),$$
$$D(a)D^*(u) = D(au), \quad D(a)D^*(b) = D(ab),$$

where D^* denotes the complex conjugate of D. The corepresentation $D(g)$ is said to be reducible if for a fixed matrix α, the matrices $D'(u)$, $D'(a)$,

$$D'(u) = \alpha^{-1}D(u)\alpha, \quad D'(a) = \alpha^{-1}D(a)\alpha^*,$$

all have the reduced structure

$$\begin{array}{|c|c|}\hline \xi & 0 \\ \hline 0 & \eta \\ \hline \end{array},$$

ξ and η being nonzero submatrices. If no such α can be found, $D(g)$ is said to be irreducible. In the above discussion, the matrix α may be assumed to be unitary.

An irreducible unitary corepresentation $D(g)$ of $\mathfrak{G}$ may belong to one of three possible categories[1]:

(i) $D(u)$ is irreducible. The corepresentation is said to be of type I.

(ii) $D(u)$ reduces into two irreducible equivalent representations. The corepresentation is said to be of type II.

(iii) $D(u)$ reduces into two nonequivalent irreducible representations of equal size. The corepresentation is said to be of type III.

The above three categories may also be enumerated as follows. We choose an arbitrary fixed antiunitary operator a_0 and denote an irreducible part of $D(u)$ by $\Delta(u)$. Then:

A. If $D(u)$ is of type I, $D(u) = \Delta(u)$, then $\Delta^*(a_0^{-1}ua_0)$ and $\Delta(u)$ are equivalent, $\Delta^*(a_0^{-1}ua_0) = \beta^{-1}\Delta(u)\beta$, and $\beta\beta^* = +\Delta(a_0^2)$. These equations are valid for all a if they are valid for one of them, a_0, though β depends on the choice of a_0. Moreover, β can be assumed to be unitary.

B. If $D(u)$ is of type II, then $D(u)$ may be transformed to

$$D(u) = \begin{bmatrix}\Delta(u) & 0 \\ 0 & \Delta(u)\end{bmatrix}.$$

$\Delta^*(a_0^{-1}ua_0)$ and $\Delta(u)$ are equivalent, $\Delta^*(a_0^{-1}ua_0) = \beta^{-1}\Delta(u)\beta$, and $\beta\beta^* = -\Delta(a_0^2)$. The remark made before about the choice of a_0 applies in this case and in the following one, also.

C. If $D(u)$ is of type III, it may be transformed to

$$\begin{bmatrix}\Delta(u) & 0 \\ 0 & \Delta^*(a_0^{-1}ua_0)\end{bmatrix},$$

where $\Delta^*(a_0^{-1}ua_0)$ and $\Delta(u)$ are not equivalent.

Note that the reducibility of only $D(u)$, u an element of G, is considered in the above; the $D(g)$ is of course taken to be irreducible. The $\Delta(g)$ is defined only when g is an element of G and a symbol like $\Delta(a_0)$ is neither defined nor used anywhere in this article.

The irreducible unitary representations $\Delta(u)$ can again be classified into three categories as follows[2]:

A. The representation $\Delta(u)$ is not equivalent to its complex conjugate $\Delta^*(u)$: $\Delta(u) \approx \Delta^*(u)$ (type C).[3]

B. It is equivalent to its complex conjugate but cannot be made real by any similarity transformation (type Q).[3]

[1] E. P. Wigner, *Group Theory and Its Applications to the Quantum Mechanics of Atomic Spectra* (Academic Press Inc., New York, 1959), Chap. 26.

[2] A. Loewy, Trans. Am. Math. Soc. **4**, 171 (1903); G. Frobenius and I. Schur, Sitzber. Deut. Akad. Wiss. Berlin, Kl. Math., Phys. Tech. 186 (1906); E. P. Wigner, Ref. 1, Chap. 24.

[3] F. J. Dyson, J. Math. Phys. **3**, 1199 (1962).

C. It can be made real by a similarity transformation (type R).[3]

The letters C, Q, and R correspond, respectively, to complex, quarternion, and real.[3]

For any compact simple Lie group G, all irreducible unitary representations $\Delta(u)$ are known.[4] In what follows we have tried to answer the natural question that for each possible group $\mathfrak{G}$ having G as a subgroup of index 2, what is the type of the irreducible corepresentation $D(g)$ such that the irreducible part of $D(u)$, u an element of G, is equivalent to a given $\Delta(u)$. In other words, we try to answer the following question: For a given group $\mathfrak{G}$ having G, any compact simple Lie group, as a subgroup of index 2 and a given irreducible representation $\Delta(u)$ of G whether $\Delta(u)$ and $\bar{\Delta}(u) \equiv \Delta^*(a_0^{-1}ua_0)$ are equivalent, and in case they are, $\bar{\Delta}(u) = \beta^{-1}\Delta(u)\beta$, whether $\beta\beta^* = +\Delta(a_0^2)$ or $\beta\beta^* = -\Delta(a_0^2)$.

2. INNER AND OUTER AUTOMORPHISMS

Let a_0, as before, be a fixed antiunitary operator and $u \in G$. Then $u \to F(u) = a_0^{-1}ua_0$ is an automorphism of G, which we shall call the automorphism induced by the antiunitary operator a_0. It may be an outer automorphism, i.e., no element $v_0 \in G$ can be found such that $a_0^{-1}ua_0 = v_0^{-1}uv_0$ for every $u \in G$. Or else the automorphism may be inner, i.e., there exists a v_0 with the above property. If the automorphism induced by one antiunitary operator of $\mathfrak{G}$ in G is outer (inner), then the automorphism induced by any other antiunitary operator of $\mathfrak{G}$ in G is also outer (inner). Any automorphism of a compact simple Lie group is either an inner automorphism or an inner automorphism together with one of a particular finite set of outer automorphisms.[5] For any given group this finite set of outer automorphisms is generated by those permutations of the simple roots[6] of its Lie algebra which preserve their scalar products: $\alpha_j \to \alpha_j'$, $(\alpha_j\alpha_k) \to (\alpha_j'\alpha_k')$ for every j and k. In other words, an outer automorphism of a semisimple Lie algebra is defined uniquely, except for an inner automorphism, by a symmetry operation on its Dynkin diagram.[7] The image of any other element of the algebra may be inferred from the commutation rules and the linearity of the automorphism. The following result is useful.

Theorem 2.1[5]: For the groups A_1, B_n, C_n, E_7, E_8, F_4, and G_2, all the automorphisms are inner. For the groups A_n $(n \geq 2)$, D_n $(n \geq 5)$, and E_6, there is an outer automorphism, unique up to an inner automorphism. For the group D_4 there are five distinct outer automorphisms.

Thus for any fixed a_0 we have $a_0^{-1}ua_0 = v_0^{-1}\bar{u}v_0$, for all u in G where v_0 is a fixed element of G, $\bar{u} = u$ if G is any of the groups A_1, B_n, C_n, E_7, E_8, F_4, G_2, and $\bar{u}$ is either equal to u or obtained from u by a symmetry operation of the Dynkin diagram if G is any of the remaining simple groups A_n $(n \geq 2)$, D_n $(n \geq 4)$, E_6. A glance at the Dynkin diagrams, Table I, shows that for all compact simple Lie groups, except possibly for D_4, $\bar{\bar{u}} = u$. The five outer automorphisms of D_4 correspond to the interchanges (12), (23), (31) and the cyclic permutations (123), (132) of the simple roots. For the first three automorphisms, we have $\bar{\bar{u}} = u$. For the last two, $\bar{\bar{u}} \neq u$, but they correspond to even permutations of the simple roots and therefore cannot be induced by antiunitary operators; in other words, there is no group $\mathfrak{G}$ having D_4 as a subgroup of index 2 corresponding to these two automorphisms.

Thus, for all the cases we have to consider, $\bar{\bar{u}} = u$. Putting $a_0' = a_0v_0^{-1}$ and calling it again a_0, we see that it satisfies the equations

$$a_0^{-1}ua_0 = \bar{u}, \quad a_0^{-2}ua_0^2 = \bar{\bar{u}} = u, \tag{2.1}$$

so that a_0^2 commutes with all the elements of G. Also, a_0^2 is an element of G, since G is a subgroup of index 2. Thus a_0^2 is an element of the center $C = \{c_j\}$ of G, say $a_0^2 = c_j$, where $\bar{c}_j = c_j$. If we put $a_1 = a_0c_1$ with c_1 any element of C, then this a_1 satisfies Eqs. (2.1) and $a_1^2 = (a_0c_1)^2 = a_0^2\bar{c}_1c_1 = c_j\bar{c}_1c_1$. Thus the choices $a_0^2 = c_j$ and $a_0^2 = c_j\bar{c}_1c_1$ lead to the same group $\mathfrak{G}$. On the other hand, if two elements c_j and c_k of the center C are such that $\bar{c}_j = c_j$, $\bar{c}_k = c_k$, and $c_j \neq c_k\bar{c}_1c_1$ for any element c_1 of C, then one can convince oneself that the choices $a_0^2 = c_j$ and $a_0^2 = c_k$ lead[8] to distinct groups $\mathfrak{G}$.

In the following, we take for G the groups obtained by exponentiation of the Lie algebras corresponding to the Dynkin diagrams given in Table I. For example, A_n denotes SU_{n+1}, the $(n+1)$-dimensional unimodular unitary group, while B_n denotes the universal covering group of the $(2n+1)$-dimensional orthogonal group.

[4] E. Cartan, Bull. Soc. Math. France **41**, 53 (1913); H. Weyl Math. Zeit. **23**, 271 (1925); **24**, 328, 377 (1925).

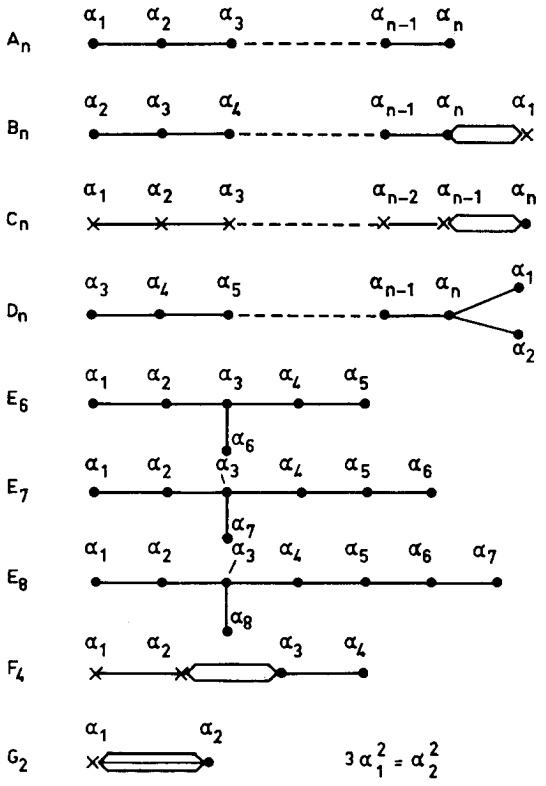
[5] E. Cartan, Bull. Soc. Math. France **49**, 361 (1925); E. B. Dynkin, Dokl. Akad. Nauk. SSSR **76**, 629 (1951); N. Jacobson, *Lie Algebras* (Interscience Publishers, Inc., New York, 1962), p. 281, Theorem 4.

[6] E. Cartan, *Oeuvres complètes* (Gauthiers-Villars, Paris, 1952), Pt. I, Vol. 1, p. 193, Theorem XI; E. B. Dynkin, Usp. Math. Nauk **20**, Ser. 2, 59 (1947) [Am. Math. Soc. transl. No. 17 (1950), paras. 6 and 7]; N. Jacobson, Ref. 5, Chap. IV, Secs. 3–6.

[7] E. B. Dynkin, Ref. 5; N. Jacobson, Ref. 5, Chap. IV, Sec. 5.

[8] L. Michel, *Group Theoretical Concepts and Methods in Elementary Particle Physics*, F. Gürsey, Ed. (Gordon and Breach, Science Publishers, Inc., New York, 1962).

TABLE I. Dynkin diagrams for the simple Lie groups. On a particular diagram, the simple roots denoted by dots are all equal in length, so also are those denoted by crosses, while the length of a root denoted by a dot is $\sqrt{2}$ times that of the one denoted by a cross. For the group G_2, the ratio of the lengths of its two simple roots is $\sqrt{3}$, as indicated.



The elements of the center of G can be written[9] as

$$c_j = \exp(4\pi i \mathbf{\Pi}_j \cdot \mathbf{H}/\boldsymbol{\alpha}_j \cdot \boldsymbol{\alpha}_j), \tag{2.2}$$

where $\boldsymbol{\alpha}_j$ is one of the l, l-dimensional simple roots, $\mathbf{\Pi}_j$ the fundamental dominant weight corresponding to it, defined by

$$2\mathbf{\Pi}_j \cdot \boldsymbol{\alpha}_k = (\boldsymbol{\alpha}_k \cdot \boldsymbol{\alpha}_k)\delta_{jk}, \quad k = 1, 2, \cdots, l, \tag{2.3}$$

and $\mathbf{H}$ is the l-dimensional element of the commuting subalgebra of the Lie algebra of G. If for every irreducible representation $\Delta(u)$, $\Delta(c_j) = \Delta(c_k)$, then c_j and c_k represent the same element of the center. The multiplication table and the behavior under outer automorphisms of the central elements c_j can all be inferred from their form (2.2). The fundamental dominant weights for various groups are listed in Table II.

[9] J. P. Serre, *Seminar Sophus Lie* (Ecole Normale Supérieure, Paris, 1954/55), exposé No. 23.

3. CLASSIFICATION OF COREPRESENTATIONS

The classification of all the irreducible corepresentations of a given group is greatly facilitated by the following theorems, which we give here without proof as they follow in much the same pattern as for the classification of irreducible representations.[10]

Theorem 3.1: Let $\Delta_j(u)$, $j = 1, 2, 3$ be three irreducible representations of G satisfying

$$\bar{\Delta}_j(u) \equiv \Delta_j^*(a_0^{-1}ua_0) = \beta_j^{-1}\Delta_j(u)\beta_j, \quad j = 1, 2, 3,$$

with

$$\beta_j\beta_j^* = \epsilon_j\Delta_j(a_0^2), \quad \epsilon_j = \pm 1, \quad j = 1, 2, 3.$$

Let $\Delta_3(u)$ occur in the reduction of the direct product $\Delta_1(u) \times \Delta_2(u)$. Then $\epsilon_1\epsilon_2\epsilon_3 = +1$.

Theorem 3.2: Let two irreducible representations $\Delta_1(u)$ and $\Delta_2(u)$ of G be such that $\Delta_1(u)$ and $\bar{\Delta}_1(u) \equiv \Delta_1^*(a_0^{-1}ua_0)$ are not equivalent, while $\Delta_2(u)$ and $\bar{\Delta}_2(u)$ are equivalent, i.e., $\Delta_2^*(a_0^{-1}ua_0) = \beta^{-1}\Delta_2(u)\beta$. If $\Delta_2(u)$ occurs in the reduction of $\Delta_1(u) \times \bar{\Delta}_1(u)$, then $\beta\beta^* = +\Delta_2(a_0^2)$.

An irreducible representation $\Delta(u)$ of any compact simple Lie group can be characterized by l nonnegative integers $\lambda_1, \cdots, \lambda_l$ so that

$$\mathbf{\Pi} = \sum_{j=1}^{l} \lambda_j \mathbf{\Pi}_j$$

is the highest weight of $\Delta(u)$, and $\Delta(u)$ occurs in the reduction of the direct product

$$\Delta_1^{\lambda_1} \times \Delta_2^{\lambda_2} \times \cdots \times \Delta_l^{\lambda_l},$$

where

$$\Delta_j^{\lambda_j} = \underbrace{\Delta_j \times \Delta_j \times \cdots \times \Delta_j}_{\lambda_j \text{ times}},$$

and $\Delta_j \equiv \Delta_j(u)$ is the irreducible representation having the highest weight $\mathbf{\Pi}_j$, the fundamental dominant weight defined by Eq. (2.3). In view of theorems 3.1 and 3.2, we need to study only the fundamental irreducible representations $\Delta_j(u)$ in order to make a statement about any irreducible representation $\Delta(u)$.

Theorem 3.3[10,11]: Any irreducible unitary representation $\Delta(u)$ of G falls into the following three classes depending on λ_i and G. These results are also summarized in Table III.

[10] M. L. Mehta, J. Math. Phys. **7**, 1824 (1966); M. L. Mehta and P. K. Srivastava, J. Math. Phys. **7**, 1833 (1966).

[11] A. I. Maltčev, Izv. Akad. Nauk SSSR. Ser. Mat. **8**, 143 (1944) [Am. Math. Soc. transl. No. 33 (1950)]; E. B. Dynkin, Tr. Mosk. Mat. Obsč. **1**, 39 (1952) [Am. Math. Soc. transl. **6**, Ser. 2, 245 (1957)]; N. Haruo, Sci. Rept. Tokyo Kyoiku Daigaku **A9**, Nos. 202–208, p. 32 (1965).

TABLE II. The fundamental dominant weights for the simple Lie groups corresponding to the simple roots as marked in Table I. They are expressed as linear combinations of the simple roots themselves and possess the property that $2(\Pi_j \cdot \alpha_k) = (\alpha_k \cdot \alpha_k)\delta_{jk}$, where the dot means the scalar product.

G	Fundamental Dominant Weights
A_n	$\Pi_j = \frac{1}{n+1}\left\{\sum_{k=1}^{j-1} k(n+1-j)\alpha_k + \sum_{k=j}^{n} j(n+1-k)\alpha_k,\ \ j = 1, 2, \cdots, n$
B_n	$\Pi_1 = \frac{1}{2}n\alpha_1 + \frac{1}{2}\sum_{k=2}^{n}(k-1)\alpha_k$ $\Pi_j = (j-1)\alpha_1 + \sum_{k=2}^{j-1}(k-1)\alpha_k + \sum_{k=j}^{n}(j-1)\alpha_k,\ \ j = 2, 3, \cdots, n$
C_n	$\Pi_j = \frac{1}{2}j\alpha_n + \sum_{k=1}^{j-1} k\alpha_k + \sum_{k=j}^{n-1} j\alpha_k, j = 1, 2, \cdots, n$
D_n	$\Pi_1 = \frac{1}{4}n\alpha_1 + \frac{1}{4}(n-2)\alpha_2 + \frac{1}{2}\sum_{k=3}^{n}(k-2)\alpha_k$ $\Pi_2 = \frac{1}{4}(n-2)\alpha_1 + \frac{1}{4}n\alpha_2 + \frac{1}{2}\sum_{k=3}^{n}(k-2)\alpha_k$ $\Pi_j = \frac{1}{2}(j-2)(\alpha_1+\alpha_2) + \sum_{k=3}^{j-1}(k-2)\alpha_k + \sum_{k=j}^{n}(j-2)\alpha_k,\ \ j = 3, 4, \cdots, n$
E_6	$\Pi_1 = \frac{1}{3}(4, 5, 6, 4, 2, 3) \equiv \frac{1}{3}(4\alpha_1 + 5\alpha_2 + 6\alpha_3 + 4\alpha_4 + 2\alpha_5 + 3\alpha_6)$ $\Pi_2 = \frac{1}{3}(5, 10, 12, 8, 4, 6),\ \ \Pi_3 = (2, 4, 6, 4, 2, 3)$ $\Pi_4 = \frac{1}{3}(4, 8, 12, 10, 5, 6),\ \ \Pi_5 = \frac{1}{3}(2, 4, 6, 5, 4, 3)$ $\Pi_6 = (1, 2, 3, 2, 1, 2)$
E_7	$\Pi_1 = (2, 3, 4, 3, 2, 1, 2) \equiv 2\alpha_1 + 3\alpha_2 + 4\alpha_3 + 3\alpha_4 + 2\alpha_5 + \alpha_6 + 2\alpha_7$ $\Pi_2 = (3, 6, 8, 6, 4, 2, 4),\ \ \Pi_3 = (4, 8, 12, 9, 6, 3, 6)$ $\Pi_4 = \frac{1}{2}(6, 12, 18, 15, 10, 5, 9),\ \ \Pi_5 = (2, 4, 6, 5, 4, 2, 3)$ $\Pi_6 = \frac{1}{2}(2, 4, 6, 5, 4, 3, 3),\ \ \Pi_7 = \frac{1}{2}(4, 8, 12, 9, 6, 3, 7)$
E_8	$\Pi_1 = (4, 7, 10, 8, 6, 4, 2, 5) \equiv 4\alpha_1 + 7\alpha_2 + 10\alpha_3 + 8\alpha_4 + 6\alpha_5 + 4\alpha_6 + 2\alpha_7 + 5\alpha_8$ $\Pi_2 = (7, 14, 20, 16, 12, 8, 4, 10)\ \ \Pi_3 = (10, 20, 30, 24, 18, 12, 6, 15)$ $\Pi_4 = (8, 16, 24, 20, 15, 10, 5, 12)\ \ \Pi_5 = (6, 12, 18, 15, 12, 8, 4, 9)$ $\Pi_6 = (4, 8, 12, 10, 8, 6, 3, 6)\ \ \Pi_7 = (2, 4, 6, 5, 4, 3, 2, 3)$ $\Pi_8 = (5, 10, 15, 12, 9, 6, 3, 8)$
F_4	$\Pi_1 = (2, 3, 2, 1) \equiv 2\alpha_1 + 3\alpha_2 + 2\alpha_3 + \alpha_4$ $\Pi_2 = (3, 6, 4, 2)\ \ \Pi_3 = (4, 8, 6, 3)\ \ \Pi_4 = (2, 4, 3, 2)$
G_2	$\Pi_1 = 2\alpha_1 + \alpha_2\ \ \Pi_2 = 3\alpha_1 + 2\alpha_2$

A. It is of type C, $\Delta^*(u) \not\cong \Delta(u)$; if G is:

(i) either A_n $(n \geq 2)$ and $\lambda_j \neq \lambda_{n+1-j}$ for some j;
(ii) or D_{2k+1} and $\lambda_1 \neq \lambda_2$;
(iii) or E_6 and at least one of the equalities $\lambda_1 = \lambda_5$, $\lambda_2 = \lambda_4$ is not satisfied.

B. It is of type Q, if G is:

(i) either A_{2n-1}, $\lambda_j = \lambda_{2n-j}$, $j = 1, 2, \cdots, 2n-1$, and $n\lambda_n$ is odd;
(ii) or B_{4m+1} or B_{4m+2}, with λ_1 odd;
(iii) or C_n and $\lambda_1 + \lambda_3 + \lambda_5 + \cdots$ is odd;
(iv) or D_{4k+2} and $\lambda_1 + \lambda_2$ is odd;
(v) or E_7 and $\lambda_4 + \lambda_6 + \lambda_7$ is odd.

C. It is of type R, if G is:

(i) either of the B_{4m}, B_{4m-1}, D_{4k}, G_2, F_4, E_8;
(ii) or A_{2n} with $\lambda_j = \lambda_{2n+1-j}$, $j = 1, 2, \cdots, 2n$;
(iii) or A_{2n-1} with $\lambda_j = \lambda_{2n-j}$, $j = 1, 2, \cdots, 2n-1$, and $n\lambda_n$ even;
(iv) or B_{4m+1} or B_{4m+2} with λ_1 even;

TABLE III. The types of the irreducible unitary representations of a simple Lie group having the highest weight $\sum \lambda_j \mathbf{\Pi}_j$, where $\mathbf{\Pi}_j$ are the fundamental dominant weights listed in Table II and λ_j are nonnegative integers.

G	Representation Types		
	R	Q	C
A_{2n-1}	$\lambda_j = \lambda_{2n-j}$, $j = 1, 2, \cdots, n$ and $n\lambda_n$ even	$\lambda_j = \lambda_{2n-j}$, $j = 1, 2, \cdots, n$ and $n\lambda_n$ odd	$\lambda_j \neq \lambda_{2n-j}$, for some j, $j = 1, 2, \cdots, n$
A_{2n}	$\lambda_j = \lambda_{2n+1-j}$, $j = 1, 2, \cdots, n$	No	$\lambda_j \neq \lambda_{2n+1+j}$ for some j, $j = 1, 2, \cdots, n$
B_n	$\frac{1}{2}n(n+1)\lambda_1$ even	$\frac{1}{2}n(n+1)\lambda_1$ odd	No
C_n	$\sum_1^{\leq \frac{1}{2}n+1} \lambda_{2j-1}$ even	$\sum_1^{\leq \frac{1}{2}n+1} \lambda_{2j-1}$ odd	No
D_{2n-1}	$\lambda_1 = \lambda_2$	No	$\lambda_1 \neq \lambda_2$
D_{2n}	$n(\lambda_1 + \lambda_2)$ even	$n(\lambda_1 + \lambda_2)$ odd	No
E_6	$\lambda_j = \lambda_{6-j}$, $j = 1, 2$	No	either $\lambda_1 \neq \lambda_5$ or $\lambda_2 \neq \lambda_4$ or both
E_7	$\lambda_4 + \lambda_6 + \lambda_7$ even	$\lambda_4 + \lambda_6 + \lambda_7$ odd	No
E_8	Yes	No	No
F_4	Yes	No	No
G_2	Yes	No	No

(v) or C_n with $\lambda_1 + \lambda_3 + \lambda_5 + \cdots$ even;
(vi) or D_{4k+2} with $\lambda_1 + \lambda_2$ even;
(vii) or D_{2k+1} with $\lambda_1 = \lambda_2$;
(viii) or E_6 with $\lambda_1 = \lambda_5$ and $\lambda_2 = \lambda_4$;
(ix) or E_7 with $\lambda_4 + \lambda_6 + \lambda_7$ even.

This list exhausts all possibilities for G.

The discussion in Sec. 2 gives us the possibility of defining all the group extensions $\mathfrak{G}$ of G, and classifying their various corepresentations. One of the possible extensions is always defined by adding an a_0 which induces an inner automorphism on G and $a_0^2 = 1$. For this case one may make a general remark.

Remark 3.1: For those group extensions $\mathfrak{G}$ of G, where $a_0^{-1} u a_0 = u$ and $a_0^2 = 1$, the reality classes R, Q, and C coincide with the corepresentation types I, II, and III. Henceforth we leave this case from our discussions; the summary in Table IV, however, includes all these results as well.

A_{n-1} or SU_n: Let G be SU_n, and let the automorphism induced by the antiunitary operators be inner. We choose a_0 such that $\bar{u} \equiv a_0^{-1} u a_0 = u$. The center of SU_n is the discrete cyclic group Z_n and consists of the elements

$$1 \quad \text{and} \quad c_k = \exp\left[4\pi i (\mathbf{\Pi}_k \cdot \mathbf{H})/(\boldsymbol{\alpha}_k \cdot \boldsymbol{\alpha}_k)\right], \quad k = 1, 2, \cdots, n-1, \tag{3.1}$$

$$c_k = c_1^k, \quad c_1^n = 1. \tag{3.2}$$

For finding the number of distinct group extensions we have to find the order of the factor group Z_n/Z_n^2.

For n odd, $Z_n^2 = Z_n$, i.e., every element of the center C is the square of some other element of C; $c_{2j} = c_j^2$, $c_{2j-1} = c_{j+\frac{1}{2}(n-1)}^2$. Thus from the discussion after Eq. (2.1), there is only one group extension $\mathfrak{G}$ for SU_n, n odd, and may be characterized by $a_0^2 = 1$. For n even, there are two such possibilities characterized by $a_0^2 = 1$ and $a_0^2 = c_1$; every other element of Z_n can be written either as c_j^2 or as $c_1 c_j^2$. The cases $a_0^2 = 1$ are covered by remark 3.1. For the case n even, $n = 2m$ and $a_0^2 = c_1$, the corepresentation is of type III if $\Delta(u)$ is of class C. If $\Delta(u)$ is of class R or Q, i.e., if $\lambda_j = \lambda_{2m-j}$, $j = 1, 2, \cdots, m$, then $\Delta(c_1) = (-1)^{\lambda_m}$. Thus the corepresentation is of type I or II, accordingly as $(m+1)\lambda_m$ is even or odd, provided that $\lambda_j = \lambda_{2m-j}$, $j = 1, 2, \cdots, m$.

A familiar example in this case is $G = SU_2$. The automorphisms are all inner. The group $\mathcal{A}_{\text{in}}^{(1)}(1)$

TABLE IV. The types of irreducible unitary corepresentations D of a group $\mathfrak{G}$ such that a compact simple Lie group G is a subgroup of $\mathfrak{G}$ with index 2, while an irreducible part of the representation obtained by restricting D to G is equivalent to a given irreducible unitary representation Δ of G. The highest weight of Δ is $\sum \lambda_j \mathbf{\Pi}_j$ as in Table III. The words "inner" or "outer" refer to the automorphism of G induced by any of the elements of the coset of G in $\mathfrak{G}$. The automorphism induced by an operator a_0, given by a symmetry operation on the Dynkin diagram, and the value of a_0^2 fix $\mathfrak{G}$. In particular, for $G = D_4$, the outer automorphism induced by a_0 is $\alpha_1 \leftrightarrow \alpha_2$; the results for the automorphisms $\alpha_2 \leftrightarrow \alpha_3$ and $\alpha_1 \leftrightarrow \alpha_3$ can be simply obtained by cyclically permuting the indices of λ.

G	Center	Automorphism	$\mathfrak{G}$	a_0	Corepresentation Types I	Corepresentation Types II	Corepresentation Types III
A_1	$1, c_1 = -1$	inner	$\mathcal{A}_{\text{in}}^{(1)}(1)$	1	λ_1 even	λ_1 odd	No
			$\mathcal{A}_{\text{in}}^{(c_1)}(1)$	c_1	Yes	No	No
A_{2n-1} $n \geq 2$	$1, c_1, c_2, \cdots, c_{2n-1}$; $c_i^j = c_j$, $c_1^{2n} = 1$	inner	$\mathcal{A}_{\text{in}}^{(1)}(2n-1)$	1	$\lambda_j = \lambda_{2n-j}, j = 1, 2, \cdots, n$; and $n\lambda_n$ even	$\lambda_j = \lambda_{2n-j}, j = 1, 2, \cdots, n$; and $n\lambda_n$ odd	$\lambda_j \neq \lambda_{2n-j}$ for some $j, j = 1, 2, \cdots, n$
			$\mathcal{A}_{\text{in}}^{(c_1)}(2n-1)$	c_1	$\lambda_j = \lambda_{2n-j}, j = 1, 2, \cdots, n$; and $(n+1)\lambda_n$ even	$\lambda_j = \lambda_{2n-j}, j = 1, 2, \cdots, n$; and $(n+1)\lambda_n$ odd	$\lambda_j \neq \lambda_{2n-j}$ for some $j, j = 1, 2, \cdots, n$
		outer	$\mathcal{A}_{\text{out}}^{(1)}(2n-1)$	1	$\sum_1^n \lambda_{2j-1}$ even	$\sum_1^n \lambda_{2j-1}$ odd	No
			$\mathcal{A}_{\text{out}}^{(c_n)}(2n-1)$	c_n	Yes	No	No
A_{2n} $n \geq 1$	$1, c_1, c_2, \cdots, c_{2n}$; $c_i^j = c_j$, $c_1^{2n+1} = 1$	inner	$\mathcal{A}_{\text{in}}^{(1)}(2n)$	1	$\lambda_j = \lambda_{2n+1-j}, j = 1, 2, \cdots, n$;	No	$\lambda_j \neq \lambda_{2n+1-j}$, for some $i, j = 1, 2, \cdots, n$,
		outer	$\mathcal{A}_{\text{out}}^{(1)}(2n)$	1	Yes	No	No
B_n	$1, c_1 = -1$	inner	$\mathcal{B}_{\text{in}}^{(1)}(n)$	1	$\frac{1}{2}n(n+1)\lambda_1$ even	$\frac{1}{2}n(n+1)\lambda_1$ odd	No
			$\mathcal{B}_{\text{in}}^{(c_1)}(n)$	c_1	$\frac{1}{2}n(n+3)\lambda_1 + \sum_1^{\leq \frac{1}{2}n+1} \lambda_{2j}$ even	$\frac{1}{2}n(n+3)\lambda_1 + \sum_1^{\leq \frac{1}{2}n+1} \lambda_{2j}$ odd	No
C_n	$1, c_1 = -1$	inner	$\mathcal{C}_{\text{in}}^{(1)}(n)$	1	$\sum_1^{\leq \frac{1}{2}n+1} \lambda_{2j-1}$ even	$\sum_1^{\leq \frac{1}{2}n+1} \lambda_{2j-1}$ odd	No
			$\mathcal{C}_{\text{in}}^{(c_1)}(n)$	c_1	Yes	No	No

D_{2n-1} $n \geq 3$	$1, c_1, c_2, c_3,$ $c_1^2 = c_2^2 = c_3,$ $c_1c_2 = c_3^2 = 1$	inner	$\mathfrak{D}_{\text{in}}^{(1)}(2n-1)$	1	$\lambda_1 = \lambda_2$	No	$\lambda_1 \neq \lambda_2$
			$\mathfrak{D}_{\text{in}}^{(c_1)}(2n-1)$	c_1	$\lambda_1 = \lambda_2, \sum_2^n \lambda_{2j-1}$ even	$\lambda_1 = \lambda_2, \sum_2^n \lambda_{2j-1}$ odd	$\lambda_1 \neq \lambda_2$
		outer	$\mathfrak{D}_{\text{out}}^{(1)}(2n-1)$	1	$(n-1)(\lambda_1 + \lambda_2)$ even	$(n-1)(\lambda_1 + \lambda_2)$ odd	No
			$\mathfrak{D}_{\text{out}}^{(c_3)}(2n-1)$	c_3	$n(\lambda_1 + \lambda_2)$ even	$n(\lambda_1 + \lambda_2)$ odd	No
D_{2n} $n \geq 2$	$1, c_1, c_2, c_3,$ $c_1^2 = c_2^2 = c_3^2 = 1,$ $c_1c_2c_3 = 1$	inner	$\mathfrak{D}_{\text{in}}^{(1)}(2n)$	1	$n(\lambda_1 + \lambda_2)$ even	$n(\lambda_1 + \lambda_2)$ odd	No
			$\mathfrak{D}_{\text{in}}^{(c_1)}(2n)$	c_1	$\lambda_2 + \sum_2^n \lambda_{2j-1}$ even	$\lambda_2 + \sum_2^n \lambda_{2j-1}$ odd	No
			$\mathfrak{D}_{\text{in}}^{(c_2)}(2n)$	c_2	$\lambda_1 + \sum_2^n \lambda_{2j-1}$ even	$\lambda_1 + \sum_2^n \lambda_{2j-1}$ odd	No
			$\mathfrak{D}_{\text{in}}^{(c_3)}(2n)$	c_3	$(n+1)(\lambda_1 + \lambda_2)$ even	$(n+1)(\lambda_1 + \lambda_2)$ odd	No
		outer	$\mathfrak{D}_{\text{out}}^{(1)}(2n)$	1	$\lambda_1 = \lambda_2$	No	$\lambda_1 \neq \lambda_2$
E_6	$1, c_1, c_2 = c_1^2,$ $c_1^3 = 1$	inner	$\mathfrak{E}_{\text{in}}^{(1)}(6)$	1	$\lambda_j = \lambda_{6-j}, j = 1, 2$	No	either $\lambda_1 \neq \lambda_5$ or $\lambda_2 \neq \lambda_4$ or both
		outer	$\mathfrak{E}_{\text{out}}^{(1)}(6)$	1	Yes	No	No
E_7	$1, c_4 = -1$	inner	$\mathfrak{E}_{\text{in}}^{(1)}(7)$	1	$\lambda_4 + \lambda_6 + \lambda_7$ even	$\lambda_4 + \lambda_6 + \lambda_7$ odd	No
			$\mathfrak{E}_{\text{in}}^{(c_4)}(7)$	c_4	Yes	No	No
E_8	1	inner	$\mathfrak{E}_{\text{in}}^{(1)}(8)$	1	Yes	No	No
F_4	1	inner	$\mathfrak{F}_{\text{in}}^{(1)}(4)$	1	Yes	No	No
G_2	1	inner	$\mathfrak{G}_{\text{in}}^{(1)}(2)$	1	Yes	No	No

consists of the elements u, a_0u, where u is the set of all 2×2 unitary matrices with determinant 1. The multiplication table is

	u_2	a_0u_2
u_1	u_1u_2	$a_0u_1u_2$
a_0u_1	$a_0u_1u_2$	u_1u_2

(3.3)

and has the $4m \times 4m$ irreducible unitary corepresentation of type II:

$$D(u) = \begin{bmatrix} \Delta(u) & 0 \\ 0 & \Delta(u) \end{bmatrix}, \quad D(a_0) = \begin{bmatrix} 0 & -\beta \\ \beta & 0 \end{bmatrix}, \tag{3.4}$$

where $\Delta(u)$ is the $2m \times 2m$ irreducible unitary representation of SU_2 and the $2m \times 2m$ matrix β is

$$\beta_{jk} = (-1)^j \delta_{j+k,2m+1}. \tag{3.5}$$

It is easy to convince oneself that it is a corepresentation. To see that it is irreducible, let, if possible, α reduce it. Writing α also in the partitioned form of $2m \times 2m$ blocks, we have

$$\alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \tag{3.6}$$

This means that the nonsingular α satisfies

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} A(u) & 0 \\ 0 & B(u) \end{bmatrix} = \begin{bmatrix} \Delta(u) & 0 \\ 0 & \Delta(u) \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

and

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} \xi & 0 \\ 0 & \eta \end{bmatrix} = \begin{bmatrix} 0 & -\beta \\ \beta & 0 \end{bmatrix}\begin{bmatrix} a^* & b^* \\ c^* & d^* \end{bmatrix}; \tag{3.7}$$

or

$$aA(u) = \Delta(u)a, \quad cA(u) = \Delta(u)c, \tag{3.8}$$

$$bB(u) = \Delta(u)b, \quad dB(u) = \Delta(u)d, \tag{3.9}$$

$$a\xi = -\beta c^*, \qquad c\xi = \beta a^*, \tag{3.10}$$

$$b\eta = -\beta d^*, \qquad d\eta = \beta b^*. \tag{3.11}$$

As $A(u)$ and $\Delta(u)$ are irreducible representations of the group SU_2, we see by Schur's Lemma that c is either nonsingular or zero. The same statement is true for the matrices a, b, and d. If $c = 0$, then from Eqs. (3.10) $a = 0$ and α is singular. Thus c is nonsingular. Similarly, each of the matrices a, b, and d are nonsingular. Now from Eqs. (3.8) one sees that ac^{-1} commutes with every $\Delta(u)$ and therefore, by Schur's Lemma, $a = kc$, k a constant. Substituting this in Eqs. (3.10), one gets after a little manipulation $kk^* = -1$, an absurdity.

The multiplication table for $\mathcal{A}_{\text{in}}^{(c_1)}(1)$ is

	u_2	a_0u_2
u_1	u_1u_2	$a_0u_1u_2$
a_0u_1	$a_0u_1u_2$	$-u_1u_2$

and can be realized by choosing $a_0 = \beta K$, where K is the complex conjugation operator and

$$\beta = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{3.12}$$

is an element of SU_2. It is well known[1] that all corepresentations of this $\mathcal{A}_{\text{in}}^{(c_1)}(1)$ are of type I.

Consider now the possible group extensions with the outer automorphism, which we take as the reflection of the Dynkin diagram $\boldsymbol{\alpha}_j \rightleftharpoons \boldsymbol{\alpha}_{n-j}$, $j = 1, 2, \cdots, n-1$. The elements of the center (3.1) undergo the same automorphism and $\bar{c}_j = c_{n-j}$. In the case of n being odd, there is no invariant element in the center besides 1, and hence the value of a_0^2 can only be 1. However, for n even, $n = 2m$. In addition to 1, c_m is invariant. Also, as $\bar{c}_k c_k = c_{2m-k}c_k = 1$ for all k, $a_0^2 = 1$ and $a_0^2 = c_m$ give rise to different extensions of SU_{2m}. We now determine the corepresentation type for the irreducible n-dimensional representation with the highest weight $\boldsymbol{\Pi}_1$. The n weights $\boldsymbol{\Pi}_1^k$ of the representation are given by

$$\boldsymbol{\Pi}_1^1 = \boldsymbol{\Pi}_1, \boldsymbol{\Pi}_1^{k+1} = \boldsymbol{\Pi}_1 - \sum_{j=1}^{k} \boldsymbol{\alpha}_j, \quad k = 1, 2, \cdots, n-1.$$

We choose a basis so that the matrices of $\mathbf{H}$ and E_{α_i} are

$$(\mathbf{H})_{kk'} = \boldsymbol{\Pi}^k \delta_{kk'}, \quad (E_{\alpha_i})_{kk'} = \delta_{ki}\delta_{k',i+1}; \tag{3.13}$$

the other matrices of the representation can be found by calculating their commutators. We want to determine whether there exists a nonsingular β such that

$$\beta \Delta_1^*(a_0^{-1}ua_0) = \Delta_1(u)\beta. \tag{3.14}$$

This implies for the matrices of E_{α_i} the following:

$$-\beta E_{\alpha_i}^T = E_{\alpha_{n-i}}\beta. \tag{3.15}$$

Using explicit expressions for $(E_{\alpha_i})_{kk'}$, we get the following relations:

$$\beta_{0,n-1} = -\beta_{1,n-2} = \beta_{2,n-3} = \cdots = (-1)^{(n-1)}\beta_{n-1,0},$$
$$\beta_{jk} = 0, \quad \text{if} \quad j + k \neq n - 1. \tag{3.16}$$

Thus β is symmetric or antisymmetric, accordingly as the sign $(-1)^{(n-1)}$ is even or odd. Also, as β is unitary,

$$\beta\beta^* = (-1)^{(n-1)}. \tag{3.17}$$

Thus if $a_0^2 = 1$, we have

$$\beta\beta^* = (-1)^{(n-1)}\Delta(a_0^2). \tag{3.18}$$

For $n = 2m$ the value of the center element c_m can be calculated in the above representation as

$$\Delta_1(c_m) = \exp[2\pi_i(2\mathbf{\Pi}_m \cdot \mathbf{\Pi}_1/\boldsymbol{\alpha}_m \cdot \boldsymbol{\alpha}_m)] = (-1). \tag{3.19}$$

Using these two facts, one gets that all the corepresentations are of type I, when n is odd ($a_0^2 = 1$) or when n is even and $a_0^2 = c_m$. For $n = 2m$ and $a_0^2 = 1$, the corepresentation containing $\Delta(u)$ is of type I or II accordingly as $(\lambda_1 + \lambda_3 + \lambda_5 + \cdots)$ is even or odd.

Remark 3.2: The above result seems to be in contradiction with what one may have naively guessed. The outer automorphism may have been defined as that of complex conjugation; then one would have immediately implied $\Delta_1^*(\bar{u}) = \Delta_1(u)$, $\Delta_1(u)$ being identical to u. Thus Δ is of type I for $a_0^2 = 1$ regardless of n being even or odd. This apparent contradiction is easily removed by the following consideration.

Let a_0 and a_1 be two antiunitary operators of the same group $\mathfrak{G}$ and let $a_0^{-1}ua_0 = \bar{u}$ and $a_1^{-1}ua_1 = u'$, with $\bar{u} = u'' = u$. The element $v_0 = a_0^{-1}a_1$ belongs to G and $v_0^{-1}\bar{u}v_0 = u'$ for every u in G. Also, $a_1^2 = (a_0v_0)^2 = a_0^2\bar{v}_0v_0$. Thus, if there is an element v_0 of G such that $\bar{v}_0v_0$, an element of the center of G, is different from unity, then one may choose a_1^2 different from a_0^2. For n even, this is the case, because there exist antisymmetric unitary matrices v_0 with $v_0^*v_0 = -1$. One can in fact verify that the reflection of the Dynkin diagram and complex conjugation of the self-representation of A_{n-1} or SU_n differ exactly by such a transformation.

B_n: For these groups all the automorphisms are inner and the center consists of two elements 1 and c_1. The representative of c_1 in $\Delta(u)$ is

$$\Delta(c_1) = (-1)^{n\lambda_1+\lambda_2+\lambda_4+\cdots}. \tag{3.20}$$

For $a_0^2 = 1$ the case is covered by remark 3.1, and for $a_0^2 = c_1$, the corepresentation is of type I or II accordingly as

$$\epsilon = (-1)^{\frac{1}{2}n(n+3)\lambda_1+\lambda_2+\lambda_4+\cdots} \tag{3.21}$$

is $+1$ or -1.

C_n or Sp_{2n}: For these groups all the automorphisms are inner and the center consists of two elements, 1 and c_1. The representative of c_1 in the representation with highest weight $\sum \lambda_j\mathbf{\Pi}_j$ is

$$\Delta(c_1) = (-1)^{\lambda_1+\lambda_3+\lambda_5+\cdots}. \tag{3.22}$$

Combining with the results in Table III, we see that for the case $a_0^2 = c_1$, all the corepresentations are of type I, while the case $a_0^2 = 1$ is covered by remark 3.1.

D_n: The center consists of four elements 1, c_1, c_2, and c_3. The representatives of these in a representation with the highest weight $\sum \lambda_k\mathbf{\Pi}_k$ are

$$\begin{aligned}
\Delta(c_1) &= \exp[i\tfrac{1}{2}\pi(\lambda_1 n + \lambda_2(n-2))](-1)^{\lambda_3+\lambda_5+\cdots},\\
\Delta(c_2) &= \exp[i\tfrac{1}{2}\pi(\lambda_1(n-2) + \lambda_2 n)](-1)^{\lambda_3+\lambda_5+\cdots},\\
\Delta(c_3) &= (-1)^{\lambda_1+\lambda_2}.
\end{aligned} \tag{3.23}$$

The multiplication laws for n even and n odd are different, as can be verified from Eqs. (3.23). For n odd the group is Z_4,

$$c_1^2 = c_3, \quad c_1^3 = c_2, \quad c_1^4 = 1, \tag{3.24}$$

and for n even, the group is $Z_2 \times Z_2$,

$$c_1^2 = c_2^2 = c_3^2 = 1 \quad \text{and} \quad c_1c_2c_3 = 1. \tag{3.25}$$

First consider the inner automorphisms. Let n be odd. From (3.24) and the discussion after Eq. (2.1), one sees that there are two extensions $\mathfrak{G}$, characterized by $a_0^2 = 1$ and $a_0^2 = c_1$. The case $a_0^2 = 1$ is already covered by remark 3.1. For $a_0^2 = c_1$, $\Delta(u)$ and $\Delta^*(u)$ are equivalent only when $\lambda_1 = \lambda_2$. Thus, if $\lambda_1 \neq \lambda_2$, the corepresentation is of type III. If $\lambda_1 = \lambda_2$, then $\Delta^*(u) = \beta^{-1}\Delta(u)\beta$ with $\beta\beta^* = 1, \Delta(c_1) = (-1)^{\lambda_3+\lambda_5+\cdots}$, and therefore the corepresentation is of type I or II accordingly as $\lambda_3 + \lambda_5 + \cdots$ is even or odd. Now let n be even, $n = 2m$. All irreducible representations $\Delta(u)$ satisfy $\Delta^*(u) = \beta^{-1}\Delta(u)\beta$ with

$$\beta\beta^* = (-1)^{m(\lambda_1+\lambda_2)}.$$

As the square of each element of the center is 1, each one of them used as a_0^2 gives a different extension $\mathfrak{G}$. Comparing $\beta\beta^* = (-1)^{m(\lambda_1+\lambda_2)}$ and $\Delta(a_0^2)$, we see that the corepresentation is of type I or II accordingly as ϵ, as given below, is $+1$ or -1:

$$\begin{aligned}
a_0^2 &= 1, \quad \epsilon = (-1)^{m(\lambda_1+\lambda_2)},\\
a_0^2 &= c_1, \quad \epsilon = (-1)^{\lambda_2+\lambda_3+\lambda_5+\cdots},\\
a_0^2 &= c_2, \quad \epsilon = (-1)^{\lambda_1+\lambda_3+\lambda_5+\cdots},\\
a_0^2 &= c_3, \quad \epsilon = (-1)^{(m+1)(\lambda_1+\lambda_2)}.
\end{aligned} \tag{3.26}$$

Next consider the outer automorphisms. We take $\bar{u}$ to be obtained from u by interchanging the simple roots $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$, so that $\bar{c}_1 = c_2$ and $\bar{c}_2 = c_1$, while $\bar{c}_3 = c_3$, $\bar{1} = 1$. In case n is odd, a_0^2 can be chosen to be either 1 or c_3, giving different group extensions $\mathfrak{G}$. On the other hand, when n is even, $\bar{c}_1c_1c_3 = c_2c_1c_3 = 1$ and there is only one extension $\mathfrak{G}$, characterized by $a_0^2 = 1$.

Now we have to see whether $\Delta^*(a_0^{-1}ua_0)$ and $\Delta(u)$ are equivalent and, in case they are, $\Delta^*(a_0^{-1}ua_0) = \beta^{-1}\Delta(u)\beta$, what is the value of $\beta\beta^*$? As any $\Delta(u)$ occurs in the reduction of a direct product of certain direct (or Kronecker) powers of $\Delta_1(u)$, $\Delta_2(u)$, and $\Delta_3(u)$, in view of Theorems 3.1 and 3.2, it is sufficient for the above purpose to study only these last three

fundamental representations. Let us begin with $\Delta_3(u)$ with the highest weight

$$\Pi_3 = \tfrac{1}{2}(\alpha_1 + \alpha_2) + \sum_{j=3}^{n} \alpha_j .$$

All the weights are simply calculated to be

$$\begin{aligned} &\Pi_3^k = \tfrac{1}{2}(\alpha_1 + \alpha_2) + \sum_{j=k+2}^{n} \alpha_j, \quad k = 1, 2, \cdots, n-2, \\ &\Pi_3^{n-1} = \tfrac{1}{2}(\alpha_1 + \alpha_2), \qquad \Pi_3^n = \tfrac{1}{2}(\alpha_1 - \alpha_2), \\ &\Pi_3^{n+k} = -\Pi_3^{n-k+1}, \qquad k = 1, 2, \cdots, n. \end{aligned} \tag{3.27}$$

Setting a representation in which $\mathbf{H}$ is diagonal, the representatives of $E_i \equiv E_{\alpha_i}$ can be simply written as

$$\begin{aligned} (E_i)_{jk} &= 1, \quad \text{if} \quad \Pi_3^j - \Pi_3^k = \alpha_i, \\ &= 0, \quad \text{otherwise.} \end{aligned} \tag{3.28}$$

The automorphism $\alpha_1 \leftrightarrow \alpha_2$ just interchanges Π_3^n and Π_3^{n+1}. The equations $\beta\Delta^*(a_0^{-1}ua_0) = \Delta(u)\beta$ in terms of the E_j read

$$\begin{aligned} -\beta E_1^T &= E_2\beta, \quad -\beta E_2^T = E_1\beta, \\ -\beta E_j^T &= E_j\beta, \quad j = 3, 4, \cdots, n. \end{aligned} \tag{3.29}$$

Equations (3.27)–(3.29) give

$$\begin{aligned} &\beta_{n-1,n+2} = \beta_{n+2,n-1} = -\beta_{nn} = -\beta_{n+1,n+1}, \\ &\beta_{j,2n-j+1} = -\beta_{j+1,2n-j}, \\ &\qquad j = 1, \cdots, n-2, n+2, \cdots, 2n-1. \end{aligned}$$

All other elements of β are zero. Thus $\beta\beta^* = 1$.

The representations $\Delta_1(u)$ and $\Delta_2(u)$ are conveniently given in a basis labelled by $|\epsilon\rangle = |\epsilon_1, \epsilon_2, \cdots, \epsilon_n\rangle$, where $\epsilon_j = \pm 1$ independently of each other except for their product $\epsilon_1, \cdots, \epsilon_n$, which is $+1$ for Δ_1 and -1 for Δ_2. Thus

$$\langle\epsilon| H_j |\epsilon'\rangle = \tfrac{1}{2}\epsilon_j\delta(\epsilon, \epsilon') \equiv \tfrac{1}{2}\epsilon_j\delta(\epsilon_1, \epsilon_1') \cdots \delta(\epsilon_n, \epsilon_n'), \quad j = 1, 2, \cdots, n, \tag{3.30}$$

$$\langle\epsilon| E_1 |\epsilon'\rangle = \delta(\epsilon_1, \epsilon_1') \cdots \delta(\epsilon_{n-2}, \epsilon_{n-2}') \times \delta(\epsilon_{n-1}, \epsilon_{n-1}' + 2)\delta(\epsilon_n, \epsilon_n' + 2), \tag{3.31}$$

$$\langle\epsilon| E_2 |\epsilon'\rangle = \delta(\epsilon_1, \epsilon_1') \cdots \delta(\epsilon_{n-2}, \epsilon_{n-2}') \times \delta(\epsilon_{n-1}, \epsilon_{n-1}' + 2)\delta(\epsilon_n, \epsilon_n' - 2), \tag{3.32}$$

$$\langle\epsilon| E_j |\epsilon'\rangle = \delta(\epsilon_1, \epsilon_1') \cdots \delta(\epsilon_{j-2}, \epsilon_{j-2}' + 2) \times \delta(\epsilon_{j-1}, \epsilon_{j-1}' - 2) \cdots \delta(\epsilon_n, \epsilon_n'), \quad j = 3, 4, \cdots, n. \tag{3.33}$$

The automorphism $\alpha_1 \leftrightarrow \alpha_2$ interchanges E_1 and E_2 and, as a consequence, the matrices for H_j, given in terms of the commutator of the E's, change as

$$\bar{H}_n \equiv a_0^{-1}H_n a_0 = -H_n, \quad \bar{H}_j = H_j, \quad j = 1, 2, \cdots, n-1. \tag{3.34}$$

The equation $\beta\Delta^*(a_0^{-1}ua_0) = \Delta(n)\beta$ implies for the algebra

$$-\beta H_n = -H_n\beta, \quad -\beta H_j = H_j\beta, \quad j = 1, 2, \cdots, n-1, \tag{3.35}$$

$$-\beta E_1^T = E_2\beta, \quad -\beta E_2^T = E_1\beta, \tag{3.36}$$

$$-\beta E_j^T = E_j\beta, \quad j = 3, 4, \cdots, n. \tag{3.37}$$

From (3.30) and (3.35) one sees that the only possible nonzero elements of β are

$$\beta_{\epsilon_1 \cdots \epsilon_n, \bar{\epsilon}_1 \cdots \bar{\epsilon}_{n-1}\epsilon_n}, \tag{3.38}$$

where $\bar{\epsilon}_j = -\epsilon_j$. As the product of all the ϵ_j is fixed, this implies that $\beta = 0$ unless n is odd. When n is odd, we indicate the element (3.38) by $\beta(\epsilon_1, \cdots, \epsilon_n)$. Equation (3.37) then shows that

$$\beta(\epsilon_1, \cdots, \epsilon_{j-2}, 1, -1, \epsilon_{j+1}, \cdots, \epsilon_n) = -\beta(\epsilon_1, \cdots, \epsilon_{j-2}, -1, 1, \epsilon_{j+1}, \cdots, \epsilon_n) \tag{3.39}$$

and (3.36) gives

$$\beta(\epsilon_1, \cdots, \epsilon_{n-2}, 1, 1) = -\beta(\epsilon_1, \cdots, \epsilon_{n-2}, -1, -1). \tag{3.40}$$

These equations give after a little manipulation

$$\beta(\epsilon_1, \cdots, \epsilon_n) = (-1)^{\frac{1}{2}(n-1)}\beta(\bar{\epsilon}_1, \cdots, \bar{\epsilon}_{n-1}, \epsilon_n), \tag{3.41}$$

i.e.,

$$\beta^T = (-1)^{\frac{1}{2}(n-1)}\beta \quad \text{or} \quad \beta\beta^* = (-1)^{\frac{1}{2}(n-1)}, \tag{3.42}$$

taking β unitary. Thus for n odd, $n = 2m + 1$, the representation $\Delta(u)$ satisfies $\Delta^*(a_0^{-1}ua_0) = \beta^{-1}\Delta(u)\beta$ with

$$\beta\beta^* = (-1)^{m(\lambda_1+\lambda_2)}\Delta(a_0^2), \quad \text{if} \quad a_0^2 = 1, \tag{3.43}$$

and

$$\beta\beta^* = (-1)^{(m+1)(\lambda_1+\lambda_2)}\Delta(a_0^2), \quad \text{if} \quad a_0^2 = c_3. \tag{3.44}$$

For n even, the automorphism $\alpha_1 \leftrightarrow \alpha_2$ interchanges Δ_1 and Δ_2. Thus the corepresentation is of type III unless $\lambda_1 = \lambda_2$. When $\lambda_1 = \lambda_2$, it is of type I.

For D_4 and the outer automorphism $\alpha_1 \leftrightarrow \alpha_2$, the above conclusion holds. For the automorphism $\alpha_1 \leftrightarrow \alpha_3 (\alpha_2 \leftrightarrow \alpha_3)$, the corepresentation is of type III unless $\lambda_1 = \lambda_3$ ($\lambda_2 = \lambda_3$). In case $\lambda_1 = \lambda_3$ ($\lambda_2 = \lambda_3$) it is of type I.

E_6: For the group E_6, the center consists of three elements $Z_3 \equiv \{1, c_1, c_2 = c_1^2;\ c_1c_2 = 1\}$. As $c_1^2 = c_2$, there is only one extension $\mathfrak{G}$ for inner automorphisms, characterized by $a_0^2 = 1$. This case is covered by remark 3.1. The outer automorphism may be taken as $1 \leftrightarrow 5$, $2 \leftrightarrow 4$, $3 \to 3$, $6 \to 6$. As $\bar{c}_1 = c_2$, $\bar{c}_2 = c_1$, there is only one extension characterized by $a_0^2 = 1$. Therefore the outer automorphism may also be taken as the complex conjugation of the 27-dimensional

representation $\Delta_1(u)$ without any complication of the type encountered in SU_{2n}. All the corepresentations are of type I, since $\Delta(u)$ occurs in the reduction of some direct product of direct powers of $\Delta_1(u)$ and $\Delta_5(u)$, both being of type I.

E_7: All automorphisms are inner. The center has two elements, 1 and c_4. The representative of c_4 in $\Delta(u)$ is

$$\Delta(c_4) = (-1)^{\lambda_4+\lambda_6+\lambda_7}.$$

Thus, for the choice $a_0^2 = c_4$, all corepresentations are of type I. For the choice $a_0^2 = 1$, see remark 3.1.

E_8, F_4, G_2: For these groups, all automorphisms are inner and the center has only one element, the identity. Thus $a_0^2 = 1$. See remark 3.1.

Remark 3.3: A glance at Tables III and IV shows that class III, C occurs only when the automorphism induced by a_0 is inner. Hence, in accordance with Dyson,[12] using his notation, only class CC2 (and not CC1) occurs when G is a compact simple Lie group.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor F. J. Dyson for suggesting the problem, and for many useful discussions. Thanks are also due to Professor E. P. Wigner and to Professor L. Michel for several illuminating discussions.

[12] F. J. Dyson, Ref. 3. See the discussion for corepresentation of a factorizable group, p. 1207, 2nd para.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Invariants of Nearly Periodic Hamiltonian Systems. II

K. J. WHITEMAN AND B. MCNAMARA
Culham Laboratory, Abingdon, Berkshire, England

(Received 6 June 1967)

In a previous paper the first few terms of the adiabatic invariant of a particular class of dynamical systems were found by solving Liouville's equation. The system considered was a periodic motion to which small perturbations were applied. The period of the unperturbed orbits was a constant and the perturbations were time-independent. In this paper similar methods are used to find the invariant for the more general system, in which the period of the unperturbed orbits is a function of the coordinates and in which the perturbation varies slowly with time. The results are applied to a simple example, the Lorentz pendulum.

1. INTRODUCTION

In a previous paper[1] (referred to in the text as I) a study was made of a dynamical system with a Hamiltonian of the form

$$H = p_1 + \epsilon\Omega(q_i, p_i), \tag{1.1}$$

where q_i and p_i $(i = 1, 2, \cdots, N)$ are canonical coordinates, Ω is periodic in q_1 period 2π, and ϵ is a small parameter. When ϵ is zero, the orbits are curves along which q_1 varies linearly with the time and the other coordinates remain constant. If q_1 is an anglelike variable, then the orbits form closed loops, the time taken to pass once around being the same for all loops. The dynamical system with ϵ nonzero then consists of a slow drift superimposed on these periodic motions. Such a nearly periodic system possesses an adiabatic invariant and the first few terms in the series representation of it can be found.[1] Examples of this kind arise in the study of nonlinearly coupled oscillators and in various problems in celestial mechanics.

A more general system with somewhat similar properties is one with a Hamiltonian of the form

$$H = \Psi + \epsilon\Omega, \tag{1.2}$$

where, as before, Ω is periodic in q_1 period 2π. Ψ is a function that is independent of q_1, depends on p_1, and varies slowly with all the other coordinates and time t. Ω depends on all the coordinates and varies slowly with time, i.e.,

$$\Psi = \Psi(\epsilon q_2, \cdots, \epsilon q_N, p_1, \epsilon p_2, \cdots, \epsilon p_N, \epsilon t), \tag{1.3}$$

$$\Omega = \Omega(q_1, \cdots, q_N, p_1, \cdots, p_N, \epsilon t). \tag{1.4}$$

When ϵ is zero, the equations of motion are

$$\frac{dq_1}{dt} = (\lambda)_{\epsilon=0}, \quad \text{where} \quad \lambda = \frac{\partial\Psi}{\partial p_1},$$

[1] B. McNamara and K. J. Whiteman, J. Math. Phys. 8, 2029 (1967).

representation $\Delta_1(u)$ without any complication of the type encountered in SU_{2n}. All the corepresentations are of type I, since $\Delta(u)$ occurs in the reduction of some direct product of direct powers of $\Delta_1(u)$ and $\Delta_5(u)$, both being of type I.

E_7: All automorphisms are inner. The center has two elements, 1 and c_4. The representative of c_4 in $\Delta(u)$ is

$$\Delta(c_4) = (-1)^{\lambda_4+\lambda_6+\lambda_7}.$$

Thus, for the choice $a_0^2 = c_4$, all corepresentations are of type I. For the choice $a_0^2 = 1$, see remark 3.1.

E_8, F_4, G_2: For these groups, all automorphisms are inner and the center has only one element, the identity. Thus $a_0^2 = 1$. See remark 3.1.

Remark 3.3: A glance at Tables III and IV shows that class III, C occurs only when the automorphism induced by a_0 is inner. Hence, in accordance with Dyson,[12] using his notation, only class CC2 (and not CC1) occurs when G is a compact simple Lie group.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor F. J. Dyson for suggesting the problem, and for many useful discussions. Thanks are also due to Professor E. P. Wigner and to Professor L. Michel for several illuminating discussions.

[12] F. J. Dyson, Ref. 3. See the discussion for corepresentation of a factorizable group, p. 1207, 2nd para.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Invariants of Nearly Periodic Hamiltonian Systems. II

K. J. WHITEMAN AND B. MCNAMARA
Culham Laboratory, Abingdon, Berkshire, England

(Received 6 June 1967)

In a previous paper the first few terms of the adiabatic invariant of a particular class of dynamical systems were found by solving Liouville's equation. The system considered was a periodic motion to which small perturbations were applied. The period of the unperturbed orbits was a constant and the perturbations were time-independent. In this paper similar methods are used to find the invariant for the more general system, in which the period of the unperturbed orbits is a function of the coordinates and in which the perturbation varies slowly with time. The results are applied to a simple example, the Lorentz pendulum.

1. INTRODUCTION

In a previous paper[1] (referred to in the text as I) a study was made of a dynamical system with a Hamiltonian of the form

$$H = p_1 + \epsilon\Omega(q_i, p_i), \tag{1.1}$$

where q_i and p_i $(i = 1, 2, \cdots, N)$ are canonical coordinates, Ω is periodic in q_1 period 2π, and ϵ is a small parameter. When ϵ is zero, the orbits are curves along which q_1 varies linearly with the time and the other coordinates remain constant. If q_1 is an anglelike variable, then the orbits form closed loops, the time taken to pass once around being the same for all loops. The dynamical system with ϵ nonzero then consists of a slow drift superimposed on these periodic motions. Such a nearly periodic system possesses an adiabatic invariant and the first few terms in the series representation of it can be found.[1] Examples of this kind arise in the study of nonlinearly coupled oscillators and in various problems in celestial mechanics.

A more general system with somewhat similar properties is one with a Hamiltonian of the form

$$H = \Psi + \epsilon\Omega, \tag{1.2}$$

where, as before, Ω is periodic in q_1 period 2π. Ψ is a function that is independent of q_1, depends on p_1, and varies slowly with all the other coordinates and time t. Ω depends on all the coordinates and varies slowly with time, i.e.,

$$\Psi = \Psi(\epsilon q_2, \cdots, \epsilon q_N, p_1, \epsilon p_2, \cdots, \epsilon p_N, \epsilon t), \tag{1.3}$$

$$\Omega = \Omega(q_1, \cdots, q_N, p_1, \cdots, p_N, \epsilon t). \tag{1.4}$$

When ϵ is zero, the equations of motion are

$$\frac{dq_1}{dt} = (\lambda)_{\epsilon=0}, \quad \text{where} \quad \lambda = \frac{\partial\Psi}{\partial p_1},$$

[1] B. McNamara and K. J. Whiteman, J. Math. Phys. 8, 2029 (1967).

and

$$\frac{d\mathbf{u}}{dt} = 0,$$

where

$$\mathbf{u} = (q_2, \cdots, q_N, p_1, \cdots, p_N).$$

Along the orbits $\mathbf{u}$ is constant and q_1 varies with a rate that is a function of p_1. (It is assumed in the following that $\lambda \neq 0$.) Again, if q_1 is an angle variable, the orbits are closed loops, but now the periodic time varies from orbit to orbit. This system, too, possesses an adiabatic invariant and it is the purpose of this paper to obtain the first few terms in its series. Examples arise in the study of the motion of charged particles in a slowly varying magnetic field and the motion of a satellite about a slightly nonspherical earth. The question of whether magnetic fields possess magnetic surfaces can also be posed in a similar canonical form.[2]

In I two methods were used to calculate the invariant. In the Poisson-bracket method, Liouville's equation was solved by expanding in ϵ and using an algebra of operators to reduce the equations to forms which had an obvious solution. In the second, Kruskal's averaging procedure[3] was used to evaluate the action integral as an asymptotic series in ϵ. Both methods have again been used to calculate the invariant for the more general system discussed here. In Sec. 2 the Poisson-bracket method is outlined and the modification to Kruskal's method indicated in Sec. 3. Section 4 contains some general comments. In Sec. 5 the invariant is calculated for a simple example, the Lorentz pendulum.

2. THE POISSON BRACKET METHOD

A constant of the motion J is sought for the dynamical system described by the Hamiltonian (1.2)–(1.4). J must satisfy Liouville's equation

$$\frac{dJ}{dt} \equiv \frac{\partial J}{\partial t} - [J, H] = 0, \tag{2.1}$$

where the Poisson bracket is defined by

$$[A, B] = \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i} - \frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i}. \tag{2.2}$$

In this paper no attempt is made to find a general solution of (2.1), as this would depend on the precise form of $H(\mathbf{q}, \mathbf{p}, t)$, which is not specified. What can be found is a particular solution making use of the known properties of the Hamiltonian. Following I, J is first expanded as a power series in ϵ:

$$J = \sum_{n=0}^{\infty} \epsilon^n J_n. \tag{2.3}$$

Furthermore, J is required to have the same dependence on the coordinates as the Hamiltonian H. That is, J is to be a function of q_i and p_i, but only a slowly varying function of the time, varying as ϵt. H is periodic in q_1 period 2π and J is required to be periodic also. This latter restriction on J has important consequences, since it prevents the occurrence of secular terms (like $q_1^n \sin q_1$, for example) and leads to a series that can be an adequate representation of the invariant even for large values of q_1.

The expansion scheme is complicated by the slow variations of ψ, and therefore of J, and the most compact expression of the scheme is achieved by regrouping the expansion of J as

$$J = J^{(n-1)} + \epsilon^n J_n + O(\epsilon^{n+1}),$$

where

$$J^{(n-1)} = \sum_{m=0}^{n-1} \epsilon^m J_m. \tag{2.4}$$

Equation (2.1) becomes

$$\epsilon^n \lambda \frac{\partial J_n}{\partial q_1} = [J^{(n-1)}, H] - \frac{\partial J^{(n-1)}}{\partial t} + O(\epsilon^{n+1}). \tag{2.5}$$

The condition that J_n be periodic is that no constant terms appear on the right-hand side of (2.5), i.e.,

$$\frac{\partial J^{(n-1)}}{\partial t} - \overline{[J^{(n-1)}, H]} = O(\epsilon^{n+1}). \tag{2.6}$$

As in paper I, the average $\bar{f}$ of a periodic function of q_1 is

$$\bar{f} = \frac{1}{2\pi}\int_0^{2\pi} f\, dq_1. \tag{2.7}$$

Also, as before,

$$\hat{f} = \int (f - \bar{f})\, dq_1, \tag{2.8}$$

where the constant of integration is chosen to make $\hat{f} = 0$. Equation (2.5) can now be integrated to give

$$\epsilon^n J_n = \frac{1}{\lambda}\widehat{[J^{(n-1)}, H]} - \frac{1}{\lambda}\widehat{\frac{\partial J^{(n-1)}}{\partial t}} + \epsilon^n G_n, \tag{2.9}$$

where the G_n are constants of integration, i.e., they can depend on all the coordinates except q_1. The J_n and G_n are found by solving Eqs. (2.6) and (2.9) order by order. From (2.5), J_0 is independent of q_1 and so the lowest-order periodicity condition (2.6) becomes

$$\frac{\partial J_0}{\partial t} - [J_0, \bar{H}] = O(\epsilon^2). \tag{2.10}$$

[2] B. McNamara and K. J. Whiteman (to be published).
[3] M. Kruskal, J. Math. Phys. 3, 806 (1962).

The only obvious solution is $J_0 = J_0(p_1)$, since $\bar{H}$ is independent of q_1. As any function of an invariant is also invariant, the solution can be chosen to be $J_0 = p_1$. The same solution for J_0 could have been chosen in I, but in that case would merely have added the constant Hamiltonian to the invariant found. In order to pursue the expansion scheme further, the operator algebra of I [Eqs. (3.11)–(3.19)] is required. Complications arise in the general case because λ is a function of the coordinates and is outside the Poisson bracket in Eq. (2.9). The following additional relations are useful and easily demonstrated from the definition (2.2):

$$[fg, h] = f[g, h] + g[f, h] = [f, gh] + [g, fh] \quad (2.11)$$

and

$$[fg, f] = \tfrac{1}{2}[g, f^2]. \quad (2.12)$$

Using (2.9) and (2.4), one now finds

$$J^{(1)} = p_1 + \epsilon\frac{\Omega - \bar{\Omega}}{\lambda} + \epsilon G_1 \quad (2.13)$$

and the first-order periodicity condition gives

$$\frac{\partial G_1}{\partial t} - [G_1, \bar{H}] + \overline{\frac{\partial}{\partial t}\left(\frac{\Omega - \bar{\Omega}}{\lambda}\right)} - \overline{\left[\frac{\Omega - \bar{\Omega}}{\lambda}, H\right]} = O(\epsilon^3). \quad (2.14)$$

The last term can be rewritten, using I(3.18), I(3.16), and (2.12):

$$\epsilon\overline{\left[\frac{\Omega - \bar{\Omega}}{\lambda}, \Omega\right]} = \epsilon\left[\frac{1}{\lambda}, \frac{\overline{\Omega^2} - \bar{\Omega}^2}{2}\right] = O(\epsilon^3).$$

The third term is zero, using I(3.12), and so (2.14) becomes

$$\frac{\partial G_1}{\partial t} - [G_1, \bar{H}] = O(\epsilon^2). \quad (2.15)$$

This is the same equation as (2.10) and the appropriate solution is $G_1 = G_1(p_1) = 0$. The methods now become much harder to operate than in I, and a further operator has to be introduced to handle the slow variations of ψ and λ. We define the slow bracket of ψ with $f = f(p_i, q_i, t)$ by

$$[\psi, f] = \lambda\frac{\partial f}{\partial q_1} + \epsilon\{\psi, f\}. \quad (2.16)$$

Using (2.5), one now finds

$$\epsilon^2 J_2 = \frac{1}{\lambda}[p_1, H] + \frac{\epsilon}{\lambda}\left[\frac{\Omega - \bar{\Omega}}{\lambda}, H\right] - \frac{\epsilon}{\lambda}\frac{\partial}{\partial t}\left(\frac{\hat{\Omega}}{\lambda}\right) + \epsilon^2 G_2$$
$$= \frac{\epsilon^2}{\lambda}\left\{\frac{\hat{\Omega}}{\lambda}, \psi\right\} + \frac{\epsilon^2}{\lambda}\left[\frac{1}{\lambda}, \frac{\widehat{\Omega^2}}{2}\right] - \frac{\epsilon^2}{\lambda}\left[\frac{\bar{\Omega}}{\lambda}, \hat{\Omega}\right]$$
$$- \frac{\epsilon^2}{\lambda}\frac{\partial}{\partial \epsilon t}\left(\frac{\hat{\Omega}}{\lambda}\right) + \epsilon^2 G_2. \quad (2.17)$$

The only new relation needed to solve (2.6) for G_2 is for the average of the Jacobi identity involving a slow bracket:

$$\overline{[\{\psi, f\}, g]} + \overline{[\{g, \psi\}, f]} + \overline{\{[f, g], \psi\}} = \left\{\lambda, \overline{f\frac{\partial g}{\partial q_1}}\right\}. \quad (2.18)$$

The equation for G_2 can now be rearranged to read

$$\frac{\partial}{\partial t}\left(G_2 - \frac{1}{2}\overline{\left[\frac{\Omega}{\lambda}, \frac{\hat{\Omega}}{\lambda}\right]}\right) - \left[G_2 - \frac{1}{2}\overline{\left[\frac{\Omega}{\lambda}, \frac{\hat{\Omega}}{\lambda}\right]}, \bar{H}\right] = O(\epsilon^2). \quad (2.19)$$

As before, the particular solution of this equation is chosen to be

$$G_2 = \frac{1}{2}\overline{\left[\frac{\Omega}{\lambda}, \frac{\hat{\Omega}}{\lambda}\right]}. \quad (2.20)$$

The adiabatic invariant correct to $O(\epsilon^2)$ for the Hamiltonian of (1.2)–(1.4) is therefore

$$J = p_1 + \epsilon\left(\frac{\Omega - \bar{\Omega}}{\lambda}\right)$$
$$+ \epsilon^2\left(\frac{1}{\lambda}\left\{\frac{\hat{\Omega}}{\lambda}, \psi\right\} + \left[\frac{1}{2\lambda^2}, \frac{\widehat{\Omega^2}}{2}\right] - \frac{1}{\lambda}\left[\frac{\bar{\Omega}}{\lambda}, \hat{\Omega}\right]\right.$$
$$\left. - \frac{1}{\lambda}\frac{\partial}{\partial \epsilon t}\left(\frac{\hat{\Omega}}{\lambda}\right) + \frac{1}{2}\overline{\left[\frac{\Omega}{\lambda}, \frac{\hat{\Omega}}{\lambda}\right]}\right) + O(\epsilon^3). \quad (2.21)$$

The slow dependence of ψ and λ on the coordinates greatly increases the work involved in using this method. It would be a lengthy task to obtain higher-order terms, but we believe that no further operator algebra would be needed.

3. KRUSKAL'S AVERAGING METHOD

Kruskal[3] deals with a set of equations

$$\mathbf{x}_t \equiv \frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}, \epsilon) \quad (3.1)$$

such that for $\epsilon = 0$ the point $\mathbf{x}(t)$ traces out closed curves as t increases. He has shown that one can introduce a transformation to new coordinates

$$\mathbf{x} \to (\mathbf{y}, \nu)$$

such that (3.1) becomes

$$\mathbf{y}_t = \epsilon\mathbf{g}(\mathbf{y}, \nu) \quad (3.2)$$

and

$$\nu_t = \chi(\mathbf{y}) + \epsilon f(\mathbf{y}, \nu), \quad (3.3)$$

where f and $\mathbf{g}$ are periodic in ν period 2π and χ is independent of ν. In I the particular case was considered where χ was a constant (taken equal to unity) and the averaging method was applied. That is, the transformation of coordinates

$$\mathbf{y}, \nu \leftrightarrow \mathbf{z}, \varphi$$

was found (as far as terms of second order in ϵ) such that $\mathbf{z}_i$ and φ_i were functions of $\mathbf{z}$ alone and not of φ.

Taking $\chi = 1$ was sufficient to allow the adiabatic invariant to be found for the dynamical systems considered in I. For the more general systems considered in this paper, it would seem necessary to carry out the averaging procedure for the general case where $\chi = \chi(\mathbf{y})$. However, we remark that if one changes the independent variable from t to s where

$$\frac{ds}{dt} = \chi(\mathbf{y}),$$

then (3.2) and (3.3) become

$$\mathbf{y}_s = \epsilon(\mathbf{g}/\chi),$$
$$\nu_s = 1 + \epsilon(f/\chi).$$

These are now of the form considered previously.

The equations of motion for the Hamiltonian system described by (1.2)–(1.4) are

$$\begin{aligned} \frac{dq_1}{dt} &= \lambda + \epsilon\Omega_{p_1}, \\ \frac{dq_i}{dt} &= \epsilon(\psi_{\epsilon p_i} + \Omega_{p_i}), \qquad i \neq 1, \\ \frac{dp_i}{dt} &= -\epsilon(\psi_{\epsilon q_i} + \Omega_{q_i}), \quad i = 1, 2, \cdots, N. \end{aligned} \tag{3.4}$$

Choosing $ds/dt = \lambda$, Eqs. (3.4) can be written as

$$\frac{d\mathbf{Y}}{ds} = \mathbf{a} + \epsilon\mathbf{G}, \tag{3.5}$$

where

$$\mathbf{Y} = (\mathbf{q}, \mathbf{p}, \epsilon t), \tag{3.6}$$

$$\mathbf{a} = (1, 0, 0, \cdots, 0), \tag{3.7}$$

and

$$\mathbf{G} = \left(\frac{\Omega_{p_i}}{\lambda}, \frac{1}{\lambda}(\psi_{\epsilon p_i} + \Omega_{p_i}), \cdots, -\frac{1}{\lambda}(\psi_{\epsilon q_i} + \Omega_{q_i}), \cdots, \frac{1}{\lambda}\right). \tag{3.8}$$

Denoting the averaged coordinates by

$$\mathbf{Z} = (\mathbf{Q}, \mathbf{P}, \epsilon T),$$

the averaging transformations are given as before by I(4.25) and I(4.26). The only change arises when representing these expressions in terms of Poisson brackets.

Using the notation of I(Sec. 4), one finds

$$\tilde{\tilde{\mathbf{G}}} \cdot S_{\mathbf{Y}} = \frac{1}{\lambda}[\tilde{\tilde{\Omega}}, S],$$

$$\bar{\mathbf{G}} \cdot \tilde{\tilde{S}}_{\mathbf{Y}} = \frac{1}{\lambda}\frac{\partial \tilde{\tilde{S}}}{\partial \epsilon t} + \frac{1}{\lambda}[\bar{\Omega}, \tilde{\tilde{S}}] + \frac{1}{\lambda}\{\psi, \tilde{\tilde{S}}\},$$

and so on. It is now possible to make calculations similar to those in I(Sec. 4) and evaluate the invariant $\oint \mathbf{p} \cdot d\mathbf{q}$ taken around a curve on which Q_1 varies and the other coordinates Q_i, P_i, and T remain constant. The working is not reproduced here, but gives the same expression as (2.21) of this paper.

4. DISCUSSION

The important feature that allows an invariant to be found by the methods of this paper is that the dominant term in

$$[J_n, H]$$

is the term containing the derivative $\partial J_n/\partial q_1$. Another form of Hamiltonian different from (1.2) to (1.4) that has this same property is one with

$$\psi = \psi(\epsilon q_2, \cdots, \epsilon q_N, p_1, p_2, \epsilon p_3, \cdots, \epsilon p_N, \epsilon t) \tag{4.1}$$

and

$$\Omega = \Omega(q_1, \epsilon q_2, q_3, \cdots, q_N, p_1, \cdots, p_N, \epsilon t). \tag{4.2}$$

That is, ψ depends both on p_1 and p_2 instead of p_1 and ϵp_2, and Ω depends on ϵq_2 instead of q_2. The total Hamiltonian is a function of ϵq_2 and we therefore seek an invariant having the same dependence. As a result, the lowest-order term in (2.1) is just $\lambda\partial J_0/\partial q_1$. The analysis proceeds exactly as above and the invariant can be found. In fact, ψ can be allowed to depend on any number of the q_i and p_i, as well as on p_1, provided that the total Hamiltonian varies only slowly in the conjugate coordinates; an invariant can still be found and it has precisely the form given in (2.21).

The motion of a charged particle in a magnetic field has a Hamiltonian that can be written in this form. One finds that $\psi = Bp_1 + p_2$, where $p_1 = v_\perp^2/B$, $p_2 = v_\parallel^2$, and B is the magnitude of the magnetic field, where $v_\parallel$ and $v_\perp$ are the components of particle velocity along and perpendicular to the field. Provided that B is assumed to vary slowly in the direction of the field (as ϵq_2), an invariant can be found. The lowest-order invariant from (2.21) is just $p_1 = v_\perp^2/B$. The next term,

$$\epsilon\frac{(\Omega - \bar{\Omega})}{\lambda},$$

gives the usual form[4] of the first-order correction to the magnetic moment.

5. THE TIME-DEPENDENT OSCILLATOR

An example for which the adiabatic invariant is simply found by the methods described is the oscillator with a time-dependent frequency, the Lorentz pendulum. (The frequency of the lowest-order orbits is constant, but the perturbations are time-dependent.)

[4] T. G. Northrop, *The Adiabatic Motion of Charged Particles* (Interscience Publishers, Inc., New York, 1963).

The Hamiltonian for this system can be written

$$H = \tfrac{1}{2}(\dot{x}^2 + \omega^2 x^2), \tag{5.1}$$

where

$$\omega = \omega(\epsilon t).$$

Defining a canonical transformation to new coordinates (PQ) by means of the generating function

$$W(x, P) = \frac{x}{2}(2\omega P - \omega^2 x^2)^{\frac{1}{2}} + P \sin^{-1}\left(\frac{\omega}{2P}\right)^{\frac{1}{2}} x \tag{5.2}$$

[found by solving the Hamilton–Jacobi equation $H(x, \partial W/\partial x) = \omega P$], one obtains

$$x = \left(\frac{2P}{\omega}\right)^{\frac{1}{2}} \sin Q$$

and

$$\dot{x} = (2P\omega)^{\frac{1}{2}} \cos Q. \tag{5.3}$$

The new Hamiltonian H^* is given by

$$H^* = H + \frac{\partial W}{\partial t} = \omega P + \epsilon \frac{\omega'}{2\omega} P \sin 2Q, \tag{5.4}$$

where $\omega' = d\omega/d\epsilon t$. This is of the standard form (1.2)–(1.4) and the invariant is, from (2.21),

$$J = P + \epsilon \frac{\Omega}{\lambda} + \tfrac{1}{2}\epsilon^2 \overline{\left[\frac{\Omega}{\lambda}, \frac{\Omega}{\lambda}\right]} - \frac{\epsilon^2}{\lambda} \frac{\partial}{\partial \epsilon t}\left(\frac{\Omega}{\lambda}\right) + O(\epsilon^3), \tag{5.5}$$

since $\bar{\Omega} = 0$ and λ is independent of P and Q. Evaluating this, one obtains

$$J = P + \epsilon \frac{\omega'}{2\omega} P \sin 2Q + \frac{\epsilon^2}{8}\left(\frac{\omega'}{\omega^2}\right) P + \frac{\omega P \cos 2Q}{2}\left(\frac{\omega''}{\omega^4} - \frac{2(\omega')^2}{\omega^5}\right) + O(\epsilon^3) \tag{5.6}$$

$$= \frac{(\dot{x}^2 + \omega^2 x^2)}{2\omega}\left(1 + \frac{\epsilon^2}{8}\left(\frac{\omega'}{\omega^2}\right)^2\right) + \frac{\epsilon\omega'}{2\omega^2} x\dot{x} + \frac{\epsilon^2}{8}(\dot{x}^2 - \omega^2 x^2)\left(\frac{\omega''}{\omega^4} - \frac{2(\omega')^2}{\omega^5}\right) + O(\epsilon^3). \tag{5.7}$$

To lowest order this is just H/ω, the familiar form of the invariant for the Lorentz-pendulum problem. The result of (5.7) to second order agrees with an expression given by Littlewood[5] [there is a typographical error in Eq. (11) of Ref. 5]. Although H/ω varies as ω varies slowly in time, one can see that the change in H/ω between two states for which $\omega' = 0$ is zero. This result has been obtained to all orders by Kulsrud.[6]

ACKNOWLEDGMENT

The authors wish to thank R. T. P. Whipple for pointing out an error in a draft copy of this paper.

[5] J. E. Littlewood, Ann. Phys. (N.Y.) **21**, 233 (1963).
[6] R. M. Kulsrud, Phys. Rev. **106**, 205 (1957).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Asymptotic Estimates of Feynman Integrals*

JAMES P. FINK†‡
Department of Mathematics and Stanford Linear Accelerator Center, Stanford University
Stanford, California

(Received 18 December 1967)

In this paper, we consider the problem of determining logarithmic, as well as polynomial, asymptotic estimates for certain convergent integrals containing parameters. We state and prove an asymptotic theorem which gives the logarithmic asymptotic behavior of a convergent integral where any subset of the parameters becomes large while the remaining parameters remain bounded. This theorem is then applied to the photon and electron self-energy graphs of quantum electrodynamics.

I. INTRODUCTION

The renormalization procedures of Dyson[1,2] and Salam[3,4] depend upon a certain convergence criterion for integrals which was proved in a paper by Weinberg.[5] In addition to stating and proving conditions under which a Feynman integral converges, Weinberg developed a method for determining a polynomial bound on the value of the integral as subsets of the external momenta become large, provided the usual rotations of energy contours can be performed. The value of his technique is that one need not evaluate the integrals under consideration. The bound on the

* The preparation of this manuscript was supported in part by U.S. Atomic Energy Commission.
† National Science Foundation predoctoral fellow in Mathematics at Stanford University during the years 1963–67.
‡ Present address is Department of Mathematics and Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania.
[1] F. Dyson, Phys. Rev. **75**, 486 (1949).
[2] F. J. Dyson, Phys. Rev. **75**, 1736 (1949).
[3] A. Salam, Phys. Rev. **82**, 217 (1951).
[4] A. Salam, Phys. Rev. **84**, 426 (1951).
[5] S. Weinberg, Phys. Rev. **118**, 838 (1960).

The Hamiltonian for this system can be written

$$H = \tfrac{1}{2}(\dot{x}^2 + \omega^2 x^2), \tag{5.1}$$

where

$$\omega = \omega(\epsilon t).$$

Defining a canonical transformation to new coordinates (PQ) by means of the generating function

$$W(x, P) = \frac{x}{2}(2\omega P - \omega^2 x^2)^{\frac{1}{2}} + P \sin^{-1}\left(\frac{\omega}{2P}\right)^{\frac{1}{2}} x \tag{5.2}$$

[found by solving the Hamilton–Jacobi equation $H(x, \partial W/\partial x) = \omega P$], one obtains

$$x = \left(\frac{2P}{\omega}\right)^{\frac{1}{2}} \sin Q$$

and

$$\dot{x} = (2P\omega)^{\frac{1}{2}} \cos Q. \tag{5.3}$$

The new Hamiltonian H^* is given by

$$H^* = H + \frac{\partial W}{\partial t} = \omega P + \epsilon \frac{\omega'}{2\omega} P \sin 2Q, \tag{5.4}$$

where $\omega' = d\omega/d\epsilon t$. This is of the standard form (1.2)–(1.4) and the invariant is, from (2.21),

$$J = P + \epsilon \frac{\Omega}{\lambda} + \tfrac{1}{2}\epsilon^2 \overline{\left[\frac{\Omega}{\lambda}, \frac{\Omega}{\lambda}\right]} - \frac{\epsilon^2}{\lambda} \frac{\partial}{\partial \epsilon t}\left(\frac{\Omega}{\lambda}\right) + O(\epsilon^3), \tag{5.5}$$

since $\bar{\Omega} = 0$ and λ is independent of P and Q. Evaluating this, one obtains

$$J = P + \epsilon \frac{\omega'}{2\omega} P \sin 2Q + \frac{\epsilon^2}{8}\left(\frac{\omega'}{\omega^2}\right) P + \frac{\omega P \cos 2Q}{2}\left(\frac{\omega''}{\omega^4} - \frac{2(\omega')^2}{\omega^5}\right) + O(\epsilon^3) \tag{5.6}$$

$$= \frac{(\dot{x}^2 + \omega^2 x^2)}{2\omega}\left(1 + \frac{\epsilon^2}{8}\left(\frac{\omega'}{\omega^2}\right)^2\right) + \frac{\epsilon\omega'}{2\omega^2} x\dot{x} + \frac{\epsilon^2}{8}(\dot{x}^2 - \omega^2 x^2)\left(\frac{\omega''}{\omega^4} - \frac{2(\omega')^2}{\omega^5}\right) + O(\epsilon^3). \tag{5.7}$$

To lowest order this is just H/ω, the familiar form of the invariant for the Lorentz-pendulum problem. The result of (5.7) to second order agrees with an expression given by Littlewood[5] [there is a typographical error in Eq. (11) of Ref. 5]. Although H/ω varies as ω varies slowly in time, one can see that the change in H/ω between two states for which $\omega' = 0$ is zero. This result has been obtained to all orders by Kulsrud.[6]

ACKNOWLEDGMENT

The authors wish to thank R. T. P. Whipple for pointing out an error in a draft copy of this paper.

[5] J. E. Littlewood, Ann. Phys. (N.Y.) **21**, 233 (1963).
[6] R. M. Kulsrud, Phys. Rev. **106**, 205 (1957).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Asymptotic Estimates of Feynman Integrals*

JAMES P. FINK†‡
Department of Mathematics and Stanford Linear Accelerator Center, Stanford University
Stanford, California

(Received 18 December 1967)

In this paper, we consider the problem of determining logarithmic, as well as polynomial, asymptotic estimates for certain convergent integrals containing parameters. We state and prove an asymptotic theorem which gives the logarithmic asymptotic behavior of a convergent integral where any subset of the parameters becomes large while the remaining parameters remain bounded. This theorem is then applied to the photon and electron self-energy graphs of quantum electrodynamics.

I. INTRODUCTION

The renormalization procedures of Dyson[1,2] and Salam[3,4] depend upon a certain convergence criterion for integrals which was proved in a paper by Weinberg.[5] In addition to stating and proving conditions under which a Feynman integral converges, Weinberg developed a method for determining a polynomial bound on the value of the integral as subsets of the external momenta become large, provided the usual rotations of energy contours can be performed. The value of his technique is that one need not evaluate the integrals under consideration. The bound on the

* The preparation of this manuscript was supported in part by U.S. Atomic Energy Commission.
† National Science Foundation predoctoral fellow in Mathematics at Stanford University during the years 1963–67.
‡ Present address is Department of Mathematics and Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania.
[1] F. Dyson, Phys. Rev. **75**, 486 (1949).
[2] F. J. Dyson, Phys. Rev. **75**, 1736 (1949).
[3] A. Salam, Phys. Rev. **82**, 217 (1951).
[4] A. Salam, Phys. Rev. **84**, 426 (1951).
[5] S. Weinberg, Phys. Rev. **118**, 838 (1960).

integral is determined simply from the asymptotic properties of the integrand alone.

Weinberg's analysis, however, does not determine the logarithmic asymptotic behavior of convergent integrals. A method which provides some clue to the logarithmic asymptotic behavior of the photon and electron self-energy graphs of quantum electrodynamics is the renormalization group approach (cf. Bjorken and Drell,[6] Bogoliubov and Shirkov,[7] and Landau[8]). The renormalization group is, by definition, the group of transformations which, when applied to the propagators, charges, and masses of a theory, yields new propagators, charges, and masses which do not change the expressions for observable quantities. The arguments of the renormalization group approach rely upon several fundamental assumptions which lead to anomalous results which in turn make one suspect the original assumptions.

In this paper, we develop a technique for determining the logarithmic asymptotic behavior of a certain class of convergent integrals and apply it to various Feynman integrals of quantum electrodynamics. We use Weinberg's results[5] as a basis, although we are required to modify and extend them.

II. ASYMPTOTIC THEOREMS FOR INTEGRALS

A. Introduction

In this section we are concerned with extending the results of Weinberg.[5] Before undertaking this task, however, we briefly summarize his results and, in so doing, we use essentially the notation used by Weinberg.

B. Summary of Weinberg's Results

Let $f(p_1, \cdots, p_n)$ be a complex-valued function of the n real variables $p_1, \cdots, p_n$. We will consider the variables $p_1, \cdots, p_n$ as the components of a vector $\mathbf{P}$ in R^n, and we will be concerned only with those functions $f(\mathbf{P})$ which belong to a certain class A_n defined as follows:

Definition: A function $f(\mathbf{P})$ is an element of the class A_n if and only if, for each subspace $S \subset R^n$, there exist coefficients $\alpha(S)$, $\beta(S)$ such that, for any choice of $m \leq n$ independent vectors $\mathbf{L}_1, \cdots, \mathbf{L}_m$ and bounded region $W \subset R^n$, we have

$$f(\mathbf{L}_1\eta_1 \cdots \eta_m + \mathbf{L}_2\eta_2 \cdots \eta_m + \cdots + \mathbf{L}_m\eta_m + \mathbf{C}) = O\{\eta_1^{\alpha(\{\mathbf{L}_1\})}(\log \eta_1)^{\beta(\{\mathbf{L}_1\})} \cdots \eta_m^{\alpha(\{\mathbf{L}_1, \cdots, \mathbf{L}_m\})} \times (\log \eta_m)^{\beta(\{\mathbf{L}_1, \cdots, \mathbf{L}_m\})}\}$$

when $\eta_1, \cdots, \eta_m$ tend independently to infinity and $\mathbf{C} \in W$. The notation $\{\mathbf{L}_1, \cdots, \mathbf{L}_r\}$ denotes the subspace spanned by the vectors $\mathbf{L}_1, \cdots, \mathbf{L}_r$.

Let I be a subspace of R^n spanned by some set of orthonormal vectors $\mathbf{L}_1', \cdots, \mathbf{L}_k'$, and consider the integral

$$f_I(\mathbf{P}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dy_1 \cdots dy_k \times f(\mathbf{P} + \mathbf{L}_1'y_1 + \cdots + \mathbf{L}_k'y_k) = \int_{\mathbf{P}' \in I} d^k\mathbf{P}' f(\mathbf{P} + \mathbf{P}').$$

Provided this integral exists, $f_I(\mathbf{P})$ is a function which depends only on the projection of $\mathbf{P}$ along the subspace I; that is, $f_I(\mathbf{P})$ depends only upon the component of $\mathbf{P}$ in the subspace complementary to I.

The following theorem was proved by Weinberg.[5]

Theorem 1: Suppose $f(\mathbf{P}) \in A_n$ with asymptotic coefficients $\alpha(S)$ and $\beta(S)$ for any nonzero subspace S of R^n. Let $f(\mathbf{P})$ be integrable over any bounded region in R^n (local integrability), and let

$$D_I = \max_{S' \subset I} \{\alpha(S') + \dim S'\},$$

where $\dim S'$ is the dimension of S'. If $D_I < 0$, then:

(a) $f_I(\mathbf{P})$ exists;

(b) $f_I(\mathbf{P}) \in A_{n-k}$ with asymptotic coefficient $\alpha_I(S)$ for $S \subset E$, where $R^n \equiv I \oplus E$, given by

$$\alpha_I(S) = \max_{\Lambda(I)S'=S} \{\alpha(S') + \dim S' - \dim S\}.$$

$\Lambda(I)$ is the operation of projection along the subspace I and $\max_{\Lambda(I)S'=S}$ means that the maximum is taken over all those subspaces S' which project onto S.

C. Definition of the Subclass B_n

Let $f(\mathbf{P}) \in A_n$ with asymptotic coefficients $\alpha(S)$ and $\beta(S)$. Let $\mathbf{L}_1, \cdots, \mathbf{L}_m$ be $m \leq n$ independent vectors and W a finite region in R^n. We arrange the logarithmic asymptotic coefficients $\beta(\{\mathbf{L}_1\}), \cdots, \beta(\{\mathbf{L}_1, \cdots, \mathbf{L}_m\})$ in increasing order, and suppose that

$$\beta(\{\mathbf{L}_1, \cdots, \mathbf{L}_{\pi_1}\}) \leq \beta(\{\mathbf{L}_1, \cdots, \mathbf{L}_{\pi_2}\}) \leq \cdots \leq \beta(\{\mathbf{L}_1, \cdots, \mathbf{L}_{\pi_m}\}),$$

where $\pi_1, \cdots, \pi_m$ is a permutation of the integers $1, \cdots, m$.

Definition: A function $f(\mathbf{P})$ is an element of the subclass B_n if and only if $f(\mathbf{P}) \in A_n$ with asymptotic

[6] J. Bjorken and S. Drell, *Relativistic Quantum Fields* (McGraw-Hill Book Co., New York, 1965).

[7] N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (Interscience Publ., Inc., New York, 1959).

[8] L. D. Landau, in *Niels Bohr and the Development of Physics*, W. Pauli, Ed. (McGraw-Hill Book Co., New York, 1955).

coefficients $\alpha(S)$ and $\beta(S)$ such that $\beta(S)$ is a nonnegative integer for all $S \subset R^n$ and

$$f(\mathbf{L}_1\eta_1 \cdots \eta_m + \mathbf{L}_2\eta_2 \cdots \eta_m + \cdots + \mathbf{L}_m\eta_m + \mathbf{C}) = O\Big\{\eta_1^{\alpha(\{\mathbf{L}_1\})} \cdots \eta_m^{\alpha(\{\mathbf{L}_1, \cdots, \mathbf{L}_m\})} \times \sum_{\gamma_1, \cdots, \gamma_m} (\log \eta_{\pi_1})^{\gamma_1} (\log \eta_{\pi_2})^{\gamma_2}, \cdots, (\log \eta_{\pi_m})^{\gamma_m}\Big\},$$

when $\eta_1, \cdots, \eta_m$ tend independently to infinity and $\mathbf{C} \in W$, where the sum ranges over all nonnegative integers $\gamma_1, \cdots, \gamma_m$ satisfying

$$\begin{aligned} \gamma_1 &\leq \beta(\{\mathbf{L}_1, \cdots, \mathbf{L}_{\pi_1}\}), \\ \gamma_1 + \gamma_2 &\leq \beta(\{\mathbf{L}_1, \cdots, \mathbf{L}_{\pi_2}\}), \\ &\cdots \\ \gamma_1 + \cdots + \gamma_m &\leq \beta(\{\mathbf{L}_1, \cdots, \mathbf{L}_{\pi_m}\}). \end{aligned}$$

Since $B_n \subset A_n$, Theorem 1 applies to the subclass B_n.

D. Generalization for One-Dimensional Integrals

Our goal is to obtain a formula for $\beta_I(S)$ for integrable functions in the subclass B_n similar to the formula for $\alpha_I(S)$ given in Theorem 1 for integrable functions in A_n. We begin with a definition based on this theorem.

Definition: A subspace S' is said to be a *maximizing subspace for the I integration* (relative to a given subspace $S \subset E$) if

$$\Lambda(I)S' = S \quad \text{and} \quad \alpha_I(S) = \alpha(S') + \dim S' - \dim S.$$

The proof of Theorem 1 given in Ref. 5 shows that maximizing subspaces always exist.

Let us first consider the case when $\dim I = 1$. For this case, the maximizing subspaces fall into two categories—those for which $\dim S' = \dim S$ and those for which $\dim S' = \dim S + 1$. Let p be the number of nonempty categories of maximizing subspaces; that is, $p = 1$ if all maximizing subspaces have the same dimension and $p = 2$ otherwise. By repeating the proof of Theorem 1 for the subclass B_n, we arrive at the following theorem:

Theorem 2: Let $f(\mathbf{P}) \in B_n$ satisfy all the conditions of Theorem 1 and suppose that $\dim I = 1$. Then $f_I(\mathbf{P}) \in B_{n-1}$ with asymptotic coefficients $\alpha_I(S)$ given by Theorem 1 and $\beta_I(S)$ given by

$$\beta_I(S) = \max_{S' \in M} \beta(S') + p - 1,$$

where M is the set of all maximizing subspaces.

E. Generalization for Two-Dimensional Integrals

In order to generalize this result when $\dim I > 1$, let us next examine the case $\dim I = 2$. We write $I = I_1 \oplus I_2$, where $\dim I_1 = \dim I_2 = 1$, and integrate first with respect to the I_1 variable and then with respect to the variable in I_2 and vice versa. Since we will be dealing only with integrable functions in B_n in the following, Fubini's theorem applies and we conclude that the integral is independent of the order of integration and of the particular choice of I_1 and I_2.

Let us perform the I_2 integration first and then the I_1 integration. We have

$$\begin{aligned} \alpha_{I_2}(S') &= \max_{\Lambda(I_2)S'' = S'} \{\alpha(S'') + \dim S'' - \dim S'\}, \\ \alpha_I(S) &= \max_{\Lambda(I_1)S' = S} \{\alpha_{I_2}(S') + \dim S' - \dim S\} \\ &= \max_{\Lambda(I)S'' = S} \{\alpha(S'') + \dim S'' - \dim S\}, \end{aligned}$$

where

$$\begin{aligned} &S \subset E \quad \text{with} \quad R^n = I \oplus E, \\ &S' \subset E_2 \quad \text{with} \quad R^n = I_2 \oplus E_2, \\ &S'' \subset R^n. \end{aligned}$$

Let $S'_\mu \subset E_2$ be the maximizing subspaces for the I_1 integration relative to S after performing the I_2 integration. For each S'_μ, let $S''_{\mu\nu} \subset R^n$ be the maximizing subspaces for the I_2 integration relative to S'_μ. We have the relations

$$\Lambda(I_2)S''_{\mu\nu} = S'_\mu, \quad \alpha_{I_2}(S'_\mu) = \alpha(S''_{\mu\nu}) + \dim S''_{\mu\nu} - \dim S'_\mu,$$
$$\Lambda(I_1)S'_\mu = S, \qquad \alpha_I(S) = \alpha_{I_2}(S'_\mu) + \dim S'_\mu - \dim S.$$

We now want to determine the maximizing subspaces for the full I integration relative to S; that is, we want to determine the subspaces $S'' \subset R^n$ for which $\Lambda(I)S'' = S$ and $\alpha_I(S) = \alpha(S'') + \dim S'' - \dim S$.

Lemma 1: The $S''_{\mu\nu}$ are precisely the maximizing subspaces for the I integration relative to S; that is, each $S''_{\mu\nu}$ is a maximizing subspace for the I integration relative to S and any such maximizing subspace for the I integration relative to S must be one of the $S''_{\mu\nu}$.

Proof: To show that each $S''_{\mu\nu}$ is a maximizing subspace for the I integration relative to S, we note that $\Lambda(I)S''_{\mu\nu} = S$ and

$$\begin{aligned} \alpha_I(S) &= \alpha_{I_2}(S'_\mu) + \dim S'_\mu - \dim S \\ &= \alpha(S''_{\mu\nu}) + \dim S''_{\mu\nu} - \dim S'_\mu \\ &\quad + \dim S'_\mu - \dim S \\ &= \alpha(S''_{\mu\nu}) + \dim S''_{\mu\nu} - \dim S. \end{aligned}$$

Conversely, suppose that S''_0 is a maximizing subspace for the I integration relative to S and let

$S_0' = \Lambda(I_2)S_0''$. We have

$$S = \Lambda(I_1)S_0' = \Lambda(I)S_0'',$$
$$\alpha_I(S) = \alpha(S_0'') + \dim S_0'' - \dim S.$$

Now

$$\begin{aligned}\alpha_I(S) &= \max_{\Lambda(I_1)S'=S} \{\alpha_{I_2}(S') + \dim S' - \dim S\} \\ &\geq \alpha_{I_2}(S_0') + \dim S_0' - \dim S \\ &= \max_{\Lambda(I_2)S''=S_0'} \{\alpha(S'') + \dim S'' - \dim S_0'\} \\ &\quad + \dim S_0' - \dim S \\ &\geq \alpha(S_0'') + \dim S_0'' - \dim S \\ &= \alpha_I(S),\end{aligned}$$

where the last equality follows from the assumptions on S_0''. Since the first and last terms in this chain are the same quantity, all inequalities must be equalities, and hence

$$\alpha_I(S) = \alpha_{I_2}(S_0') + \dim S_0' - \dim S, \quad \Lambda(I_1)S_0' = S,$$
$$\alpha_{I_2}(S_0') = \alpha(S_0'') + \dim S_0'' - \dim S_0', \quad \Lambda(I_2)S_0'' = S_0'.$$

Thus, S_0' is a maximizing subspace for the I_1 integration relative to S after performing the I_2 integration and so must be one of the S_μ'. S_0'' is a maximizing subspace for the I_2 integration relative to S_0' (which is one of the S_μ') and consequently must be one of the $S_{\mu\nu}''$.

We observe that Lemma 1 does not depend upon the fact that we are assuming $\dim I_1 = \dim I_2 = 1$.

Let p_1 be the number of different dimensions among the maximizing subspaces for the I_1 integration relative to S after performing the I_2 integration, and let $p_{2\mu}$ be the number of different dimensions among the maximizing subspaces for the I_2 integration relative to S_μ'.

Lemma 2: $p_{2\mu}$ is independent of μ.

Proof: Suppose not. Then there exist two maximizing subspaces for the I_1 integration relative to S after performing the I_2 integration, say S_1' and S_2', such that $p_{21} = 1$ and $p_{22} = 2$.

There are several cases to be considered. We work out the details for one case only because the others are all similar.

Let S_{11}'' be maximizing for the I_2 integration relative to S_1' and let S_{21}'' and S_{22}'' be maximizing for the I_2 integration relative to S_2'. Suppose $\dim S_1' = \dim S_2' = \dim S_{11}'' = \dim S_{21}'' = \dim S_{22}'' - 1$. Performing the I_2 integration first and then the I_1 integration, we obtain

$$\begin{aligned}\beta_I(S) &= \max\{\beta_{I_2}(S_1'), \beta_{I_2}(S_2')\} \\ &= \max\{\beta(S_{11}''), \max\{\beta(S_{21}''), \beta(S_{22}'')\} + 1\}.\end{aligned}$$

On the other hand, reversing the order of integration gives

$$\beta_I(S) = \max\{\beta(S_{11}''), \beta(S_{21}''), \beta(S_{22}'')\} + 1.$$

These two expressions are not equal for all nonnegative integral values of $\beta(S'')$ and hence we have a contradiction.

Since $p_{2\mu}$ is independent of μ, we will denote it simply by p_2.

Now let $\tilde{I}_1$ and $\tilde{I}_2$ be two one-dimensional subspaces of I different from I_1 and I_2, respectively, such that $I = \tilde{I}_1 \oplus \tilde{I}_2$. Just as with I_1 and I_2, we let T_ρ' be the maximizing subspaces for the $\tilde{I}_1$ integration relative to S after performing the $\tilde{I}_2$ integration, and, for each T_ρ', we let $T_{\rho\sigma}''$ be the maximizing subspaces for the $\tilde{I}_2$ integration relative to T_ρ'. Let $\tilde{p}_1$ be the number of different dimensions among the maximizing subspaces for the $\tilde{I}_1$ integration relative to S after performing the $\tilde{I}_2$ integration, and let $\tilde{p}_2$ be the number of different dimensions among the maximizing subspaces for the $\tilde{I}_2$ integration relative to T_ρ'. By Lemma 2, $\tilde{p}_2$ is independent of ρ and we have the following lemma:

Lemma 3: Let $I = I_1 \oplus I_2$ and $I = \tilde{I}_1 \oplus \tilde{I}_2$ be two decompositions of the two-dimensional space of integration I into one-dimensional components. Let $p_1, p_2, \tilde{p}_1$, and $\tilde{p}_2$ be defined as above. Then

$$p_1 + p_2 = \tilde{p}_1 + \tilde{p}_2.$$

Proof: Consider the decomposition $I = I_1 \oplus I_2$. Performing first the I_2 integration and then the I_1 integration, we obtain, by Theorem 2,

$$\beta_{I_2}(S_\mu') = \max_\nu \beta(S_{\mu\nu}'') + p_2 - 1,$$
$$\beta_I(S) = \max_\mu \beta_{I_2}(S_\mu') + p_1 - 1.$$

Combining these two expressions,

$$\begin{aligned}\beta_I(S) &= \max_\mu \left\{\max_\nu \beta(S_{\mu\nu}'') + p_2 - 1\right\} + p_1 - 1 \\ &= \max_{\mu,\nu} \beta(S_{\mu\nu}'') + p_1 + p_2 - 2,\end{aligned}$$

since p_2 does not depend upon μ.

Similarly, for the decomposition $I = \tilde{I}_1 \oplus \tilde{I}_2$, we have

$$\beta_I(S) = \max_{\rho,\sigma} \beta(T_{\rho\sigma}'') + \tilde{p}_1 + \tilde{p}_2 - 2.$$

By Lemma 1, however, the $S_{\mu\nu}''$ are precisely the maximizing subspaces for the I integration relative to S and so are the $T_{\rho\sigma}''$. Consequently, the $T_{\rho\sigma}''$ are merely the $S_{\mu\nu}''$ relabeled. Thus,

$$\max_{\mu,\nu} \beta(S_{\mu\nu}'') = \max_{\rho,\sigma} \beta(T_{\rho\sigma}'')$$

and we obtain the desired result.

We see that the proof of Lemma 3 provides us with a formula for $\beta_I(S)$ when dim $I = 2$.

Theorem 3: Let $f(\mathbf{P}) \in B_n$ satisfy all the conditions of Theorem 1 and suppose that dim $I = 2$. Then

$$\beta_I(S) = \max_{S' \in M} \beta(S') + p_1 + p_2 - 2,$$

where M is the set of all maximizing subspaces for the I integration relative to S.

F. General Asymptotic Theorem

The generalization to the case dim $I = k$ is now reasonably straightforward. We write $I = I_1 \oplus \cdots \oplus I_k$, where each component subspace I_j has dimension one.

Definition: Let $I = I_1 \oplus \cdots \oplus I_k$ with dim $I_j = 1$. The *dimension numbers* $p_1, \cdots, p_k$ are defined inductively as follows: p_1 is the number of dimensions among the maximizing subspaces for the I_1 integration relative to S after performing the $I_2 \oplus \cdots \oplus I_k$ integration. $p_j, j = 2, \cdots, k$, is the number of dimensions among the maximizing subspaces for the I_j integration after performing the $I_{j+1} \oplus \cdots \oplus I_k$ integration relative to any one of the maximizing subspaces for the I_{j-1} integration after performing the $I_j \oplus \cdots \oplus I_k$ integration.

By definition, the dimension numbers p_j can take on only the values 1 and 2. The definition of $p_j, j = 2, \cdots, k$, appears to be ambiguous, however, because it does not specify the maximizing subspace for the I_{j-1} integration relative to which p_j is computed. The next lemma shows that this ambiguity actually does not exist.

Lemma 4: The dimension numbers $p_j, j = 2, \cdots, k$, are independent of the maximizing subspaces for the I_{j-1} integration relative to which they are computed.

Proof: The result for dim $I = k = 2$ was proved already as Lemma 2 in Sec. IIE. Therefore, if $k > 2$, we assume that the $p_j, j = 2, \cdots, k-1$, are independent of the maximizing subspaces for the I_{j-1} integration relative to which they are computed.

Suppose that p_k does not enjoy this property. Then there exist two maximizing subspaces S_1' and S_2' for the I_{k-1} integration after performing the I_k integration for which $p_k = p_{k1} = 1$ and $p_k = p_{k2} = 2$. Using Lemma 1, S_1' and S_2' are maximizing subspaces for the $I_1 \oplus \cdots \oplus I_{k-1}$ integration relative to S after performing the I_k integration. Let $S_{1\nu}''$ and $S_{2\nu}''$ be the maximizing subspaces for the I_k integration relative to S_1' and S_2', respectively.

Since we are assuming that S_1' and S_2' are two different maximizing subspaces for the $I_1 \oplus \cdots \oplus I_{k-1}$ integration relative to S after performing the I_k integration, there exists a one-dimensional subspace J_1 of $I_1 \oplus \cdots \oplus I_{k-1}$ and its orthogonal complement J_2 in $I_1 \oplus \cdots \oplus I_{k-1}$ $(I_1 \oplus \cdots \oplus I_{k-1} = J_1 \oplus J_2)$ such that the subspaces $\Lambda(J_2)S_1'$ and $\Lambda(J_2)S_2'$ are different. (See Proposition A1 of the Appendix.) We now integrate out the J_2 subspace leaving the J_1 subspace. Let

$$\Lambda(J_2)S_1' = T_1', \quad \Lambda(J_2)S_{1\nu}'' = T_{1\nu}'',$$
$$\Lambda(J_2)S_2' = T_2', \quad \Lambda(J_2)S_{2\nu}'' = T_{2\nu}''.$$

By Lemma 1, T_1' and T_2' are maximizing subspaces for the J_1 integration relative to S after performing the $I_k \oplus J_2$ integration, and $T_{1\nu}''$ and $T_{2\nu}''$ are maximizing subspaces for the I_k integration relative to T_1' and T_2', respectively, after performing the J_2 integration.

Let p_{k1}' and p_{k2}' be the numbers of different dimensions among the subspaces $T_{1\nu}''$ and $T_{2\nu}''$, respectively. Then $p_{k1}' = p_{k1}$ and $p_{k2}' = p_{k2}$. Therefore, $p_{k1} \neq p_{k2}$ implies that $p_{k1}' \neq p_{k2}'$, which contradicts Lemma 2 because dim $J_1 =$ dim $I_k = 1$.

Lemma 5: Let $I = I_1 \oplus I_2 \oplus \cdots \oplus I_k$ and $I = \tilde{I}_1 \oplus \tilde{I}_2 \oplus \cdots \oplus \tilde{I}_k$ be two decompositions of the k-dimensional space of integration I into one-dimensional components. Let $p_1, p_2, \cdots, p_k$ and $\tilde{p}_1, \tilde{p}_2, \cdots, \tilde{p}_k$ be the corresponding dimension numbers as defined above. Then

$$\sum_{i=1}^{k} p_i = \sum_{i=1}^{k} \tilde{p}_i.$$

The proof of this lemma is almost identical to that of Lemma 3, where it is assumed that dim $I = 2$.

The proof of Lemma 5 now gives us the general asymptotic theorem.

Theorem 4: Let $f(\mathbf{P}) \in B_n$ satisfy all the conditions of Theorem 1 and suppose that dim $I = k$. Let $p_1, p_2, \cdots, p_k$ be the dimension numbers corresponding to any decomposition of I into one-dimensional components. Then $f_I(\mathbf{P}) \in B_{n-k}$ with asymptotic coefficients $\alpha_I(S)$ given by Theorem 1 and $\beta_I(S)$ given by

$$\beta_I(S) = \max_{S' \in M} \beta(S') + \sum_{i=1}^{k} p_i - k,$$

where M is the set of all maximizing subspaces for the I integration relative to S.

III. ASYMPTOTIC ESTIMATES FOR SELF-ENERGY GRAPHS

A. Introduction

We now apply Theorems 1 and 4 to photon and electron self-energy graphs in order to obtain asymptotic bounds for the corresponding renormalized

Feynman integrals. We remark that, although our discussion centers around photon self-energy graphs, the same results apply to electron self-energy graphs with the obvious modifications.

B. Degree of Divergence of a Subgraph

Weinberg shows in his article[5] that the integrand of any Feynman integral corresponding to a certain Feynman diagram is an element of the class A_{4N}, where N is the number of independent four-momenta in the diagram, provided the energy contour can be rotated from the real to the imaginary axis. Thus, if q is a four-momentum, the hyperbolic metric

$$q^2 = q_0^2 - q_1^2 - q_2^2 - q_3^2$$

becomes negative-definite for q_0 purely imaginary. We therefore assume that this well-known energy-contour rotation[2,6] has always been carried out. Furthermore, since the logarithmic asymptotic coefficients of any Feynman integrand are zero, the integrands belong to the subclass B_{4N}, defined in Sec. IIC.

For a detailed discussion of the connection between subgraphs of a Feynman graph and the corresponding subspaces of R^{4N}, where N is the number of independent four-momenta in the Feynman graph, we again refer to Weinberg[5] and also to Bjorken and Drell.[6] In the following, the subspace S of Theorems 1 and 4 is always the subspace associated with the external momenta of the Feynman diagram, which, for the case of a photon self-energy graph, is simply the photon four-momentum q. The maximizing subspaces $S' \in M$ correspond to those subgraphs of the original Feynman graph with maximum degree of divergence. For a subgraph g' corresponding to a subspace S', the degree of divergence $D_I(g')$ is defined as

$$D_I(g') = \alpha(S') + \dim S' - \dim S, \qquad (1)$$

where $\alpha(S')$ is the asymptotic coefficient for the integrand corresponding to the original graph. In renormalizable field theories, it turns out that

$$D_I(g') = 4 - \tfrac{3}{2}F(g') - B(g'),$$

where $F(g')$ and $B(g')$ are the numbers of fermion and Boson lines, respectively, attached to the subgraph g', including external lines belonging to g'. (See, for example, Dyson[2] and Bjorken and Drell.[6])

Rules for determining the degree of divergence of a subgraph in which there are subtraction terms are given in Bjorken and Drell.[6] By a simple counting technique, we can determine the degree of divergence $D_I(g')$ of a subgraph g', which, according to Eq. (1) and Theorems 1 and 4, is the quantity we need to know in order to calculate the asymptotic coefficients $\alpha_I(S)$ and $\beta_I(S)$ of the integral.

In order to calculate the logarithmic asymptotic coefficient $\beta_I(S)$, we must first determine the dimension numbers p_j defined in Sec. IIF. Before we can do this, however, we need some facts concerning maximizing subspaces of convergent Feynman integrals.

C. Irreducible Subspaces of the Space of Integration

Suppose that the space of integration I of a convergent integral has dimension $4k$, as is the case for Feynman integrands. Let $I = I_1 \oplus \cdots \oplus I_{4k}$ be a decomposition of I into one-dimensional components I_j, and let $I_1', \cdots, I_k'$ be the four-dimensional subspaces of I defined as

$$\begin{aligned} I_1' &= I_1 \oplus I_2 \oplus I_3 \oplus I_4, \\ I_2' &= I_5 \oplus I_6 \oplus I_7 \oplus I_8, \\ &\cdots \\ I_k' &= I_{4k-3} \oplus I_{4k-2} \oplus I_{4k-1} \oplus I_{4k}. \end{aligned}$$

Furthermore, suppose that the maximizing subspaces for the I integration relative to S are of the form

$$\begin{aligned} &S, \\ &S \oplus I_{l_1}', \\ &S \oplus I_{l_1}' \oplus I_{l_2}', \\ &\cdots \\ &S \oplus \bigoplus_{i=1}^{j} I_{l_i}', \end{aligned}$$

where

$$j = 1, \cdots, k, \quad l_i = 1, \cdots, k, \quad l_{i_1} < l_{i_2} \text{ if } i_1 < i_2.$$

Definition: Suppose that the maximizing subspaces for the I integration relative to S are of the form just given. A direct sum

$$J = \bigoplus_{i=1}^{j} I_{l_i},$$

where $j = 1, \cdots, k$, $l_i = 1, \cdots, k$, and $l_{i_1} < l_{i_2}$ if $i_1 < i_2$, is called an *irreducible subspace* of I if every maximizing subspace for the I integration relative to S which contains one or more of the components I_{l_i} of J actually contains the entire sum J.

For example, if $I = I_1' \oplus I_2' \oplus I_3'$ with I_1' and $I_2' \oplus I_3'$ irreducible, then the possible maximizing

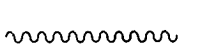


FIG. 1. The subgraph of a photon self-energy graph corresponding to the subspace S.

subspaces for the I integration relative to S are

$$S, \quad S \oplus I_1', \quad S \oplus I_2' \oplus I_3',$$
$$S \oplus I_1' \oplus I_2' \oplus I_3' = S \oplus I.$$

The subspace $S \oplus I_2'$, in particular, could not be maximizing because, by the irreducibility of $I_2' \oplus I_3'$, the subspaces I_2' and I_3' cannot be split up.

D. Maximizing Subspaces and Dimension Numbers of Convergent Feynman Integrals

We begin with a lemma which applies to any photon self-energy graph.

Lemma 6: For any renormalized photon self-energy graph, the subgraph shown in Fig. 1 has degree of divergence equal to 2. In other words, the subspace S is itself maximizing for the I integration relative to S.

Proof: Clearly, $\Lambda(I)S = S$.

For a given photon self-energy graph, let the corresponding Feynman integral be denoted by

$$\Pi_{\mu\nu}(q) = \int dP_I R_{\mu\nu}(P_I, q),$$

where P_I denotes the integration variables in the space of integration I. Suppose that $R_{\mu\nu}$ is the integrand which results after all subtractions have been performed with the exception of the overall subtractions. Performing the overall subtractions, we obtain

$$\Pi^c_{\mu\nu}(q) = \int dP_I \Big\{ R_{\mu\nu}(P_I, q) - R_{\mu\nu}(P_I, 0) - q_\rho \frac{\partial R_{\mu\nu}(P_I, 0)}{\partial q_\rho} - \frac{q_\rho q_\sigma}{2} \frac{\partial^2 R_{\mu\nu}(P_I, 0)}{\partial q_\rho \partial q_\sigma} \Big\}.$$

For this new integrand, we have that

$$\alpha(S) + \dim S - \dim S = \alpha(S) = 2,$$

which is equal to $\alpha_I(S)$ for a photon self-energy graph. We recall that S is the subspace associated with the external momentum q. In S, all the variables denoted by P_I are zero.

Thus, S is maximizing for the I integration relative to S.

The counting technique for determining the degree of divergence of a subgraph gives the value of the expression

$$D_I(g') = \alpha(S') + \dim S' - \dim S,$$

and hence we can determine the maximizing subspaces for the I integration relative to S. In order to calculate the dimension numbers p_j, however, we must be able to determine the subspaces which maximize subintegrations of the full I integration.

We again write $I = I_1 \oplus \cdots \oplus I_{4k}$, where dim $I = 4k$, and define the four-dimensional subspaces $I_1', \cdots, I_k'$, as in Sec. IIIC. Let p_j, $j = 1, \cdots, 4k$ be the corresponding dimension numbers defined in Sec. IIF.

Theorem 5: Suppose there exists a decomposition of the space of integration I such that the irreducible subspaces of I can be written as

$$I_1'' = I_1' \oplus \cdots \oplus I_{k_1}',$$
$$I_2'' = I_{k_1+1}' \oplus \cdots \oplus I_{k_2}',$$
$$\vdots \qquad\qquad \vdots$$
$$I_m'' = I_{k_{m-1}+1}' \oplus \cdots \oplus I_{k_m}', \quad k_1 < k_2 < \cdots < k_m = k,$$

for some integers $m, k_1, k_2, \cdots, k_m = k$. In other words, we assume that the maximizing subspaces for the I integration relative to S are of the form

$$S,$$
$$S \oplus I_{l_1}'',$$
$$S \oplus I_{l_1}'' \oplus I_{l_2}'',$$
$$\vdots$$
$$S \oplus \bigoplus_{i=1}^{j} I_{l_i}'',$$

where

$$j = 1, \cdots, m, \quad l_i = 1, \cdots, m,$$
$$l_{i_1} < l_{i_2} \quad \text{if} \quad i_1 < i_2.$$

Furthermore, suppose that the subspaces

$$S, \quad S \oplus I_1'', \quad S \oplus I_1'' \oplus I_2'', \cdots,$$
$$S \oplus I_1'' \oplus \cdots \oplus I_m'' = S \oplus I$$

are included in the set of all maximizing subspaces. Then

$$\sum_{j=1}^{4k} p_j = 4k + m,$$

where m, defined implicitly above, is the number of irreducible subspaces of I.

Proof: The proof of this theorem, although somewhat long, is not difficult. It amounts to calculating each of the dimension numbers p_j, and this is done by determining the maximizing subspaces for the subintegrations of the I integration.

Consider the sequence of maximizing subspaces

$$S,\quad S\oplus I_1'',\quad S\oplus I_1''\oplus I_2'',\cdots,$$
$$S\oplus I_1''\oplus\cdots\oplus I_m''=S\oplus I,$$

and take any two adjacent subspaces from this sequence:

$$S\oplus I_1''\oplus\cdots\oplus I_r''\quad\text{and}\quad S\oplus I_1''\oplus\cdots\oplus I_{r+1}'',\qquad 0\le r\le m-1.$$

In terms of the I_j', these two subspaces are

$$S\oplus I_1'\oplus\cdots\oplus I_{k_r}'\quad\text{and}\quad S\oplus I_1'\oplus\cdots\oplus I_{k_{r+1}}',$$

and in terms of the I_j, they are

$$S\oplus I_1\oplus\cdots\oplus I_{4k_r}\quad\text{and}\quad S\oplus I_1\oplus\cdots\oplus I_{4k_{r+1}}.$$

(For $r=0$, we define $k_r=0$. Then $S\oplus I_1''\oplus\cdots\oplus I_r''$ and $S\oplus I_1'\oplus\cdots\oplus I_{k_r}'$ refer to the subspace S.) Set

$$S_0=S,$$
$$S_j=S\oplus I_1\oplus\cdots\oplus I_j,\quad j=1,\cdots,4k.$$

The two maximizing subspaces we are considering are then denoted S_{4k_r} and $S_{4k_{r+1}}$. Now

$$\Lambda(I_{4k_r+1})S_{4k_r}=S_{4k_r},$$
$$\Lambda(I_{4k_r+1})S_{4k_{r+1}}=S_{4k_r}.\qquad(2)$$

Also,

$$\begin{aligned}\alpha_I(S)&=\max_{\substack{\Lambda(I_1\oplus\cdots\oplus I_{4k_r})S'=S\\ S'\subset S_{4k_r}}}\{\alpha_{I_{4k_r+1}\oplus\cdots\oplus I_{4k}}(S')\\&\qquad+\dim S'-\dim S\}\\&\ge\alpha_{I_{4k_r+1}\oplus\cdots\oplus I_{4k}}(S_{4k_r})+\dim S_{4k_r}-\dim S\\&=\max_{\substack{\Lambda(I_{4k_r+1})S''=S_{4k_r}\\ S''\subset S_{4k_r+1}}}\{\alpha_{I_{4k_r+2}\oplus\cdots\oplus I_{4k}}(S'')\\&\qquad+\dim S''-\dim S_{4k_r}\}+\dim S_{4k_r}-\dim S\\&\ge\alpha_{I_{4k_r+2}\oplus\cdots\oplus I_{4k}}(S_{4k_r})+\dim S_{4k_r}-\dim S\\&=\max_{\Lambda(I_{4k_r+2}\oplus\cdots\oplus I_{4k})S'''=S_{4k_r}}\{\alpha(S''')+\dim S'''\\&\qquad-\dim S_{4k_r}\}+\dim S_{4k_r}-\dim S\\&\ge\alpha(S_{4k_r})+\dim S_{4k_r}-\dim S\\&=\alpha_I(S),\end{aligned}$$

where this last step follows because $S_{4k_r}=S\oplus I_1''\oplus\cdots\oplus I_r''$ is a maximizing subspace for the I integration relative to S. Thus,

$$\alpha_{I_{4k_r+1}\oplus\cdots\oplus I_{4k}}(S_{4k_r})=\alpha_{I_{4k_r+2}\oplus\cdots\oplus I_{4k}}(S_{4k_r})+\dim S_{4k_r}-\dim S_{4k_r}.\qquad(3)$$

Similarly,

$$\begin{aligned}\alpha_I(S)&\ge\alpha_{I_{4k_r+1}\oplus\cdots\oplus I_{4k}}(S_{4k_r})+\dim S_{4k_r}-\dim S\\&\ge\alpha_{I_{4k_r+2}\oplus\cdots\oplus I_{4k}}(S_{4k_r+1})+\dim S_{4k_r+1}-\dim S\\&=\max_{\Lambda(I_{4k_r+2}\oplus\cdots\oplus I_{4k})S'''=S_{4k_r+1}}\{\alpha(S''')+\dim S'''\\&\qquad-\dim S_{4k_r+1}\}+\dim S_{4k_r+1}-\dim S\\&\ge\alpha(S_{4k})+\dim S_{4k}-\dim S\\&=\alpha_I(S)\end{aligned}$$

by the maximizing property of $S_{4k}=S\oplus I$. Hence,

$$\alpha_{I_{4k_r+1}\oplus\cdots\oplus I_{4k}}(S_{4k_r})=\alpha_{I_{4k_r+2}\oplus\cdots\oplus I_{4k}}(S_{4k_r+1})+\dim S_{4k_r+1}-\dim S_{4k_r}.\qquad(4)$$

Relations (2), (3), and (4) together imply that both S_{4k_r} and S_{4k_r+1} are maximizing subspaces for the I_{4k_r+1} integration relative to S_{4k_r} after performing the $I_{4k_r+2}\oplus\cdots\oplus I_{4k}$ integration. Since $\dim S_{4k_r+1}=\dim S_{4k_r}+1$, the corresponding dimension number has the value 2; that is,

$$p_{4k_r+1}=2,\quad r=0,1,\cdots,m-1.$$

We next consider the dimension numbers p_{4k_r+l} for $r=0,\cdots,m-1$ and $l=2,\cdots,4k_{r+1}-4k_r$. Our task is to determine the maximizing subspaces for the I_{4k_r+l} integration relative to S_{4k_r+l-1} after performing the $I_{4k_r+l+1}\oplus\cdots\oplus I_{4k}$ integration. S_{4k_r+l} is such a subspace because

$$\Lambda(I_{4k_r+l})S_{4k_r+l}=S_{4k_r+l-1}$$

and

$$\begin{aligned}\alpha_I(S)&\ge\alpha_{I_{4k_r+l}\oplus\cdots\oplus I_{4k}}(S_{4k_r+l-1})+\dim S_{4k_r+l-1}-\dim S\\&\ge\alpha_{I_{4k_r+l+1}\oplus\cdots\oplus I_{4k}}(S_{4k_r+l})+\dim S_{4k_r+l}-\dim S\\&\ge\alpha(S_{4k})+\dim S_{4k}-\dim S\\&=\alpha_I(S).\end{aligned}$$

Since $\dim S_{4k_r+l}=\dim S_{4k_r+l-1}+1$, any other maximizing subspace for the I_{4k_r+l} integration relative to S_{4k_r+l-1} after performing the $I_{4k_r+l+1}\oplus\cdots\oplus I_{4k}$ integration must have the same dimension as S_{4k_r+l-1}. Let T be such a subspace; that is, assume that T is a maximizing subspace for the I_{4k_r+l} integration relative to S_{4k_r+l-1} after performing the $I_{4k_r+l+1}\oplus\cdots\oplus I_{4k_r+l}$ integration with

$$T\subsetneq S_{4k_r+l}\quad\text{and}\quad\dim T=\dim S_{4k_r+l-1}.$$

Then T is also maximizing for the $I_1\oplus\cdots\oplus I_{4k_r+l}$ integration relative to S after performing the $I_{4k_r+l+1}\oplus\cdots\oplus I_{4k}$ integration because

$$\Lambda(I_1\oplus\cdots\oplus I_{4k_r+l})T=S,$$
$$\begin{aligned}\alpha_I(S)&=\alpha_{I_{4k_r+l}\oplus\cdots\oplus I_{4k}}(S_{4k_r+l-1})+\dim S_{4k_r+l-1}-\dim S\\&=\alpha_{I_{4k_r+l+1}\oplus\cdots\oplus I_{4k}}(T)+\dim T-\dim S.\end{aligned}$$

Now take any maximizing subspace T' for the $I_{4k_r+l+1}\oplus\cdots\oplus I_{4k}$ integration relative to T. By

Lemma 1, T' is a maximizing subspace for the full I integration relative to S.

By assumption, T does not contain the subspace I_{4k_r+l} and, consequently, neither does T'. Since $I_{4k_r+l} \subset I''_{r+1}$ for $l = 2, \cdots, 4k_{r+1} - 4k_r$, the hypotheses of the Theorem 5 imply that we can write T' in the form

$$T' = S \quad \text{or} \quad T' = S \oplus \bigoplus_{i=1}^{j} I''_{l_i} \tag{5}$$

for some $j = 1, \cdots, m$, where $l_i = 1, \cdots, m, l_{i_1} < l_{i_2}$ if $i_1 < i_2$, and $l_i \neq r + 1$.

Now we also have that

$$\begin{aligned} \Lambda(I_{4k_r+l} \oplus \cdots \oplus I_{4k})T' &= \Lambda(I_{4k_r+l})T \\ &= S_{4k_r+l-1}, \end{aligned} \tag{6}$$

and we recall that

$$S_{4k_r+l-1} = S \oplus I''_1 \oplus \cdots \oplus I''_r \oplus I_{4k_r+1} \oplus \cdots \oplus I_{4k_r+l-1}.$$

Thus, if $2 \leq l \leq 4k_{r+1} - 4k_r$, S_{4k_r+l-1} contains a nontrivial part of I''_{r+1}. The statements (5) and (6) are, therefore, not compatible for $2 \leq l \leq 4k_{r+1} - 4k_r$ because no l_i in the direct sum in (5) can take the value $r + 1$, and we have a contradiction.

Hence, there are no maximizing subspaces for the I_{4k_r+l} integration relative to S_{4k_r+l-1} after performing the $I_{4k_r+l+1} \oplus \cdots \oplus I_{4k}$ integration other than S_{4k_r+l}. Thus,

$$p_{4k_r+l} = 1 \quad \text{for} \quad r = 0, 1, \cdots, m-1, \quad l = 2, \cdots, 4k_{r+1} - 4k_r.$$

Therefore, we have that

$$\begin{aligned} \sum_{j=1}^{4k} p_j &= \sum_{r=0}^{m-1} \left(p_{4k_r+1} + \sum_{l=2}^{4k_{r+1}-4k_r} p_{4k_r+l} \right) \\ &= \sum_{r=0}^{m-1} (4k_{r+1} - 4k_r + 1) \\ &= 4k + m. \end{aligned}$$

E. An Asymptotic Theorem for Self-Energy Graphs

We now turn to the case of an arbitrary photon or electron self-energy graph of electrodynamics.

Definition: The *order* of a self-energy graph is defined as the number of vertex points in the graph. With this definition, the order of a photon or electron self-energy graph is always an even number.

For photon self-energy graphs, we have the following theorem:

Theorem 6: Any nth-order photon self-energy graph with m irreducible insertions ($m \leq n/2$) has asymptotic coefficients

$$\alpha_I(\{q\}) = 2, \quad \beta_I(\{q\}) = m,$$

where q is the momentum of the photon.

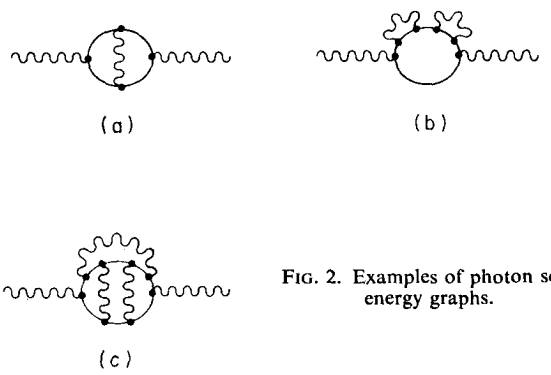


FIG. 2. Examples of photon self-energy graphs.

Proof: The fact that $\alpha_I(\{q\}) = 2$ follows directly from Theorem 1.

Consider a photon self-energy graph with m irreducible insertions. Using Lemma 6, the counting technique for determining the degree of divergence of a subgraph, Theorem 4, and Theorem 5, we obtain

$$\begin{aligned} \beta_I(\{q\}) &= \max_{S' \in M} \beta(S') + \sum_{j=1}^{4k} p_j - 4k \\ &= 0 + 4k + m - 4k \\ &= m. \end{aligned}$$

Thus, the logarithmic asymptotic coefficient of the graph shown in Fig. 2(a) is $\beta_I(\{q\}) = 2$, and the logarithmic asymptotic coefficient of the graph shown in Fig. 2(b) is $\beta_I(\{q\}) = 3$. However, the graph shown in Fig. 2(c) has the logarithmic asymptotic coefficient $\beta_I(\{q\}) = 2$ due to the irreducibility of the vertex insertion shown in Fig. 3.

An analogous theorem for electron self-energy graphs is the following:

Theorem 7: Any nth-order electron self-energy graph with m irreducible insertions ($m \leq n/2$) has asymptotic coefficients

$$\alpha_I(\{q\}) = 1, \quad \beta_I(\{q\}) = m,$$

where q is the momentum of the electron.

Theorem 7 for electron self-energy graphs is proved in exactly the same way as Theorem 6 for photon self-energy graphs. The only difference is that the

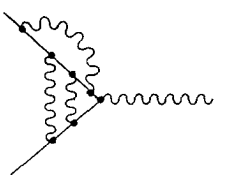


FIG. 3. Irreducible vertex insertion in the graph shown in Fig. 2 (c).

maximum degree of divergence of a subgraph is 1 instead of 2.

IV. SUMMARY AND CONCLUSIONS

A. Conclusions about the Perturbation-Expansion Parameter

In Sec. III we showed that if

(a) the energy contours of the Feynman integral corresponding to a photon or electron self-energy graph are rotated from the real axis to the imaginary axis, and

(b) the momentum q of the photon or electron is replaced by tq, where t is a real scalar,

then the asymptotic behavior of the photon or electron self-energy graph as $t \to \infty$ is given by

$$ct^{\alpha}(\log t)^{\beta},$$

where c is a constant, $\alpha = 1$ for electron self-energy graphs and 2 for photon self-energy graphs, and $\beta = m$, the number of irreducible insertions in the graph. For a given order n, the maximum value of the logarithmic asymptotic coefficient β is $n/2$. Consequently, in a perturbation expansion of the total photon propagator or electron propagator, we would expect the expansion parameter to involve, not only the square of the charge e^2, but the quantity

$$e^2 \log \frac{q^2}{\lambda^2},$$

where renormalization is carried out by subtracting at the point $q^2 = \lambda^2 < 0$. In perturbation expansions and renormalization-group arguments, one usually assumes that the expansion parameter is $e^2 \log (q^2/\lambda^2)$. (See, for example, Bjorken and Drell,[6] Bogoliubov and Shirkov,[7] and Landau.[8]) That this assumption is the correct one is supported by our results.

B. Summing Different Graphs

Although the maximum value of the logarithmic asymptotic coefficient β for nth-order self-energy graphs is $n/2$, it may be that the sum of all the nth-order graphs has a logarithmic asymptotic coefficient less than $n/2$. For example, consider the three fourth-order photon self-energy graphs in Fig. 4. Each one of these graphs has logarithmic asymptotic coefficient $\beta = 2$, but when the three graphs are summed together, the $\log^2 (q^2/\lambda^2)$ terms cancel (see Bjorken and Drell[6]). Thus, the total fourth-order photon propagator has logarithmic asymptotic coefficient $\beta = 1$.

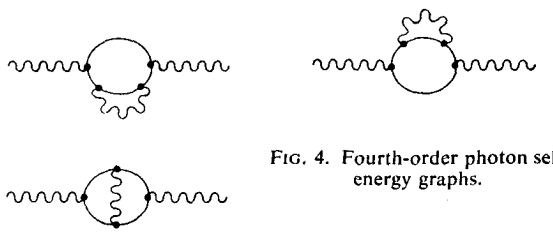


FIG. 4. Fourth-order photon self-energy graphs.

The arguments of the renormalization group predict this cancellation at least for the fourth-order and sixth-order graphs in the perturbation expansion of the photon propagator (cf. Bjorken and Drell,[6] Bogoliubov and Shirkov[7]), and perhaps a similar cancellation occurs for the graphs of other orders. (There is, of course, no cancellation for the single second-order self-energy graph.) This question is unanswered by our results as they stand. We are able to give the asymptotic behavior of any self-energy graph of arbitrary order, but we do so without regard for multiplicative constants.

The problem of summing and determining asymptotic estimates for the entire perturbation expansion remains open. In the first place, it is not even clear that the perturbation series of quantum electrodynamics actually converge. Assuming they do converge, it may turn out that the individual terms have an asymptotic behavior quite unlike that of their sum.

C. Graphs Other than Self-Energy Graphs; the Problem of Unphysical Momenta

We point out that the general theorems of Sec. II and the theorems about maximizing subspaces in Sec. III are applicable to any convergent Feynman integral. Although we have concentrated on self-energy graphs, one could just as well determine the asymptotic behavior of a graph like that shown in Fig. 5, a contribution to eighth-order electron–proton scattering. Three of the four external momenta are independent, say p_1, p_2, and p_1'. Therefore, the asymptotic behavior of this graph will depend upon which subset of p_1, p_2, and p_1' becomes large.

In all of these results there remains one undesirable feature, the necessity of performing energy-contour rotations in order to avoid the singularities associated with the hyperbolic metric. Quantities of the type

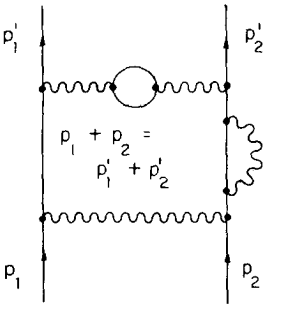


FIG. 5. A contribution to eight-order electron–proton scattering.

$p^2 - m^2 \equiv p_0^2 - p_1^2 - p_2^2 - p_3^2 - m^2$, where m is a constant, appear in the denominators of Feynman integrals; however, if the energy contours can be rotated from the real up to the imaginary axis ($p_0 \to ip_0$), then the expression $p^2 - m^2$ never vanishes.

By performing this rotation, we are restricting ourselves to unphysical momenta. It would be useful to determine the asymptotic behavior of graphs like that shown in Fig. 5 when a certain subset of the external momenta remains on the mass shell while others become large. In other words, one would like to apply asymptotic estimates to real physical experiments. This more difficult problem is not yet solved.

ACKNOWLEDGMENT

The author wishes to thank Professor Sidney Drell for his valuable suggestions and encouragement throughout the course of this work, which forms part of a Ph. D. dissertation submitted to Stanford University (1967).

APPENDIX

In this appendix, we prove a statement which was used in the proof of Lemma 4. Let S and I be two subspaces of R^n whose direct sum $S \oplus I$ is R^n.

Proposition A1: Let S_1 and S_2 be two different subspaces of R^n satisfying $\Lambda(I)S_i = S$, $i = 1, 2$, where $\Lambda(I)$ is the operation of projection along the subspace I, and suppose that $\dim I > 1$. Then there exists a one-dimensional subspace J_1 of I such that $\Lambda(J_2)S_1 \neq \Lambda(J_2)S_2$, where J_2 is the orthogonal complement of J_1 in I ($I = J_1 \oplus J_2$).

The proof of this proposition will follow from the next three lemmas.

If σ is a point in S, we let

$$\Lambda(I)^{-1}\sigma = \{x : x \in R^n, \Lambda(I)x = \sigma\}.$$

Lemma A1: Suppose $\Lambda(I)S_i = S$, $i = 1, 2$. Then $S_1 = S_2$ if and only if $\Lambda(I)^{-1}\sigma \cap S_1 = \Lambda(I)^{-1}\sigma \cap S_2$ for all $\sigma \in S$.

Proof: Since $\Lambda(I)S_i = S$, then

$$S_i = \bigcup_{\sigma \in S} \{\Lambda(I)^{-1}\sigma \cap S_i\}, \quad i = 1, 2.$$

If $S_1, = S_2$, then clearly

$$\Lambda(I)^{-1}\sigma \cap S_1 = \Lambda(I)^{-1}\sigma \cap S_2 \quad \text{for all} \quad \sigma \in S.$$

Conversely, if $\Lambda(I)^{-1}\sigma \cap S_1 = \Lambda(I)^{-1}\sigma \cap S_2$ for all $\sigma \in S$, then we obviously have $S_1 = S_2$ because

$$S_i = \bigcup_{\sigma \in S} \{\Lambda(I)^{-1}\sigma \cap S_i\}, \quad i = 1, 2.$$

Lemma A2: Suppose $\sigma \in S$ and $\Lambda(I)S_i = S$, $i = 1, 2$. Then

$$\Lambda(I)^{-1}\sigma \cap S_1 = \Lambda(I)^{-1}\sigma \cap S_2$$

if, and only if,

$$\Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S_1\} = \Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S_2\}.$$

Proof: If $\Lambda(I)^{-1}\sigma \cap S_1 = \Lambda(I)^{-1}\sigma \cap S_2$, then clearly

$$\Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S_1\} = \Lambda(S)\{\Lambda(I)^{-1}\sigma \cup S_2\}.$$

Conversely, suppose that $\Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S_1\} = \Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S_2\}$. Let $x \in \Lambda(I)^{-1}\sigma \cap S_1$. Then x has the same S coordinates as σ. Now

$$\Lambda(S)x \in \Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S_1\} = \Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S_2\}.$$

Thus, there exists a point $y \in \Lambda(I)^{-1}\sigma \cap S_2$ such that $\Lambda(S)y = \Lambda(S)x$. Since $y \in \Lambda(I)^{-1}\sigma \cap S_2$, y has the same S coordinates as σ and hence as x. Since $\Lambda(S)y = \Lambda(S)x$, y has the same I coordinates as x. Consequently, $x = y$ and so $x \in \Lambda(I)^{-1}\sigma \cap S_2$. Thus, we have that $\Lambda(I)^{-1}\sigma \cap S_1 \subset \Lambda(I)^{-1}\sigma \cap S_2$. A similar argument gives containment the other way.

Now suppose that $\dim I = k$ and write

$$I = I_1 \oplus \cdots \oplus I_k,$$
$$S = I_{k+1} \oplus \cdots \oplus I_n,$$

where each of the component subspaces $I_j, j = 1, \cdots, n$, is one-dimensional.

Lemma A3: Let I_j be one of the component subspaces of I; that is, take $j = 1, \cdots, k$. Let S_0 be a subspace of R^n for which $\Lambda(I)S_0 = S$. Then for any $\sigma \in S$,

$$\Lambda(I_j)\{\Lambda(I)^{-1}\sigma \cap S_0\} = \{\Lambda(I_j)\Lambda(I)^{-1}\sigma\} \cap \Lambda(I_j)S_0.$$

We remark that, in general, if f maps X into Y and A and B are two subsets of X, then

$$f(A \cap B) \subset f(A) \cap f(B).$$

Lemma A3 says that, in our special case, we actually have equality.

Proof: Let $x \in \Lambda(I_j)\{\Lambda(I)^{-1}\sigma \cap S_0\}$. Then there exists a point $y \in \Lambda(I)^{-1}\sigma \cap S_0$ such that $\Lambda(I_j)y = x$. We have

(i) $y \in \Lambda(I)^{-1}\sigma$,
(ii) $y \in S_0$,
(iii) $\Lambda(I_j)y = x$.

Now (i) and (ii) imply that $x \in \Lambda(I_j)\Lambda(I)^{-1}\sigma$, and (ii) and (iii) imply that $x \in \Lambda(I_j)S_0$. Thus

$$x \in \{\Lambda(I_j)\Lambda(I)^{-1}\sigma\} \cap \Lambda(I_j)S_0.$$

Conversely, suppose that $x \in \{\Lambda(I_j)\Lambda(I)^{-1}\sigma\} \cap \Lambda(I_j)S_0$. Then $x \in \Lambda(I_j)\Lambda(I)^{-1}\sigma$ and $x \in \Lambda(I_j)S_0$. Since $x \in \Lambda(I_j)S_0$, there exists a point $z \in S_0$ such that $\Lambda(I_j)z = x$. Then

$$\Lambda(I)z = \Lambda(I_1)\cdots\widehat{\Lambda(I_j)}\cdots\Lambda(I_k)\Lambda(I_j)z$$
$$= \Lambda(I_1)\cdots\widehat{\Lambda(I_j)}\cdots\Lambda(I_k)x,$$

where the hat over $\Lambda(I_j)$ means that $\Lambda(I_j)$ does not appear in the product. But since $x \in \Lambda(I_j)\Lambda(I)^{-1}\sigma$, we have that $\Lambda(I)z = \sigma$, and so $z \in \Lambda(I)^{-1}\sigma$. Thus,

$$z \in \Lambda(I)^{-1}\sigma \cap S_0$$

and

$$x = \Lambda(I_j)z \in \Lambda(I_j)\{\Lambda(I)^{-1}\sigma \cap S_0\}.$$

We can now prove the proposition.

Proof of Proposition A1: Since $S_1 \neq S_2$ and $\Lambda(I)S_i = S$, $i = 1, 2$, Lemmas A1 and A2 imply that there exists a point $\sigma \in S$ such that

$$\Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S_1\} \neq \Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S_2\}.$$

Thus, there exists a point

$$x \in \Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S_1\},$$

but

$$x \notin \Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S_2\}$$

(or vice versa). We can, therefore, find a point $y \in \Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S_2\}$ such that the I_j component of x is not equal to the I_j component of y for some $j = 1, \cdots, k$.

Now consider $S \oplus I_j$. Since the I_j components of x and y are unequal,

$$\Lambda(I)\cdots\widehat{\Lambda(I_j)}\cdots\Lambda(I_k)\Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S_1\}$$
$$\neq \Lambda(I_1)\cdots\widehat{\Lambda(I_j)}\cdots\Lambda(I_k)\Lambda(S)\{\Lambda(I)^{-1}\sigma \cap S\}$$

or

$$\Lambda(S)\Lambda(I_1)\cdots\widehat{\Lambda(I_j)}\cdots\Lambda(I_k)\{\Lambda(I)^{-1}\sigma \cap S_1\}$$
$$\neq \Lambda(S)\Lambda(I_1)\cdots\widehat{\Lambda(I_j)}\cdots\Lambda(I_k)\{\Lambda(I)^{-1}\sigma \cap S_2\}.$$

Using Lemma A3,

$$\Lambda(S)\{\Lambda(I_j)^{-1}\sigma \cap \Lambda(I_1)\cdots\widehat{\Lambda(I_j)}\cdots\Lambda(I_k)S_1\}$$
$$\neq \Lambda(S)\{\Lambda(I_j)^{-1}\sigma \cap \Lambda(I_1)\cdots\widehat{\Lambda(I_j)}\cdots\Lambda(I_k)S_2\}.$$

By Lemmas A1 and A2 again, this last statement implies that

$$\Lambda(I_1)\cdots\widehat{\Lambda(I_j)}\cdots\Lambda(I_k)S_1$$
$$\neq \Lambda(I_1)\cdots\widehat{\Lambda(I_j)}\cdots\Lambda(I_k)S_2$$

or

$$\Lambda(I_1 \oplus \cdots \oplus \hat{I}_j \oplus \cdots \oplus I_k)S_1$$
$$\neq \Lambda(I_1 \oplus \cdots \oplus \hat{I}_j \oplus \cdots \oplus I_k)S_2.$$

Therefore, we can take $J_1 = I_j$ and

$$J_2 = I_1 \oplus \cdots \oplus \hat{I}_j \oplus \cdots \oplus I_k.$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Scalar Product for Harmonic Functions of the Group $SU(2)$

DAVE PANDRES, JR., AND DONALD A. JACOBSON
Douglas Advanced Research Laboratories, Huntington Beach, California

(Received 28 April 1967)

A scalar product is defined which results in the single- and double-valued spherical harmonics spanning a seminormed linear vector space that carries all of the irreducible unitary representations of the group $SU(2)$. The possibility of defining such a scalar product was indicated in a previous paper. A Hilbert space is derived from the seminormed space through a further construction involving equivalence classes of vectors.

INTRODUCTION

In a previous paper,[1] hereafter designated I, one of us showed that the double-valued spherical harmonics provide a basis for the irreducible spinor representations of the three-dimensional rotation group $O(3)$, or, more precisely, its covering group $SU(2)$. Two difficulties arise when such a construction is attempted. The first difficulty is that the vector space V_j spanned by the double-valued spherical harmonics Y_{jm} is not closed with respect to the generators J_α of $O(3)$; to be specific,

$$J^- Y_{j,-j} = Q_{j,-j-1} \tag{1}$$

does not vanish and is not a linear combination of the Y_{jm}. This difficulty was met in I by observing first that, although one can be led out of V_j according to Eq. (1), one cannot be led back in since

$$J^+ J^- Y_{j,-j} = 0, \tag{2}$$

and, secondly, $Q_{j,-j-1}$ is orthogonal to all Y_{jm} so that, in the spirit of Dirac,[2] $Q_{j,-j-1}$ is in some sense a "representation" of the zero vector.

The second difficulty which arises is that the usual scalar product consisting of an integration over the unit sphere is not well defined for all of the Y_{jm} due to existence of nonintegrable singularities at $\theta = 0, \pi$. This difficulty was only partially met in I by observing that a proper scalar product must exist[3] and giving an outline of how such a scalar product might be found.

The purpose of this paper is to explicitly display a scalar product which is well defined not only for the Y_{jm}, but also for $Q_{j,-j-1}$ and the additional functions obtained from it by further applications of the lowering operator J^-. This scalar product has the desirable property of producing Hermitian generators J_α, thus insuring unitary representations for all half-integer j values. In addition, it turns out that all functions of the type $Q_{j,-j-1}$ have zero norm, giving a precise meaning to the statement that the functions

$$Q_{jm} = (J^-)^{|m|-j} Y_{j,-j}, \quad m = -j-1, -j-2, \cdots, \tag{3}$$

"represent" the zero vector.

THE SCALAR PRODUCT AND ITS PROPERTIES

The double-valued spherical harmonics are defined iteratively as follows:

$$\begin{aligned} Y_{jj} &= N_j (\sin\theta)^j e^{ij\varphi}, \\ Y_{j,m-1} &= N_{jm}^{-1} J^- Y_{jm}, \quad m = j, j-1, \cdots, -j+1, \\ N_{jm} &= [(j+m)(j-m+1)]^{\frac{1}{2}}, \\ N_j &= \frac{1}{2\pi^2} \frac{2 \cdot 4 \cdots 2j+1}{1 \cdot 3 \cdots 2j}. \end{aligned} \tag{4}$$

The constant N_j was chosen so that

$$\int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta \, d\theta \, Y_{jj}^*(\theta, \varphi) Y_{jj}(\theta, \varphi) = 1. \tag{5}$$

The complex vector space spanned by the Y_{jm} defined above has already been denoted by V_j. It is convenient also to designate the vector space spanned by the Q_{jm} defined in Eq. (3) by W_j, and by $\bar{V}_j$ the combined vector space

$$\bar{V}_j = V_j \cup W_j. \tag{6}$$

All of these vector spaces have a fixed angular momentum j. The scalar product will actually be defined for the larger space $\bar{V}$ defined by

$$\bar{V} = \bar{V}_{\frac{1}{2}} \cup \bar{V}_{\frac{3}{2}} \cup \bar{V}_{\frac{5}{2}} \cup \cdots. \tag{7}$$

The definition of the scalar product is as follows: Let Ψ and Ψ' be any two functions in $\bar{V}$; the scalar product of Ψ and Ψ' denoted by (Ψ, Ψ') is then

$$(\Psi, \Psi') = \int_0^{\pi} d\theta \left[\sin\theta \int_0^{2\pi} d\varphi \Psi^*(\theta, \varphi) \Psi'(\theta, \varphi) - f(\theta) \right], \tag{8}$$

[1] D. Pandres, Jr., J. Math. Phys. 6, 1098 (1965).
[2] P. A. M. Dirac, *Quantum Mechanics* (Oxford University Press, New York, 1958), 4th ed., p. 20.
[3] I. M. Gel'fand and Y. Ya Sapiro, Trans. Am. Math. Soc. No. 2 (1956).

where $f(\theta)$ is a singular function of the form

$$f(\theta) = \sum_{m=1}^{\infty} a_m(\sin\theta)^{-m} + \cos\theta \sum_{m=1}^{\infty} b_m(\sin\theta)^{-m}, \quad (9)$$

chosen so that the term in brackets in Eq. (8) becomes integrable. Speaking loosely, one expands the odd and even part of the usual integrand $\Psi^*\Psi' \sin\theta$ into powers of $\sin\theta$, drops those terms in the expansion which would give divergent integrals, and then integrates the remaining integrand over the sphere. This modification of the usual scalar product is reminiscent of the renormalization formalism for the removal of divergences in electrodynamics. The functional (Ψ, Ψ') satisfies the following identities necessary for a scalar product:

$$\begin{aligned} (\Psi, \Psi) &\geq 0, \\ (c\Psi, \Psi') &= c^*(\Psi, \Psi'), \\ (\Psi, \Psi')^* &= (\Psi', \Psi), \\ (\Psi + \Psi', \Psi'') &= (\Psi, \Psi'') + (\Psi', \Psi''), \end{aligned} \quad (10)$$

for all Ψ, Ψ', and Ψ'' in $\bar{V}$ with c an arbitrary complex constant. The last three of these identities follow in a straightforward manner from the definition of the scalar product in Eq. (8). However, the first relation, stating that the scalar product is nonnegative, is not so easily proved. In fact, it is not even true if Ψ is not restricted to lie in $\bar{V}$ (for example, take $\Psi = \cot\theta$). We shall therefore postpone its proof until the scalar products of the basis functions of $\bar{V}$ are determined.

To be precise, the bilinear functional (Ψ, Ψ') should not be called a scalar product, since positive-definiteness is usually taken as a necessary requirement.[4] For the same reason, the vector space $\bar{V}$ taken with the "scalar product" (Ψ, Ψ') is not precisely a Hilbert space. It is what mathematicians call a semi-normed[5] vector space. However, since these slight inadequacies will be cleared up presently by a further construction, we will retain the name "scalar product" for the bilinear functional.

The scalar product (Ψ, Ψ') has the following two important properties: When Ψ and Ψ' are non-singular, the scalar product reduces to a simple integral over the unit sphere; J^+ and J^- are mutual Hermitian adjoints

$$(J^+\Psi, \Psi') = (\Psi, J^-\Psi') \quad (11)$$

for all Ψ and Ψ' in $\bar{V}$. This latter fact is proved in the Appendix. Armed with the Hermitian property of Eq. (11), we can prove a number of important facts.

[4] P. R. Halmos, *Introduction to Hilbert Space* (Chelsea Publishing Co., New York, 1957), p. 13.

[5] A. E. Taylor, *Introduction to Functional Analysis* (John Wiley & Sons, Inc., New York, 1958), p. 143.

For example, consider the scalar product of $Q_{j,-j-1}$ and $Q_{j'm'}$:

$$\begin{aligned} &(Q_{j'm'}, Q_{j,-j-1}) \\ &\quad = [(J^-)^{|m'|-j'}Y_{j',-j'}, J^-Y_{j,-j}] \\ &\quad = [(J^-)^{|m'|-j'-1}Y_{j',-j'}, J^+J^-Y_{j',-j'}] = 0, \end{aligned} \quad (12)$$

when use is made of Eq. (2). This can be easily extended to the more general result

$$(Q_{jm}, Q_{j'm'}) = 0, \quad (13)$$

so that each Q_{jm} is a nonzero function of $\bar{V}$ with zero norm, proving the nondefiniteness of the scalar product. Using the same technique, the following further identities may be derived:

$$(Y_{jm}, Y_{j'm'}) = \delta_{jj'}\delta_{mm'}, \quad (14)$$

$$(Y_{jm}, Q_{j'm'}) = 0. \quad (15)$$

The scalar product therefore has the necessary property of producing orthonormal Y_{jm} (with the usual choice of constants N_{jm}), thus yielding the usual unitary representations of $SU(2)$ when the matrices

$$(Y_{jm}, J_\alpha Y_{jm'})$$

are formed. Also, Eqs. (13)–(15) immediately lead to the validity of the nonnegative condition in Eq. (10).

According to Eqs. (13) and (15), the Q_{jm} are orthogonal to all Ψ in $\bar{V}$. This strongly suggests that all of the Q_{jm} are essentially zero. We shall now proceed to show in just what sense this is true and, at the same time, construct a Hilbert space with a strictly positive scalar product.

We begin by defining an equivalence relation in $\bar{V}$: Two functions Ψ and Ψ' in $\bar{V}$ will be called equivalent if $\Psi - \Psi'$ is in $W = W_{\frac{1}{2}} \cup W_{\frac{3}{2}} \cup \cdots$.

Let $|\Psi\rangle$ denote the set of all functions in $\bar{V}$ which are equivalent to Ψ, called the equivalence class of Ψ (the class could also be labeled by any other function equivalent to Ψ). Consider the set of all equivalence classes in $\bar{V}$, usually denoted by $\bar{V}/W$. It is easy to see that $\bar{V}/W$ is in fact a vector space spanned by the vectors $|Y_{jm}\rangle$, which consists of all functions in $\bar{V}$ equivalent to Y_{jm}. A scalar product can be defined for the vector space of equivalence classes as follows: Let $|1\rangle$ and $|2\rangle$ be vectors in $\bar{V}/W$ with Ψ_1 and Ψ_2 two elements of the respective equivalence classes. The scalar product of $|1\rangle$ and $|2\rangle$, denoted by $\langle 1 \mid 2\rangle$, is defined by

$$\langle 1 \mid 2\rangle = (\Psi_1, \Psi_2). \quad (16)$$

Equations (13)–(15) insure that the scalar product is independent of which representative elements Ψ_1 and Ψ_2 are chosen. The zero equivalence class $|0\rangle$ is precisely W, so that, again according to Eq. (13), $\langle \Psi \mid \Psi\rangle$

vanishes if and only if $|\Psi\rangle = |0\rangle$, and the scalar product $\langle\Psi|\Psi'\rangle$ is strictly positive. The statement that W is the zero vector in the space of equivalence classes is the precise way of saying that all Q_{jm} are essentially zero.

The vector space $\bar{V}/W$ along with the scalar product $\langle\Psi|\Psi'\rangle$ now constitute a Hilbert space

$$H = \bar{V}/W, \tag{17}$$

which can clearly be decomposed into the subspaces

$$H = H_{\frac{1}{2}} \cup H_{\frac{3}{2}} \cup \cdots, \tag{18}$$

where

$$H_j = \bar{V}_j/W_j. \tag{19}$$

Each of the Hilbert spaces H_j are $(2j+1)$-dimensional, spanned by the orthonormal basis $|Y_{jm}\rangle$, forming the basis for irreducible unitary representations of $SU(2)$ with half-odd integer angular momentum.

Finally, we note that if one enlarges the space to include the usual single-valued spherical harmonics and modifies the scalar product defined in Eq. (8) by integrating φ from 0 to 4π, one obtains a single Hilbert space defined on the double sphere which carries all of the irreducible unitary representations of $SU(2)$.

APPENDIX

In the main body of this paper, we gave no proof of Eq. (11), which states that the matrices representing J^+ and J^- are Hermitian adjoints of each other. Note first of all that it is sufficient to prove this equation when Ψ and Ψ' are a pair of basis functions Y_{jm} or Q_{jm}. Secondly, if m and m' are the J_z values of Ψ and Ψ', respectively, Eq. (11) will be trivially satisfied if $m' \neq m+1$. Therefore, let us assume that $m' = m+1$. By examining the general forms for Y_{jm} and Q_{jm}, it is easy to show that Ψ will be a sum of terms of the form

$$\Phi = c(\cos\theta)^{\lambda}(\sin\theta)^{2n+m}e^{im\varphi}, \tag{A1}$$

where c is a constant, $\lambda = 0, 1$, and n is a positive integer. Similarly, Ψ' will be a sum of terms of the form

$$\Phi' = c'(\cos\theta)^{\lambda'}(\sin\theta)^{2n'+m+1}e^{i(m+1)\varphi}. \tag{A2}$$

It is clearly sufficient to prove Eq. (11) with Ψ and Ψ' replaced by Φ and Φ', respectively.

Denote the right and left side of Eq. (11) by R and L, respectively. Then substituting Φ and Φ' into the defining Eq. (8) and using the standard forms for J^+ and J^- yields

$$R = 2\pi c^*c'\int_0^{\pi} d\theta[A(\theta) - f_R(\theta)], \tag{A3}$$

$$L = 2\pi c^*c'\int_0^{\pi} d\theta[B(\theta) - f_L(\theta)], \tag{A4}$$

where

$$\begin{aligned} A(\theta) &= \frac{1}{2\pi c^*c'}\int_0^{2\pi} d\varphi\Phi^*J^-\Phi' \\ &= [\lambda'(\cos\theta)^{\lambda+\lambda'-1} - (2n'+2m+2) \\ &\quad \times (\cos\theta)^{\lambda+\lambda'+1}(\sin\theta)^{-2}](\sin\theta)^N, \end{aligned} \tag{A5}$$

$$\begin{aligned} B(\theta) &= \frac{1}{2\pi c^*c'}\int_0^{2\pi} d\varphi(J^+\Phi)^*\Phi' \\ &= [-\lambda(\cos\theta)^{\lambda+\lambda'-1} + 2n(\cos\theta)^{\lambda+\lambda'+1} \\ &\quad \times (\sin\theta)^{-2}](\sin\theta)^N, \end{aligned} \tag{A6}$$

where $N = 2n + 2n' + 2m + 3$ is an even integer. In general, $\lambda + \lambda' = 0$, 1, or 2. However, if $A(\theta)$ and $B(\theta)$ are odd functions of $\cos\theta$, R and L will both vanish. Therefore, we may assume that $\lambda + \lambda' = 1$. $A(\theta)$ and $B(\theta)$ can then be written in the form

$$\begin{aligned} A(\theta) &= [(2n'+2m+3-\lambda) \\ &\quad - (2n'+2m+2)(\sin\theta)^{-2}](\sin\theta)^N, \\ B(\theta) &= [-(2n+\lambda) + 2n(\sin\theta)^{-2}](\sin\theta)^N. \end{aligned} \tag{A7}$$

There are now three separate cases to be considered.

(a) When $N \leq -2$, it is necessary that $f_R = A$ and $f_L = B$, so that $R = L = 0$.

(b) When $N = 0$,

$$\begin{aligned} f_R &= -(2n'+2m+2)(\sin\theta)^{-2}, \\ f_L &= (2n+1)(\sin\theta)^{-2} \end{aligned} \tag{A8}$$

and

$$\begin{aligned} R &= 2\pi^2c^*c'(2n'+2m+3-\lambda), \\ L &= -2\pi^2c^*c'(2n+\lambda), \end{aligned} \tag{A9}$$

so that

$$\begin{aligned} R = L &= 2\pi^2c^*c'(2n'+2n+2m+3) \\ &= 2\pi^2c^*c'N = 0, \end{aligned} \tag{A10}$$

and again $R = L$.

(c) When $N \geq 2$, $f_R = f_L = 0$ and

$$\begin{aligned} R - L = 2\pi c^*c'\int_0^{\pi} d\theta[N(\sin\theta)^N \\ - (N-1)(\sin\theta)^{N-2}] = 0 \end{aligned} \tag{A11}$$

by direct computation. [Actually, for this last case when neither Φ of Φ' is sufficiently singular, Eq. (11) follows simply by integration by parts.]

Thus, for all three cases, $R = L$, so that Eq. (11) is valid when Ψ and Ψ' are replaced by Φ and Φ', respectively, which, as explained above, is sufficient to demonstrate the validity of Eq. (11) in general.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Analytic Renormalization

EUGENE R. SPEER*
Department of Mathematics, Princeton University, Princeton, New Jersey

(Received 1 December 1967)

Renormalized Feynman amplitudes are defined by a method of analytic continuation in subsidiary parameters. The results are shown to belong to the class of renormalized amplitudes defined by Boguliubov, Parasiuk, and Hepp.

1. INTRODUCTION

In the perturbation-series expansion of the S matrix or the time-ordered vacuum expectation values in a Lagrangian field theory, there occur formal expressions of the form

$$\prod_{l\in\mathfrak{L}} \Delta_F^l(x_{i_l} - x_{f_l}), \tag{1.1}$$

where $\mathfrak{L}$ is the collection of lines of a certain Feynman graph $G(V_1, \cdots, V_n; \mathfrak{L})$, with vertices $\{V_i\}$, and V_{i_l} and V_{f_l} are the initial and final vertices of the lth line. Δ_F^l is given in p space by

$$\tilde{\Delta}_F^l(p) = iP_l(p)(p^2 - m_l^2 + i0)^{-1}, \tag{1.2}$$

with $P_l(p)$ a polynomial of degree r_l. In general, however, (1.1) is not well defined (even as a distribution) because the convolutions in p space diverge. In the theory of renormalization, (1.1) is given a well-defined meaning by a variety of methods, among which that of Hepp[1] is distinguished by its mathematical coherence.

In this paper we apply to (1.1) a method of defining divergent quantities which was originated by Riesz[2] and has been used in various contexts by many authors.[3] To define a formally divergent quantity I, these authors introduce a function $I(\lambda)$, analytic in some region Ω of the complex plane, and defined by an expression which is formally equal to I for $\lambda = \lambda_0$. I is then defined as the analytic continuation of $I(\lambda)$ from Ω to $\lambda = \lambda_0$. In some cases $I(\lambda)$ has a pole at λ_0; an acceptable definition of I may then be obtained as the constant term of the Laurent series of $I(\lambda)$ about λ_0.

To apply these techniques to (1.1) we find it necessary to consider functions of several complex variables $\lambda_1, \cdots, \lambda_L$, one associated with each line of the Feynman graph. The main difficulty is the extension of the above treatment of poles to the more complicated singularities which occur in several complex variables. Such an extension is given and a renormalized value of (1.1) is defined. It is shown that this definition is one of the class of renormalized values of (1.1) defined by Boguliubov, Parasiuk, and Hepp.[1]

We remark that we are interested only in defining (1.1) as a tempered distribution in $\mathcal{S}'(R^{4n})$. We restrict attention to the case of $m_l > 0$, and without loss of generality assume that $G(V_1, \cdots, V_n; \mathfrak{L})$ is connected.

2. ANALYTIC PROPERTIES

We generalize (1.2) by defining, for any complex λ_l,[4]

$$\tilde{\Delta}_{\lambda_l}^l(p) = P_l(p)e^{\frac{1}{2}i\pi\lambda_l}(p^2 - m_l^2 + i0)^{-\lambda_l}, \tag{2.1}$$

and use Hepp's regularization to write, for $\operatorname{Re}\lambda_l > 0$,

$$\Delta_{\lambda_l}^l = \lim_{\epsilon\to 0+} \lim_{r\to 0+} \Delta_{\lambda_l,\epsilon,r}^l,$$

where

$$\tilde{\Delta}_{\lambda_l,\epsilon,r}^l(p) = P_l(p)\Gamma(\lambda_l)^{-1} \times \int_r^\infty d\alpha_l \alpha_l^{\lambda_l - 1} \exp i\alpha_l(p^2 - m_l^2 + i\epsilon). \tag{2.2}$$

The distributions $\Delta_{\lambda_l}^l$ and $\Delta_{\lambda_l,\epsilon,r}^l$ are entire functions of λ_l. Moreover, when $\epsilon > 0$ and $r > 0$, $\tilde{\Delta}_{\lambda_l,\epsilon,r}^l$ is in is in $\mathcal{O}_C'(R^4)$ (the space of rapidly decreasing distributions), and its Fourier transform $\Delta_{\lambda_l,\epsilon,r}^l$ is in $\mathcal{O}_M(R^4)$ (the space of polynomially bounded infinitely differentiable functions).[5] Thus we may define unambiguously

$$\mathcal{C}_{\lambda_1,\cdots,\lambda_L,\epsilon,r}(V_1, \cdots, V_n; \mathfrak{L}) = \prod_{l\in\mathfrak{L}} \Delta_{\lambda_l,\epsilon,r}^l(x_{i_l} - x_{f_l}). \tag{2.3}$$

---

* Supported by a National Science Foundation Graduate Fellowship.

[1] K. Hepp, Commun. Math. Phys. **2**, 301 (1966). See also N. N. Bogoliubov and O. S. Parasiuk, Acta Math. **97**, 227 (1957); O. S. Parasiuk, Ukr. Math. J. **12**, 287 (1960).

[2] M. Riesz, Acta Math. **81**, 1, 1949.

[3] See, e.g., N. E. Fremberg, Proc. Roy. Soc. (London) **A188**, 18 (1946); T. Gustafson, Arkiv Mat. Astron. Fysik **34A** No. 2 (1947); S. B. Nilsson, Arkiv Fysik **1**, 369 (1950); G. Källen, Arkiv Fysik **5**, 130 (1951); E. Karlson, Arkiv Fysik **7**, 221 (1954); I. M. Gel'fand and G. E. Shilov, *Generalized Functions, Vol. I* (Academic Press Inc., New York, 1964), Chap. 3; and C. G. Bollini, J. J. Ciambiagi, and A. Gonzalez Dominguez, Nuovo Cimento **31**, 550 (1964).

[4] See I. M. Gel'fand and G. E. Shilov, Ref. 3, Chap. 3, Sec. 2.4. This is a good basic reference for the properties of distributions depending analytically on a parameter.

[5] These spaces are discussed in L. Schwarts, *Théorie des distributions* (Hermann & Cie., Paris, 1966), pp. 243–244.

In this section we investigate the analytic properties of (2.3) after the limit $r \to 0+$. For convenience we write $(\lambda_1, \cdots, \lambda_L) = \boldsymbol{\lambda}$.

We remark that our results in this section would not be changed if, in (2.1), we also generalized $P_l(p)$ to $P_l(\lambda_l, p)$. Here $P_l(\lambda_l, p)$ is a covariant polynomial in p of degree r_l, whose coefficients are entire functions of λ_l which satisfy $P_l(1, p) = P_l(p)$. Consistent renormalization of a theory would require in addition that $P_l(\lambda_l, p)$ depend only on the particle associated with the lth line. Such a change in P_l would result in a finite change in the renormalization constants.

Theorem 1: Let $G(V_1, \cdots, V_n; \mathfrak{L})$ be a connected Feynman graph, as above. Define $N = L - n + 1$ to be the number of loops of G, and $\Omega = \{\boldsymbol{\lambda} \in \mathbb{C}^L \mid \operatorname{Re} \lambda_l > M, l = 1, \cdots, L\}$, where $M = N(2 + \sum_1^L r_l)$. For $\boldsymbol{\lambda} \in \Omega$, define

$$\mathcal{C}_{\boldsymbol{\lambda},\epsilon}(V_1, \cdots, V_n; \mathfrak{L}) = \lim_{r \to 0+} \mathcal{C}_{\boldsymbol{\lambda},\epsilon,r}(V_1, \cdots, V_n; \mathfrak{L}). \quad (2.4)$$

Then: (a) The limit (2.4) exists [in $\mathcal{S}'(R^{4n})$] and $\mathcal{C}_{\boldsymbol{\lambda},\epsilon}(V_1, \cdots, V_n; \mathfrak{L})$ is holomorphic in Ω.

(b) $\mathcal{C}_{\boldsymbol{\lambda},\epsilon}(V_1, \cdots, V_n; \mathfrak{L})$ may be analytically continued to a function meromorphic in $\mathbb{C}^L$. If we use the same notation for the continued function, then

$$\mathcal{C}_{\boldsymbol{\lambda},\epsilon}(V_1, \cdots, V_n; \mathfrak{L}) \prod_A \Gamma\Big[\sum_{l \in A} (\lambda_l - M)\Big]^{-1} \quad (2.5)$$

is holomorphic in $\mathbb{C}^L$. Here $\prod_A$ is taken over all subsets A of $\{1, \cdots, L\}$.

We remark that a more detailed discussion of the singularities of $\mathcal{C}_{\boldsymbol{\lambda},\epsilon}$ is possible but is not needed in this paper.

Proof: Let p_j be the momentum dual to x_j. We may evaluate (2.3) in p space by attaching to each vertex V_j an external line directed into the diagram and carrying momentum p_j, and then applying the integration methods of Chisholm.[6] That is, we assign paths through the diagram for the external momenta and choose loops and loop momenta $k_1, \cdots, k_N$, so that the lth line is assigned momentum

$$q_l = \sum_{i=1}^{N} a_{li} k_i + \sum_{j=1}^{n} b_{lj} p_j. \quad (2.6)$$

Then (2.3) becomes

$$\mathcal{C}_{\boldsymbol{\lambda},\epsilon,r}(V_1, \cdots, V_n; \mathfrak{L}) = \delta\Big(\sum_1^n p_j\Big) \int dk_1 \cdots dk_N \prod_{l=1}^{L} \tilde{\Delta}^l_{\lambda_l,\epsilon,r}(q_l). \quad (2.7)$$

[6] J. S. R. Chisholm, Proc. Cambridge Phil. Soc. **48**, 300 (1952). See, e.g., R. J. Eden, P. V. Landshoff, D. I. Olive, and J. C. Polkinghorne, *The Analytic S-Matrix* (Cambridge University Press, Cambridge, England, 1966), pp. 31–34.

If we interchange the k and α integrations and use $k_i = -i(\partial/\partial S_i) e^{ik_i S_i}\big|_{S_i = 0}$ in the factors $P_l(q_l)$, we may write (2.7) as a sum of terms of the form

$$\begin{aligned}(\text{const})\delta(\textstyle\sum p_j) A(p) \int_r^\infty \cdots \int_r^\infty \prod_1^L [d\alpha_l \alpha_l^{\lambda_l - 1} \Gamma(\lambda_l)^{-1}] \\ \times \Big\{ \int dk_1 \cdots dk_N A'(-i\nabla_S) \exp i\Big[\sum_{i,j=1}^{N} \theta_{ij} k_i k_j \\ + \sum_1^N (2\phi_i + S_i) k_i + \psi + i\epsilon \sum_1^L \alpha_l \Big] \Big\}\Big|_{S=0}. \quad (2.8)\end{aligned}$$

Here A and A' are monomials of degree $\leq \rho = \sum_1^L r_l$, and

$$\theta_{ij} = \sum_{l=1}^{L} \alpha_l a_{li} a_{lj}, \quad (2.9a)$$

$$\phi_i = \sum_{l=1}^{L} \sum_{j=1}^{n} \alpha_l a_{li} b_{lj} p_j, \quad (2.9b)$$

$$\psi = \sum_{l=1}^{L} \sum_{j,k=1}^{n} \alpha_l b_{lj} b_{lk} p_j p_k - \sum_{l=1}^{L} \alpha_l m_l^2. \quad (2.9c)$$

When all α_l are positive, θ_{ij} is positive-definite. Thus, if we now do the k integrations, the bracket in (2.8) becomes, up to a constant factor,

$$\begin{aligned}(\det \theta)^{-2} A'(-i\nabla_S) \exp i\Big[\psi - \frac{1}{4} \sum_{i,j=1}^{N} (2\phi_i + S_i) \\ \times (\theta^{-1})_{ij} (2\phi_j + S_j) + i\epsilon \sum_1^L \alpha_l \Big].\end{aligned}$$

Using $\theta^{-1} = \theta^{\mathrm{Ad}}/\det \theta$, where θ^{Ad} is the transpose of the matrix of cofactors, performing the S derivatives, and setting $S = 0$, we may finally write

$$\begin{aligned}&\mathcal{C}_{\boldsymbol{\lambda},\epsilon,r}(V_1, \cdots, V_n; \mathfrak{L}) \\ &= \sum_{m=0}^{\rho} \delta(\textstyle\sum p_j) \int_r^\infty \cdots \int_r^\infty \prod_1^L [d\alpha_l \alpha_l^{\lambda_l - 1} \Gamma(\lambda_l)^{-1}] \\ &\quad \times B_m(p, \alpha) C(\alpha)^{-(m+2)} \exp i[D(\alpha, p)/C(\alpha) + i\epsilon \textstyle\sum \alpha_l],\end{aligned} \quad (2.10)$$

where B_m is a polynomial, $C(\alpha) = \det \theta$, and $D(\alpha, p) = \det \chi$, with

$$\chi = \begin{vmatrix} \theta_{11} & \cdots & \theta_{1N} & \phi_1 \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \theta_{N1} & & & \phi_N \\ \phi_1 & \cdots & \phi_N & \psi \end{vmatrix}. \quad (2.11)$$

The "ultraviolet divergences" occur in the limit $r \to 0+$ because $C(\alpha)$ vanishes when certain $\alpha_l \to 0$. We now investigate this behavior in a region $0 \leq \alpha_{l_1} \leq \cdots \leq \alpha_{l_L}$; for simplicity, we consider

$$0 \leq \alpha_1 \leq \alpha_2 \cdots \leq \alpha_L. \quad (2.12)$$

Within this region we introduce new variables $t_1, \cdots, t_L$, defined by $\alpha_l = t_L t_{L-1} \cdots t_l$, so that (2.12) becomes

$$0 \leq t_L \leq \infty,$$
$$0 \leq t_l \leq 1 \quad \text{if} \quad l = 1, \cdots, L-1. \tag{2.13}$$

Let G_l be the graph consisting of lines 1 through l with their vertices, and let N_l be the number of loops of G_l.

Lemma 1: For α in (2.12),

$$C(\alpha) = \prod_1^L t_l^{N_l} E(t_1, \cdots, t_{L-1}), \tag{2.14a}$$

$$D(\alpha, p) = t_L \prod_1^L t_l^{N_l} F(t_1, \cdots, t_{L-1}, p), \tag{2.14b}$$

where E and F polynomials, and E does not vanish in (2.13).

Proof: Since $N_1 = 0$, $N_L = N$, and $(N_{l+1} - N_l)$ is always 0 or 1, there exist integers $1 < l_1 < \cdots < l_N \leq L$ such that $N_{l_i} = N_{(l_i - 1)} + 1$. Thus we may choose loop variables so that the ith loop is contained in G_{l_i}, that is, so that $a_{l_i} = 0$ unless $l \leq l_i$ [see (2.6)]. From (2.9) and (2.11) we see that the ith row and column $(1 \leq i \leq N)$ of θ and χ contain a factor $t_L \cdots t_{l_i}$, and the $(N+1)$th row and column of D contain a factor t_L. We remove these factors from the rows to produce new matrices θ' and χ'; this gives (2.14) with $E = \det \theta'$, $F = \det \chi'$.

To show that E does not vanish, we consider instead of θ' the matrix θ'', obtained from θ by removing a factor $(t_L \cdots t_{l_i})^{\frac{1}{2}}$ from the ith row *and* column of θ. θ'' is symmetric, and $E = \det \theta''$. Suppose $E(t) = 0$ at some point $t = \tau$ in (2.13). Then there exist numbers $\delta_1, \cdots, \delta_N$ such that $\sum \delta_i \theta''_{ij}(\tau) \delta_j = 0$, or

$$\sum_{l=1}^{L} \left[\sum_{i=1}^{N} \delta_i a_{li} \prod_{l \leq l' < l_i} \tau_{l'}^{\frac{1}{2}} \right]^2 = 0. \tag{2.15}$$

Each term in the sum over l must vanish. Let $I = \max \{i \mid \delta_i \neq 0\}$ and consider the term with $l = l_I$. $\delta_i = 0$ for $i > I$, while $a_{l_i i} = 0$ for $i < I$. Thus we must have $\delta_I a_{l_I I} = 0$. But $\delta_I \neq 0$, and the Ith loop must go through the l_Ith line, so $a_{l_I I} \neq 0$. This contradiction proves the lemma.

Now consider an integrand of (2.10) in the region (2.12) and change variables to $t_1, \cdots, t_L$. The Jacobian of this change is $\prod_1^L t_l^{l-1}$, so that (2.10) becomes a sum of terms of the form

$$\delta(\textstyle\sum p_j) \displaystyle\int_r^\infty dt_L \int^1 dt_{L-1} \cdots \int^1 dt_1 \times \prod_1^L [\Gamma(\lambda_l)^{-1} t_l^{(\mu_l - (m+2)N_l - 1)}] B'_m(p, t) E(t)^{-(m+2)} \times \exp i t_L [F/E + i\epsilon(1 + t_{L-1} + \cdots)], \tag{2.16}$$

where $\mu_l = \sum_{l'=1}^{l} \lambda_{l'}$, and B'_m is a polynomial. The lower limits of the $t_{L-1}, \cdots, t_1$ integrations in (2.16) are complicated functions whose only relevant property is that they approach 0 when $r \to 0+$. For $\lambda \in \Omega$, $\operatorname{Re} \mu_l > (m+2)N$, so the integrand of (2.16) is absolutely integrable in all of the region (2.13). This justifies the limit $r \to 0+$ in Ω; the analyticity is clear. Thus $\mathcal{C}_{\lambda,\epsilon}(V_1, \cdots, V_n; \mathfrak{L})$ for $\boldsymbol{\lambda} \in \Omega$ is a sum of terms of the form (2.16) with 0 as the lower limit on all integrals, and, in general, with $\mu_l = \sum_{l' \in A} \lambda_{l'}$ for some $A \subset \{1, \cdots, L\}$.

We now prove part (b) of the theorem. Given a positive integer M', we may construct a continuation of (2.16) into the region

$$\Omega_{M'} = \{\boldsymbol{\lambda} \in \mathbb{C}^L \mid \operatorname{Re} \lambda_l > X_{M'}, \quad l = 1, \cdots, L\},$$

where

$$X_{M'} = \begin{cases} M - M', & \text{if } M - M' \geq 0, \\ (M - M')/L, & \text{if } M - M' < 0, \end{cases}$$

as follows. We do M' integrations by parts with respect to each of $t_1, \cdots, t_{L-1}$, integrating the factor $t_l^{[\mu_l - (m+2)N_l - 1]}$ (or the higher powers of t arising from this) and differentiating the rest. This is permissible for $\boldsymbol{\lambda} \in \Omega$; in each partial integration the integrated terms vanish as the lower limit. Finally, the t_L integration may be done explicitly with the use of the formula

$$\int_0^\infty dt\, t^{\mu-1} e^{it\kappa} = e^{\frac{1}{2}i\pi\mu} \Gamma(\mu) \kappa^{-\mu}, \tag{2.17}$$

valid for $\operatorname{Re} \mu > 0$, $\operatorname{Im} \kappa > 0$. Thus $\mathcal{C}_{\lambda,\epsilon}(V_1, \cdots, V_n; \mathfrak{L})$ may be written as a sum of terms of the form

$$H(\boldsymbol{\lambda}) \int_0^1 \cdots \int_0^1 \prod_{l=1}^{L'} \{dt'_l t'^{[\mu_l - (m+2)N_l + M' - 1]}_l\} G(t', p, \epsilon) E(t')^i \times [F/E + i\epsilon(1 + t_{L-1} + \cdots)]^{(j - \sum_1^L \lambda_l)}. \tag{2.18}$$

Here $\{t'_1, \cdots, t'_L\}$ is a subset of $\{t_1, \cdots, t_{L-1}\}$ (the rest having been set equal to 1 during some partial integration), G is a polynomial, i and j are integers, and $H(\boldsymbol{\lambda})$ contains factors from (2.17) as well as factors $(\mu_l - k)^{-1}$ arising from the partial integrations. Since $[\operatorname{Re} \mu_l - (m+2)N_l + M'] > 0$ for $\boldsymbol{\lambda} \in \Omega_{M'}$, (2.18) provides a continuation of $\mathcal{C}_{\lambda,\epsilon}(V_1, \cdots, V_n; \mathfrak{L})$ to the region $\Omega_{M'}$; moreover,

$$H(\boldsymbol{\lambda}) \prod_{A \subset \{1 \cdots L\}} \Gamma\left[\sum_{l \in A} (\lambda_l - M) \right]^{-1}$$

is an entire function of $\boldsymbol{\lambda}$. Since $\Omega_{M'}$ increases to $\mathbb{C}^L$ as M' approaches infinity, part (b) of the theorem is proved.

3. RENORMALIZATION

It would now be natural to define (1.1) as

$$\lim_{\epsilon\to 0} \mathcal{C}_{1,\cdots 1,\epsilon}(V_1,\cdots,V_n;\mathfrak{L});$$

however, Theorem I implies that $\mathcal{C}_{\lambda,\epsilon}$ may have a complicated singularity at $\boldsymbol{\lambda}=(1,\cdots,1)$. In one complex variable we could discard the singular part by using the constant term of the Laurent series. In this section we generalize this procedure to several variables.

Definition: Let $U\subset\mathbb{C}^L$ $(L\geq 1)$ be an open neighborhood of $(1,\cdots,1)$. Let $\mathcal{A}_L(U)=\{f(\boldsymbol{\lambda})\mid f(\boldsymbol{\lambda})\prod_{A\subset\{1,\cdots,L\}}[\sum_{l\in A}(\lambda_l-1)]^m$ is analytic in U for some integer $m\geq 0\}$, and let $\mathcal{A}_L=\cup\mathcal{A}_L(U)$, the union taken over all neighborhoods U. Then a family of maps $\mathcal{F}=\{\mathcal{F}_L\}_{L=1}^{\infty}$, $\mathcal{F}_L:\mathcal{A}_L\to\mathbb{C}$, is a *generalized evaluator* [at $(1,\cdots,1)$] if the following conditions are satisfied for each L:

(1) $\mathcal{F}_L$ is linear;

(2) if $f\in\mathcal{A}_L$ is analytic at $(1,\cdots,1)$, then $\mathcal{F}_L f=f(1,\cdots,1)$;

(3) if $f_n\in\mathcal{A}_L(U)$, for

$$n=0,1,\cdots,g_n(\boldsymbol{\lambda})=f_n(\boldsymbol{\lambda})\prod_A[\textstyle\sum_{l\in A}(\lambda_l-1)]^m,$$

is analytic in U, and $g_n\to g_0$ uniformly on U, then $\mathcal{F}_L f_n\to\mathcal{F}_L f_0$;

(4) if σ is a permutation of $\{1,\cdots,L\}$, $f\in\mathcal{A}_L$, and $f_\sigma\in\mathcal{A}_L$ is defined by

$$f_\sigma(\lambda_1,\cdots,\lambda_L)=f(\lambda_{\sigma(1)},\cdots,\lambda_{\sigma(L)}),$$

then $\mathcal{F}_L f_\sigma=\mathcal{F}_L f$;

(5) if $f\in\mathcal{A}_L$ depends only on $\lambda_1,\cdots,\lambda_{L'}$, where $L'<L$, then $\mathcal{F}_{L'}f=\mathcal{F}_L f$;

(6) if $f_1,f_2\in\mathcal{A}_L$, and f_1 depends only on $\lambda_1,\cdots,\lambda_{L'}$, f_2 only on $\lambda_{L'+1},\cdots,\lambda_L$, then $\mathcal{F}_L(f_1f_2)=(\mathcal{F}_L f_1)\times(\mathcal{F}_L f_2)$.

If $f\in\mathcal{A}_L$, we use Conditions (4) and (5) to write without ambiguity $\mathcal{F}f=\mathcal{F}_L f=\mathcal{F}_{L'}f$ for any $L'\geq L$. Conditions (1)–(5) are rather natural; the utility of (6) will be shown in Sec. 5. It is this condition which would be violated by setting $\lambda_1=\cdots=\lambda_L=\lambda$ and defining $\mathcal{F}f$ as the constant term of the Laurent series of $f(\lambda,\lambda,\cdots,\lambda)$ at $\lambda=1$.

Example: Suppose $f\in\mathcal{A}_L(U)$, and let U contain the poly disc $|\lambda_l-1|<R$. Choose $0<R_1<\cdots<R_L<R$, in such a way that $R_i>\sum_{j=1}^{i-1}R_j$, and let C_i be the contour $|z-1|=R_i$ oriented counterclockwise. Define

$$\mathcal{F}_L f=\frac{1}{L!}\sum_\sigma\frac{1}{(2\pi i)^L}\int_{C_{\sigma(1)}}d\lambda_1\cdots\times\int_{C_{\sigma(L)}}d\lambda_L f(\lambda)\prod_1^L(\lambda_l-1)^{-1},\quad(3.1)$$

where $\sum_\sigma$ runs over all permutations σ of $\{1,\cdots,L\}$. One easily checks that $\mathcal{F}$ is well defined, independent of the choice of $\{R_i\}$, and satisfies (1)–(6).

We want to be able to apply a generalized evaluator to meromorphic distributions. Consider such a distribution:

$$S(\boldsymbol{\lambda})=S'(\boldsymbol{\lambda})\prod_{A\subset\{1\cdots L\}}\Big[\sum_{l\in A}(\lambda_l-1)\Big]^{-m},$$

where $S'(\boldsymbol{\lambda})$ is an analytic function of $(\lambda_1,\cdots,\lambda_L)$ in some neighborhood U of $(1,\cdots,1)$, taking values in $\mathcal{S}'(R^n)$. Then the formula $(\mathcal{F}_L S)(\psi)=\mathcal{F}_L(S(\psi))$ defines a linear functional $\mathcal{F}_L S$ on $\mathcal{S}(R^n)$. Now $S':U\to\mathcal{S}'(R^n)$ is continuous [when $\mathcal{S}'(R^n)$ is given the usual weak topology], so that if $K\to U$ is compact, $S'(K)$ is (weakly) compact in $\mathcal{S}'(R^n)$, and hence is strongly bounded.[7] That is, there is a constant C_K and a norm $\|\ \|$ on $\mathcal{S}(R^n)$ (one of the norms defining the topology) such that $|S'(\boldsymbol{\lambda})(\psi)|\leq C_K\|\psi\|$ for any $\boldsymbol{\lambda}\in K$ and any $\psi\in\mathcal{S}(R^n)$. So for any sequence $\{\psi_i\}$ of elements of $\mathcal{S}(R^n)$, converging to an element ψ_0, the sequence $\{S'(\boldsymbol{\lambda})(\psi_i)\}$ converges uniformly for $\boldsymbol{\lambda}\in K$ to $S'(\boldsymbol{\lambda})(\psi_0)$. Then property (3) of $\mathcal{F}$ implies that $\mathcal{F}S$ as defined above is continuous.

Definition: The renormalized value of (1.1) is defined to be

$$\mathcal{C}(V_1,\cdots,V_n;\mathfrak{L})=\lim_{\epsilon\to 0+}\mathcal{F}\mathcal{C}_{\lambda,\epsilon}(V_1,\cdots,V_n;\mathfrak{L}).\quad(3.2)$$

The existence of the $\epsilon\to 0+$ limit follows from the theorem we prove in Sec. 5: the agreement of this definition with that of Boguliubov, Parasiuk, and Hepp. It may also be proved directly that:

(a) $\lim_{\epsilon\to 0}\mathcal{C}_{\lambda,\epsilon}=\mathcal{C}_\lambda$ exists and is a meromorphic function of $\boldsymbol{\lambda}$ with the same singularities as $\mathcal{C}_{\lambda,\epsilon}$;

(b) $\mathcal{C}=\mathcal{F}\mathcal{C}_\lambda$.

We remark that a change in the generalized evaluator used in (3.2) is reflected in a finite change in the renormalization constants.

4. BOGULIUBOV–PARASIUK–HEPP RENORMALIZATION

We now review the renormalization methods of Boguliubov, Parasiuk, and Hepp,[1] and extend their results slightly. We follow the notation of Hepp.

Definition: A graph $G(V_1,\cdots,V_n;\mathfrak{L})$ is *one-particle irreducible* (OPI) if, for any $l\in\mathfrak{L}$ and $\mathfrak{L}'=\mathfrak{L}-\{l\}$, $G(V_1,\cdots,V_n;\mathfrak{L}')$ is connected. Otherwise G is *one-particle reducible* (OPR). A *generalized vertex* of G is a nonempty subset $U=\{V_1'\cdots V_m'\}$ of $\{V_1\cdots V_n\}$.

[7] I. M. Gel'fand and G. E. Schilow, *Verallgemeinerte Funktionen II* (VEB Deutscher Verlad der Wissenschaften, Berlin, 1962), Chap. I, Sec. 5.

If $U_1, \cdots, U_m$ are pairwise-disjoint generalized vertices, with $\bigcup_{i=1}^m U_i = \{V_1', \cdots, V_s'\}$, the graph $G(U_1, \cdots, U_m; \mathfrak{L})$ is obtained from $G(V_1', \cdots, V_s'; \mathfrak{L})$ by collapsing each generalized vertex U_i, and any lines which join two vertices in U_i, to a single point. The *superficial divergence* of $U = \{V_1', \cdots, V_m'\}$ is defined by

$$\nu(V_1', \cdots, V_m') = \sum_{\text{conn}} (r_l + 2) - 4(m-1), \quad (4.1)$$

where $\sum_{\text{conn}}$ runs over all lines of $\mathfrak{L}$ connecting different vertices of $\{V_1', \cdots, V_m'\}$. We do not distinguish between the vertex V_i and the generalized vertex $\{V_i\}$.

Definition: A *finite renormalization* is a map assigning to each generalized vertex $U = \{V_1', \cdots, V_m'\}$ a distribution $\hat{\mathfrak{X}}_\epsilon(V_1', \cdots, V_m'; \mathfrak{L})$ [also written $\hat{\mathfrak{X}}_\epsilon(U; \mathfrak{L})$] in $\mathcal{S}'(R^{4m})$ such that

$$\hat{\mathfrak{X}}_\epsilon(V_1', \cdots, V_m'; \mathfrak{L}) = \begin{cases} 1, & \text{for } m = 1, \\ 0, & \text{for IPR } G(V_1', \cdots, V_m'; \mathfrak{L}), \\ \delta\left(\sum_1^m p_j'\right) P_\epsilon(p_1', \cdots, p_m'), & \text{otherwise.} \end{cases} \quad (4.2)$$

Here P_ϵ is a covariant polynomial of degree $\leq \nu(V_1' \cdots V_m')$, whose coefficients approach finite limits as $\epsilon \to 0$, and which depends only on the structure of the graph $G(V_1', \cdots, V_m'; \mathfrak{L})$.

Definition: Given a finite renormalization $\hat{\mathfrak{X}}_\epsilon$, $U_1, \cdots, U_\tau$ pairwise-disjoint generalized vertices, define recursively for $\{U_1', \cdots, U_m'\} \subset \{U_1, \cdots, U_\tau\}$:

$$\mathfrak{X}_{\lambda,\epsilon,r}(U_1', \cdots U_m'; \mathfrak{L}) = \begin{cases} \mathfrak{X}_\epsilon(U_1'; \mathfrak{L}), & \text{if } m = 1, \quad (4.3a) \\ 0, & \text{for OPR } G(U_1', \cdots, U_m'; \mathfrak{L}), \quad (4.3b) \\ -\mathcal{M}\bar{\mathfrak{R}}_{\lambda,\epsilon,r}(U_1', \cdots, U_m'; \mathfrak{L}), & \text{otherwise,} \quad (4.3c) \end{cases}$$

$$\bar{\mathfrak{R}}_{\lambda,\epsilon,r}(U_1', \cdots, U_m'; \mathfrak{L}) = \sum_P{}' \prod_{j=1}^{k(P)} \mathfrak{X}_{\lambda,\epsilon,r}(U_{j1}^P, \cdots, U_{jr(j)}^P; \mathfrak{L}) \prod_{\text{conn}} \Delta_{\lambda_l,\epsilon,r}^l, \quad (4.4)$$

$$\mathfrak{R}_{\lambda,\epsilon,r}(U_1', \cdots, U_m'; \mathfrak{L}) = \bar{\mathfrak{R}}_{\lambda,\epsilon,r}(U_1', \cdots, U_m'; \mathfrak{L}) + \mathfrak{X}_{\lambda,\epsilon,r}(U_1', \cdots, U_m'; \mathfrak{L}). \quad (4.5)$$

Here $\sum_P'$ in (4.4) runs over all partitions of $\{U_1', \cdots, U_m'\}$ into $k(P) \geq 2$ disjoint subsets

$$\{U_{j1}^P, \cdots, U_{jr(j)}^P\}$$

and $\prod_{\text{conn}}$ runs over those $l \in \mathfrak{L}$ which connect different subsets of the partition. When

$$G(U_1', \cdots, U_m'; \mathfrak{L})$$

is OPI, and

$$\bigcup_{i=1}^m U_i' = \{V_1', \cdots, V_s'\},$$

then $\bar{\mathfrak{R}}$ has in p space the form $\delta(\sum_{i=1}^s p_i')F(p_1', \cdots, p_s')$, and $\mathcal{M}$ is the operation of truncating the Taylor series of F about the origin at order $\nu(V_1', \cdots, V_s')$ [$\mathcal{M} = 0$ if $\nu(V_1', \cdots, V_s') < 0$].

In the case where each U_i is a single vertex V_i, we also define

$$\mathfrak{X}'_{\lambda,\epsilon,r}(V_1', \cdots, V_m'; \mathfrak{L}) = \begin{cases} 1, & \text{if } m = 1, \quad (4.3a') \\ 0, & \text{for OPR } G(V_1', \cdots, V_m'; \mathfrak{L}), \quad (4.3b') \\ -\mathcal{M}\bar{\mathfrak{R}}'_{\lambda,\epsilon,r}(V_1', \cdots, V_m'; \mathfrak{L}) & \\ \quad +\hat{\mathfrak{X}}_\epsilon(V_1', \cdots, V_m'; \mathfrak{L}), & \text{otherwise,} \quad (4.3c') \end{cases}$$

$$\bar{\mathfrak{R}}'_{\lambda,\epsilon,r}(V_1', \cdots, V_m'; \mathfrak{L}) = \sum_P{}' \prod_{j=1}^{k(P)} \mathfrak{X}'_{\lambda,\epsilon,r}(V_{j1}^P, \cdots, V_{jr(j)}^P; \mathfrak{L}) \prod_{\text{conn}} \Delta_{\lambda_l,\epsilon,r}^l, \quad (4.4')$$

$$\mathfrak{R}'_{\lambda,\epsilon,r}(V_1', \cdots, V_m'; \mathfrak{L}) = \bar{\mathfrak{R}}'_{\lambda,\epsilon,r}(V_1', \cdots, V_m'; \mathfrak{L}) + \mathfrak{X}'_{\lambda,\epsilon,r}(V_1', \cdots, V_m'; \mathfrak{L}), \quad (4.5')$$

with $\sum_P'$, $\prod_{\text{conn}}$, and $\mathcal{M}$ as above. The following lemma may be proved by straightforward manipulation of these definitions.

Lemma 2: With the above definitions, we have

$$\bar{\mathfrak{R}}'_{\lambda,\epsilon,r}(V_1', \cdots, V_m'; \mathfrak{L}) = \sum_P \bar{\mathfrak{R}}_{\lambda,\epsilon,r}(U_1^P, \cdots, U_{m(P)}^P; \mathfrak{L}),$$

$$\mathfrak{X}'_{\lambda,\epsilon,r}(V_1', \cdots, V_m'; \mathfrak{L}) = \sum_P \mathfrak{X}_{\lambda,\epsilon,r}(U_1^P, \cdots, U_{m(P)}^P; \mathfrak{L}),$$

and hence

$$\mathfrak{R}'_{\lambda,\epsilon,r}(V_1', \cdots, V_m'; \mathfrak{L}) = \sum_P \mathfrak{R}_{\lambda,\epsilon,r}(U_1^P, \cdots, U_{m(P)}^P; \mathfrak{L}),$$

where $\sum_P$ runs over all partitions of $\{V_1', \cdots, V_m'\}$ into $m(P)$ generalized vertices $\{U_j^P\}$.

Now Boguliubov, Parasiuk, and Hepp define the renormalized value of (1.1) to be

$$\lim_{\epsilon \to 0+} \lim_{r \to 0+} \mathfrak{R}'_{1,\cdots,1,\epsilon,r}(V_1, \cdots, V_n; \mathfrak{L}); \quad (4.6)$$

that is, they define a class of values of (1.1) which depend on the finite renormalization used. The main result of Hepp is the existence of the $r \to 0+$ limit in (4.6); it may be generalized as follows.

Theorem 2: Let

$$\Omega' = \{\boldsymbol{\lambda} \in \mathbb{C}^L \mid \operatorname{Re} \lambda_l > 1 - 1/2L, \, l = 1, \cdots, L\}.$$

Then

$$\mathfrak{R}'_{\lambda,\epsilon}(V_1, \cdots, V_n; \mathfrak{L}) = \lim_{r \to 0+} \mathfrak{R}'_{\lambda,\epsilon,r}(V_1, \cdots, V_n; \mathfrak{L}) \quad (4.7)$$

exists in $\mathcal{S}'(R^{4n})$ and is analytic for $\boldsymbol{\lambda} \in \Omega'$.

Proof: Hepp actually proves the existence of

$$\lim_{r\to 0+} \mathcal{R}_{1,\cdots,1,\epsilon,r}(V_1, \cdots, V_n; \mathfrak{L}),$$

that is, the existence of (4.7) for $\lambda = (1, \cdots, 1)$ when $\mathcal{R}'$ is defined using zero finite renormalization. However, it is a trivial modification of his proof to show the existence and analyticity in Ω' of

$$\lim_{r\to 0+} \mathcal{R}_{\lambda,\epsilon,r}(U_1, \cdots, U_r; \mathfrak{L}),$$

for any pairwise-disjoint generalized vertices $U_1, \cdots, U_r$. The theorem then follows from Lemma 2.

5. EQUIVALENCE OF THE DEFINITIONS

In this section we show that our definition (3.2) of the renormalized amplitude agrees with the Boguliubov definition (4.6), calculated using a certain finite renormalization.

Definition: We write

$$J_L(\lambda) = \prod_{A\subset\{1,\ldots,L\}} \Gamma\Big[\sum_{l\in A}(\lambda_l - M)\Big] \Big[\text{recall } M = N\Big(2 + \sum_1^L r_l\Big)\Big].$$

Let $\mathcal{B}(L, m)$ be the set of mappings $\phi: \mathbb{C}^L \to \mathcal{S}'(R^{4m})$ with the form

$$\phi(\lambda)(p_1, \cdots, p_m) = \delta\Big(\sum_{i=1}^m p_i\Big) J_L(\lambda) f(\lambda, p_1, \cdots, p_m), \tag{5.1}$$

where

(a) $f \in C^\infty(R^{2L+4m})$;

(b) f is analytic in λ for fixed p;

(c) if D is a monomial in the p derivatives and $K \subset \mathbb{C}^L$ a compact set, there are positive constants C_1 and C_2 such that

$$|Df(\lambda, p_1, \cdots, p_m)| \le C_1(1 + \|p\|^2)^{C_2}$$

uniformly for $\lambda \in K$.

For any integer ν, define $\mathcal{M}_\nu: \mathcal{B}(L, m) \to \mathcal{B}(L, m)$ by

$$[\mathcal{M}_\nu(\phi)](\lambda)(p_1, \cdots, p_m) = \delta\Big(\sum_1^m p_i\Big) J_L(\lambda) F_\nu(\lambda, p_1, \cdots, p_m),$$

where ϕ is given by (5.1) and F_ν is the Taylor series of f in p about the origin up to order ν ($\mathcal{M}_\nu = 0$ if $\nu < 0$).

Lemma 3: Let $\mathcal{F}$ be a generalized evaluator. Then $\mathcal{F}: \mathcal{B}(L, m) \to \mathcal{B}(L, m)$, and $\mathcal{F}$ commutes with $\mathcal{M}_\nu$ on $\mathcal{B}(L, m)$.

Proof: $\mathcal{F}$ is defined on an element $\phi \in \mathcal{B}(L, m)$ by $(\mathcal{F}\phi)(\psi) = \mathcal{F}[\phi(\psi)]$, for any $\psi \in \mathcal{S}(R^{4m})$. We claim that, if ϕ has the form (5.1),

$$\mathcal{F}\phi(p) = \delta\Big(\sum_1^m p_i\Big)\mathcal{F}[J(\lambda)f(\lambda, p)]. \tag{5.2}$$

Note first that the difference quotient defining a p derivative of f converges uniformly in λ (on compact sets), so that property (3) of $\mathcal{F}$ implies that

$$\mathcal{F}[J(\lambda)f(\lambda, p)] \in C^\infty(R^{4m}).$$

Moreover, for $\lambda \in K$, $f(\lambda, p) \times (1 + \|p\|^2)^{-(C_2+1)} \to 0$ as $\|p\| \to \infty$, so that (3) implies $\mathcal{F}[J(\lambda)f(\lambda, p)] \in \mathcal{O}_M(R^{4m})$, that is, (5.2) is indeed in $\mathcal{B}(L, m)$ (as a constant function of λ). Now

$$\phi(\lambda)(\psi) = \int_{\Sigma p_i = 0} \psi(p) J(\lambda) f(\lambda, p)\, dp,$$

and this integral may be approximated uniformly in compact subsets of $\mathbb{C}^L$ by Riemann sums. The linearity and continuity of $\mathcal{F}$ then imply (5.2). The fact that $\mathcal{M}_\nu$ and $\mathcal{F}$ commute follows again from the uniformity of the limit defining a p derivative.

The results of Sec. 2 imply that

$$\mathcal{C}_{\lambda,\epsilon}(V_1', \cdots, V_m'; \mathfrak{L}) \in \mathcal{B}(L, m)$$

for any $\{V_1', \cdots, V_m'\}$. Thus we may define

$$\hat{\mathfrak{X}}_\epsilon(V_1', \cdots, V_m'; \mathfrak{L}) = \begin{cases} 1, & \text{for } m = 1, \\ 0, & \text{for OPR } G(V_1', \cdots, V_m'; \mathfrak{L}), \\ \mathcal{F}\mathcal{M}\mathcal{C}_{\lambda,\epsilon}(V_1', \cdots, V_m'; \mathfrak{L}), & \text{otherwise.} \end{cases} \tag{5.3}$$

Here $\mathcal{M} = \mathcal{M}_{\nu(V_1', \cdots V_{m'})}$.

Lemma 4: $\hat{\mathfrak{X}}_\epsilon(V_1', \cdots, V_m'; \mathfrak{L})$ as given by (5.3) is a finite renormalization.

Proof: $\mathfrak{X}_\epsilon$ clearly has the correct form (4.2); property (4) guarantees that $\hat{\mathfrak{X}}_\epsilon$ depends only on the structure of the graph $G(V_1', \cdots, V_m'; \mathfrak{L})$. The existence of the $\epsilon \to 0+$ limit follows from the explicit form of $\mathcal{C}_{\lambda,\epsilon}$ given in (2.18).

Now we may define $\mathfrak{X}'_{\lambda,\epsilon,r}(V_1', \cdots, V_m'; \mathfrak{L})$,

$$\overline{\mathcal{R}}'_{\lambda,\epsilon,r}(V_1'; \cdots, V_m'; \mathfrak{L}),$$

and $\mathcal{R}'_{\lambda,\epsilon,r}(V_1', \cdots, V_m'; \mathfrak{L})$ by formulas (4.3′)–(4.5′), using (5.3) as finite renormalization. We have already discussed the behavior of $\lim_{r\to 0+} \mathcal{R}'_{\lambda,\epsilon,r}$.

Lemma 5: Let Ω be as in Theorem 1. Then

$$\mathfrak{X}'_{\lambda,\epsilon}(V'_1, \cdots, V'_m; \mathcal{L}) = \lim_{r\to 0+} \mathfrak{X}'_{\lambda,\epsilon,r}(V'_1, \cdots, V'_m; \mathcal{L}),$$
$$\overline{\mathcal{R}}'_{\lambda,\epsilon}(V'_1, \cdots, V'_m; \mathcal{L}) = \lim_{r\to 0+} \overline{\mathcal{R}}'_{\lambda,\epsilon,r}(V'_1, \cdots, V'_m; \mathcal{L})$$

exist for $\lambda \in \Omega$ and may be analytically continued to $\mathbb{C}^L$; they are in $\mathcal{B}(L, m)$.

Proof: Similar to Theorem 1. We note in particular that $\mathfrak{X}'_{\lambda,\epsilon}(V'_1, \cdots, V'_m; \mathcal{L})$ has the form

$$\delta\left(\sum_1^m p'_j\right) \sum_{|i| < v(V'_1 \cdots V'_{m'})} f_{(i)}(\lambda, \epsilon) p'^{(i)}, \tag{5.4}$$

where (i) is a multi-index,

$$p'^{(i)} = \prod_{j=1}^{m} \prod_{\mu=0}^{4} p'^{i_j\mu}_{j\mu},$$

and $f_{(i)}(\lambda, \epsilon) \in \mathcal{A}_L$.

Theorem 3: Let $\mathcal{R}'_{\lambda,\epsilon,r}(V_1, \cdots, V_n; \mathcal{L})$ be defined using (5.3) as finite renormalization. Then

$$\mathcal{F}\mathcal{C}_{\lambda,\epsilon}(V_1, \cdots, V_n; \mathcal{L}) = \lim_{r\to 0+} \mathcal{R}'_{1,\cdots,1,\epsilon,r}(V_1, \cdots, V_n; \mathcal{L}). \tag{5.5}$$

We remark that Hepp has shown that the $\epsilon \to 0$ limit of the right-hand side of (5.5) exists. This justifies our definition (3.2) of $\mathcal{C}(V_1, \cdots, V_n; \mathcal{L})$, and the $\epsilon \to 0$ limit of (5.5) is just the equality of the two definitions of the renormalized amplitudes.

Proof: We first show that, for $m' > 1$,

$$\mathcal{F}\mathfrak{X}'_{\lambda,\epsilon}(V'_1, \cdots, V'_{m'}; \mathcal{L}) = 0. \tag{5.6}$$

The statement is, of course, true (vacuously) for $m' = 1$; we assume it for all $1 \le m' < m$, and consider an OPI graph $G(V'_1, \cdots, V'_m; \mathcal{L})$.

From (4.3C'),

$$\mathfrak{X}'_{\lambda,\epsilon,r}(V'_1, \cdots, V'_m; \mathcal{L}) = -\left\{\sum_P{}' \prod_{j=1}^{k(P)} \mathfrak{X}'_{\lambda,\epsilon,r}(V^P_{j1}, \cdots, V^P_{jr(j)}; \mathcal{L}) \prod_{\text{conn}} \Delta^l\right\} + \hat{\mathfrak{X}}_\epsilon(V'_1, \cdots, V'_m; \mathcal{L}). \tag{5.7}$$

Consider a term from $\sum'_P$ in (5.7) in which $r(j) > 1$ for some j, say $j = 1$ [note $k(P) \ge 2$, so we must have $r(j) < m$]. From (5.4) this has the form in p space

$$W_P(\lambda, \epsilon, r) = \sum_{(i)} f_{(i)}(\lambda, \epsilon, r)\{(\delta(\textstyle\sum p) p^{(i)}) * V\}, \tag{5.8}$$

where V is the Fourier transform of

$$\prod_{j=2}^{k(P)} \mathfrak{X}'_{\lambda,\epsilon,r}(V^P_{j1}, \cdots) \prod_{\text{conn}} \Delta^l.$$

For $\lambda \in \Omega$, we can let $r \to 0+$ in (5.8). The bracketed term converges to an element in $\mathcal{B}(L, m)$, and $f_{(i)}(\lambda, \epsilon, r)$ converges to $f_{(i)}(\lambda, \epsilon) \in \mathcal{A}_L$. Actually, however, $f_{(i)}(\lambda, \epsilon)$ depends only on those λ_l such that lth line joins two vertices of $\{V^P_{11}, \cdots, V^P_{1r(1)}\}$, while the bracket in (5.8) depends on those λ_l such that the lth line has at least one end point outside this set. Thus property (6) of $\mathcal{F}$ implies

$$\mathcal{F}\left[\lim_{r\to 0+} W_P\right] = \sum_{(i)} [\mathcal{F} f_{(i)}(\lambda, \epsilon)]\left[\mathcal{F} \lim_{r\to 0} \{\quad\}\right].$$

But by the induction assumption

$$\mathcal{F}\mathfrak{X}'_{\lambda,\epsilon}(V^P_{11}, \cdots, V^P_{1r(1)}; \mathcal{L}) = 0,$$

so that $\mathcal{F} f_{(i)}(\lambda, \epsilon) = 0$ and hence

$$\mathcal{F}\left[\lim_{r\to 0+} W_P(\lambda, \epsilon, r)\right] = 0. \tag{5.9}$$

Now, using Lemma 3,

$$\mathcal{F}\mathfrak{X}'_{\lambda,\epsilon}(V_1, \cdots, V_m; \mathcal{L}) = -\mathcal{M}\mathcal{F}\left(\sum_P{}'\left[\lim_{r\to 0+} W_P\right]\right) + \hat{\mathfrak{X}}_\epsilon(V_1, \cdots, V_m; \mathcal{L}), \tag{5.10}$$

since property (2) of $\mathcal{F}$ implies $\mathcal{F}^2 = \mathcal{F}$. But by (5.9), all terms of $\sum'_P$ in (5.10) vanish except for that partition in which $r(j) = 1$ for all j. However, this term is exactly cancelled by $\hat{\mathfrak{X}}_\epsilon(V_1, \cdots, V_m; \mathcal{L})$; this proves (5.6).

Equation (4.5'), defining $\mathcal{R}'$, may be written

$$\mathcal{R}'_{\lambda,\epsilon,r}(V_1, \cdots, V_n; \mathcal{L}) = \prod_{\mathcal{L}} \Delta^l_{\lambda_l,\epsilon,r} + \sum_P{}'' \prod_{j=1}^{k(P)} \mathfrak{X}'_{\lambda,\epsilon,r}(V^P_{j1}, \cdots, V^P_{jr(j)}; \mathcal{L}) \prod_{\text{conn}} \Delta^l, \tag{5.11}$$

where $\sum''_P$ is over all partitions of $\{V_1, \cdots, V_n\}$ with $1 \le k(P) < n$. For $\lambda \in \Omega$, we let $r \to 0+$ in (5.11) and then apply $\mathcal{F}$ to both sides. Equation (5.6) and another use of property (6) show that $\mathcal{F}$ annihilates the second term on the right-hand side. But the first term on this side is just $\mathcal{C}_{\lambda,\epsilon}(V_1, \cdots, V_n; \mathcal{L})$, so that (5.11) becomes

$$\mathcal{F}\mathcal{R}'_{\lambda,\epsilon}(V_1, \cdots, V_n; \mathcal{L}) = \mathcal{F}\mathcal{C}_{\lambda,\epsilon}(V_1, \cdots, V_n; \mathcal{L}).$$

Theorem 2 and property (2) of $\mathcal{F}$ show that

$$\mathcal{F}\mathcal{R}'_{\lambda,\epsilon}(V_1, \cdots, V_n; \mathcal{L}) = \mathcal{R}'_{1,\cdots,1,\epsilon}(V_1, \cdots, V_n; \mathcal{L});$$

this completes the proof of the theorem.

ACKNOWLEDGMENT

I would like to thank Professor Arthur S. Wightman for suggesting this problem and for encouraging and guiding its development.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Violation of the Quantum Ordering of Propositions in Hidden-Variable Theories

JAMES E. TURNER
Department of Mathematical Physics, University of Birmingham, Birmingham, England

(Received 20 January 1968)

A general definition of hidden-variable theories in terms of the dual structure of states and propositions is proposed. As a consequence of a theorem due to Zierler and Schlessinger, this definition implies a violation of the quantum ordering of propositions in the corresponding hidden-variable theory. This violation is shown explicitly for the theory of measurement due to Bohm and Bub.

1. INTRODUCTION

What is a hidden-variable theory of quantum mechanics? In this paper a simple answer is proposed in the form of a physically motivated definition. Some consequences of this definition are examined in relation to the investigations of Zierler and Schlessinger[1] and in relation to a recent hidden-variable theory of measurement due to Bohm and Bub.[2]

The early proof of von Neumann[3] of the impossibility of hidden-variable theories is now generally recognized not to be relevant to the discussion except as a starting point. The principal justification for the assumptions that von Neumann makes is that they reproduce the usual Hilbert-space structure of quantum mechanics. His principal assumption, the linear additivity of eigenvalues, turns out not to be experimentally verifiable in the framework of quantum mechanics. There is, therefore, the possibility of violating this postulate but still producing a theory with the same experimental predictions as quantum theory. One interest of this approach is to produce a physical theory of the measurement process.[2] There is an excellent short discussion of von Neumann's theorem in the review article by Bell.[4]

The approach adopted here is based on the lattice structure of the set of propositions (yes–no experiments) in quantum theory. This method was introduced by von Neumann and Birkhoff[5] and has been developed by Jauch and Piron[6–8] and applied to the problem of hidden variables. In terms of this assumed lattice structure, a violation of the quantum ordering of propositions is proved to be a general feature of hidden-variable theories. This is shown explicitly for the theory of measurement proposed by Bohm and Bub.

2. BOOLEAN EMBEDDINGS OF NONDISTRIBUTIVE LATTICES

A few definitions must be introduced. *A partially ordered set* is a system X in which a binary relation $x > y$ is defined which satisfies three postulates:

(Reflexive) $x > x$,

(Antisymmetric) $x > y, y > x \Rightarrow x = y$,

(Transitive) $x > y, y > z \Rightarrow x > z$.

A *lattice* is a partially ordered set P such that any two elements have an "intersection" $x \cap y$ and a "union" $x \cup y$, with the usual properties relative to the partial order.

It is useful to define special lattices satisfying additional assumptions:

(i) *Distributive lattice:*

$$a \cup (b \cap c) = (a \cup b) \cap (a \cup c),$$
$$a \cap (b \cup c) = (a \cap b) \cup (a \cap c);$$

(ii) *Modular lattice:*

$$a < c \Rightarrow a \cup (b \cap c) = (a \cup b) \cap c;$$

(iii) *Weakly modular lattice:*

$$a < b \Rightarrow (a \cup b') \cap b = a.$$

These three lattice types are those which arise in relation to usual physical theories. They can be used to give a convenient characterization of classical or quantum theories. Note that (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii).

There is a simple diagrammatic representation of the lattice structure. For example, the lattice of subsets of a set of three elements has the following representation (Fig. 1). The vertices are interpreted as the elements of the lattice and the joining lines give the partial order. This is a distributive lattice. Figure 2 gives an example of a nonmodular lattice. This lattice

[1] N. Zierler and M. Schlessinger, Duke Math. J. **32**, 251 (1965).
[2] D. Bohm and J. Bub, Rev. Mod. Phys. **38**, 453, 470 (1966).
[3] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, N.J., 1955).
[4] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[5] J. von Neumann and G. Birkhoff, Ann. Math. **37**, 823 (1936).
[6] C. Piron, Helv. Phys. Acta **37**, 439 (1964).
[7] J. M. Jauch and C. Piron, Helv. Phys. Acta **36**, 827 (1963).
[8] J. M. Jauch, Helv. Phys. Acta **37**, 293 (1964).

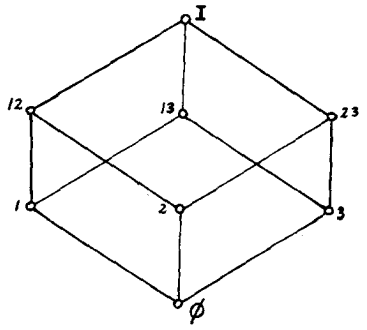


FIG. 1. Distributive lattice.

is used as an example in the fundamental paper of Birkhoff and von Neumann.[5]

Distributive lattices are the most familiar. They arise as lattices of subsets as in the first example. Modular lattices can be considered as lattices of closed linear subspaces of finite-dimensional Hilbert spaces. The corresponding lattices for infinite-dimensional Hilbert spaces are weakly modular. These are the structures that arise in connection with the physical theories considered in this paper. The basis of this investigation is that every physical theory has a corresponding calculus or lattice of propositions. The propositions correspond to those observables of the physical system which are associated with two possible values: yes or no, true or false, 1 or 0. In quantum mechanics, these propositions correspond to projection operators, Hermitian operators with eigenvalues 1 and 0. This makes obvious the identification of propositions with the closed linear subspaces of Hilbert space. The various assumptions of distributivity, modularity, and weak modularity characterize not only the lattice of propositions of the physical theory but also the theory itself.

The underlying problem in hidden-variable theories is to embed the usual quantum theory in a larger framework which has the characteristics of a classical theory. Lattices provide a natural mathematical language for this problem.

Physical theories are characterized by the dual structure of states and propositions, as well as by the lattice structure. Quantum-mechanical states are defined as measures on the closed linear subspaces of Hilbert space and classical states are measures on the Borel subsets of phase space.[9] The dualities are defined by the action of the states on the propositions mapping them to the unit interval of the real line. These are separating dualities in the sense that two propositions are identical if no state can be found to distinguish between them. Similarly, the states are separated by the duality. This dual structure of states and propositions gives an effective framework in which to discuss hidden-variable theories. The dual structure is called the logic of the physical theory.

In a general notation, the problem can be expressed by the following diagram:

$$S \xrightarrow{\tau} S'$$
$$P \xrightarrow{\sigma} P',$$

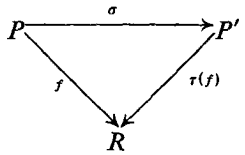
where $S(S')$ is the space of quantum (classical) states and $P(P')$ is the lattice of quantum (classical) propositions. S and P are dual spaces as are S' and P' with the structures defined above. The embedding is achieved by the two maps $\tau: S \to S'$ and $\sigma: P \to P'$. There is a relation between τ and σ given by the dualities.

There is only one condition on the embedding that seems necessary from physical considerations. The classical theory which proposes to replace quantum theory must predict the same expectation values as the quantum theory at least in those cases which are accessible to experiment. Embeddings which do not have this property have no relevance to the present discussion.

Consider

$$f \in S, \quad a \in P \to f_\tau = \tau(f) \in S', \quad \sigma(a) \in P'.$$

The condition is that $f(a) = f_\tau(\sigma(a))$ or $f = (\tau(f)) \circ \sigma$. The following diagram is commutative:



A typical question arises immediately. If $\sigma(P)$ does not span P', an extension of f_τ to all of P' must be defined. The simplest assumptions will be made on τ and σ, namely, the 1–1 property and $f = (\tau(f)) \circ \sigma$. for $f \in S$. The first property is usually understood in the definition of an embedding. It is a requirement of nontriviality—different quantum states are mapped to different classical states. The definition of a hidden-variable theory can now be given.

An embedding of the quantum logic (S, P) into the

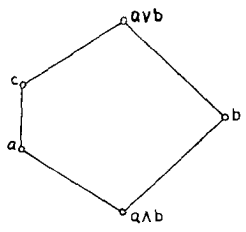


FIG. 2. Nonmodular lattice.

[9] G. W. Mackey, *The Mathematical Foundations of Quantum Mechanics* (W. A. Benjamin Co., Inc., New York, 1963).

distributive logic (S', P') is a pair (τ, σ) of 1–1 maps

$$\tau: S \to S' \quad \text{and} \quad \sigma: P \to P',$$

such that expectation values are preserved, i.e., $f \in S \Rightarrow f = (\tau(f)) \circ \sigma$. This will be called a *Boolean embedding* or a *hidden-variable* theory. The problem now is to establish the properties of these embeddings and to relate these properties to specific hidden-variable systems.

The rest of this section is devoted to a discussion of the work of Zierler and Schlessinger.[1] They consider the problem of the construction of hidden-variable theories under an additional assumption. They assume that the quantum ordering of propositions is preserved in the embedding $\sigma: P \to P'$. Now, the quantum ordering of propositions is defined in terms of the duality

$$a < b \Leftrightarrow f(a) < f(b) \quad (\forall f \in S).$$

Therefore the assumption that Zierler and Schlessinger adopt is a very strong condition on τ and σ. They are able to prove that the embedding must be a trivial one.

This result is a new proof of von Neumann's theorem. However, it is subject to the same objections. Their assumption (which is called the isotone property) is the following:

$$f(a) < f(b)\ (\forall f \in S) \Rightarrow \tau(f)(\sigma(a)) < \tau(f)(\sigma(b))\ [\forall \tau(f) \in S'].$$

This need not be true in general. There may be states in S' which violate the second ordering. The isotone property is an assumption which severely limits the structure of S'. In fact, the assumption limits the embedding to the trivial case.

It seems more profitable to consider the alternative definition of embedding and to allow the possibility of violations of the quantum ordering of propositions. From the above argument, this is seen to be a general feature of hidden-variable theories. In Sec. 3 this is shown explicitly for the theory due to Bohm and Bub.[2]

3. A HIDDEN-VARIABLE THEORY OF MEASUREMENT

There have been numerous hidden-variable theories proposed, notably by Bohm and co-workers, to illustrate shortcomings of quantum mechanics, to produce counter examples to various versions of von Neumann's theorem, and to suggest alternative developments. From the general considerations of Sec. 2, every such theory provides an example of a Boolean embedding. The recent theory of Bohm and Bub[2] gives a hidden-variable theory of measurement for a particle with spin without translational motion.

For the spin-$\frac{1}{2}$ particle, the states are given by a wavefunction represented as a vector in a two-dimensional Hilbert space:

$$|\psi\rangle = \psi_1 |S_1\rangle + \psi_2 |S_2\rangle, \quad |\psi_1|^2 + |\psi_2|^2 = 1.$$

To complete the description of the state, a dual Hilbert space is postulated:

$$\langle\xi| = \xi_1 \langle S_1| + \xi_2 \langle S_2|, \quad |\xi_1|^2 + |\xi_2|^2 = 1.$$

These are the hidden variables assumed to be randomly distributed on the hypersphere of unit radius.

The measurement process is described by nonlinear equations of motion relating ψ and ξ. These equations are such that the initial values of the parameters determine the result of a measurement:

$$\frac{d\psi_1}{dt} = \gamma\psi_1\left\{\frac{|\psi_1|^2}{|\xi_1|^2} - \frac{|\psi_2|^2}{|\xi_2|^2}\right\}|\psi_2|^2,$$

$$\frac{d\psi_2}{dt} = \gamma\psi_2\left\{\frac{|\psi_2|^2}{|\xi_2|^2} - \frac{|\psi_1|^2}{|\xi_1|^2}\right\}|\psi_1|^2.$$

It will be noticed from the equations of measurement that the phases of the wavefunctions do not enter the theory. Furthermore, the hidden variables have no clear physical interpretation or direct correspondence with well-known classical observables.

The phase space of the classical theory is six-dimensional, the Cartesian product of two three-dimensional hyperspheres. The Hilbert space of the corresponding quantum theory is two-dimensional. The lattice of subspaces of a two-dimensional Hilbert space has the representation shown in Fig. 3. Here α is an angle variable that parametrizes the one-dimensional subspaces. This lattice is modular. The distributive lattice of the hidden-variable theory is the lattice of Borel subsets of the six-dimensional phase space. This has many more degrees of freedom than the corresponding lattice parametrized by one angle variable.

It is completely trivial to study the action of the embedding on the ordering of propositions in the two-dimensional case. This follows from the observation that there is essentially only one order relation

$$\varnothing < a_\alpha < I \quad (0 \leq \alpha < 2\pi)$$

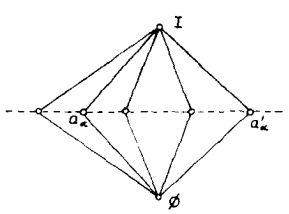


FIG. 3. Modular lattice.

and this is preserved under the embedding

$$\sigma(\varnothing) = \varnothing < \sigma(a_\alpha) < I = \sigma(I) \quad (0 \leq \alpha < 2\pi).$$

The theory for a particle of spin 1 (corresponding to a three-dimensional Hilbert space) provides the first example of the violation of the quantum ordering. This is shown very simply by the following argument. In the theory of Bohm and Bub, if the initial values of the variables are such that

$$\frac{|\psi_3|^2}{|\xi_3|^2} > \frac{|\psi_1|^2}{|\xi_1|^2}, \quad \frac{|\psi_2|^2}{|\xi_2|^2},$$

the result of the measurement is spin-polarized in the 3 direction in the Hilbert space. In the quantum theory, this measurement process corresponds to the projection operator P_{12} onto the 3 direction. If the two-dimensional subspace is rotated through an angle θ,

$$\begin{aligned}
\psi_1' &= \psi_1 \cos\theta + \psi_2 \sin\theta, \\
\psi_2' &= -\psi_1 \sin\theta + \psi_2 \cos\theta, \\
\psi_3' &= \psi_3, \\
\xi_1' &= \xi_1 \cos\theta - \xi_2 \sin\theta, \\
\xi_2' &= \xi_1 \sin\theta + \xi_2 \cos\theta, \\
\xi_3' &= \xi_3,
\end{aligned}$$

this gives the same projection operator in quantum mechanics, i.e., $P_{1'2'} = P_{12}$. However, in the hidden-variable theory, the two corresponding propositions are distinct. After embedding, the projection operator $P_{1'2'}$ corresponds to the subset of phase space given by

$$\frac{|\psi_3|^2}{|\xi_3|^2} > \frac{|\psi_1 + \psi_2 \tan\theta|^2}{|\xi_1 - \xi_2 \tan\theta|^2}, \quad \frac{|\psi_1 \tan\theta - \psi_2|^2}{|\xi_1 \tan\theta + \xi_2|^2}.$$

This is not, in general, the same subset of phase space. There is a dependence on θ. For example, if $\tan\theta = 1$, there is a neighborhood (of positive measure) of the point $|\psi_3|^2 = \frac{2}{3}$, $|\psi_1|^2 = |\psi_2|^2 = \frac{1}{3}$, $|\xi_1|^2 = |\xi_2|^2 = \frac{1}{4}$, $|\xi_3|^2 = \frac{1}{2}$, which satisfies the first inequality but not the second. Therefore, the propositions P_{12} and $P_{1'2'}$ are distinct in the hidden-variable theory.

The violation of the quantum ordering of propositions is illustrated explicitly by the following example with $\tan\theta = 1$. Four quantum propositions are introduced with the corresponding subsets of phase space after embedding:

P_{23}, spin in 1 direction,

$$\left\{\frac{|\psi_1|^2}{|\xi_1|^2} > \frac{|\psi_2|^2}{|\xi_2|^2}\right\} \wedge \left\{\frac{|\psi_1|^2}{|\xi_1|^2} > \frac{|\psi_3|^2}{|\xi_3|^2}\right\};$$

$P_{2'3}$, spin in 1′ direction,

$$\left\{\frac{|\psi_1 + \psi_2|^2}{|\xi_1 - \xi_2|^2} > \frac{|\psi_1 - \psi_2|^2}{|\xi_1 + \xi_2|^2}\right\} \wedge \left\{\frac{|\psi_1 + \psi_2|^2}{|\xi_1 - \xi_2|^2} > \frac{|\psi_3|^2}{|\xi_3|^2}\right\};$$

$\bar{P}_{12}$, spin in 12 plane,

$$\left\{\frac{|\psi_1|^2}{|\xi_1|^2} > \frac{|\psi_3|^2}{|\xi_3|^2}\right\} \vee \left\{\frac{|\psi_2|^2}{|\xi_2|^2} > \frac{|\psi_3|^2}{|\xi_3|^2}\right\};$$

$\bar{P}_{1'2'}$, spin in 1′2′ plane,

$$\left\{\frac{|\psi_1 + \psi_2|^2}{|\xi_1 - \xi_2|^2} > \frac{|\psi_3|^2}{|\xi_3|^2}\right\} \vee \left\{\frac{|\psi_1 - \psi_2|^2}{|\xi_1 + \xi_2|^2} > \frac{|\psi_3|^2}{|\xi_3|^2}\right\}.$$

In the quantum theory:

$$\begin{aligned}
P_{23} &< \bar{P}_{12} = \bar{P}_{1'2'}, \\
P_{2'3} &< \bar{P}_{12} = \bar{P}_{1'2'}.
\end{aligned}$$

In the hidden-variable theory:

$$\begin{aligned}
P_{23} &< \bar{P}_{12} \quad \text{but} \quad P_{23} \not< \bar{P}_{1'2'}, \\
P_{2'3} &< \bar{P}_{1'2'} \quad \text{but} \quad P_{2'3} \not< \bar{P}_{12}.
\end{aligned}$$

For example, if $|\xi_1|^2$ and $|\psi_1|^2$ are both nearly zero but with

$$\frac{|\psi_1|^2}{|\xi_1|^2} > \frac{|\psi_2|^2}{|\xi_2|^2} \quad \text{and} \quad \frac{|\psi_1|^2}{|\xi_1|^2} > \frac{|\psi_3|^2}{|\xi_3|^2},$$

then verifying $\bar{P}_{1'2'}$ reduces to verifying

$$\frac{|\psi_2|^2}{|\xi_2|^2} > \frac{|\psi_3|^2}{|\xi_3|^2},$$

which need not be true. Therefore subsets of phase space of positive measure can be found satisfying P_{23} but not $\bar{P}_{1'2'}$,

There are other embeddings which could have been chosen for this theory. As an extreme case, consider embeddings which map the quantum-pure states to points of phase space. This choice clearly violates the quantum ordering of propositions. The violation occurs for every such choice of embedding.

This is a general feature of hidden-variable theories (if the Hilbert space of the quantum theory has dimension greater than two). It arises as a result of the interaction between the measurement apparatus and the system under observation. For the theory of Bohm and Bub, this interaction has been shown explicitly by the noninvariance of the measurement process under change of coordinates. The violation of quantum ordering can also be considered as a result of the nonlinearity of the equations of measurement. This is illustrated by the following simple argument.

Suppose that the equation of measurement, written $d\psi/dt = L(\psi, \xi)$, is linear-invariant in the sense that it is invariant under linear coordinate transformations, i.e., $d\psi'/dt = L(\psi', \xi')$. The pure states of the quantum theory when embedded in the hidden-variable theory correspond to subsets of phase space with boundaries

$L(\psi, \xi) = 0$. Consider again the three-dimensional case (although the result will be quite general). Suppose the three regions of phase space corresponding to quantum-pure states are given by the labels A_1, A_2, and A_3. Now, the assumption is that the boundaries of these regions given by $L(\psi, \xi) = 0$ remain the same under coordinate transformations. This implies that the above regions are the same after coordinate transformations

$$A_1' = A_1, \quad A_2' = A_2, \quad A_3' = A_3.$$

Therefore there is no violation of the quantum ordering of propositions and the hidden-variable theory is equivalent to the quantum theory. This result shows the relation of nonlinearity to the violation of quantum ordering.

ACKNOWLEDGMENTS

The suggestions and encouragement of Dr. J. Gunson are gratefully acknowledged.

The work was supported by a Commonwealth Scholarship.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Lorentz Covariant Distributions*

F. J. YNDURÁIN†
Department of Physics, New York University

(Received 21 April 1967)

Tensor distributions of several four-vector variables which transform according to a finite representation of the Lorentz group are considered. We give a canonical classification of all possible forms of such objects; this is used to show that, in the relevant cases, it is sufficient to regularize them with respect to the invariants that may be formed out of the variables to obtain analytic functions. We apply this result to Wightman functions, showing a result similar to a theorem proved in position space by Borchers under different assumptions.

1. INTRODUCTION

The fundamental role played in field theory by tensor distributions of several four-vector variables makes it very interesting to consider the general features of these objects. The fact that such distributions possess definite transformation properties under the Lorentz group puts severe restrictions on their possible form; thus, for instance, it seems natural (and this has been widely used) that one may project them onto invariant ones. The last are a special case of the former, and it is also interesting to inquire whether the statement that they really only depend on the invariants formed with their arguments is true. To put forth an example, given $T_\mu(x, y)$, the question is posed whether one can expand it as $x_\mu t_1 + y_\mu t_2$, t_i invariant, and whether it is true that $t_i = t_i(x^2, y^2, x \cdot y)$. That this is in fact the case for analytic functions has been shown by several people[1]; but one may expect some trouble for distributions as the existence of counterexamples such as, e.g., $\partial_\mu \delta(x + \alpha y)$, any α, indicates. One is thus led to the problem of giving a characterization of the "well behaved" as well as a description of the "pathological" ones, i.e., to a classification of tensor distributions. For invariant distributions of one variable this has been done by Methée[2] who proved that if $T(x)$ is invariant, then

$$T(x) = t(x^2) + [\text{polynomial } (\partial^\mu \partial_\mu)]\delta(x),$$

and his analysis has been partially extended to more variables by several people[1–3] (in special cases). In the present article we give a generalization of these results, presenting a complete classification of all tensor (not necessarily invariant) distributions in several variables.

Related problems are relevant in different contexts. Thus, Borchers[3] has shown (using support properties and translational invariance) that the Wightman functions in position space need only be tested in the

* This paper was supported in part by the National Science Foundation.

† Present address: CERN European Organization for Nuclear Research, 1211 Geneva 23 Switzerland.

[1] K. Hepp, Helv. Phys. Acta **37**, 55 (1964); **36**, 355 (1963); D. N. Williams, UCRL-1113 (1963) (unpublished); G. C. Fox, Cavendish Laboratory Preprint (unpublished) (1966); D. Hall and A. S. Wightman, Kgl. Dansk. Vid. Selsk. Matt.-Fys. Medd., **31**, No. 5. (1957).

[2] P. Methée, Comm. Math. Helv. **28**, 225 (1954); **32**, 153 (1957); C. R. Paris, **240**, 1179 (1955); L. Gårding and J. L. Lions, Nuovo Cimento Suppl. **14**, 9, (1959).

[3] H. J. Borchers, Nuovo Cimento **33**, 1600 (1964).

$L(\psi, \xi) = 0$. Consider again the three-dimensional case (although the result will be quite general). Suppose the three regions of phase space corresponding to quantum-pure states are given by the labels A_1, A_2, and A_3. Now, the assumption is that the boundaries of these regions given by $L(\psi, \xi) = 0$ remain the same under coordinate transformations. This implies that the above regions are the same after coordinate transformations

$$A_1' = A_1, \quad A_2' = A_2, \quad A_3' = A_3.$$

Therefore there is no violation of the quantum ordering of propositions and the hidden-variable theory is equivalent to the quantum theory. This result shows the relation of nonlinearity to the violation of quantum ordering.

ACKNOWLEDGMENTS

The suggestions and encouragement of Dr. J. Gunson are gratefully acknowledged.

The work was supported by a Commonwealth Scholarship.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Lorentz Covariant Distributions*

F. J. YNDURÁIN†
Department of Physics, New York University

(Received 21 April 1967)

Tensor distributions of several four-vector variables which transform according to a finite representation of the Lorentz group are considered. We give a canonical classification of all possible forms of such objects; this is used to show that, in the relevant cases, it is sufficient to regularize them with respect to the invariants that may be formed out of the variables to obtain analytic functions. We apply this result to Wightman functions, showing a result similar to a theorem proved in position space by Borchers under different assumptions.

1. INTRODUCTION

The fundamental role played in field theory by tensor distributions of several four-vector variables makes it very interesting to consider the general features of these objects. The fact that such distributions possess definite transformation properties under the Lorentz group puts severe restrictions on their possible form; thus, for instance, it seems natural (and this has been widely used) that one may project them onto invariant ones. The last are a special case of the former, and it is also interesting to inquire whether the statement that they really only depend on the invariants formed with their arguments is true. To put forth an example, given $T_\mu(x, y)$, the question is posed whether one can expand it as $x_\mu t_1 + y_\mu t_2$, t_i invariant, and whether it is true that $t_i = t_i(x^2, y^2, x \cdot y)$. That this is in fact the case for analytic functions has been shown by several people[1]; but one may expect some trouble for distributions as the existence of counterexamples such as, e.g., $\partial_\mu \delta(x + \alpha y)$, any α, indicates. One is thus led to the problem of giving a characterization of the "well behaved" as well as a description of the "pathological" ones, i.e., to a classification of tensor distributions. For invariant distributions of one variable this has been done by Methée[2] who proved that if $T(x)$ is invariant, then

$$T(x) = t(x^2) + [\text{polynomial } (\partial^\mu \partial_\mu)]\delta(x),$$

and his analysis has been partially extended to more variables by several people[1–3] (in special cases). In the present article we give a generalization of these results, presenting a complete classification of all tensor (not necessarily invariant) distributions in several variables.

Related problems are relevant in different contexts. Thus, Borchers[3] has shown (using support properties and translational invariance) that the Wightman functions in position space need only be tested in the

* This paper was supported in part by the National Science Foundation.

† Present address: CERN European Organization for Nuclear Research, 1211 Geneva 23 Switzerland.

[1] K. Hepp, Helv. Phys. Acta **37**, 55 (1964); **36**, 355 (1963); D. N. Williams, UCRL-1113 (1963) (unpublished); G. C. Fox, Cavendish Laboratory Preprint (unpublished) (1966); D. Hall and A. S. Wightman, Kgl. Dansk. Vid. Selsk. Matt.-Fys. Medd., **31**, No. 5. (1957).

[2] P. Methée, Comm. Math. Helv. **28**, 225 (1954); **32**, 153 (1957); C. R. Paris, **240**, 1179 (1955); L. Gårding and J. L. Lions, Nuovo Cimento Suppl. **14**, 9, (1959).

[3] H. J. Borchers, Nuovo Cimento **33**, 1600 (1964).

timelike directions, proving that

$$\int dx_0 \phi(x_0) A(x) = \langle \phi(x_0), A(x) \rangle$$

is differentiable infinitely many times in spacelike directions. As a result of our analysis it follows that Lorentz covariance alone is sufficient to ensure a similar result for Wightman functions in momentum space.

In Secs. 2–4 we describe the mathematical tools that are needed later; some results, previous to the analysis itself and concerning chiefly the description of invariant variables, are presented in Secs. 2 and 3, while in Sec. 5 we deal with tensor distributions. In this context, we show that (apart from certain pathological cases) tensor distributions need only be tested in the invariants to get infinitely differentiable (and even analytic) functions. A canonical representation of the pathological cases is described in Sec. 6, thus obtaining a general classification of all tensor distributions, which is used to give (Sec. 7) some applications; we conclude with some remarks and comments, in Sec. 7, as well as in the appendices where a few auxiliary questions are discussed.

Finally, we remark that, although we use loose language, our results are mathematically rigorous; they also possess the advantage of using intrinsic methods, so that they are straightforwardly extendable to more general situations.

2. LITTLE GROUPS, ORBITS, AND CURVILINEAR COORDINATES

Let $\mathfrak{L}$ be a connected Lie group with parameters $\xi_1, \cdots, \xi_l$, and let its continuous representation $v^{(I)} \to \Lambda v^{(I)}$ in the linear space $M^{(I)}$ of dimension $n+1$ be given. We denote with the same letter the element Λ in $\mathfrak{L}$ and its representer, since no confusion may arise. We assume the representation to be irreducible, form the direct sum of N spaces identical to $M^{(I)}, M^{(1)}, \cdots, M^{(N)}$, and define a new representation acting on $E = M^{(1)} \oplus \cdots \oplus M^{(N)}$ by setting

$$\Lambda : v = v^{(1)} \oplus \cdots \oplus v^{(N)} \to \Lambda v = \Lambda v^{(1)} \oplus \cdots \oplus \Lambda v^{(N)}, \tag{1}$$

which we still denote with the same letter.[4]

If v is a vector in E, we define the *little group*[5] $\mathfrak{W}_v$ as the subgroup of all Γ in $\mathfrak{L}$ such that $\Gamma v = v$; it is clear that, if the Γ_n are in $\mathfrak{W}_v$, and $\Gamma_n \to \Gamma$, also $(\Gamma_n v - v) \to (\Gamma v - v) = 0$, so that $\mathfrak{W}_v$ is closed in $\mathfrak{L}$ and therefore is a Lie subgroup of $\mathfrak{L}$ (Ref. 6, Chap. IV, Secs. IV and V); we take $\xi_{r+1}, \cdots, \xi_l$ to be its coordinates.

Together with the little group, we will consider the *orbit* $R(v)$; it is defined as the set of all elements of the form Λv, where Λ ranges over $\mathfrak{L}$. We now have our first theorem.

Theorem 2.1: The homogeneous space $\mathfrak{L}/\mathfrak{W}_v$ is an analytic manifold, isomorphic to $R(v)$. $R(v)$ may be parametrized with the parameter $\xi_1, \cdots, \xi_r$; it then becomes an analytic manifold, analytically embedded into E. In other words, if v' is in $R(v)$ and its Cartesian coordinates are $v'^{(I)}_\mu$, then these depend analytically and nonsingularly on the curvilinear ones $\xi_1, \cdots, \xi_r$ and vice versa.

Proof: The analyticity and construction of the isomorphism $\mathfrak{L}/\mathfrak{W}_v \leftrightarrow R(v)$ are shown in Ref. 7 (p. 111, Theorem 32). Now, $\mathfrak{W}_v$ being closed, the projection $\mathfrak{L} \to \mathfrak{L}/\mathfrak{W}_v$ is analytic (Ref. 6, pp. 109–111; Ref. 8, p. 43); since any representation of a Lie group is analytic,[6] we have the analytic chain $\mathfrak{L} \to$ representation of $\mathfrak{L}$ in $E \to \mathfrak{L}/\mathfrak{W}_v \leftrightarrow R(v)$. Q.E.D.

As to the introduction of coordinates in $R(v)$, let L (W_v) be the Lie algebra of $\mathfrak{L}$ ($\mathfrak{W}_v$). If $M_1 \cdots M_l$ are the generators of L, $M_{r+1} \cdots M_l$ those of W, then the mapping

$$\xi_1, \cdots, \xi_r \to \exp(\xi_1 M_1 + \cdots + \xi_r M_r) v \tag{2}$$

maps (analytically) a neighborhood of $\mathfrak{L}/\mathfrak{W}_v$ onto a neighborhood of v in $R(v)$ (Ref. 7, p. 113, Lemma 4.1). This is the desired (local) parametrization of $R(v)$. The action of $\mathfrak{L}$ on the coordinates $\xi_1, \cdots, \xi_r$ is then the natural one, viz., if v' has coordinates $\xi'_1, \cdots, \xi'_r$, and v'' has $\xi''_1, \cdots, \xi''_r$, and if $v'' = \Lambda v'$, we define

$$\Lambda(\xi'_1, \cdots, \xi'_r) = (\xi''_1, \cdots, \xi''_r); \tag{3}$$

it may be shown that this action is analytic (Ref. 8, pp. 42ff; see, alternatively, Refs. 6 and 7). These induced nonlinear "representations" have been considered in physics in connection with quite different problems.[9] Although we do not use this, we note that

[4] This may be thought of as a representation reducible into N equivalent representations acting on the $M^{(I)}$. The case of the $M^{(I)}$ being unequivalent could also be considered, but our construction suffices for applications.

[5] E. P. Wigner, Ann. Math. **40**, No. 1 (1939); F. J. Ynduráin, Nuovo Cimento **45**, 239 (1966); E. Salusti and F. J. Ynduráin (unpublished.)

[6] C. Chevalley, *Lie Groups, Vol. I* (Princeton Math. Series, Princeton, 1946).

[7] S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic Press Inc., New York, 1962).

[8] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Interscience Publ. Inc., New York, 1963).

[9] L. Michel, in *Axiomatic Field Theory, Brandeis Lectures 1965* (Gordon and Breach, Science Publ., Inc., New York, 1966); S. Wenberg, Phys. Rev. Letters **18**, 188 (1967).

$R(v)$ is, in some cases, even a Riemann symmetric manifold.[7,8]

Two analytic manifolds are locally isomorphic whenever they have the same dimension; it is then natural to split E as the union of disjoint sets E_r such that if v is in E_r, $\dim R(v) = r$. If $\dim E_r = a_r$, then E_r/R is an analytic manifold of dimension $a_r - r$; locally,

$$E_r = (E_r/R) \times R.$$

Denoting by P_r to E_r/R, this shows that we may introduce analytical coordinates $\rho_1, \cdots, \rho_{a_r - r}$ in P_r, and the ρ are invariant under $\mathfrak{L}$ (see Refs. 5–8; an explicit construction of the ρ is given in Sec. 4 and Appendix A). Collecting the results, we have the following theorem.

Theorem 2.2: For every point $\bar{v}$ in E_r there exists a neighborhood $U(\bar{v})$ of $\bar{v}$ in E_r and a corresponding neighborhood (that we may take to be cylindric) $U_P \times U_R$, in $P_r \times R$ such that if v is in E_r, v has curvilinear coordinates (ρ, ξ) in $P_r \times R$, the relation between the Cartesian and curvilinear coordinates is analytic and the Jacobian $J(v_\mu^{(I)}; \rho, \xi)$ is analytic and nonsingular over $U(\bar{v})$, $U_P \times U_R$. The action of Λ in $\mathfrak{L}$ on the ξ is as in Eq. (3), and the ρ are unchanged.

Explicit examples are found in Sec. 4.

3. CURVILINEAR COVARIANT COORDINATES

Consider the Minkowski space M_{n+m} with n space and m time coordinates. The connected part of the group of linear transformations of M_{n+m} that leave invariant the metric

$$v \cdot w = \sum_{\mu\nu} v_\mu w_\nu g_{\mu\nu},$$

$$g_{\mu\nu} = \operatorname{diag}(+1, \overset{m}{\cdots}, +1; -1, \overset{n}{\cdots}, -1),$$

where v_μ, w_ν are the (Cartesian) coordinates of v, w, is called the $(n + m)$ Lorentz group[5] and denoted by $\mathfrak{L}_n^m$. The space E is defined by taking the direct sum of N Minkowski spaces M_{n+m}:

$$E = M_{n+m} \oplus \overset{N}{\cdots} \oplus M_{n+m}.$$

If we are given N' Minkowski vectors $v^{(I)}$ and we select a set of linearly independent ones $v'^{(I'')}$, $I'' = 1, \cdots, N'' \leq n + m - 1$, we know (Hall and Wightman, Ref. 1) that the invariants that may be formed out of the $v^{(I)}$ are functions of the invariants of the $v'^{(I'')}$. Moreover, these invariants are simply the scalar products $v'^{(I'')} \cdot v'^{(J'')} = \rho_{I''J''}$. In view of that, we always assume $N' < n + m$ and, consequently, $N < n + m$. The general case (that may be treated along similar lines) is left for the moment (in all the following we take the physical case $n = 3$, $m = 1$; the generalization for arbitrary n, m is straightforward).

In virtue of the discussions of Sec. 2, the first step for studying objects $T(v)$ defined on E is to reexpress them in terms of canonical coordinates ρ, ξ. For this we have to classify the little groups

$$\mathfrak{W}_v = \mathfrak{W}_{v^{(1)} \cdots v^{(N)}}.$$

This has already been done (Ref. 5; see also Appendix B) and, in regard to their dimensionality (which is the result that is relevant here in view of the analysis of Sec. 2 and the fact that $\dim R = \dim \mathfrak{L}/\mathfrak{W} = \dim \mathfrak{L} - \dim \mathfrak{W}$) the answer is:

Theorem 3.1.

Case (1), $N = 1$. (i) If $v \neq 0$, $\dim \mathfrak{W}_v = 3$ and $\dim R(v) = 3$. (ii) If $v = 0$, then $\mathfrak{W}_v = \mathfrak{L}_3^1$, $\dim \mathfrak{W}_v = 3$, and $\dim R(v) = 0$.

Case (2), $N = 2$. (i) If $v^{(1)}$ and $v^{(2)}$ are linearly independent, $\dim \mathfrak{W}_v = 1$ and $\dim R(v) = 5$. (ii) If they are parallel, but $v \neq 0$, $\dim \mathfrak{W}_v = 3$ and $\dim R(v) = 3$. (iii) If $v = 0$, then $\mathfrak{W}_v = \mathfrak{L}_3^1$, $\dim \mathfrak{W}_v = 6$, and $\dim R(v) = 0$. Here $R(0) = \{0\}$.

Case (3), $N = 3$. (i) If the three $v^{(1)}$, $v^{(2)}$, and $v^{(2)}$ are linearly independent, $\mathfrak{W}_v = 1$ and $\dim R(v) = \dim \mathfrak{L}_3^1 = 6$. (ii) If two are linearly independent, $\dim \mathfrak{W}_v = 1$ and $\dim R(v) = 5$. (iii) If all are multiples of a nonzero vector, $\dim \mathfrak{W}_v = 3$ and $\dim R(v) = 3$. (iv) If $v = 0$ and $\mathfrak{W}_v = \mathfrak{L}_3^1$, $\dim \mathfrak{W}_v = 6$ and $\dim R(v) = 0$; in fact, $R(0) = \{0\}$.

We may then form the corresponding E_r, $r = \dim R(v)$, as in Sec. 2; applying Theorem 2.2, we get our next theorem.

Theorem 3.2: In situation (3i) of Theorem 3.1 we have the coordinates $\xi_1, \cdots, \xi_6$ and the $\rho_1, \cdots, \rho_6$ that may be taken to be the $\rho_{IJ} = v^{(I)} \cdot v^{(J)}$. In (2i) of Theorem 3.1, the coordinates are $\xi_1, \cdots, \xi_5$ and $\rho_1, \cdots, \rho_3$ that again may be taken to be the ρ_{IJ}. In situation (1i) of Theorem 3.1, the parameters are $\xi_1, \cdots, \xi_3$ and $\rho = v \cdot v$.

The "singular" situations are more difficult to handle. The simplest of the singular cases are the last of each case, i.e., (1ii), (2iii), and (3iv) of Theorem 3.1, since $R(0)$ reduces to a point. The invariant coordinates (the ξ do not exist here!) are simply the coordinates of $v = 0$. In the remaining cases, we relabel the components $v^{(I)}$ of v in such a manner that the last vectors $v^{(A+1)}, \cdots, v^{(N)}$ are expressible as linear

combinations of the first, which are taken to be independent; i.e.,

$$v^{(A+B)} = \sum_{A'=1}^{A} \alpha_{A'}^{(A+B)} v^{(A')}.$$

We may then show, by direct checking, that the following is true.

Theorem 3.3: In situation (3ii) of Theorem 3.1, the coordinates are $\xi_1, \cdots, \xi_5$ and the ρ's are formed by the $\rho_{A'A''} = v^{(A')} \cdot v^{(A'')}$; $A', A'' = 1, 2$, and $\alpha_{A'}^{(3)}$, $A' = 1, 2$. This gives a total of five $\rho_1, \cdots, \rho_5$, and $\dim E_5 = 10$. In (3iii), we have $\xi_1, \cdots, \xi_3$; $\rho_1 = v^{(1)} \cdot v^{(1)}$, $\rho_2 = \alpha_1^{(2)}$, and $\rho_3 = \alpha_1^{(3)}$. Here $\dim E_3 = 6$. Finally, in situation (2ii), the parameters are $\xi_1, \cdots, \xi_3$ and $\rho_1 = v^{(1)} \cdot v^{(1)}$, $\rho_2 = \alpha_1^{(2)}$. Moreover, $\dim E_3 = 5$.

Note that the analysis is not yet complete since we still have to specify E_r in E. For this, let the case be N; form the matrix $\mathcal{M}_N$

$$\mathcal{M}_N = \begin{pmatrix} v_0^{(1)} & \cdots & v_0^{(N)} \\ v_1^{(1)} & \cdots & v_1^{(N)} \\ v_2^{(1)} & \cdots & v_2^{(N)} \\ v_3^{(1)} & \cdots & v_3^{(N)} \end{pmatrix}.$$

We remark that the number of linearly independent vectors is given by the rank of $\mathcal{M}_N$. A straightforward application of matrix calculus then gives the following theorem.

Theorem 3.4: The missing "invariant parameters" that specify E_r in E are: in situation (3ii) of Theorem 3.1, $d_1^{(3)} = 0$, $d_2^{(3)} = 0$, where the $d^{(3)}$ are any two different (3×3) minors of $\mathcal{M}_3$; in (3iii), $d_1^{(2)} = \cdots = d_6^{(2)} = 0$, where the $d^{(2)}$ are any six different (2×2) minors of $\mathcal{M}_3$. In situation (2ii), we have $d_1^{(2)} = d_2^{(2)} = d_3^{(2)} = 0$, and the $d^{(2)}$ are any three different (2×2) minors of $\mathcal{M}_2$.

We remark that the E_r are invariant sets, and if $r' > r''$, then $E_{r''}$ is of null measure with respect to $E_{r'}$ and lies on its boundary.

Definition: We denote by E_R the set of maximal dimension among the E_r. For 3 (respectively, 2, 1) vectors, $R = 6$ (respectively, 5, 3). Note that $\dim E_R = \dim E$. Moreover, E_R is open in E.

4. DISTRIBUTIONS AND FUNCTION SPACES

Let U be an open set in the real finite-dimensional vector space E; let F be a topological linear space, and let F' be its dual.[10] We define the following[10–13]:

$C_n^F(U)$, the space of n-times differentiable functions in U with values in F;

$C_\omega^F(U)$, the space of analytic functions in U with values in F;

$\mathcal{S}^F(U)$, the space of functions in $C_\infty^F(U)$ of fast decrease;

$\mathcal{S}_\omega^F(U)$, the intersection of $C_\omega^F(U)$ and $\mathcal{S}^F(U)$;

$\mathcal{S}'^F(U)$, the space of tempered distributions with values in F.

If we do not write the superscript, it should be understood that F is the field of complex numbers. Let $A^F(U)$ denote any of these objects; then,[13] X belongs to $A^F(U)$ if, for every e in F', $\langle e, X\rangle$ belongs to $A(U)$. Now we have the following lemmas.

Lemma 4.1: If, under the change of variables $v \to \beta$, the image of U_v is U_β, and if the Jacobian $J(v; \beta)$ is analytic, bounded, and nonsingular, $X(v)$ belongs to $A^F(U_v)$ whenever $X(v(\beta))$ belongs to $A^F(U_\beta)$.

Lemma 4.2: If $\dim V < \infty$, if $T(\rho, \xi)$ is in

$$\mathcal{S}'^V(U_\rho \times U_\beta),$$

and if, for every $\psi(\rho)$ in $\mathcal{S}(U_\rho)$,

$$\int d\rho \psi(\rho) T(\rho, \xi) = \langle \psi T\rangle(\xi) \tag{4}$$

is in $C_\infty^V(U_\xi)$, then $T(\rho, \xi)$ is in $C_\infty^{V \otimes \mathcal{S}'(U_\rho)}(U_\xi)$, and whenever $\phi(\rho)$ is in $\mathcal{S}(U_\rho)$, the convolution product

$$(\phi * T)(\rho, \xi) = \int d\rho' \phi(\rho - \rho') T(\rho', \xi) \tag{5}$$

is in $C_\infty^V(U_\rho \times U_\xi)$.

Lemma 4.3: $T_n(\rho, \xi)$ in $C_\infty^V(U_\rho \times U_\xi)$ converges in $C_\infty^{V \times \mathcal{S}'(U_\rho)}(U_\xi)$ if, for any $\psi(\rho)$ in $\mathcal{S}(U_\rho)$, $\langle \psi T\rangle(\xi)$ converges in $C_\infty^V(U_\xi)$. The same is true with the substitutions

$$\mathcal{S}'(U_\rho) \leftrightarrow C_\infty(U_\xi), \quad \psi \in \mathcal{S}(U_\rho) \leftrightarrow \psi \in C_\infty'(U_\xi). \tag{6}$$

Proof: The first lemma is proved by directly checking the definitions; as for the last two, they are straightforward consequences of standard distribution theory[11,13]; they may also be found (with slight alterations) in Ref. 3.

[10] G. Köthe, *Topologische Lineare Räume* (Springer-Verlag, Berlin, 1960).

[11] L. Schwartz, *Théorie des distributions* (Hermann & Cie., Paris, 1950), Vols. I, and II.

[12] J. M. Gel'fand *et al.*, *Les distributions* (Dunod Cie., Paris, 1967), Vols. I–V.

[13] L. Schwartz, J. Anal. Math. (Jérusalem) **4**, 88 (1954).

Finally, we present two more standard results of distribution theory.

Lemma 4.4: If the U_i form a basis of E, i.e., if the union of the U_i covers E, and if A^F is either C_n^F or $\mathcal{S}'^F$, then X is in $A^F(E)$ wherever the restriction of X to U_i belongs to $A^F(U_i)$ [Schwartz's "principe du récollement des morceaux," (Ref. 11, Vol. I, Chap.I, Sec. 3)].

Lemma 4.5: If T is in $\mathcal{S}^V(U)$, and if T obeys an elliptic symmetric differential equation with analytic coefficients on U, then T is in $C_\infty^V(U)$ (Ref. 12, Vol. III, Chap. IV, Sec. 8).

Definition: If the group $\mathfrak{L}$ acts on E, i.e.,

$$\Lambda: v \in E \to \Lambda v \in E, \quad \Lambda \in \mathfrak{L}, \tag{7}$$

and if $X(v)$ is in $A^V(E)$ and verified,

$$X(\Lambda v) = D(\Lambda)X(v), \tag{8}$$

where $D(\Lambda)$ is the matrix of a representation of $\mathfrak{L}$ in V, then X is called a *tensorial A object* (tensorial analytic function, tensorial distribution, etc.), and the space of such X's will be denoted by $A^D(E)$ [rather than $A^V(E)$; clearly $A^D(E)$ is a subspace of $A^V(E)$].

5. TENSOR DISTRIBUTIONS: GENERAL PROPERTIES

Let $T(v)$ be a tensor distribution over E. If $\bar{v}$ is in E_R, and since E_R is open in E, then there exists a neighborhood $U_{\bar{v}}$ of $\bar{v}$ contained in E_R; the restriction of T to $U_{\bar{v}}$ is in $\mathcal{S}'^D(U_{\bar{v}})$. We perform the change to curvilinear coordinates and, if $U_{\bar{\rho}} \times U_{\bar{\xi}}$ is the image of $U_{\bar{v}}$ in these coordinates, then $T(v(\rho, \xi)) = T(\rho, \xi)$ is in $\mathcal{S}'^D(U_{\bar{\rho}} \times U_{\bar{\xi}})$ (Theorem 2.2 and Lemma 4.1). By virtue of the definitions (Sec. 4), for every $\psi(\rho)$ in $\mathcal{S}(U_{\bar{\rho}})$,

$$\langle \psi(\rho), T(\rho, \xi) \rangle \in \mathcal{S}'^D(U_{\bar{\xi}}). \tag{9}$$

(The only fact that needs some discussion is the conservation of the tensor character of T. But this is obvious if we notice that the ρ are invariant.) Now, if we recall the way $\mathfrak{L}$ acts in the ξ [Sec. 2, Eq. (3)], it follows that, in a neighborhood U'_∞ of $\bar{\xi}$, the $\langle \psi, T \rangle(\xi)$ satisfy an elliptic differential equation with analytic coefficients (see Appendix C for the explicit construction of this equation) so that Lemma 4.5 tells us that the restriction of $\langle \psi, T \rangle(\xi)$ to U'_∞ is in $C_\infty^D(U'_\infty)$. But $\mathfrak{L}$ acts effectively in R, that is, any ξ'' in $U_{\bar{\xi}}$ may be obtained as $\Lambda\xi'$ with ξ' in U'_∞. Combining this with the tensor character of $\langle \psi, T \rangle$, i.e., with the formula

$$\langle \psi, T \rangle(\xi'') = \langle \psi, T \rangle(\Lambda\xi') = D(\Lambda)\langle \psi, T \rangle(\xi'),$$

and with the linearity of the spaces C_∞, we obtain that the differentiability of $\langle \psi, T \rangle$ may be extended to all of $U_{\bar{\xi}}$, i.e., that $\langle \psi, T \rangle(\xi)$ is in $C_\infty^D(U_{\bar{\xi}})$, so that Lemma 4.2 tells us that whenever $\phi(\rho)$ is in $\mathcal{S}(U_\rho)$,

$$(\phi * T)(\rho, \xi) \in C_\infty^D(U_{\bar{\rho}} \times U_{\bar{\xi}}).$$

But then we may use Lemma 4.1 and Theorem 2.2 to perform back the change of variables and obtain

$$(\phi * T)(\xi) = \int d\rho' \phi(\rho' - \rho) T(v) \in C_\infty^D(U_{\bar{v}}). \tag{10}$$

Now, the $U_{\bar{v}}$ form a covering of E_R, and thus Lemma 4.4 applies. We have therefore proved our main theorem, which is stated below.

Theorem 5.1 (Main Theorem): If $T(v)$ is in $\mathcal{S}'^D(E)$ and $\phi(\rho(v))$ is in $\mathcal{S}(P_R)$, then $\langle \phi, T \rangle(v)$ and $(\phi * T)(v)$ are infinitely differentiable in E_R, in the remaining variables (the first), and in all the variables (the last). Moreover, they keep the same tensor character as T. This is mathematically expressed by formula (10) (for the last), replacing $U_{\bar{v}}$ by E_R.

Corollary 5.1: If we extend T to complex values of v by means of the complexified $\mathfrak{L}$ group $\mathfrak{L}^C$ by setting

$$T(\Lambda_C v) = D(\Lambda_C)T(v),$$

Λ_C in $\mathfrak{L}^C$, then we may replace the requirement "ϕ is in $\mathcal{S}(P_R)$" by "ϕ is in $\mathcal{S}_\omega(P_R)$" to get "analyticity" instead of "infinite differentiability" in Theorem 5.1. This follows from the fact that the $D(\Lambda_C)$ depend analytically on the parameters of Λ_C and from Lemma 4.1.

Corollary 5.2: The invariance of the ρ is decisive in all above arguments. However, once Theorem 5.1 and Corollary 5.1 are proved, we may perform a new change of variables $\rho \to T$ and still keep the properties of differentiability and analyticity (but no longer the tensor character!), whenever the change of variables is admissible (Lemma 4.1 applies). Thus, e.g., if $N = 1$ (Sec. 3), we may average $T(v)$ along any timelike direction, i.e., in v_t (the component of v along any timelike axis that may, in particular, be v_0, i.e., the time axis).

From our subsequent analysis (Sec. 6) it follows that if we define $T^{C'}$ as the restriction of T to E_R, then $T^{C'}$ may be "continuously" extended to all of E (by "continuity" across the frontier of E_R; recall the definition at the end of Sec. 3); let us call T^C such an extension. We then may define

$$T^d(v) = T(v) - T^C(v); \tag{11}$$

T^d is the "discontinuous" part of T and has support in $E - E_R$. By iterating this process, we would arrive at a decomposition

$$T(v) = T^C(v) + \sum_{r<R} T^r(v), \tag{12}$$

where T^r has support in the union of $E_{r'}$ with $r' < r$ and is "continuous" across the boundary of E_r. In any case, from (11) it follows that if we subtract at any T its discontinuous part, we obtain a "regular" T^C.

6. CLASSIFICATION OF TENSOR DISTRIBUTIONS

If, in Corollary 5.1, we take $\phi(\rho) \to \delta(\rho)$, we see that the problem of classifying tensor distribution may be reduced to the classification of tensor analytic functions. This has been solved quite generally,[1] and we simply state the results; for definiteness, we give them explicitly in the case $T = T_\mu(v^{(1)}, v^{(2)}, v^{(3)})$. We now return to the situations of Theorem 3.1.[14]

Situation (3i):

$$T^C_\mu(v^{(1)}, v^{(2)}, v^{(3)}) = \sum_{I=1}^{3} v_\mu^{(I)} t_I(v^{(I)}, v^{(J)}), \tag{13}$$

and the t_I are invariant distributions. Thus, the usual decomposition is valid on E_R.

As for the T^r, we have to consider distributions with support on the surface E_r. Their general form is well known (cf., e.g., Refs. 11 and 12) and we only have to apply the general theory. We do it in increasing order of difficulty.

Situation (3ii): E_0 is the point $v^{(1)} = v^{(2)} = v^{(3)} = 0$, so that T^0_μ is of the form

$$T^0_\mu(v^{(1)}, v^{(2)}, v^{(3)}) = \sum_{I=1}^{3} p_I(\Box) \frac{\partial}{\partial v_\mu^{(I)}} \delta^4(v^{(1)})\delta^4(v^{(2)})\delta^4(v^{(3)}), \tag{14}$$

where the p_I are polynomials in the d'Alembertian $\Box = \sum g_{\mu\nu}\partial_\mu\partial_\nu$.

Situation (3iii): Taking into account the results of Sec. 3, Theorem 3.4, we may represent T^3_μ as

$$T^3_\mu(d_1^{(2)}; \alpha_1^{(2)}, \alpha_1^{(3)}; v^{(1)}),$$

and the fact that T^3 has support in E_3 is exhibited by decomposing it as

$$p\left(\frac{\partial}{\partial d_1^{(2)}}\right)\delta^1(d_2^{(2)}) \cdots \delta^1(d_6^{(2)}) T'^{(3)}_\mu(\alpha, v^{(1)}).$$

It is not difficult, although cumbersome, to find the transformation properties of T^3. We leave this to the reader and work explicitly only the case $p \equiv 1$. The $\delta(d)$ are then invariant, and T'^3_μ is still a vector. The final result is then

$$T^3_\mu(v^{(1)}, v^{(2)}, v^{(3)}) = \delta^1(d_1^{(2)}) \cdots \delta^1(d_6^{(2)}) v_\mu^{(1)} t(\alpha_1^{(2)}, \alpha_1^{(3)}, v^{(1)} \cdot v^{(2)}), \tag{15}$$

and t is again invariant. We remark that $\alpha_1^{(1)}$ may be rewritten as $v^{(I)} \cdot v^{(I)}/v^{(1)} \cdot v^{(1)}$ (if $v^{(1)}$ is lightlike, we define the quotient by a limiting procedure), so that $t(\alpha, v^{(2)} \cdot v^{(1)}) = t'(v^{(I)} \cdot v^{(J)})$; this is true whenever neither of the v is zero and t is continuous across the light cone.

Situation (3iv): The analysis is similar to the former situation, and the result is likewise:

$$T^3_\mu(v^{(1)}, v^{(2)}, v^{(3)}) = \delta^1(d_1^{(3)})\delta^1(d_2^{(3)}) \sum_{I=1}^{2} v_\mu^{(I)} t_I(\alpha_1^{(3)}, \alpha_2^{(3)}, v^{(I')} \cdot v^{(J')}), \quad I', J' \leq 2, \tag{16}$$

and here again we may reexpress the α's in terms of the invariants $v^{(I)} \cdot v^{(J)}$ whenever the $v^{(I)} \neq 0$.

If T satisfies suitable support properties, we may set still stronger results. Thus, if, e.g., T vanishes unless $v^{(J)} = 0$ or $v^{(I)} \cdot v^{(I)} > 0$, the condition $d_1^{(2)} = \cdots = d_6^{(2)} = 0$ may be reexpressed as [we take the situation (3iii), for example]

$$[v^{(2)} \cdot v^{(1)}]^2 = [v^{(2)} \cdot v^{(2)}][v^{(1)} \cdot v^{(1)}],$$
$$[v^{(3)} \cdot v^{(1)}]^2 = [v^{(3)} \cdot v^{(3)}][v^{(1)} \cdot v^{(1)}],$$

i.e., in terms of the invariants ρ_{IJ}. We are thus allowed to separate the part of t in (15) that is "continuous" across $\alpha = 0$ and the part with support in $\alpha = 0$. The first can be rewritten in terms of the invariants in a form analogous to the second member of Eq. (13); the second is, by virture of arguments familiar by now, of the type

$$p_1(\Box^{(3)}) \frac{\partial}{\partial v^{(3)}} \delta^4(v^{(3)}) t_1(\rho_{IJ}) + p_2(\Box^{(3)}) v^{(1)} t_2(\rho_{IJ})$$

or the similar one containing the product $\delta^4(v^{(3)})\delta^4(v^{(2)})$ and the same with an arbitrary permutation of the superscripts 1, 2, 3 (since we have to consider all possible labelings).

Quite generally, the following general result may be easily obtained by putting together all these conditions.

Theorem 6.1: If the support of T is contained in the union of the open cones $v^{(I)} \cdot v^{(I)} > 0$ and the origins

[14] We will only consider proper vectors; otherwise, terms of the form $\epsilon_{\mu\sigma\tau\rho} v_\sigma v'_\tau v''_\rho$, where ϵ is the Levi-Civita symbol, should appear. We thank Professor O. Steinmann for this remark.

$v^{(2)} = 0$, and $T_\mu(v^{(1)}, \cdots, v^{(N)})$ is in $\mathcal{S}'^D(E)$, then T_μ may be canonically decomposed as

$$T_\mu(v^{(1)}, \cdots, v^{(N)}) = \sum_{S(1,2,3)} \Bigg\{ \sum_{I=0}^{N} \Bigg[\sum_{1 \le I' \le N} p_{I',S}(\Box^{(1)}, \cdots, \Box^{(1)}) \times \frac{\partial}{\partial v_\mu^{(I')}} \prod_{I'=1} \delta^4(v^{(I')}) t_{II',\nu}(\rho_{AB}) + \sum_{I \le I'' \le N} p_{I'',S}(\Box^{(1)}, \cdots, \Box^{(I)}) \times \prod_{I'=1} \delta^4(v^{(I')} v_\mu^{(I'')}) t_{II'',S}(\rho_{AB}) \Bigg] \Bigg\}, \tag{17}$$

where $\sum_{S(1,2,3)}$ is extended to all permutations of (1, 2, 3) and the term $I = 0$ in the sum $\sum_{I=0}^{N}$ is to be interpreted as causing the corresponding $\delta(v^{(0)})$, $p_0(\Box^{(0)})$, $\partial/\partial v^{(0)}$ to disappear, i.e., it gives the regular part [Eq. (13)] of T. $\rho_{AB} = v^{(A)} \cdot v^{(B)}$, $A, B = 1, \cdots, N$.

Corollary 6.1: The above theorem extends the result of Theorem 5.1 from E_R to the set E' consisting of v's such that neither of the $v^{(I)}$ $(I = 1, \cdots, N)$ vanishes.

Remark: We have taken the case of a vector distribution as an example. The general tensor distribution T_i may be expressed, *mutatis mutandi*, in the same manner; one only has to change the vectors $\partial/\partial v_r$, v_μ by the corresponding tensor functions (Refs. 1 and 15) $\pi_i(\partial/\partial v_\mu)$ $\pi_i(v_\mu)$. We also note that the t may be found from the T in a canonical manner with the standard procedures (see, e.g., Ref. 1).

7. SOME APPLICATIONS AND COMMENTS

If $W_{N+1}(v^{(1)}, \cdots, v^{(N)})$ are the Wightman functions in momentum space and, after having separated the translation invariance,[16] they have the suitable support properties and are tensor distributions. Our analysis applies to them, and consequently the results of Theorems 5.1, 6.1, and Corollary 6.1 hold. Thus, e.g., we obtain that the regular part of W need only be spread out in the invariants ρ_{IJ}. This is to be compared with a similar result obtained by Borchers[3] for x space; our results extend his to momentum space if we get rid of the singular part W^d of W. This is easily done in, e.g., a theory of massive particles where W^d is simply the contribution from the vacuum. Note that no analyticity properties have been used in obtaining these results.

Similar comments could be made concerning scattering amplitudes. We do not, however, give the pertinent analysis here.

Finally, a few more questions are considered.

1. *Discrete transformations.* Space reversal does not add anything new to what has been said, but this is not the case for time reversal ($p_0 \to -p_0$). The manifold of tensor distributions may be split invariantly into even and odd distributions with respect to time reversal, depending on $T(-v_0, v) = \pm T(v_0, v)$. For the (+) choice, the analysis remains unchanged; if the (−) sign holds, then sgn $v_0 T$ is even. We may thus write any T as

$$T = T^{\text{even}} + (\text{sgn } v_0) T'^{\text{even}}.$$

These considerations are well known (see, e.g., Gårding and Lions, Ref. 2).

2. *Distributions of the invariants.* To complete the description of Sec. 6, one has to give a description of the spaces of the t's, i.e., of the spaces of distributions of the invariants. This has been done by several people and we refer to the corresponding literature (Gårding and Lions,[2] Hepp,[1,17] etc.).

ACKNOWLEDGMENTS

The author would like to thank Professor B. Zumino for the hospitality extended to him at the Physics Department of New York University. Thanks are also due to Professor L. Ehrenpreis, Professor A. Galindo, Professor J.-L. Gervais, Professor E. Lurçat, Dr. E. Salusti, and Professor K. Symanzik for illuminating discussions and suggestions.

APPENDIX A

We give an explicit construction of the ρ, ξ. Let $\bar{v}$ be a point in E_r and consider $R(\bar{v})$. Let U be a neighborhood of $\bar{v}$ in E_r. If we denote by $R^\perp$ the plane orthogonal to $R(\bar{v})$ at $\bar{v}$, by virtue of the continuity of the R, we may suppose that U is small enough to guarantee that if $v' \neq v''$ are in the intersection of U and $R^\perp$, then $R(v') \neq R(v'')$; then, if ρ' are the coordinates of points in $R^\perp$, to every value of these there corresponds one single $R(v'(\rho'))$ so that we have a parametric family of surfaces $R_{\rho'}$. We may then take the system of orthogonal trajectories of the $R_{\rho'}$, and obtain a curvilinear system of coordinates. [This

[15] K. Hepp, Ann. Math. **152**, 149 (1963).

[16] R. Streater and A. S. Wightman, *TCP, Spin, and Statistics and All That* (W. A. Benjamin, Inc., New York, 1964); R. Jost, *The General Theory of Quantized Fields* (Am. Math. Soc. Publ., Providence, R.I., 1963).

[17] K. Hepp, Helv. Phys. Acta **37**, 639 (1964).

construction is classical; see, for instance, Ref. 18; the ξ's are the coordinates of R, the ρ those of the orthogonal family. In the concrete situation it is not difficult to show the ρ's explicitly (Theorems 3.1–3.3, Sec. 3).] Alternative group-theoretical constructions may also be easily given.[5]

APPENDIX B

We give, for completeness, the list of the different little groups. A detailed discussion with proofs may be found in Ref. 5.

(1) Three linearly independent vectors:

$$\mathfrak{W}_{v^{(1)}v^{(2)}v^{(3)}} = \{1\}.$$

(2) Two linearly independent vectors: (a) The plane $v^{(1)}v^{(2)}$ is spacelike; then $\mathfrak{W}_{v^{(1)}v^{(2)}} = \mathfrak{L}_1^1$. (b) The plane $v^{(1)}v^{(2)}$ is tangent to the light cone; then $\mathfrak{W}_{v^{(1)}v^{(2)}} = \mathfrak{E}_1$, where $\mathfrak{E}_1$ effects dilations along the tangent to the light cone. (c) The plane $v^{(1)}v^{(2)}$ cuts the light cone; then $\mathfrak{W}_{v^{(1)}v^{(2)}} = O_2$.

(3) One nonzero vector. (a) Spacelike: $\mathfrak{W}_v = \mathfrak{L}_2^1$. (b) Lightlike: $\mathfrak{W}_v = \mathfrak{E}_2$ (Euclidean group in two dimensions). (c) Timelike: $\mathfrak{W}_v = O_3$.

(4) $v \equiv 0$; then, $\mathfrak{W}_0 = \mathfrak{L}_3^1$.

APPENDIX C

Let us show that if $T(v^{(1)}, \cdots, v^{(N)})$ is a tensor distribution, then $\langle \psi T \rangle$ obeys an elliptic-symmetric (e-s, for short) differential equation in the ξ. First of all, we remark that if the $M_{\mu\nu}$ are the usual generators of the Lorentz group $\mathfrak{L}_3^1$, then, by going over to infinitesimal transformations, it follows that if T is a tensorial object [Definition, Sec. 4, formulas (7) and (8)], then

$$\sum_{I=1}\left\{v_\mu^{(I)}\frac{\partial}{\partial v_\nu^{(I)}} - v_\nu^{(I)}\frac{\partial}{\partial v_\mu^{(I)}}\right\}T(v^{(2)}, \cdots, v^{(N)}) = dD(M_{\mu\nu})T(v^{(1)}, \cdots, v^{(N)}), \quad \text{(C1)}$$

where dD is the representation of L_3^1 induced[6,7] by the representation D of $\mathfrak{L}_3^1$. We want to prove that, in terms of the ξ, Eq. (C1) becomes e-s for $\langle \psi, T \rangle$; it is clear that, since the ρ and ξ are independent, it is sufficient to show this for T as a function of ξ. Let $\bar{v}$ be a fixed point in E, $\mathfrak{W}_{\bar{v}}$ its little group, and $M_1, \cdots, M_r$ the generators of $\mathfrak{L}_3^1/\mathfrak{W}_{\bar{v}}$ (Sec. 2); then the dependence of T in the ξ is given by

$$T(v) \equiv T(\cdot, \xi_1, \cdots, \xi_r) = T(\cdot, \exp(\xi_1 M_1 + \cdots + M_r\xi_r)\bar{v},) \quad \text{(C2)}$$

and the action of the Λ on the ξ is as in formula (3), Sec. 2, i.e.,

$$\Lambda(\eta_1, \cdots, \eta_l) = \exp\left(\sum_1^l \eta_i M_i\right) : T\left(\cdot, \exp\sum_1^r \xi_j M_j\right) \to T\left(\cdot, \left[\exp\sum_1^l \eta_i M_i\right]\left[\exp\sum_1^r \xi_j M_j\right]\right). \quad \text{(C3)}$$

For small ξ, η, and for $\eta_i = \delta_{ji}\eta$,

$$[\exp \eta M_j]\left[\exp\sum_i^r \xi_i M_i\right] \simeq \exp[\xi_1 M_1 + \cdots + (\xi_j + \eta)M_j + \cdots + \xi_r M_r],$$

so that, near $\xi = 0$, Eq. (C3) gives

$$\frac{\partial}{\partial \xi_j} T(\cdot, \xi) = dD(M_j)T(\cdot, \xi). \quad \text{(C4)}$$

In general, the equation is of the form

$$\sum_i \alpha_{ji}(\xi)\frac{\partial}{\partial \xi_i} T(\cdot, \xi) = dD(M_j)T(\cdot, \xi), \quad \text{any } i, \quad \text{(C5)}$$

and the α_{ji} are analytic functions with $\alpha_{ji}(0) = \delta_{ji}$. Equation (C4) is clearly elliptic [i.e., Eq. (C5) is elliptic at $\xi = 0$]; by virtue of the analyticity of the α, Eq. (C5) is also elliptic in a neighborhood of the point $\xi = 0$. Q.E.D.

Let us remark that this argument contains as particular examples the ones currently used[1]; thus, e.g., the proof of Methée[2] is a specialization of ours for the simple case $N = 1$ and T scalar, i.e., $dD(M) \equiv 0$.

[18] L. P. Eisenhart, *Differential Geometry of Curves and Surfaces* (reprinted by Dover Publ., Inc., N.Y. 1909).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

"Haag Theorem" for the Point-Coupling Relativistic Lee Model*

F. J. YNDURÁIN†
New York University, New York, New York‡

(Received 28 November 1967)

We prove that the free and total (renormalized) Hamiltonians exist as essentially self-adjoint operators in the relativistic Lee model without cutoff. Nevertheless, their domains of definition only have the zero vector in common and thus the interaction Hamiltonian is meaningless ("Haag's theorem").

In a previous paper[1] [hereafter referred to as (I)], it has been shown that, by using a limit procedure, it is possible to define a total Hamiltonian H for the relativistic Lee model even in the limit of point coupling (no cutoff). The properties of this Hamiltonian are further analyzed in the present note, where we are able to prove the following: (1) The total Hamiltonian H, as defined in the domain introduced in (I) [Eq. (4.5)], which we henceforth denote by $\mathfrak{D}$, is an essentially self-adjoint operator. It governs the dynamics of the model; in particular, it possesses the correct spectral properties. (This last statement was proved in Ref. 2, by constructing a Schrödinger equation.) (2) If $\mathfrak{D}_0$ is the maximal domain of definition of the free (renormalized) Hamiltonian H_0, then $\mathfrak{D}_0$ and $\mathfrak{D}$ only have the null vector in common. As a consequence, the interaction Hamiltonian $H_{\text{int}} \equiv H - H_0$ becomes meaningless. These results, of course, complete and substantiate those of (I); in particular, they explain why the Møller operators are nonexistent, thus showing that the relativistic Lee model "satisfies" the theorem of Haag[3] and confirms (in a sense) the conjecture of Van Hove.[4]

1. We recall[5] that $\mathfrak{D}$ consists of vectors $|\chi\rangle$ of the form

$$|\chi\rangle = \int d^3p\, \Phi(p) V_p^* |0\rangle + \int d^3q\, d^3k \times \left\{\Psi(q,k) - \frac{\lambda g(q+k,k)\Phi(q+k)}{E_N(q)+\omega_k - E_V(q+k)}\right\} N_q^* a_k^* |0\rangle.$$

For them, the action of H is defined as

$$H|\chi\rangle = \lim_{f\to 1} H^f |\chi\rangle,$$

where [cf., (I), Eqs. (2.1), (2.2), (2.3), and (2.4)]

$$H^f = H_0 + \int d^3p\, \delta E_V^f(p) V_p^* V_p + \lambda \int d^3p\, d^3k\, \{[8E_V(p)E_N(p-k)\omega_k]^{-\frac{1}{2}} f(p,k) V_p^* N_{p-k} a_k + \text{H.c.}\},$$

$$H_0 = \int d^3p\, E_V(p) V_p^* V_p + \int d^3p\, E_N(p) N_p^* N_p + \int d^3k\, \omega_k a_k^* a_k,$$

$$\delta E_V^f(p) = |\lambda|^2 \int \frac{d^3k\, f(p,k)^2}{\{[E_N(p-k)+\omega_k - E_V(p)][8E_V(p)E_N(p-k)\omega_k]\}}.$$

The result, as is easy to check, is the last equation in (I), viz.,

$$H|\chi\rangle = \int d^3p \left[E_V(p)\Phi(p) + \lambda \int d^3k \frac{\Psi(p-k,k)}{(8E_V(p)E_N(p-k)\omega_k)^{\frac{1}{2}}}\right] V_p^* |0\rangle + \int d^3q\, d^3k \times \left\{[E_N(q)+\omega_k]\Psi(q,k) - \frac{\lambda E_V(q+k)\Phi(q+k)}{[E_N(q)+\omega_i - E_V(q+k)][8E_V(q+k)E_N(q)\omega_k]^{\frac{1}{2}}}\right\} N_q^* a_k^* |0\rangle. \quad (1)$$

We will presently show that, so defined, H is essentially self-adjoint in $\mathfrak{D}$. For this, since H is symmetric, it is sufficient[6] to prove that the only solution for $H^*|\zeta\rangle = \pm i|\zeta\rangle$ is the trivial one, $|\zeta\rangle = 0$; i.e., we have to check that if

$$\langle \zeta, H\chi\rangle = \pm i\langle \zeta, H\chi\rangle, \quad \text{for all } |\chi\rangle \text{ in } \mathfrak{D},$$

then $|\zeta\rangle = 0$. Computing explicitly, we get that, if

$$|\zeta\rangle = \int d^3p\, \xi(p) V_p |0\rangle + \int d^3q\, d^3k\, \eta(q,k) N_q^* a_k^* |0\rangle,$$

* Research supported in part by the National Science Foundation.
† Supported in part by Comisaría de Protección Escolar (Spain).
‡ Address from January, 1968: Theoretical Division, CERN, Geneva 23, Switzerland.

[1] F. J. Ynduráin, J. Math. Phys. **7**, 1133 (1966).
[2] F. J. Ynduráin, Anales Real Soc. Espan. Fis. Quim. (Madrid) **62**, A 317 (1966).
[3] R. Haag, Kgl. Dansk. Vid. Selsk. Mat.-Fys. Medd. **29**, No. 12 (1955). We use quotation marks for the word "satisfies" because Haag's theorem is a relativistic local theorem, whereas the Lee model is not.
[4] L. Van Hove, Physica **18**, 145 (1952); **22**, 343 (1956).
[5] We use consistently the notation of (I), which also coincides with the standard notation.
[6] See, e.g., N. I. Achieser and I. M. Glasmann, *Theorie der Linearen Operatoren in Hilbertraum* (Akademie-Verlag, Berlin, 1954).

then we should have

$$\int d^3p\, \bar{\xi}(p)\, E_V(p)\Phi(p) + \lambda \int d^3p\, d^3k\, \bar{\xi}(p) q(p,k)\Psi(p-k,k) + \int d^3q\, d^3k \times \left\{ \bar{\eta}(q,k)[E_N(q)+\omega_k]\Psi(q,k) - \frac{\lambda \bar{\eta}(q,k) E_V(q+k) g(q+k,k)\Phi(q+k)}{E_N(q)+\omega_k - E_V(q+k)} \right\}$$

$$= \pm i \left\{ \int d^3p\, \bar{\xi}(p)\Phi(p) + \int d^3q\, d^3k \left[\bar{\eta}(q,k)\Psi(q,k) - \lambda \frac{g(q+k,k)\bar{\eta}(q,k)\Phi(q+k)}{E_N(q)+\omega_k - E_V(q+k)} \right] \right\},$$

$$g(p,k) = [8E_V(p)E_N(p-k)\omega_k]^{-\frac{1}{2}}. \qquad (2)$$

From this, it already follows that both ξ and η must be different from zero or vanish at the same time. Since (2) must hold for all Φ and Ψ, we may take $\Psi \equiv 0$ and Φ in the manifold orthogonal to the function

$$\eta'(p) = \int \frac{d^3k\, \eta(p-k,k) E_V(p) g(p,k)}{[E_N(p-k)+\omega_k - E_V(q)]},$$

from which Eq. (2) gives that ξ must be parallel to $\eta'/[E_V \mp i]$, i.e.,

$$\xi(p) = \frac{c_1 \eta'(p)}{[E_V(p) \mp i]}. \qquad (3a)$$

Analogously, if $\Phi \equiv 0$ and Ψ is orthogonal to $\xi'(q,k) = \xi(q+k)g(q+k,k)$, then we get

$$\eta(q,k) = \frac{c_2 \xi'(q,k)}{[E_N(q)+\omega_k \mp i]}. \qquad (3b)$$

From Eqs. (3) we conclude that either $c_1 c_2 = 0$ or, on the support of ξ and η,

$$E_V(p) \mp i = c_1 c_2 E_V(p) \int d^3k\, \frac{|g(p,k)|^2}{E_N(p-k)+\omega_k \mp i} = c_1 c_2 E_V(p) \times \infty.$$

But this last equation only is satisfied if $p = \infty$, so that $|\zeta\rangle$ must vanish. Q.E.D.

2. We will show that $H_{\text{int}} = H - H_0$ is not defined by proving that, for any $|\chi\rangle$ in $\mathfrak{D}$, $H_0|\chi\rangle$ lies outside of the Hilbert space of (normalizable) superpositions of states with V and $N\theta$ particles. This result is obtained by direct computation:

$$H_0|\chi\rangle = |\chi_1\rangle + |\chi_2\rangle + |\chi_3\rangle,$$

$$|\chi_1\rangle = \int d^3p\, E_V(p)\Phi(p) V_p^* |0\rangle + \int d^3q\, d^3k [E_N(q)+\omega_k]\Psi(q,k) N_q^* a_k^* |0\rangle,$$

$$|\chi_2\rangle = -\lambda \int d^3q\, d^3k \times \frac{E_V(q+k) g(q+k,k)\Phi(q+k)}{E_N(q)+\omega_k - E_V(q+k)} N_q^* a_k^* |0\rangle,$$

$$|\chi_3\rangle = -\lambda \int d^3q\, d^3k\, g(q+k,k)\Phi(q+k) N_q^* a_k^* |0\rangle.$$

It is easy to check that $|\chi_1\rangle$ and $|\chi_2\rangle$ are normalizable. However,

$$\langle \chi_3 | \chi_3 \rangle = \int d^3q\, d^3k\, |g(q+k,k)\Phi(q+k)|^2 = \infty,$$

whatever Φ may be. Hence $H|\chi\rangle$ is outside the Hilbert space. Q.E.D.

ACKNOWLEDGMENTS

In conclusion we would like to acknowledge interesting correspondence with Dr. J. M. Lévy-Leblond, and the hospitality extended to us by Professor B. Zumino at the Department of Physics of New York University, where part of this work was performed.

Thanks are also due to the P.I.O., Comisaría de Protección Escolar (Spain), for partial economic support.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Generalized Second-Order Relativistic Wave Equations. I*

KENNETH RAFANELLI†
TRW Systems, Redondo Beach, California

(Received 9 November 1967)

A covariant second-order wave equation, free of subsidiary conditions, is deduced from the familiar linear relativistic wave equation for a free particle of arbitrary spin by use of the representation-invariant Lie algebra of $O(4, 1)$. The correspondence principle is used to interpret the physical content of this generalized equation, which explicitly admits zitterbewegung and implies an inverse spin dependence for the rest energy. Without further assumption this generalized second-order equation is equivalent to the Klein–Gordon equation for the particular Lie algebras of the Dirac and Duffin–Kemmer rings. For higher spins the imposition of a subsidiary condition, understood via the Bargmann–Wigner analysis, extends the equivalence with the Klein–Gordon equation and explicitly displays the above mass spectrum.

I. INTRODUCTION

The linear relativistic wave equation

$$(i\Gamma_\mu P_\mu + \kappa)\psi = 0, \tag{1}$$

where κ is a constant parameter related to the rest mass $P_\mu = -i\hbar\partial_\mu$, and the four-matrix operators Γ_μ satisfy the representation-invariant Lie algebra of the de Sitter group $O(4, 1)$,[1] i.e.,

$$(\Gamma_\mu, \Gamma_\nu) = \Gamma_{\mu\nu} \tag{2}$$

and

$$(\Gamma_{\mu\nu}, \Gamma_\sigma) = \Gamma_\mu\delta_{\nu\sigma} - \Gamma_\nu\delta_{\mu\sigma}, \tag{3}$$

which together imply

$$(\Gamma_{\mu\nu}, \Gamma_{\rho\sigma}) = -(\Gamma_{\mu\rho}\delta_{\nu\sigma} + \Gamma_{\nu\sigma}\delta_{\mu\rho} - \Gamma_{\mu\sigma}\delta_{\nu\rho} - \Gamma_{\nu\rho}\delta_{\mu\sigma}), \tag{4}$$

constitutes an accepted description of a free particle with arbitrary spin. There is, however, some discussion in the literature on the completeness of this description without demanding that ψ, a solution of (1), also be a solution of the Klein–Gordon (KG) equation.[2] This demand generally necessitates the introduction of subsidiary conditions, which can cause inconsistencies for the usual inclusion of electromagnetic interactions. It is partly to the question of subsidiary conditions that we address ourselves here, although interactions are not explicitly considered in this first paper.

We adopt the point of view that (1) does indeed constitute a complete description of a free-spinning particle and show that this does not preclude the existence of a second-order wave equation. In fact, the well-known Lie algebra of $O(4, 1)$, based on Eqs. (2) and (3), permits us to deduce a general, representation-invariant, second-order wave equation without resorting to the use of subsidiary conditions. This equation is developed in Sec. II. The remainder of this paper is then devoted to analysis of the properties of this equation.

It is first shown that the covariant Hamiltonian operator defined by this general second-order equation has as its correspondence limit a form of the already studied classical Hamiltonian for a free-spinning particle.[3] The solutions of the admitted Poisson-bracket equations of motion are known to predict helical trajectories, the superimposed oscillatory motion corresponding to zitterbewegung. These solutions are also known to yield a rest energy which decreases as the observed spin increases.[4] Thus comparison with classical theory indicates that in this general form the second-order wave equation displays an energy and angular momentum which are not divided into separately conserved orbital and intrinsic spin contributions and that the coupling gives rise to an implicit mass spectrum.

Finally, we examine the equivalence of this general second-order wave equation and the KG equation. The Lie algebra of the Dirac and Duffin–Kemmer rings[5] suffices to reduce our second-order equation to the familiar KG equation for these particular representations. Generally, however, for higher spins it is necessary to postulate a subsidiary condition in order to arrive at an equation of the KG type. To understand the nature of the subsidiary condition for the finite-dimensional representations, our second-order wave equation is first recast into the language of

* This work supported in part by the Office of Naval Research.

† Permanent address: Queens College of the City University of New York, Flushing, N.Y.

[1] Literature on the algebra of $O(4, 1)$ is extensive. For the notation adopted here and an exhaustive reference list, see H. C. Corben, *Classical and Quantum Theories of Spinning Particles* (Holden Day Publishing Co., San Francisco, 1968), Sec. 13.

[2] P. A. M. Dirac, Proc. Roy. Soc. (London) **A155**, 447 (1936); M. Fierz and W. Pauli, Proc. Roy. Soc. (London) **A173**, 211 (1939); H. J. Bhabha, Rev. Mod. Phys. **17**, 200 (1945); Harish-Chandra, Phys. Rev. **71**, 793 (1947).

[3] H. C. Corben, see Ref. 1, Sec. 7; K. Rafanelli, Phys. Rev. **5**, **155**, 1420 (1967); K. Rafanelli, Nuovo Cimento **52A**,342 (1967).

[4] H. C. Corben, see Ref. 1, Sec. 8; K. Rafanelli, J. Math. Phys. **8**, 1440 (1967).

[5] N. Kemmer, Proc. Roy. Soc. (London) **A173**, 91 (1939); E. M. Corson, *Introduction to Tensors, Spinors, and Relativistic Wave Equations* (Hafner Publishing Co., New York, 1953), Sec. 39.

Kramers, Belinfante, and Lubanski (KBL).[6] In this form imposing the subsidiary condition is equivalent to assuming the Bargmann–Wigner (BW) equations.[7] With this assumption our second-order equation reduces to an equation of the KG type (orbital and spin contributions separately conserved), with the previously implicit mass spectrum now explicitly displayed.

II. GENERALIZED SECOND-ORDER EQUATION

Adopting the usual procedure for constructing a second-order equation,[8] we multiply (1) on the left by $(-i\Gamma_\mu P_\mu + \kappa)$, yielding

$$(\Gamma_\mu\Gamma_\nu P_\mu P_\nu + \kappa^2)\psi = 0. \tag{5}$$

Then with

$$\Gamma_\mu\Gamma_\nu = \tfrac{1}{2}(\Gamma_\mu\Gamma_\nu + \Gamma_\nu\Gamma_\mu) + \tfrac{1}{2}(\Gamma_\mu\Gamma_\nu - \Gamma_\nu\Gamma_\mu) \tag{6}$$

we note that, since $(P_\mu, P_\nu) = 0$, the antisymmetric part of (6) does not contribute to (5).

Introducing the four-vector

$$\chi_\mu \equiv \tfrac{1}{2}(\Gamma_{\mu\sigma}\Gamma_\sigma + \Gamma_\sigma\Gamma_{\mu\sigma}), \tag{7}$$

the following representation-invariant identity is easily obtained, using (2) and (3):

$$(\Gamma_\mu\Gamma_\nu + \Gamma_\nu\Gamma_\mu) = 2\Gamma_\sigma\Gamma_\sigma\delta_{\mu\nu} - (\Gamma_{\mu\sigma}\Gamma_{\sigma\nu} + \Gamma_{\nu\sigma}\Gamma_{\sigma\mu}) + 2(\chi_\mu, \Gamma_\nu). \tag{8}$$

This last relation (8) is not to be considered as a subsidiary condition, since it follows directly from the Lie algebra of $O(4, 1)$. Since

$$C_0^s = -\tfrac{1}{2}\Gamma_{\mu\nu}\Gamma_{\mu\nu} + \Gamma_\mu\Gamma_\mu \tag{9}$$

is a Casimir operator of $O(4, 1)$, and

$$C_0^L = -\tfrac{1}{2}\Gamma_{\mu\nu}\Gamma_{\mu\nu} \tag{10}$$

is a Casimir operator of the homogeneous Lorentz group $SL(2, C)$, it follows that $\Gamma_\sigma\Gamma_\sigma = C_0^s - C_0^L$ separately commutes with all the elements of $O(4, 1)$ only for those representations for which $\chi_\mu = 0$ (e.g., the finite-dimensional Dirac ring, and the infinite-dimensional Majorana representation).[4,9] However, it is worth noting at this point that

$$(\Gamma_\sigma\Gamma_\sigma, \Gamma_{\mu\nu}) = 0 \tag{11}$$

for all representations.

With relations (6) and (8) we may write the second-order wave equation (5) as

$$\{\Gamma_\sigma\Gamma_\sigma P_\mu P_\mu - (\Gamma_{\mu\sigma}\Gamma_{\sigma\nu})P_\mu P_\nu + (\chi_\mu, \Gamma_\nu)P_\mu P_\nu + \kappa^2\}\psi = 0. \tag{12}$$

Equation (12) constitutes the representation-invariant second-order equation, free of subsidiary conditions, implied by (1) and the Lie algebra of $O(4, 1)$. An analysis of some properties of this equation is taken up in the next two sections.

III. THE CORRESPONDENCE PRINCIPLE

In order to interpret the physical content of Eq. (12), we turn to the relation between the quantum and classical theory of spinning particles. To facilitate this discussion we first examine the representation-invariant relation between velocity and momentum operators implied by (1).

If we multiply (1) on the left by $-i\Gamma_\mu$ and use (8), then, for $\kappa \neq 0$,

$$i\Gamma_\mu\psi = \frac{1}{2\kappa}\{2\Gamma_\sigma\Gamma_\sigma P_\mu + 2(\chi_\mu, \Gamma_\nu)P_\nu - (\Gamma_{\mu\sigma}\Gamma_{\sigma\nu} + \Gamma_{\nu\sigma}\Gamma_{\sigma\mu})P_\nu + \Gamma_{\mu\nu}P_\nu\}\psi. \tag{13}$$

This relation is not to be considered a subsidiary condition on ψ, since again it follows directly from the Lie algebra of $O(4, 1)$. If (13) is multiplied by P_μ, then comparison shows that (1) and (12) are equivalent statements, the latter arising when the velocity operator $i\Gamma_\mu$, which appears explicitly in (1), is replaced by its functional dependence on momentum. Thus, as is customary, if we consider the wave equation (1) to define the covariant Hamiltonian operator

$$H = i\Gamma_\mu P_\mu + \kappa, \tag{14}$$

then (12) defines the Hamiltonian operator re-expressed quadratically in the momentum as

$$H' = \Gamma_\sigma\Gamma_\sigma P_\mu P_\mu - (\Gamma_{\mu\sigma}\Gamma_{\sigma\mu})P_\mu P_\nu + (\chi_\mu, \Gamma_\nu)P_\mu P_\nu + \kappa^2. \tag{15}$$

The classical Hamiltonian H_{cl} corresponding to (15) may be obtained from

$$H_{\text{cl}} = \langle H' \rangle = \int \bar{\psi} H' \psi d^4x. \tag{16}$$

If we write out the third term of (15) explicitly, using (2) and (7), then

$$\langle(\chi_\mu, \Gamma_\nu)P_\mu P_\nu\rangle = \tfrac{1}{2}\int \bar{\psi}\{2\Gamma_\mu\Gamma_\sigma\Gamma_\sigma\Gamma_\nu - \Gamma_\sigma\Gamma_\sigma\Gamma_\mu\Gamma_\nu - \Gamma_\mu\Gamma_\nu\Gamma_\sigma\Gamma_\sigma\}P_\mu P_\nu\psi,$$

[6] H. A. Kramers, F. J. Belinfante, and J. K. Lubanski, Physica **8**, 597 (1941).

[7] V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. **34**, 211 (1948).

[8] M. E. Rose, "*Relativistic Electron Theory*" (John Wiley & Sons, Inc., New York, 1961), Sec. 22.

[9] E. Majorana, Nuovo Cimento **9**, 335 (1932).

which vanishes when use is made of Eq. (1), the adjoint equation $-i(P_\mu\bar{\psi})\Gamma_\mu + \kappa\bar{\psi} = 0$, and surface terms are neglected. Then, with (9), (10), and (16),

$$H_{\rm cl} = \int \bar{\psi}\Big\{(C_0^s - C_0^L)P_\mu P_\mu + 2\frac{C_0^L(\Gamma_{\mu\sigma}\Gamma_{\sigma\nu})P_\mu P_\nu}{(\Gamma_{\alpha\beta}\Gamma_{\alpha\beta})} + \kappa^2\Big\}\psi = 0. \quad (17)$$

With the bilinear associations[10]

$$v_\mu = \langle i\Gamma_\mu\rangle, \quad s_{\mu\nu} = \langle -ih\Gamma_{\mu\nu}\rangle, \quad (18)$$

the classical Hamiltonian predicted by (12) is

$$H_{\rm cl} = aP_\mu P_\mu + 2b\frac{s_{\mu\sigma}s_{\sigma\nu}P_\mu P_\nu}{s_{\alpha\beta}s_{\alpha\beta}} + \kappa^2 = 0, \quad (19)$$

where

$$a = C_0^s - C_0^L, \quad b = C_0^L. \quad (20)$$

Similarly, the above procedure may be used to obtain the relation between the classical variables corresponding to (13):

$$\kappa v_\mu = aP_\mu + 2b\frac{s_{\mu\sigma}s_{\sigma\nu}P_\nu}{s_{\alpha\beta}s_{\alpha\beta}}. \quad (21)$$

For the special choice $b = a$, (19) and (21) define the Hamiltonian formulation of the classical-relativistic pure gyroscope. This classical theory and its relation to the quantum theory of spinning particles has been studied elsewhere.[3,4] The classical theory predicted by (12) then is a generalization of the pure gyroscope with the supplementary condition $s_{\mu\nu}v_\nu = 0$ now replaced by

$$s_{\mu\nu}v_\nu = \left(\frac{\eta}{\kappa}\right)s_{\mu\nu}P_\nu, \quad (22)$$

where $\eta = a - b$. The Poisson-bracket equations of motions admitted by (19) are nevertheless identical to those of the pure gyroscope, because of (11). The solutions to these equations are known to predict helical trajectories, and the superimposed circular motion corresponds to the classical equivalent of zitterbewegung.[3,4] Further, it has been shown that the energy in the "momentum-rest" frame varies inversely with the magnitude of the observed spin.[3,4]

This analysis provides us with an interpretation of the dynamical content of the generalized second-order wave Eq. (12). It indeed describes a free-spinning particle, but in terms of variables from which the zitterbewegung or, equivalently, the coupling of positive and negative energy states has not been removed.[11] In other words, the energy and angular momentum are not generally divided into separately conserved orbital and intrinsic spin contributions. The question of equivalence to a covariantly uncoupled description is taken up next.

IV. RELATION TO THE KG EQUATION

Transposing the mass term, (12) becomes

$$\{\Gamma_\sigma\Gamma_\sigma P_\mu P_\mu + \kappa^2\}\psi = \{(\Gamma_{\mu\sigma}\Gamma_{\nu\sigma})P_\mu P_\nu - (\chi_\mu, \Gamma_\nu)P_\mu P_\nu\}\psi. \quad (23)$$

If we use (2), (3), and (7), the right-hand side of (23) may be written as

$$\tfrac{1}{2}(\Gamma_\sigma\Gamma_\sigma\Gamma_\mu\Gamma_\nu + \Gamma_\mu\Gamma_\nu\Gamma_\sigma\Gamma_\sigma + 2\Gamma_\sigma\Gamma_\mu\Gamma_\nu\Gamma_\sigma - 2\Gamma_\mu\Gamma_\sigma\Gamma_\nu\Gamma_\sigma - 2\Gamma_\sigma\Gamma_\mu\Gamma_\sigma\Gamma_\nu)P_\mu P_\nu\psi.$$

So that after some lengthy but straightforward algebra, based on (2), (3), and (5), Eq. (23) becomes

$$\{\Gamma_\sigma\Gamma_\sigma P_\mu P_\mu + \kappa^2(1 + \Gamma_\sigma\Gamma_\sigma)\}\psi = \{\Gamma_\sigma\Gamma_\mu\Gamma_\nu\Gamma_\sigma - 2\Gamma_\sigma\Gamma_\mu\Gamma_\sigma\Gamma_\nu\}P_\mu P_\nu\psi. \quad (24)$$

Equation (24) is still a representation-invariant statement, free of subsidiary conditions. It is not difficult to see that further attempts to simplify the right-hand side of (24), $R_{(24)}$, based on representation-invariant algebra, yield empty identities. In fact, unless

$$R_{(24)} = \kappa^2 f\psi, \quad (25)$$

where f is some constant or some function of $\Gamma_\sigma\Gamma_\sigma$, an equation at all resembling the KG equation is not recovered.

If we turn to particular representations of $O(4, 1)$, then, for the cases of spin $\frac{1}{2}$ and spin (0, 1), (25) is a direct consequence of the Lie algebra of the Dirac and Duffin–Kemmer rings. This is verified as follows.

A. *The Dirac Ring:* For this case, in our notation[1]

$$\Gamma_\mu = \tfrac{1}{2}\gamma_\mu, \quad \Gamma_\mu\Gamma_\nu + \Gamma_\nu\Gamma_\mu = \tfrac{1}{2}\delta_{\mu\nu}, \quad \Gamma_\sigma\Gamma_\sigma = 1. \quad (26)$$

Thus $R_{(24)} = -2\kappa^2$, and Eq. (24) becomes

$$\{P_\mu P_\mu + (2\kappa)^2\}\psi = 0, \quad (27)$$

and since by (26) we must have $m_0c = 2\kappa$ for the rest mass, (27) is the correct KG equation.

B. *The Duffin–Kemmer Ring:* For this case, in our notation[1,5]

$$\Gamma_\mu = \beta_\mu, \quad \Gamma_\mu\Gamma_\rho\Gamma_\nu = \Gamma_\nu\Gamma_\rho\Gamma_\mu = \Gamma_\mu\delta_{\nu\rho} + \Gamma_\nu\delta_{\mu\rho}, \quad P_\sigma\psi = \Gamma_\mu\Gamma_\sigma P_\mu\psi. \quad (28)$$

Thus $R_{(24)} = \kappa^2\psi$, and Eq. (24) becomes

$$\Gamma_\sigma\Gamma_\sigma\{P_\mu P_\mu + \kappa^2\}\psi = 0, \quad (29)$$

which, nontrivially, yields the correct KG equation with $m_0c = \kappa$.

[10] K. Rafanelli and R. Schiller, Phys. Rev. **135**, B279 (1964).
[11] See Ref. 8, Sec. 18.

Thus, for the above two cases, (25) is an identity in virtue of the Lie algebra characterizing the representations. For more general cases, (25) must be assumed and hence becomes a subsidiary condition. To justify this condition and evaluate f for arbitrary spin, we use as a tool the construction due to KBL[6] for the finite-dimensional representations. The wavefunction ψ is considered to be a four-spinor of rank $N = 2s$, where s is the spin, and in our notation

$$\Gamma_\mu = \tfrac{1}{2}\sum_{l=1}^{N}\gamma_\mu^l. \tag{30}$$

The γ matrices satisfy

$$\left.\begin{aligned}\gamma_\mu^l\gamma_\nu^k &= \gamma_\nu^k\gamma_\mu^l; \quad (l \neq k)\\ \gamma_\mu^l\gamma_\nu^l + \gamma_\nu^l\gamma_\mu^l &= 2\delta_{\mu\nu}; \quad (\text{each } l)\end{aligned}\right\}. \tag{31}$$

Using (30), our general second-order wave equation (12) becomes

$$\Big\{\sum_{k,l=1}^{N}\{\tfrac{1}{4}\gamma_\sigma^k\gamma_\sigma^l(P_\mu P_\mu + \kappa^2)\} + \kappa^2 - \tfrac{1}{16}\sum_{i,j=1}^{N}(\gamma_\sigma^k\gamma_\mu^i\gamma_\nu^j\gamma_\sigma^l - 2\gamma_\sigma^k\gamma_\mu^i\gamma_\sigma^l\gamma_\nu^j)P_\mu P_\nu\Big\}\psi = 0. \tag{32}$$

Again lengthy but straightforward algebra, now based on (31) and (5), allows simplification of (32) to

$$\Big\{\kappa^2 + \tfrac{1}{4}\sum_{k,l=1}^{N}\gamma_\sigma^k\gamma_\sigma^l P_\mu P_\mu\Big\}\psi = \tfrac{1}{4}\sum_{k,l=1}^{N}\{\gamma_\sigma^k\gamma_\sigma^l\gamma_\mu^l\gamma_\nu^l - \gamma_\mu^k\gamma_\nu^l\}P_\mu P_\nu\psi. \tag{33}$$

Although (33) is not representation-invariant, because of the completely reducible construction (30), it is nevertheless free of subsidiary conditions. As one might expect, further attempts to simplify (33), relying solely on (31) and (5), lead to empty identities. Thus, in order to arrive at an equation resembling the KG equation, we must assume something about the right-hand side of (33) [except for $N = 1$, in which case (27) is recovered]. Now, however, thanks to the BW analysis,[7] the choice is a natural one. If, in addition to (1), the BW equations are obeyed, i.e.,

$$i\gamma_\mu^l P_\mu\psi = -a\kappa\psi, \tag{34}$$

where $a\kappa$ is the rest mass, then (33) becomes

$$\{\kappa^2 + \Gamma_\sigma\Gamma_\sigma P_\mu P_\mu\}\psi = \Big\{-(a\kappa)^2\Gamma_\sigma\Gamma_\sigma + \frac{N^2}{4}(a\kappa)^2\Big\}\psi, \tag{35}$$

and we have returned to our original notation for $C_0^s - C_0^L$. Consistency between (30) and (34) requires $a = 2/N$. Finally, we have

$$\Gamma_\sigma\Gamma_\sigma\Big\{P_\mu P_\mu + \frac{4\kappa^2}{N^2}\Big\}\psi = 0. \tag{36}$$

Since the curly bracket is separately zero in virtue of (30) and (34) alone, we see that for the finite-dimensional representations our general second-order wave equation (12) reduces to a KG equation, provided the BW equations are assumed. The reduction explicitly displays the spin dependence of the rest mass, since

$$m_0 c = \frac{2\kappa}{N} = \frac{\kappa}{s}, \tag{37}$$

which is in accord with Bhabha's conclusions,[2] and corresponds to the classical spin dependence.[4]

It should also be noted that while the BW analysis applies only to the finite-dimensional representations, Eq. (36) correctly gives the well-known mass-spectrum characteristic of the infinite-dimensional Majorana representation if $N = 2s + 1$ and s is allowed to be $0, \frac{1}{2}, 1, \cdots$.

V. CONCLUSION

We have used the Lie algebra of $O(4, 1)$ to construct a representation-invariant second-order wave equation, free of subsidiary conditions, from the linear relativistic wave equation for a free particle of arbitrary spin. Both the first-order (linear) and second-order wave equations define the same Hamiltonian operator. The corresponding classical Hamiltonian describes a free-spinning particle in terms of variables possessing zitterbewegung and predicting a certain mass spectrum. In general, to obtain equivalence with a second-order equation free of zitterbewegung, it is necessary to impose a subsidiary condition. The condition is equivalent to assuming the BW equations and results in the explicit display of the same mass spectrum implicitly predicted by the general form of the equation. The spectrum exhibits an expected accumulation point at zero mass.

ACKNOWLEDGMENT

I wish to thank Dr. H. C. Corben for valuable suggestions and discussion.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Ensemble of Random Matrices with a Random Bias*

L. D. FAVRO† AND J. F. MCDONALD‡
Department of Physics, Wayne State University, Detroit, Michigan

(Received 22 September 1967)

The ensemble exp $[-\gamma \operatorname{Tr}(H - H_0)^2]$, where the eigenvalues λ_n of H_0 are given by some distribution $f(\lambda)$, is investigated. In particular, the limits of large and small γ for the orthogonal case are considered. Formal expressions are obtained for various distributions of the eigenvalues in the two limits. The approximation developed in the large-γ limit is also applied to the thermodynamics of an incompletely specified system. Further, it is shown that this approximation is easily extended to include the unitary and symplectic ensembles.

1. INTRODUCTION

In a previous paper[1] we investigated the effect of biasing a Gaussian ensemble of random matrices with a given matrix H_0. In particular, we calculated the nearest-neighbor spacing distribution for the orthogonal ensemble given by

$$P(H, H_0, \gamma) = \eta \exp[-\gamma \operatorname{Tr}(H - H_0)^2], \quad (1)$$

$$\eta(\gamma) = (\gamma/\pi)^{N/2}(2\gamma/\pi)^{N(N-1)/4}, \quad (2)$$

in the large- and small-γ limits. There it was assumed that the eigenvalues λ_n of H_0 were known explicitly. Here we generalize our discussion to the case when the eigenvalues are not known exactly, but are instead specified by a given probability distribution.

We shall again restrict our discussion to limiting cases. In the small-γ limit we shall calculate a formal expression for the nth-order spacing distribution. In the large-γ limit we shall calculate formal expressions for the joint eigenvalue distribution and the single eigenvalue distribution. In addition, the approximation methods developed for the large-γ limit will be applied to the thermodynamics of an incompletely specified system. In particular, the free energy of such a system will be expressed in terms of the free energy of the known Hamiltonian H_0.

2. SMALL-γ LIMIT

Given that the joint distribution for the unperturbed eigenvalues λ_n is $f(\lambda)$, the joint matrix-element distribution $p(H, \gamma)$ for the perturbed system is

$$p(H, \gamma) = \int f(\lambda) P(H, H_0, \gamma)\, d\lambda, \quad (3)$$

where $d\lambda = \prod_{i=1}^{N} d\lambda_i$. In the previous paper (MF) we found that, for *fixed* values of the λ_n, the nth-order spacing distribution $p^{(n)}(S, H_0, \gamma)$ could be written to second order as

$$p^{(n)}(S, H_0, \gamma) \simeq [\exp(-\gamma \operatorname{Tr} H_0^2)]\left\{p_0^{(n)}(S, \gamma) + \frac{2\gamma \operatorname{Tr} H_0^2}{(N-1)(N+2)} \times \left[p_0^{(n)}(S, \gamma) - 2\gamma\eta \frac{\partial}{\partial \gamma}(p_0^{(n)}(S, \gamma)/\eta)\right]\right\}, \quad (4)$$

where $p_0^{(n)}(S, \gamma)$ is the nth-order distribution for an unbiased Gaussian distribution. It appeared that this series expansion in powers of γ would converge rapidly if

$$(\tfrac{1}{4}\bar{S}_0^2) \gg \gamma N, \quad (5)$$

where $\bar{S}_0$ is an average spacing associated with the particular set of λ_n under consideration. The difference between that case and the present one is that Eq. (3) contains the additional integrations over the λ_n so that $p^{(n)}(S, \gamma)$ can be written as

$$p^{(n)}(S, \gamma) = \int f(\lambda) p^{(n)}(S, H_0, \gamma)\, d\lambda. \quad (6)$$

Assuming that $f(\lambda)$ is such that the condition given by (5) is satisfied for any set of λ_n with appreciable probability, we can insert our expansion for $p^{(n)}(S, H_0, \gamma)$ into Eq. (6). To second order it follows from (4) that

$$p^{(n)}(S, \gamma) \simeq A(\gamma) p_0^{(n)}(S, \gamma) - \frac{2\gamma}{(N-1)(N-2)} \frac{\partial A(\gamma)}{\partial \gamma} \times \left\{p_0^{(n)}(S, \gamma) - 2\gamma\eta \frac{\partial}{\partial \gamma}[p_0^{(n)}(S, \gamma)/\eta]\right\}, \quad (7)$$

where

$$A(\gamma) = \int f(\lambda)[\exp(-\gamma \operatorname{Tr} H_0^2)]\, d\lambda. \quad (8)$$

Thus, in the small-γ limit, the spacing distributions again approach the corresponding results for an unbiased Gaussian distribution.

---

* Some of this material is based on portions of a thesis of one of the authors (J. F. M.) presented to Wayne State University in partial fulfillment of the requirements of the Ph.D. degree.

† Supported in part by a Faculty Research Fellowship, Wayne State University.

‡ Present address: Department of Mathematics, University of Windsor, Ontario, Canada.

[1] J. F. McDonald and L. D. Favro, J. Math. Phys. **9**, 1114 (1968). This paper will be referred to as MF.

3. GENERAL EXPANSION FOR THE MATRIX-ELEMENT DISTRIBUTION IN THE LARGE-γ LIMIT

In the small-γ limit we calculated the various spacing distributions by interchanging the order of integrations, the λ_n integrations being performed last. Thus, we simply used our results for a fixed value of the λ_n, multiplied by $f(\lambda)$, and carried out the λ_n integrations. In the large-γ limit this approach is not as fruitful. This is because the conditions on the λ_n spectrum, for the validity of the approximations used in the fixed-λ case, impose conditions on $f(\lambda)$ such that we obtain nothing really new. In particular, $f(\lambda)$ must be such that each λ_n is much more sharply peaked about some value, say e_n, than the E_n and the rotation parameters (i.e., the parameters which determine the matrix which diagonalizes H) are about their peak values. Thus we find it necessary to use an alternative approximation for the integration over the rotation parameters.

The approximation we will use involves a cluster-type expansion, analogous to those used, for example, in the theory of imperfect gases.[2] In general, the technique of making a cluster expansion consists of adding and subtracting terms in such a way as to obtain a (supposedly convergent) series expansion. That is, given some function of a set of variables x_i, $i = 1, \cdots, N$, say $h(x)$, which is of the form

$$h(x) = \prod_{i=1}^{N} f_i(x_i), \tag{9}$$

and, given that each $f_i(x_i)$ is in some sense approximated by $g_i(x_i)$, then $h(x)$ can be expressed as

$$h(x) = \prod [g_i(x_i) + l_i(x_i)], \tag{10}$$

where

$$l_i(x_i) = f_i(x_i) - g_i(x_i). \tag{11}$$

This can be rewritten as

$$h(x) = \prod_{i=1}^{N} g_i(x_i) + \sum_{j=1}^{N} l_j(x_j)\left[\prod_{i}^{j} g_i(x_i)\right] + \cdots, \tag{12}$$

where $\prod_i^j$ is the product with the term $i = j$ missing. Thus, $h(x)$ has been expressed as a power series in the l_i, which are by assumption small.

We wish to apply this approach to the function

$$P(H, H_0, \gamma) = \left[\prod_{i=1}^{N} F(H_{ii} - \lambda_i, \gamma)\right]\left[\prod_{j>k} F(H_{jk}, 2\gamma)\right], \tag{13}$$

where

$$F(x, \alpha) = (\alpha/\pi)^{\frac{1}{2}} \exp(-\alpha x^2). \tag{14}$$

Note that we have chosen to use the representation in which H_0 is diagonal. Since

$$\lim_{\alpha \to \infty} F(x, \alpha) = \delta(x), \tag{15}$$

we make the expansion

$$\begin{aligned} P(H, H_0, \gamma) &= \left[\prod_i \delta(H_{ii} - \lambda_i)\right]\left[\prod_{k>l} \delta(H_{kl})\right] \\ &\quad + \sum_j l_{jj}\left[\prod_i^j \delta(H_{ii} - \lambda_i)\right]\left[\prod_{k>l} \delta(H_{kl})\right] \\ &\quad + \left[\prod_i \delta(H_{ii} - \lambda_i)\right] \sum_{m>n} l_{mn}\left[\prod_{k>l}^{mn} \delta(H_{kl})\right] + \cdots, \end{aligned} \tag{16}$$

where $\prod_{k>l}^{mn}$ is the product with the term corresponding to $k = m$ and $l = n$ missing, and where the following definitions have been used:

$$l_{jj} = F(H_{jj} - \lambda_j, \gamma) - \delta(H_{jj} - \lambda_j) \tag{17}$$

and

$$l_{mn} = F(H_{mn}, 2\gamma) - \delta(H_{mn}). \tag{18}$$

To first order this expansion can be rewritten as

$$\begin{aligned} P(H, H_0, \gamma) \sim \delta(H - H_0) &+ \sum_j \left\{\int_{-\infty}^{\infty} [\delta(H - H_0)]_{\lambda_j = x} \right. \\ &\left. \times F(x - \lambda_j, \gamma)\, dx - \delta(H - H_0)\right\} \\ &+ \sum_{m>n} \left\{\int_{-\infty}^{\infty} [\delta(H - H_0)]_{H_{mn} = (H_{mn} - x)} \right. \\ &\left. \times F(x, 2\gamma)\, dx - \delta(H - H_0)\right\}, \end{aligned} \tag{19}$$

where

$$\delta(H - H_0) = \left[\prod_i \delta(H_{ii} - \lambda_i)\right]\left[\prod_{k>l} \delta(H_{kl})\right]. \tag{20}$$

Note that we have inserted additional delta functions and integrations. This is to facilitate the transformation of variables from the matrix elements to the eigenvalues. Some applications of this expansion are given in the next two sections.

4. EXPANSION FOR THE JOINT-EIGENVALUE AND SINGLE-EIGENVALUE DISTRIBUTIONS FOR THE LARGE-γ LIMIT

To obtain an expansion for the joint-eigenvalue distribution from the expansion developed in the last section, we change variables from the H_{ij}'s to the eigenvalues E_n and some rotation parameters ϕ_i. If H_{ij} and H'_{ij} are two Hamiltonian matrices in the same

[2] See, for example, J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (J. Wiley & Sons, Inc., New York, 1959), Chap. 13, p. 277.

representation, then

$$\prod_{i\geq j}\delta(H_{ij}-H'_{ij})\,dH = \left[\prod_i (E_i-E'_i)\right]\left[\prod_j \delta(\phi_j-\phi'_j)\right]\Big/ J\begin{pmatrix}H_{ii}, H_{ij}\\ E_i\,,\ \phi_j\end{pmatrix} dH, \tag{21}$$

where the E_i and E'_i are ordered the same way, and the limits on the ϕ_j and ϕ'_j have been chosen so that the entire H_{ij} and H'_{ij} spaces are covered only once.[3]

Using this result to change variables in Eq. (19), averaging over the rotation parameters, and symmetrizing with respect to the labeling of the E_n, one obtains the unordered joint-eigenvalue distribution for a particular set of λ_n. The result of these calculations is

$$P(E_n,\lambda,\gamma)\sim\frac{1}{N!}\sum_p\Big[\delta(E-\lambda) + \sum_j\left\{\int_{-\infty}^{\infty}F(x-\lambda_j,\gamma)[\delta(E-\lambda)]_{\lambda_j=x}\,dx-\delta(E-\lambda)\right\} + \sum_{m>n}\left\{\int_{\infty}^{\infty}F(x,2\gamma)[\delta(E-\lambda)]_{\substack{\lambda_n=Q_{nm}^+\\ \lambda_m=Q_{nm}^-}}-\delta(E-\lambda)\right\}\Big], \tag{22}$$

where $\sum_p$ is the sum over all permutations of the labeling of the E_n:

$$\delta(E-\lambda)=\prod\delta(E_i-\lambda_i) \tag{23}$$

and

$$Q_{nm}^{\pm}=\tfrac{1}{2}\{(\lambda_n+\lambda_m)\pm[(\lambda_n-\lambda_m)^2+4x^2]^{\frac{1}{2}}\}. \tag{24}$$

Note that for this correction term it has been necessary to diagonalize a 2×2 matrix. Successive terms involve diagonalization of matrices of higher dimension (i.e., the next correction requires that 3×3 and 2×2 matrices be diagonalized, etc.). The x integrations in the second and third terms can now be explicitly carried out. However, depending on the form of $f(\lambda)$, it may be convenient in particular cases to retain the x integrations in the third term. Thus we shall retain that integration. Performing the λ_n integrations and also the x integrations in the second term, for the joint eigenvalue distribution $p(E_r,\gamma)$ we obtain

$$p(E_r,\gamma)=\int d\lambda\,P(E_r,\lambda,\gamma)f_s(\lambda) \sim f_s(E_r)+\sum_j[F_j(E_r)-f_s(E_r)] + \tfrac{1}{2}\sum_{m\neq n}[F_{mn}(E_r)-f_s(E_r)], \tag{25}$$

where

$$f_s(E_r)=\frac{1}{N!}\sum_p f(E_r), \tag{26}$$

$$F_j(E_r)=\int_{-\infty}^{\infty}F(\lambda_1-E_j,\gamma)[f_s(E_r)]_{E_j=\lambda_1}\,d\lambda_1, \tag{27}$$

and

$$F_{mn}(E_r)=\iiint_{-\infty}^{\infty}dx\,d\lambda_1\,d\lambda_2 F(x,2\gamma)[f_s(E_r)]_{\substack{E_n=\lambda_1\\ E_m=\lambda_2}} \times\delta(E_m-Q_{12}^-)\delta(E_n-Q_{12}^+). \tag{28}$$

Note that $\frac{1}{2}[F_{mn}(E_r)+F_{nm}(E_r)]$ is the joint eigenvalue distribution corresponding to the ensemble

$$P_{mn}(H)=f_s(H_{ii})F(H_{mn},2\gamma)\prod_{i>j}^{mn}\delta(H_{ij}), \tag{29}$$

where $m>n$.

The result of integrating Eq. (25) over all but one of the E_n can be written as

$$p(E,\gamma)\sim P_0^1(E)+[p_1(E,\gamma)-P_0^1(E)] + (N-1)[p_2(E,\gamma)-P_0'(E)], \tag{30}$$

where $P_0^n(\lambda_1,\cdots,\lambda_n)$ is the joint distribution of the first n eigenvalues of H_0:

$$P_0^n(\lambda_1,\cdots,\lambda_n)=\int_{-\infty}^{\infty}d\lambda_{n+1}\cdots\int_{-\infty}^{\infty}d\lambda_N f_s(\lambda), \tag{31}$$

$$p_1(E,\gamma)=\int_{-\infty}^{\infty}F(u,\gamma)P_0^1(E+u)\,du, \tag{32}$$

$$p_2(E,\gamma)=\tfrac{1}{2}\iiint_{-\infty}^{\infty}F(x,2\gamma)[\delta(E-Q_{12}^-)+\delta(E-Q_{12}^+)] \times P_0^2(\lambda_1,\lambda_2)\,dx\,d\lambda_1\,d\lambda_2, \tag{33}$$

and where the remaining eigenvalue has been simply denoted as E. One can explicitly perform one more integration in (33) without explicit knowledge of P_0^2. However, depending on P_0^2, there may be some preferred order of integration, so that we shall leave $p_2(E)$ in this form. It should be noted that $p_2(E)$ is the single eigenvalue distribution for the two-dimensional ensemble

$$P(H)=F(H_{12},2\gamma)P_0^2(H_{11},H_{22}). \tag{34}$$

Thus, if the distributions $P_0^1(\lambda_1)$ and $P_0^2(\lambda_1,\lambda_2)$ are known, the single-eigenvalue distribution can be calculated to first order from (30), (31), and (33).

[3] That is, on the right-hand side we have omitted peaks of the distribution function which correspond to trivial interchanges of the labels on the eigenvalues. Also, it should be noted that if there are degenerate levels, delta functions on the ϕ_j, which correspond to rotations in degenerate subspaces, will not occur in this expression since all rotations in such a subspace are equivalent. These delta functions are to be replaced by constants such that the integration over all the angles still yields unity.

In an analogous fashion one can calculate a formal expression for the nearest-neighbor spacing distribution. The expression is rather complicated and will not be given here.

The expressions given above for the corrections to $p(E_r, \gamma)$ and $P(E, \gamma)$ (also those for the spacing distribution) may, in general, be difficult to evaluate exactly. One could, of course, resort to numerical calculations to evaluate the required integrals. However, it may also be possible, depending on $f(\lambda)$, to obtain an asymptotic expansion in powers of γ^{-1} for the required integrals. As an example consider the orthogonal Gaussian distribution for which

$$f(\lambda) = c\eta(\alpha)\left(\prod_{i>j} |\lambda_i - \lambda_j|\right) \exp(-\alpha \sum \lambda_1^2), \quad (35)$$

where c is a normalization constant which is independent of α. The exact joint-eigenvalue distribution for the biased ensemble can be calculated exactly in this case by performing the required integrations in the matrix element space. The result is

$$p(E_r, \gamma) = c\eta(\alpha')\left(\prod_{i>j} |E_i - E_j|\right) \exp(-\alpha' \sum E_i^2), \quad (36)$$

where $\alpha' = \alpha\gamma/(\alpha + \gamma)$. We can use this exact expression for $p(E, \gamma)$ to investigate the validity of our cluster-type approximation.

If we now insert (35) into (27) and (28) and use Laplace's method[4] to obtain asymptotic expansions (to order $1/\gamma$) of the integrals, we get

$$F_j(E_r) \sim f(E_r)\Big[1 - \frac{\alpha}{2\gamma} - \frac{\alpha}{\gamma}\sum_{l\neq j}\frac{E_j}{E_j - E_l} + \frac{1}{2\gamma} \times \sum_{\substack{i>k \\ l\neq j}} \frac{1}{(E_j - E_l)(E_j - E_k)} + \frac{\alpha^2}{\gamma} E_j^2\Big], \quad (37)$$

$$F_{mn}(E_r) \sim f(E_r)\theta(E_n - E_m) \times \Big[2 + \frac{\alpha}{\gamma} - \frac{1}{2\gamma}\sum_{k\neq m,n}(E_k - E_n)(E_k - E_m)\Big]. \quad (38)$$

It is now easily seen that

$$\sum_j [F_j(E_r) - f(E_r)] \sim f(E_r)[-\alpha N^2/2\gamma + (\alpha^2/\gamma)\sum E_j^2] \quad (39)$$

and

$$\tfrac{1}{2}\sum_{m\neq n} [F_{mn}(E_r) - f(E_r)] \sim f(E_r)\alpha N(N-1)/4\gamma, \quad (40)$$

so that

$$p(E_r, \gamma) \sim f(E_r)[1 - N(N+1)\alpha/4\gamma + \alpha^2(\sum E_i^2)/\gamma]. \quad (41)$$

That this is indeed the correct expansion can be seen by expanding the exact expression (36) directly in powers of γ^{-1}. It should be noted that the N-dependent term in this series arises from the binomial expansion of the normalization constant in (36). If one wishes to take the limit of this expression as $N \to \infty$, this term must be reabsorbed into the normalization.

In everything we have done above it was assumed that the joint distribution for the unperturbed *eigenvalues* was given. Alternatively, one might consider the joint distribution for the *matrix elements* of the unperturbed system $f(H_0)$ to be given. Assuming a random perturbation to this distribution (here we consider only a Gaussian perturbation of half-width $\gamma^{-\frac{1}{2}}$), the perturbed joint eigenvalue distribution is

$$P(H, \gamma) = \eta \int_{-\infty}^{\infty} dH_0 f(H_0) e^{-\gamma \operatorname{Tr}(H - H_0)^2}, \quad (42)$$

where η is the normalization constant. At this point one can make a cluster expansion similar to (16) with the exponential factor in this expression. However, the explicit calculation of terms in that expansion is made more difficult by the fact that H_0 is not diagonal.

5. THERMODYNAMICS OF AN INCOMPLETELY SPECIFIED SYSTEM

Another application of our expansion for $P(H, H_0, \gamma)$, which was derived in Sec. 3, is to the statistical thermodynamics of an incompletely specified system. The formalism of the theory of an incompletely specified system is given by Mazo,[5] and an expansion given by Leff.[6] Here we give only a brief outline of the general formalism.

If the exact Hamiltonian of a thermodynamic system is unknown, one can still obtain information about the system by considering an ensemble of Hamiltonians of which the Hamiltonian of interest is a member. The ensemble should be consistent with whatever knowledge one has concerning the Hamiltonian. For example, all members of the ensemble might be assumed to have the same volume V and temperature T.

Each system in the ensemble is assumed to be describable by ordinary statistical mechanics—that is, by an ensemble of the possible states of that system. Thus, we have the so-called dual ensemble formalism. In the following we shall denote averages over the states of a given system by a single horizontal bar and averages over the ensemble of different Hamiltonians by the brackets $\langle\ \rangle$. We shall consider only the canonical ensemble of states.

[4] N. G. de Bruijn, *Asymptotic Methods in Analysis* (North-Holland Publ. Co., Amsterdam, 1961), p. 60.

[5] R. M. Mazo, J. Chem. Phys. **39**, 1224 (1963).

[6] H. S. Leff, J. Chem. Phys. **41**, 596 (1964).

The thermodynamic behavior of a system is given by performing the dual averaging process to thermodynamic variables. That is, if X is a thermodynamic variable, we must calculate $\langle \bar{X} \rangle$, where the statistical-mechanical average must be done before the average over systems.

Generally, the quantities of interest are functions of the logarithm of the partition function, and we restrict ourselves to one of these, namely, the free energy. Leff[6] has shown that the free energy for such systems involves the average of the logarithm of the partition function, not the logarithm of the average of the partition function.

We shall calculate an expansion for the free energy F, using the expansion for $P(H, H_0, \gamma)$ derived previously. That is, we shall calculate an expansion for

$$F = -\beta^{-1} \int f(\lambda)\, d\lambda \int P(H, H_0, \gamma) \ln (\mathrm{Tr}\, e^{-\beta H})\, dH \tag{43}$$

in the large-γ limit, where $\beta = (kT)^{-1}$ and k is Boltzmann's constant. Further, we shall restrict ourselves to fixed H_0, i.e.,

$$f(\lambda) = \prod_i \delta(\lambda_i - e_i), \tag{44}$$

where the e_i are fixed.

If the expansion for $P(H, H_0, \gamma)$ given by Eq. (19) is used, it follows that

$$\begin{aligned} -\beta F = \ln Z_0^\beta + \sum_j \Big\{ \int_{-\infty}^{\infty} F(x, \gamma) \ln [Z_0^\beta + \exp(-\beta x) \\ - \exp(-\beta e_j)]\, dx - \ln Z_0^\beta \Big\} \\ + \frac{1}{2} \sum_{m \neq n} \Big\{ \int_{-\infty}^{\infty} F(x, 2\gamma) \ln [Z_0 - \exp(-\beta e_m) \\ - \exp(-\beta e_n) + \exp(-\beta e_{mn}^{+}) \\ + \exp(-\beta e_{mn}^{-})] - \ln Z_0^\beta \Big\}, \end{aligned} \tag{45}$$

where

$$Z_0^\beta = \mathrm{Tr}\, [\exp(-\beta H_0)] \tag{46}$$

and

$$e_{mn}^{\pm} = \tfrac{1}{2}\{(e_m + e_n) \pm [(e_m - e_n)^2 + 4x^2]^{\frac{1}{2}}\}. \tag{47}$$

If the levels e_n have the type of structure considered in MF (i.e., the spacings of the levels are either small or large compared with $1/\gamma^{\frac{1}{2}}$), one can obtain an asymptotic expansion for the integrals involved in Eq. (47) in powers of γ^{-1}. This can be accomplished by noting that the integrals involved are of the form $\int_{-\infty}^{\infty} f(x) e^{-\alpha(x-x_0)^2}\, dx$, where $\alpha \gg 1$. Thus, if we expand $f(x)$ about x_0 in a power series and integrate term by term, we obtain a series in powers of α^{-1}. Applying this method to Eq. (45), we obtain

$$\begin{aligned} -\beta F \sim \ln Z_0^\beta + [\beta^2/4\gamma][1 - Z_0^{2\beta}/(Z_0^\beta)^2] \\ + (\beta^2/8\gamma Z_0^\beta) \sum_{m \neq n} \delta_{e_n e_m} \exp(-\beta e_m) \\ - [\beta/8\gamma Z_0^\beta] \sum_{m \neq n} (1 - \delta_{e_n e_m}) \\ \times \{\exp[-\beta(e_n + e_m)]\}/(e_m - e_n), \end{aligned} \tag{48}$$

which can be rewritten as

$$\begin{aligned} -\beta F \sim \ln Z_0^\beta + [\beta^2/8\gamma][1 - 2Z_0^{2\beta}/(Z_0^\beta)^2] \\ + (\beta^2/8\gamma Z_0^\beta) \int_0^1 Z_0^{\beta x} Z_0^{\beta(1-x)}\, dx. \end{aligned} \tag{49}$$

Hence, given the partition function of H_0, Z_0^β as a function of temperature, the free energy of the system can be calculated from Eq. (49). Note that we need to know Z_0 for a whole range of temperatures in order to evaluate F at a given temperature. This expansion is valid only if $\beta^2 \ll \gamma$ (i.e., $kT \gg \gamma^{-\frac{1}{2}}$).

6. APPLICATION OF THE LARGE-γ APPROXIMATION TO THE UNITARY AND SYMPLECTIC ENSEMBLES

The approximation method developed in the large-γ limit for the orthogonal ensemble given by Eq. (1) is easily extended to the corresponding unitary and symplectic ensembles. In particular, if we consider the ensembles

$$P_\beta(H, H_0, \gamma) = \eta_\beta \exp[-\gamma\, \mathrm{Tr}\, (H - H_0)^2], \tag{50}$$

$$\eta_\beta = (\gamma/\pi)^{N/2} (2\gamma/\pi)^{\beta N(N-1)/4} \tag{51}$$

(where $\beta = 1$, 2, or 4 corresponding to orthogonal, unitary, or symplectic cases, respectively), the joint eigenvalue distribution, for a particular set of λ_n given by (22), is modified only in that the term arising from the off-diagonal matrix elements (i.e., the term with $\sum_{m>n}$) is multiplied by the factor β.

The distribution $f(\lambda)$ can correspond to any of the three cases. The value of β is, of course, determined by the symmetry properties of $H - H_0$. Thus, the large-γ approximation can be used for any of the three cases as well as for various mixtures. For example, suppose we consider $P_2(H, H_0, \gamma)$ and assume that $f(\lambda)$ corresponds to an orthogonal ensemble. This corresponds to an othogonal ensemble (time-reversal invariant) with a random perturbation $H - H_0$, which is unitary (not time-reversal invariant). Hence the method developed is applicable to many interesting problems.

7. SUMMARY

In this paper we considered an orthogonal Gaussian ensemble of random matrices biased by a random matrix H_0, whose eigenvalues λ_n are given by a distribution $f(\lambda)$. The limiting cases of large and small γ were considered.

The results for the small-γ limit were found to approach those for an unbiased Gaussian distribution regardless of the form of $f(\lambda)$. On the other hand, the large-γ limit results were found to approach the corresponding results for $f(\lambda)$. In each limit a perturbation method was developed and a first-order correction calculated. In the small-γ limit we considered the nth-order spacing distribution, while in the large-γ limit we considered the single-eigenvalue and nearest-neighbor spacing distribution, as well as the application of the approximation to the thermodynamics of an incompletely specified system.

It was also pointed out that the methods developed for the large-γ limit are easily intended to include the unitary and symplectic cases. Thus, many interesting problems (such as mixtures of various ensembles) can be investigated using the formalism which was developed.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Special Functions and the Complex Euclidean Group in 3-Space. III

WILLARD MILLER, JR.
University of Minnesota, Minneapolis, Minnesota

(Received 12 February 1968)

This paper is the third in a series analyzing identities for special functions which can be derived from a study of the local representations of the Euclidean group in 3-space. Here identities are derived which relate Gegenbauer polynomials, Whittaker functions, Jacobi polynomials, and Bessel functions. Among the results are generalizations of the addition theorems for solid-spherical harmonics and a group-theoretic interpretation of the Maxwell theory of poles.

INTRODUCTION

This paper is the third in a series analyzing the special function theory related to T_6, the complex Euclidean group in 3-space. In the first two papers[1,2] (which we shall refer to as I and II, respectively) it was shown that important identities relating Bessel functions, Gegenbauer polynomials, Whittaker functions, and Jacobi polynomials could be derived in a straightforward manner from the study of certain local *irreducible* representations of T_6. After a brief review of terminology (Sec. 1), this paper proceeds as follows: In Secs. 2–4 we study classes of local *reducible* representations of T_6. These representations, closely related to the solution of Laplace's equation in spherical coordinates, lead to identities for Gegenbauer polynomials, which are generalizations of the addition theorems for solid-spherical harmonics.[3,4] Also, the Maxwell pole theory for spherical harmonics appears as a byproduct of the analysis. Section 5 is devoted to an examination of a class of irreducible representations closely related to the type F factorizations of Infeld and Hull.[5] These representations yield new identities for the Whittaker functions. Finally, in Sec. 6 we apply a technique developed by Weisner[6] and use T_6 to derive identities for special functions which are not directly related to the local representations of T_6.

As usual with this kind of work, most of the special function identities that we derive are well known. Our primary interest is in systematically deriving and elucidating the group-theoretic meaning of these identities rather than in deriving new identities.

The special functions studied in this paper ordinarily arise in one of two ways: as matrix elements corresponding to a local representation of T_6, or as basis vectors in a model of such a representation. Once the matrix elements have been computed, they remain valid for any model of the representation which occurs in modern physical theories.

[1] W. Miller, J. Math. Phys. **9**, 1162 (1968).
[2] W. Miller, J. Math. Phys. **9**, 1175 (1968).
[3] R. A. Sack, J. Math. Phys. **5**, 252 (1964).
[4] Y. N. Chiu, J. Math. Phys. **5**, 283 (1964).
[5] L. Infeld and T. Hull, Rev. Mod. Phys. **23**, 21 (1951).
[6] L. Weisner, Pacific J. Math. **5**, 1033 (1955).

7. SUMMARY

In this paper we considered an orthogonal Gaussian ensemble of random matrices biased by a random matrix H_0, whose eigenvalues λ_n are given by a distribution $f(\lambda)$. The limiting cases of large and small γ were considered.

The results for the small-γ limit were found to approach those for an unbiased Gaussian distribution regardless of the form of $f(\lambda)$. On the other hand, the large-γ limit results were found to approach the corresponding results for $f(\lambda)$. In each limit a perturbation method was developed and a first-order correction calculated. In the small-γ limit we considered the nth-order spacing distribution, while in the large-γ limit we considered the single-eigenvalue and nearest-neighbor spacing distribution, as well as the application of the approximation to the thermodynamics of an incompletely specified system.

It was also pointed out that the methods developed for the large-γ limit are easily intended to include the unitary and symplectic cases. Thus, many interesting problems (such as mixtures of various ensembles) can be investigated using the formalism which was developed.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Special Functions and the Complex Euclidean Group in 3-Space. III

WILLARD MILLER, JR.
University of Minnesota, Minneapolis, Minnesota

(Received 12 February 1968)

This paper is the third in a series analyzing identities for special functions which can be derived from a study of the local representations of the Euclidean group in 3-space. Here identities are derived which relate Gegenbauer polynomials, Whittaker functions, Jacobi polynomials, and Bessel functions. Among the results are generalizations of the addition theorems for solid-spherical harmonics and a group-theoretic interpretation of the Maxwell theory of poles.

INTRODUCTION

This paper is the third in a series analyzing the special function theory related to T_6, the complex Euclidean group in 3-space. In the first two papers[1,2] (which we shall refer to as I and II, respectively) it was shown that important identities relating Bessel functions, Gegenbauer polynomials, Whittaker functions, and Jacobi polynomials could be derived in a straightforward manner from the study of certain local *irreducible* representations of T_6. After a brief review of terminology (Sec. 1), this paper proceeds as follows: In Secs. 2–4 we study classes of local *reducible* representations of T_6. These representations, closely related to the solution of Laplace's equation in spherical coordinates, lead to identities for Gegenbauer polynomials, which are generalizations of the addition theorems for solid-spherical harmonics.[3,4] Also, the Maxwell pole theory for spherical harmonics appears as a byproduct of the analysis. Section 5 is devoted to an examination of a class of irreducible representations closely related to the type F factorizations of Infeld and Hull.[5] These representations yield new identities for the Whittaker functions. Finally, in Sec. 6 we apply a technique developed by Weisner[6] and use T_6 to derive identities for special functions which are not directly related to the local representations of T_6.

As usual with this kind of work, most of the special function identities that we derive are well known. Our primary interest is in systematically deriving and elucidating the group-theoretic meaning of these identities rather than in deriving new identities.

The special functions studied in this paper ordinarily arise in one of two ways: as matrix elements corresponding to a local representation of T_6, or as basis vectors in a model of such a representation. Once the matrix elements have been computed, they remain valid for any model of the representation which occurs in modern physical theories.

[1] W. Miller, J. Math. Phys. **9**, 1162 (1968).
[2] W. Miller, J. Math. Phys. **9**, 1175 (1968).
[3] R. A. Sack, J. Math. Phys. **5**, 252 (1964).
[4] Y. N. Chiu, J. Math. Phys. **5**, 283 (1964).
[5] L. Infeld and T. Hull, Rev. Mod. Phys. **23**, 21 (1951).
[6] L. Weisner, Pacific J. Math. **5**, 1033 (1955).

1. THE LIE ALGEBRA $\mathcal{C}_6$

The 6-dimensional complex Lie algebra $\mathcal{C}_6$ is defined by the commutation relations

$$\begin{aligned}
&[\mathfrak{J}^3, \mathfrak{J}^\pm] = \pm\mathfrak{J}^\pm, \ [\mathfrak{J}^+, \mathfrak{J}^-] = 2\mathfrak{J}^3,\\
&[\mathfrak{J}^3, \mathfrak{P}^\pm] = [\mathfrak{P}^3, \mathfrak{J}^\pm] = \pm\mathfrak{P}^\pm,\\
&[\mathfrak{J}^+, \mathfrak{P}^+] = [\mathfrak{J}^-, \mathfrak{P}^-] = [\mathfrak{J}^3, \mathfrak{P}^3] = \mathcal{O},\\
&[\mathfrak{J}^+, \mathfrak{P}^-] = [\mathfrak{P}^+, \mathfrak{J}^-] = 2\mathfrak{P}^3,\\
&[\mathfrak{P}^3, \mathfrak{P}^\pm] = [\mathfrak{P}^+, \mathfrak{P}^-] = \mathcal{O}. \qquad (1.1)
\end{aligned}$$

The 6-parameter complex Lie group T_6 consists of elements $\{\mathbf{w}, g\}$, $\mathbf{w} = (\alpha, \beta, \gamma) \in C^3$, $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2)$, $ad - bc = 1$ with group multiplication

$$\{\mathbf{w}, g\}\{\mathbf{w}', g'\} = \{\mathbf{w} + g\mathbf{w}', gg'\}, \qquad (1.2)$$

$$\begin{aligned}
g\mathbf{w} = (&a^2\alpha - b^2\beta + ab\gamma, \ -c^2\alpha + d^2\beta - cd\gamma,\\
&\times 2ac\alpha - 2bd\beta + (bc + ad)\gamma). \qquad (1.3)
\end{aligned}$$

$\mathcal{C}_6$ is the Lie algebra of T_6 and a neighborhood of $\mathcal{O} \in \mathcal{C}_6$ can be mapped diffeomorphically onto a neighborhood of the identity $\{\mathbf{0}, \mathbf{e}\} \in T_6$ ($\mathbf{e}$ is the 2×2 identity matrix) by means of the relation

$$\begin{aligned}
\{\mathbf{w}, g\} = &\exp(\alpha\mathfrak{P}^+ + \beta\mathfrak{P}^- + \gamma\mathfrak{P}^3)\exp(-b/d\mathfrak{J}^+)\\
&\times \exp(-cd\mathfrak{J}^-)\exp(-2\ln d\mathfrak{J}^3), \qquad (1.4)
\end{aligned}$$

where "exp" is the exponential map.

If V is a complex abstract vector space and ρ is a representation of $\mathcal{C}_6$ by linear operators on V, we set $\rho(\mathfrak{P}^\pm) = P^\pm$, $\rho(\mathfrak{P}^3) = P^3$, $\rho(\mathfrak{J}^\pm) = J^\pm$, $\rho(\mathfrak{J}^3) = J^3$. The linear operators $P^\pm$, P^3, $J^\pm$, J^3 satisfy commutation relations analogous to (1.1), where now $[A, B] = AB - BA$ for linear operators A, B on V. The invariant operators

$$\begin{aligned}
\mathbf{P}\cdot\mathbf{P} &= -P^+P^- - P^3P^3, \ \mathbf{P}\cdot\mathbf{J}\\
&= -\tfrac{1}{2}(P^+J^- + P^-J^+) - P^3J^3
\end{aligned}$$

have the property

$$[\mathbf{P}\cdot\mathbf{P}, \rho(\alpha)] = [\mathbf{P}\cdot\mathbf{J}, \rho(\alpha)] = 0$$

for all $\alpha \in \mathcal{C}_6$.

2. SOME REDUCIBLE REPRESENTATIONS

We examine the following two classes of *reducible* representations of $\mathcal{C}_6$ on a complex vector space V: $R^+(u_0)$ and $\uparrow^+$.

A. $R^+(u_0)$

Here u_0 is a complex number of such that $0 \leq \operatorname{Re} u_0 < 1$ and $2u_0$ is not an integer. There is a countable basis $\{f_m^{(u)}\}$ for V such that $m = u, \ u - 1, \ u - 2, \cdots$, and $u = u_0, \ u_0 \pm 1, \ u_0 \pm 2, \cdots$. The action of the infinitesimal operators on the basis vectors is given by

$$J^3 f_m^{(u)} = m f_m^{(u)}, \quad J^\pm f_m^{(u)} = (-u \pm m) f_{m\pm1}^{(u)}, \qquad (2.1)$$

$$P^3 f_m^{(u)} = \frac{1}{2u+1} f_m^{(u+1)}, \quad P^+ f_m^{(u)} = \frac{1}{2u+1} f_{m+1}^{(u+1)},$$

$$P^- f_m^{(u)} = \frac{-1}{2u+1} f_{m-1}^{(u+1)}, \qquad (2.2)$$

$$\mathbf{P}\cdot\mathbf{P} f_m^{(u)} \equiv 0, \quad \mathbf{P}\cdot\mathbf{J} f_m^{(u)} \equiv 0. \qquad (2.3)$$

B. $\uparrow^+$

There is a countable basis $\{f_m^{(u)}\}$ for V such that $m = u, \ u - 1, \cdots, \ -u + 1, \ -u$; $u = 0, 1, 2, \cdots$. The action of the infinitesimal operators on the basis vectors is given by (2.1)–(2.3). [If a vector $f_m^{(u)}$ on the right-hand side of one of the expressions (2.1)–(2.3) does not belong to the representation space, we set this vector equal to zero.]

It is left to the reader to verify that $R^+(u_0)$ and $\uparrow^+$ do define reducible representations of $\mathcal{C}_6$. In fact these representations are degenerate cases of the irreducible representations $R_3(\omega, 0, u_0)$ and $\uparrow_4(\omega, 0)$, constructed in II, obtained formally by choosing a new basis $f'^{(u)}_m = \omega^u f^{(u)}_m$ and passing to the limit as $\omega \to 0$. Corresponding to a fixed value of u, the vectors $\{f_m^{(u)}\}$ form a basis for an irreducible representation of the subalgebra $sl(2)$ of $\mathcal{C}_6$. Each such representation induced by $\uparrow^+$ has dimension $2u + 1$ and is denoted by $D(2u)$, while each representation induced by $R^+(u_0)$ is infinite-dimensional and is denoted by $\downarrow u$. A detailed analysis of the representations $D(2u)$ and $\downarrow u$ is given by Miller.[7]

In accordance with the procedure developed in I and II, we search for models of these abstract representations ρ such that the infinitesimal operators $\rho(\alpha)$, $\alpha \in \mathcal{C}_6$, are linear-differential operators in n complex variables. The basis vectors $\{f_m^{(u)}\}$ are then certain functions in these variables and the relations (2.1)–(2.3) are differential equations and recursion relations for the "special" functions $\{f_m^{(u)}\}$. Furthermore, each of our Lie algebra representations of $\mathcal{C}_6$ can be extended to a local Lie group representation of T_6. Such a local representation is defined by linear operators $\mathbf{T}(h)$, $h \in T_6$, acting on V such that $\mathbf{T}(h)\mathbf{T}(h') = \mathbf{T}(hh')$ for h, h' in a sufficiently small neighborhood of the identity. Due to this group property of the $\mathbf{T}$ operators, the matrix elements of these operators with respect to the basis $\{f_m^{(u)}\}$ will satisfy a series of addition theorems.

[7] W. Miller, *Lie Theory and Special Functions* (Academic Press Inc., New York, 1968).

3. MODELS OF THE REPRESENTATIONS

To begin we look for all models of the representation $R^+(u_0)$ in $n = 1$, 2, or 3 complex variables. According to Ref. 7, no model exists for $n = 1$. For $n = 2$, there is exactly one model (a special case of the type F operators):

$$J^3 = -z\frac{\partial}{\partial z} + u_0, \quad J^- = z^2\frac{\partial}{\partial z} - zt\frac{\partial}{\partial t} - u_0 z,$$
$$J^+ = \frac{-\partial}{\partial z} - \frac{t}{z}\frac{\partial}{\partial t} + \frac{u_0}{z}, \quad P^- = -\tfrac{1}{2}zt,$$
$$P^+ = \frac{1}{2}\frac{t}{z}, \quad P^3 = \tfrac{1}{2}t. \tag{3.1}$$

Here z, t are complex variables and u_0 is a fixed complex constant. The constant $\frac{1}{2}$ has been chosen for convenience in the computations to follow. Clearly the operators (3.1) satisfy the commutation relations (1.1). Furthermore, $\mathbf{P}\cdot\mathbf{P} \equiv 0$, $\mathbf{P}\cdot\mathbf{J} \equiv 0$. The basis vectors $f_m^{(u)}(z, t)$ for this model of $R^+(u_0)$ are defined up to a multiplicative constant by expressions (2.1) and (2.2) and may be chosen as follows:

$$f_m^{(u)} = \Gamma(u + \tfrac{1}{2})z^k t^u,$$
$$k = u_0 - m = u_0 - u, \quad u_0 - u + 1, \quad u_0 - u + 2, \cdots. \tag{3.2}$$

The possible values of u_0, u, m depend on the representation $R^+(u_0)$ and are listed in Sec. 2.

Since the operators (3.1) satisfy the commutation relations of $\mathfrak{G}_6$, they induce a local-multiplier representation of T_6 by operators $\mathbf{T}(h)$, $h \in T_6$, acting on the space of analytic functions in 2 complex variables. The operators $\mathbf{T}(h)$ can easily be computed from standard results in local Lie theory. We list only the results.

Clearly $\mathbf{T}(h) = \mathbf{T}(\mathbf{w}, g) = \mathbf{T}(\mathbf{w}, \mathbf{e})\mathbf{T}(\mathbf{0}, g)$, where $h = \{\mathbf{w}, g\}$ is defined by (1.3). If f is an analytic function defined in a neighborhood of the point $(z, t) \in \mathbb{C}^2$, $(t \neq 0)$, then

$$[\mathbf{T}(\mathbf{w}, \mathbf{e})f](z, t) = \exp\left[\frac{t}{2}\left(\frac{\alpha}{z} - \beta z + \gamma\right)\right]f(z, t),$$
$$\mathbf{w} = (\alpha, \beta, \gamma), \tag{3.3}$$

$$[\mathbf{T}(\mathbf{0}, g)f](z, t) = (a + cz)^{u_0}\left(d + \frac{b}{z}\right)^{-u_0} \times f\left[\frac{dz + b}{cz + a}, t(a + cz)\left(d + \frac{b}{z}\right)\right],$$
$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2), \quad ad - bc = 1. \tag{3.4}$$

As the reader can verify, these operators satisfy the property

$$\mathbf{T}(hh')f = \mathbf{T}(h)[T(h')f] \tag{3.5}$$

whenever both sides of the expression are well defined.

The matrix elements $\{v, n|\, \mathbf{w}, g\,|u, m\}$ of our model are defined by

$$\mathbf{T}(\mathbf{w}, g)f_m^{(u)} = \sum_v \sum_n \{v, n|\, \mathbf{w}, g\,|u, m\}f_n^{(v)}, \tag{3.6}$$

where v is summed over the values u_0, $u_0 \pm 1$, $u_0 \pm 2, \cdots$, and n over the values v, $v - 1$, $v - 2, \cdots$. It is clear that the functions (3.2) form an analytic basis for the representation space in the sense of Ref. 7, Chap. 2. Therefore, the matrix elements (3.6) are uniquely determined by the Lie-algebra relations (2.1) and (2.2), and are independent of our model.

Substituting (3.2) and (3.4) into (3.6), we find

$$(a + cz)^{u+m}\left(d + \frac{b}{z}\right)^{u-m} = \sum_{k=0}^{\infty}\{u, u - k|\, \mathbf{0}, g\,|u, m\}z^k, \quad \left|\frac{cz}{a}\right| < 1, \quad \left|\frac{b}{dz}\right| < 1, \tag{3.7}$$

or

$$\{v, n|\, \mathbf{0}, g\,|u, m\} = \frac{d^{u-n}a^{u+m}b^{n-m}(u - m)!}{(u - n)!} \times \frac{F(n - u, -m - u; n - m + 1; bc/ad)}{\Gamma(n - m + 1)}\delta_{v,u}$$
$$= \frac{d^{u-m}a^{u+n}c^{m-n}\Gamma(u + m + 1)}{\Gamma(u + n + 1)} \times \frac{F(m - u, -n - u; m - n + 1; bc/ad)}{\Gamma(m - n + 1)}\delta_{v,u}, \tag{3.8}$$

where $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2)$, $ad - bc = 1$. Clearly, these matrix elements are defined only in a suitably small neighborhood of $\mathbf{e}$. Substituting (3.3) into (3.6), we find

$$\{v, n|\, \mathbf{w}, \mathbf{e}\,|u, m\} = \frac{\Gamma(u + \frac{1}{2})(-\beta)^{m-n}(\gamma)^{v-u+n-m}}{2^{v-u}(v - u)!\Gamma(v + \frac{1}{2})} \times \sum_a \frac{(-\alpha b/\gamma^2)^a}{a!(m - n + a)!(v - u + n - m - 2a)!} \quad \text{if } v - u \geq 0,$$
$$= 0 \quad \text{otherwise.} \tag{3.9}$$

Here the sum is taken over all integral values of a such that the summand is defined.

By construction the matrix elements satisfy the

addition theorem

$$\{v, n|\, \mathbf{w} + g\mathbf{w}', gg'\,|u, m\} = \sum_{l=-\infty}^{\infty} \sum_{k=0}^{\infty} \{v, n|\, \mathbf{w}, g\,|u + l, u + l - k\} \times \{u + l, u + l - k|\, \mathbf{w}', g'\,|u, m\}, \quad (3.10)$$

valid for g, g' in a suitably small neighborhood of $\mathbf{e} \in SL(2)$.

Now that we have computed the matrix elements of $R^+(u_0)$, we look for a model of this representation in three complex variables. There is only one such model:

$$J^3 = t\frac{\partial}{\partial t}, \quad J^+ = -t\frac{\partial}{\partial z},$$
$$J^- = t^{-1}\left((1 - z^2)\frac{\partial}{\partial z} - 2zt\frac{\partial}{\partial t}\right),$$
$$P^3 = z\frac{\partial}{\partial r} + \frac{(1 - z^2)}{r}\frac{\partial}{\partial z} - \frac{zt}{r}\frac{\partial}{\partial t},$$
$$P^+ = t\left(\frac{\partial}{\partial r} - \frac{z}{r}\frac{\partial}{\partial z} - \frac{t}{r}\frac{\partial}{\partial t}\right),$$
$$P^- = t^{-1}\left((1 - z^2)\frac{\partial}{\partial r} - \frac{z(1 - z^2)}{r}\frac{\partial}{\partial z} + \frac{(z^2 + 1)}{r}t\frac{\partial}{\partial t}\right) \quad (3.11)$$

(the Model B operators constructed in I). The basis vectors $f_m^{(u)}[r, z, t]$ for this model are determined up to a multiplicative constant by relations (2.1)–(2.3) and may be chosen as follows:

$$f_m^{(u)}[r, z, t] = \frac{(r/2)^{-u-1}(u - m)!}{\Gamma(-u + \frac{1}{2})\sqrt{2}}\,\Gamma(m + \tfrac{1}{2})C_{u-m}^{m+\frac{1}{2}}(z)(2t)^m. \quad (3.12)$$

[Note: The relation $\mathbf{P} \cdot \mathbf{J} \equiv 0$ is satisfied identically by the operators (3.11), while the requirement $\mathbf{P} \cdot \mathbf{P} f_m^{(v)} = 0$ is closely related to Laplace's equation in spherical coordinates.] In fact, substitution of (3.11) and (3.12) into (2.1) and (2.2) leads to the following identities for the Gegenbauer polynomials $C_n^\lambda(z)$:

$$\frac{d}{dz}C_n^\lambda(z) = 2\lambda C_{n-1}^{\lambda+1}(z),$$
$$\left[(1 - z^2)\frac{d}{dz} - 2z\lambda + z\right]C_n^\lambda(z) = \frac{(n + 1)(n + 2\lambda - 1)}{2(1 - \lambda)}C_{n+1}^{\lambda-1}(z), \quad (2.1')$$
$$\left[(z^2 - 1)\frac{d}{dz} + (2\lambda + n + \tfrac{1}{2})z\right]C_n^\lambda(z) = (n + 1)C_{n+1}^\lambda(z),$$
$$\left[z\frac{d}{dz} + (2\lambda + n)z\right]C_n^\lambda(z) = 2\lambda C_n^{\lambda+1}(z),$$
$$2(1 - \lambda)\left[-(2\lambda + n)z^2 + (n + 1) + z(1 - z^2)\frac{d}{dz}\right]C_n^\lambda(z) = (n + 1)(n + 2)C_{n+2}^{\lambda-1}(z), \quad (2.2')$$

valid for $2\lambda \in \not\subset$ not an integer and $n = 0, 1, 2, \cdots$.

Using the type F operators and the basis vectors (3.2), we find

$$f_m^{(u)} = \frac{\Gamma(u + \frac{1}{2})}{\Gamma(m + \frac{1}{2})}(2P^3)^{u-m}f_m^{(m)}. \quad (3.13)$$

Clearly, this relation must hold for any model of the representations $R^+(u_0)$ or $\uparrow^+$. In terms of the operators (3.11) and basis functions (3.12) it reads

$$k!r^{-\lambda-k-\frac{1}{2}}C_k^\lambda(z) = \left(z\frac{\partial}{\partial r} + \frac{(1 - z^2)}{r}\frac{\partial}{\partial z} - \frac{z}{r}(\lambda - \tfrac{1}{2})\right)^k (r^{-\lambda-\frac{1}{2}}), \quad k = 0, 1, 2, \cdots.$$

Using the type F operators, the reader can easily derive other similar identities for the Gegenbauer polynomials. The study of identities of this form constitutes the Maxwell theory of poles.[8]

The differential operators (3.11) which define model B can be used to construct a local representation of T_6 by operators $\mathbf{T}(h)$, $h \in T_6$, acting on the space of analytic functions in 3 complex variables. The operators $\mathbf{T}(h)$ have been computed in I:

$$[\mathbf{T}(\mathbf{0}, g)f](r, z, t) = f\left(r, z(1 + 2bc) + abt + cd\frac{(z^2 - 1)}{t}, a^2t + 2acz + c^2\frac{(z^2 - 1)}{t}\right), \quad g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2), \quad (3.14)$$

$$[\mathbf{T}(\mathbf{w}, \mathbf{e})f](r, z, t) = f(rQ, (z + \gamma/r)Q^{-1}, (t + 2\beta/r)Q^{-1}),$$
$$Q = \left[1 + \frac{2\beta(1 - z^2)}{rt} + \frac{2\alpha}{r}\left(t + \frac{2\beta}{r}\right) + \frac{\gamma^2}{r^2} + \frac{2\gamma z}{r}\right]^{\frac{1}{2}}, \quad \mathbf{w} = (\alpha, \beta, \gamma). \quad (3.15)$$

Here f is defined and analytic in some neighborhood of the point $(r, z, t) \in \not\subset^3$. We have the group multiplication property

$$\mathbf{T}(hh')f = \mathbf{T}(h)[\mathbf{T}(h')f]$$

whenever both sides of this expression are well defined as analytic functions of r, z, and t.

It is easy to verify that the basis functions $f_m^{(u)}[r, z, t]$, Eq. (3.12), form an analytic basis for the representation space V. Therefore, we immediately have the identity

$$[\mathbf{T}(\mathbf{w}, g)f_m^{(u)}][r, z, t] = \sum_v \sum_n \{v, n|\, \mathbf{w}, g\,|u, m\}f_n^{(v)}[r, z, t], \quad (3.16)$$

[8] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. Tricomi, *Higher Transcendental Functions* (McGraw-Hill Book Company, New York, 1953), Vol. 2, Chap. 11.

where the operators $\mathbf{T}(\mathbf{w}, g)$ are given by (3.14), (3.15); the matrix elements are given by (3.8), (3.9). We examine some special cases of this identity.

If $\mathbf{w} = \mathbf{0}$, (3.16) reduces to

$$\frac{k!\Gamma(u-k+\frac{1}{2})}{\Gamma(2u-k+1)}\left(\frac{x^2}{2}\right)^k C_k^{u-k+\frac{1}{2}} \times \left[z^2 - z - 1 + \frac{2z-1}{x} + \frac{1}{x^2}\right] \times (1 + 2xz + x^2(z^2-1))^{u-k} = \sum_{l=0}^{\infty} \frac{l!\Gamma(u-l+\frac{1}{2})}{\Gamma(2u-l+1)}\left(\frac{x}{2}\right)^l \times \frac{F(-k, -2u+l; l-k+1; 1-x)}{\Gamma(l-k+1)} C_l^{u-l+\frac{1}{2}}(z),$$

$$|2xz + x^2(z^2-1)| < 1, \quad k = 0, 1, 2, \cdots, \quad (3.17)$$

which was already derived in I. If $g = \mathbf{e}$, $\alpha = \beta = 0$, we obtain

$$[1 + 2\gamma z + \gamma^2]^{-\lambda-k/2} C_k^{\lambda}[(z+\gamma)(1+2\gamma z+\gamma^2)^{-1/2}] = \sum_{l=0}^{\infty} (-\gamma)^l \binom{l+k}{l} C_{k+l}^{\lambda}(z), \quad |2\gamma z + \gamma^2| < 1, \quad (3.18)$$

which, when $k = 0$, simplifies to the well-known generating function

$$[1 + 2\gamma z + \gamma^2]^{-\lambda} = \sum_{l=0}^{\infty} (-\gamma)^l C_l^{\lambda}(z).$$

If $g = \mathbf{e}$, $\alpha = \gamma = 0$, we obtain

$$[1 + \beta(1-z^2)]^{-\lambda-k/2} \times C_k^{\lambda}[z(1+\beta(1-z^2))^{-1/2}](1+\beta)^{\lambda-1/2} = \sum_{l=0}^{\infty} (\beta/4)^l \frac{(k+2l)!}{k!l!} \frac{\Gamma(\lambda-l)}{\Gamma(\lambda)} C_{k+2l}^{\lambda-l}(z), \quad k = 0, 1, 2, \cdots, \quad |\beta(1-z^2)|, \ |\beta| < 1. \quad (3.19)$$

Finally, if $g = \mathbf{e}$, $\beta = \gamma = 0$, (3.16) reduces to

$$[1+\alpha]^{-\lambda-k/2} C_k^{\lambda}[z(1+\alpha)^{-1/2}] = \sum_{l=0}^{\infty} \frac{(-\alpha)^l}{l!} \frac{\Gamma(\lambda+l)}{\Gamma(\lambda)} C_k^{\lambda+l}(z), \quad k = 0, 1, 2, \cdots, \quad |\alpha| < 1. \quad (3.20)$$

If we restrict ourselves to consideration of the representation $\uparrow^+$, we can be somewhat more specific. First of all, the matrix element

$$\{u, m|\, \mathbf{w}, \mathbf{e}\,|0, 0\} = \{u, m|\, \alpha, \beta, \gamma; \mathbf{e}\,|0, 0\}$$

can be computed by making use of the identity

$$\{u, m|\, ab\xi, -cd\xi, (1+2bc)\xi; \mathbf{e}\,|0, 0\} = \{u, m|\, \mathbf{0}, g\,|u, 0\}\{u, 0|\, 0, 0, \xi; \mathbf{e}\,|0, 0\}, \quad (3.21)$$

where $g \in SL(2)$. In terms of the new variables [$\alpha = \beta b\xi, \beta = -cd\xi, \gamma = (1+2bc)\xi, \rho^2 = \gamma^2 + 4\alpha\beta = \xi^2$] the matrix elements on the right-hand side of (3.21) are

$$\{u, m|\, \mathbf{0}, g\,|u, 0\} = \frac{\Gamma(|m|+\frac{1}{2})u!}{\pi^{\frac{1}{2}}(u+|m|)!}\left(\frac{4}{\rho}\right)^{|m|} \alpha^{(|m|+m)/2} \times (-\beta)^{(|m|-m)/2} C_{u-|m|}^{|m|+\frac{1}{2}}(\gamma/\rho),$$

$$\{u, 0|\, 0, 0, \xi; \mathbf{e}\,|0, 0\} = \left(\frac{\xi}{2}\right)^u \frac{\Gamma(\frac{1}{2})}{u!\,\Gamma(u+\frac{1}{2})}.$$

Hence

$$\{u, m|\, \alpha, \beta, \gamma; \mathbf{e}\,|0, 0\} = \frac{\Gamma(|m|+\frac{1}{2})}{(u+|m|)!\,\Gamma(u+\frac{1}{2})}\left(\frac{\rho}{2}\right)^u \left(\frac{4}{\rho}\right)^{|m|} \times \alpha^{(|m|+m)/2}(-\beta)^{(|m|-m)/2} C_{u-|m|}^{|m|+\frac{1}{2}}(\gamma/\rho). \quad (3.22)$$

Note that this matrix element is a polynomial function of α, β, γ, and ρ^2. Thus, even though our derivation was valid only if $\rho^2 \neq 0$, (3.22) is also correct in the limit as $\rho \to 0$.

Applying the identity

$$\mathbf{T}(\alpha, \beta, \gamma; \mathbf{e}) f_0^{(0)}[r, z, t] = \sum_{u=0}^{\infty} \sum_{m=-u}^{u} \{u, m\,|\, \alpha, \beta, \gamma; \mathbf{e}\,|0, 0\} f_m^{(u)}[r, z, t]$$

to the Model B operators and simplifying, we obtain

$$[1 + 2\beta(1-z^2) + 2\alpha + \rho^2 + 2\gamma z]^{-\frac{1}{2}} = \sum_{u=0}^{\infty} \sum_{m=-u}^{u} \frac{\Gamma(|m|+\frac{1}{2})\Gamma(m+\frac{1}{2})(u-m)!}{\pi(u+|m|)!} 2^{|m|+m} \times (2\alpha)^{(|m|+m)/2}(-2\beta)^{(|m|-m)/2}\rho^{u-|m|} \times C_{u-|m|}^{|m|+\frac{1}{2}}(\gamma/\rho) C_{u-m}^{m+\frac{1}{2}}(z). \quad (3.23)$$

Just as in I we can use the Clebsch–Gordan coefficients $C(\cdot;\cdot\,|\,\cdot)$ to compute the general matrix element $\{v, n|\, \alpha, \beta, \gamma; \mathbf{e}\,|u, m\}$. The result is

$$\{v, n|\, \alpha, \beta, \gamma; \mathbf{e}\,|u, m\} = \sum_s \left[\frac{\pi(u-m)!(u+m)!}{(v-n)!(v+n)!} \times (v-u+s+n-m)!(v-u+s+m-n)!\right]^{\frac{1}{2}} \times C(u, 0; v-u+s, 0\,|\,v, 0) \times C(u, m; v-u+s, n-m\,|\,v, n) \times \{v-u+s, n-m|\, \alpha, \beta, \gamma; \mathbf{e}\,|0, 0\}, \quad (3.24)$$

where s ranges over the finite set of nonnegative integer values for which the summand is defined. The computation of identities for Gegenbauer polynomials using these matrix elements is left to the reader.

4. MORE REDUCIBLE REPRESENTATIONS

In analogy with the procedure in Secs. 2 and 3, we shall briefly analyze the following two new classes of

reducible representations of $\mathcal{C}_6$ on V: $R^-(u_0)$, $0 \leq \operatorname{Re} u_0 < 1$, $2u_0$ not an integer; and $\uparrow^-$.

A. $R^-(\mu_0)$, $0 \leq \operatorname{Re} u_0 < 1$, $2u_0$ not an integer

There is a countable basis $\{f_m^{(u)}\}$ for V such that $m = u,\ u-1,\ u-2, \cdots$, and $u = u_0,\ u_0 \pm 1,\ u_0 \pm 2, \cdots$.

There is a countable basis $\{f_m^{(u)}\}$ for V such that $m = u, u-1, \cdots, -u+1, -u$ and $u = 0, 1, 2, \cdots$.

For each representation the action of the infinitesimal operators on the basis vectors is given by

$$J^3 f_m^{(u)} = m f_m^{(u)}, \quad J^{\pm} f_m^{(u)} = (-u \pm m) f_{m\pm 1}^{(u)}, \tag{4.1}$$

$$P^3 f_m^{(u)} = \frac{(u+m)(u-m)}{2u+1} f_m^{(u-1)},$$

$$P^+ f_m^{(u)} = \frac{-(u-m)(u-m-1)}{2u+1} f_{m+1}^{(u-1)}, \tag{4.2}$$

$$P^- f_m^{(u)} = \frac{(u+m)(u+m-1)}{2u+1} f_{m-1}^{(u-1)},$$

$$\mathbf{P} \cdot \mathbf{P} f_m^{(u)} \equiv 0, \quad \mathbf{P} \cdot \mathbf{J} f_m^{(u)} \equiv 0. \tag{4.3}$$

[If a vector $f_m^{(u)}$ on the right-hand side of one of the expressions (4.1)–(4.3) does not belong to the representation space, we set this vector equal to zero.]

The representations $R^-(u_0)$ and $\uparrow^-$ are degenerate cases of the irreducible representations $R_3(\omega, 0, u_0)$ and $\uparrow_4(\omega, 0)$ constructed in II, obtained formally by choosing a new basis $f'^{(u)}_m = \omega^{-u} f_m^{(u)}$ and going to the limit as $\omega \to 0$. Corresponding to each fixed value of u, the vectors $\{f_m^{(u)}\}$ form a basis for an irreducible representation of the subalgebra $sl(2)$ of $\mathcal{C}_6$. The finite-dimensional representations $D(2u)$ are induced by $\uparrow^-$, while the infinite-dimensional representations $\downarrow u$ are induced by $R^-(u_0)$.

According to our usual procedure, we search for models of these representations and compute their matrix elements. Unfortunately, $R^-(u_0)$ and $\uparrow^-$ have no models in two complex variables. However, the structure of the abstract recursion relations (4.1)–(4.3) is simple enough that we can compute the matrix elements of $R^-(u_0)$ and $\uparrow^-$ directly from the abstract relations. We then apply our results to a model in three complex variables, which does exist.

The matrix elements can be defined formally by

$$\begin{aligned} \mathbf{T}(\mathbf{w}, g) f_m^{(u)} &= \exp(\alpha P^+ + \beta P^- + \gamma P^3) \exp(-b/dJ^+) \\ &\quad \times \exp(-cdJ^-) \exp(-2 \ln d J^3) f_m^{(u)} \\ &= \sum_{v,n} \{v, n | \mathbf{w}, g | u, m\} f_n^{(v)}, \end{aligned}$$

$$\mathbf{w} = (\alpha, \beta, \gamma), \quad g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2).$$

The values assumed by the variables u, v, m, n depend on which of the representations $R^-(u_0)$ or $\uparrow^-$ we are studying. For the present we treat both representations simultaneously.

Since relations (4.1) and (2.1) are identical, it follows immediately that the matrix elements

$$\{v, n | \mathbf{0}, g | u, m\}$$

are given by Eq. (3.8). Furthermore, a simple induction argument based on (4.2) yields the results

$$\begin{aligned} &\{v, n | 0, 0, \gamma; \mathbf{e} | u, m\} \\ &= \left(\frac{\gamma}{2}\right)^{u-v} \frac{\Gamma(v+\frac{3}{2})}{\Gamma(u+\frac{3}{2})} \frac{\Gamma(u+m+1)}{\Gamma(v+m+1)} \frac{(u-m)!}{(v-m)!} \frac{\delta_{m,n}}{(u-v)!} \\ &\qquad \text{if} \quad u - v \geq 0, \\ &= 0 \qquad \text{if} \quad v - u > 0, \end{aligned} \tag{4.4}$$

$$\begin{aligned} &\{v, n | \alpha, 0, 0; \mathbf{e} | u, m\} \\ &= \left(\frac{-\alpha}{2}\right)^{u-v} \frac{\Gamma(v+\frac{3}{2})}{\Gamma(u+\frac{3}{2})} \frac{(u-m)!}{(2v-u-m)!} \frac{\delta_{m,n-u+v}}{(u-v)!} \\ &\qquad \text{if} \quad u - v \geq 0, \\ &= 0 \qquad \text{if} \quad v - u > 0, \end{aligned} \tag{4.5}$$

$$\begin{aligned} &\{v, n | 0, \beta, 0; \mathbf{e} | u, m\} \\ &= \left(\frac{\beta}{2}\right)^{u-v} \frac{\Gamma(v+\frac{3}{2})}{\Gamma(u+\frac{3}{2})} \frac{\Gamma(u+m+1)}{\Gamma(2v-u+m+1)} \frac{\delta_{m,n-u+v}}{(u-v)!} \\ &\qquad \text{if} \quad u - v \geq 0, \\ &= 0 \qquad \text{if} \quad v - u > 0. \end{aligned} \tag{4.6}$$

The expression for the general matrix element $\{v, n | \mathbf{w}; \mathbf{e} | u, m\}$ of the representation $R^-(u_0)$ is rather complicated and we need not take the time to derive it. Similarly, we do not derive an expression for the most general matrix element of $\uparrow^-$, although this is not so complicated.[4]

The Model B operators (3.11) can be used to construct models of $R^-(u_0)$ and $\uparrow^-$ in three complex variables. In fact, relations (4.1)–(4.3) will be satisfied, provided that we choose the basis vectors as follows:

$$f_m^{(u)}[r, z, t] = \left(\frac{r}{2}\right)^u \frac{(u-m)!\,\Gamma(m+\frac{1}{2})}{\sqrt{2}\,\Gamma(u+\frac{3}{2})} C_{u-m}^{m+\frac{1}{2}}(z)(2t)^m, \tag{4.7}$$

where the possible values of the variables u, m are determined by the representation space to which the basis vectors belong. To see the equivalence between our models and certain recursion relations for Gegenbauer polynomials, we substitute (4.7) into (4.2):

$$\left[(1-z^2)\frac{d}{dz} + kz\right] C_k^{\lambda}(z) = (k + 2\lambda - 1) C_{k-1}^{\lambda}(z),$$

$$\left[z\frac{d}{dz} - k\right] C_k^{\lambda}(z) = 2\lambda C_{k-2}^{\lambda+1}(z),$$

$$\begin{aligned} &2(\lambda-1)\left[-z(1-z^2)\frac{d}{dz} - z^2 k + (k + 2\lambda - 1)\right] C_k^{\lambda}(z) \\ &\quad = (k + 2\lambda - 1)(k + 2\lambda - 2) C_k^{\lambda-1}(z). \end{aligned} \tag{4.8}$$

As was shown in Sec. 3, the differential operators (3.11) define a local representation of T_6 by operators $\mathbf{T}(\mathbf{w}, g)$, Eqs. (3.14), and (3.15). Furthermore, it is easy to see that the functions (4.7) form an analytic basis for such a representation. Thus the matrix elements defined by

$$\mathbf{T}(\mathbf{w}, g)f_m^{(u)} = \sum_{v,n} \{v, n|\, \mathbf{w}, g\,|u, m\} f_n^{(v)} \tag{4.9}$$

are identical with those computed earlier in this section. In addition, we have the identity

$$\{v, n|\, \mathbf{w} + g\mathbf{w}';\, gg'\,|u, m\} = \sum_{l=-\infty}^{\infty} \sum_{k=0}^{\infty} \{v, n|\, \mathbf{w}, g\,|u + l, u + l - k\} \times \{u + l, u + l - k|\, \mathbf{w}', g'\,|u, m\}, \tag{4.10}$$

valid for g, g' in a suitably small neighborhood of $\mathbf{e}$. The following special cases of (4.9) are of interest: If $g = \mathbf{e}$, $\alpha = \beta = 0$, this identity becomes

$$[1 + \gamma^2 + 2\gamma z]^{k/2} C_k^{\lambda}((z + \gamma)[1 + \gamma^2 + 2\gamma z]^{-1/2}) = \sum_{l=0}^{k} \gamma^l \binom{2\lambda + k - 1}{l} C_{k-l}^{\lambda}(z); \tag{4.11}$$

if $g = \mathbf{e}$ and $\beta = \gamma = 0$, one obtains

$$(1 - \alpha)^{k/2} C_k^{\lambda}(z(1 - \alpha)^{-1/2}) = \sum_{l=0}^{k/2} \alpha^l \binom{\lambda + l - 1}{l} C_{k-2l}^{\lambda+l}(z); \tag{4.12}$$

if $g = \mathbf{e}$, and $\alpha = \gamma = 0$, there follows

$$[1 + \beta(1 - z^2)]^{k/2} (1 + \beta)^{\lambda - 1/2} C_k^{\lambda}(z[1 + \beta(1 - z^2)]^{-1/2}) = \frac{\Gamma(2\lambda + k)}{\Gamma(\lambda)} \sum_{l=0}^{\infty} \left(\frac{\beta}{4}\right)^l \frac{\Gamma(\lambda - l)}{l!\, \Gamma(2\lambda + k - 2l)} C_k^{\lambda - l}(z), \quad |\beta| < 1. \tag{4.13}$$

5. A CLASS OF IRREDUCIBLE REPRESENTATIONS

We now turn our attention to a new class of *irreducible* representations of $\mathcal{C}_6$ listed in Ref. 7:

$R_3'(\zeta, u_0)$, $\zeta \neq 0$, $0 \leq \operatorname{Re} u_0 < 1$, $2u_0$ not an integer.

There is a countable basis $\{f_m^{(u)}\}$ for the representation space V such that $m = u, u - 1, u - 2, \cdots$, and $u = u_0, u_0 \pm 1, u_0 \pm 2, \cdots$. The action of the infinitesimal operators on the basis vectors is given by

$$J^3 f_m^{(u)} = m f_m^{(u)}, \quad J^{\pm} f_m^{(u)} = (-u \pm m) f_{m\pm1}^{(u)}, \tag{5.1}$$

$$P^3 f_m^{(u)} = \frac{-\zeta}{(2u + 1)(u + 1)} f_m^{(u+1)} + \frac{\zeta m}{u(u + 1)} f_m^{(u)} + \frac{\zeta(u + m)(u - m)}{(2u + 1)u} f_m^{(u-1)}, \tag{5.2}$$

$$P^+ f_m^{(u)} = \frac{-\zeta}{(2u + 1)(u + 1)} f_{m+1}^{(u+1)} - \frac{\zeta(u - m)}{u(u + 1)} f_{m+1}^{(u)} - \frac{\zeta(u + m)(u - m - 1)}{(2u + 1)u} f_{m+1}^{(u-1)}, \tag{5.3}$$

$$P^- f_m^{(u)} = \frac{\zeta}{(2u + 1)(u + 1)} f_{m-1}^{(u+1)} - \frac{\zeta(u + m)}{u(u + 1)} f_{m-1}^{(u)} + \frac{\zeta(u + m)(u + m - 1)}{(2u + 1)u} f_{m-1}^{(u-1)}, \tag{5.4}$$

$$\mathbf{P} \cdot \mathbf{P} f_m^{(u)} = 0, \quad \mathbf{P} \cdot \mathbf{J} f_m^{(u)} = -\zeta f_m^{(u)}. \tag{5.5}$$

[The representations $R_3'(\zeta, u_0)$ can be obtained formally from the representations $R_3(\omega, q, u_0)$ by setting $q = -\zeta/\omega$ and passing to the limit as $\omega \to 0$.]

$R_3'(\zeta, u_0)$ has no models in two complex variables. However, in three variables the type F operators[7] provide a model:

$$J^3 = t\frac{\partial}{\partial t}, \quad J^{\pm} = t^{\pm1}\left(z\frac{\partial}{\partial z} \pm t\frac{\partial}{\partial t} \mp \frac{z}{2}\right),$$
$$P^3 = 2\zeta z^{-1}, \quad P^{\pm} = \pm 2\zeta t^{\pm1} z^{-1}, \quad \zeta \in \mathbb{C}. \tag{5.6}$$

As is easy to verify, these operators satisfy the commutation relations of $\mathcal{C}_6$. Furthermore,

$$\mathbf{P} \cdot \mathbf{P} \equiv 0, \quad \mathbf{P} \cdot \mathbf{J} = -\zeta.$$

Corresponding to this model, the basis vectors are determined up to a multiplicative constant by relations (5.1)–(5.4) and may be given by

$$f_m^{(u)}(z, t) = \frac{(-1)^{-u}}{\Gamma(-2u)} M_{m,-u-\frac{1}{2}}(z) t^m = \frac{1}{\Gamma(-2u)} M_{-m,-u-\frac{1}{2}}(-z) t^m, \tag{5.7}$$

where the functions

$$M_{\chi,\mu}(z) = e^{-z/2} z^{\mu+1/2}\, {}_1F_1(\tfrac{1}{2} + \mu - \chi;\, 1 + 2\mu;\, z)$$

are Whittaker functions.[9] In fact, expressions (5.1) are equivalent to the recursion relations

$$\left(z\frac{d}{dz} \pm m \mp \frac{z}{2}\right) M_{m,\mu}(z) = (\mu + \tfrac{1}{2} \pm m) M_{m+1,\mu}(z), \tag{5.8}$$

while expressions (5.2)–(5.4) are equivalent to the relations

$$z^{-1} M_{m,\mu}(z) = M_{m,\mu-1}(z) + \frac{m}{2(\mu + \frac{1}{2})(\mu - \frac{1}{2})} M_{m,\mu}(z) + \frac{(m - \mu - \frac{1}{2})(m + \mu + \frac{1}{2})}{4\mu(\mu + \frac{1}{2})} M_{m,\mu+1}(z), \tag{5.9}$$

[9] W. Magnus, F. Oberhettinger, and R. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer-Verlag, Berlin, 1966), 3rd ed.

$$z^{-1}M_{m,\mu}(z) = M_{m+1,\mu-1}(z) + \frac{(m+\mu+\frac{1}{2})}{2(\mu+\frac{1}{2})(\mu-\frac{1}{2})} M_{m+1,\mu}(z) + \frac{(m+\mu+\frac{1}{2})(m+\mu+\frac{3}{2})}{4\mu(\mu+\frac{1}{2})} M_{m+1,\mu+1}(z), \tag{5.10}$$

$$z^{-1}M_{m,\mu}(z) = -M_{m-1,\mu-1}(z) + \frac{(m-\mu-\frac{1}{2})}{2(\mu+\frac{1}{2})(\mu-\frac{1}{2})} M_{m-1,\mu}(z) - \frac{(m-\mu-\frac{1}{2})(m-\mu-\frac{3}{2})}{4\mu(\mu+\frac{1}{2})} M_{m-1,\mu+1}(z). \tag{5.11}$$

We now prove some auxiliary lemmas which will enable us to extend the representation $R_3'(\zeta, u_0)$ of $\mathfrak{G}_6$ to a local representation of T_6. In the following, all operators and basis vectors are assumed to satisfy relations (5.1)–(5.5).

Lemma 1:

$$(P^3)^k f_u^{(u)} = (2\zeta)^k k!\,\Gamma(2u+1) \times \sum_{n=0}^{k} \frac{(-1)^n(2u+2n+1)}{n!(k-n)!\Gamma(2u+n+k+2)} f_u^{(u+n)}, \qquad k = 0, 1, 2, \cdots.$$

Proof: Expression (5.2) and induction on k.

Corollary 1:

$$e^{+z/2}(z)^{-u-k} = \sum_{n=0}^{k} \binom{k}{n} \frac{\Gamma(2u+2n+2)}{\Gamma(2u+n+k+2)} M_{u,-u-n-1/2}(z) \qquad k = 0, 1, 2, \cdots.$$

By definition,

$$M_{u,-u-k-1/2}(z) = e^{z/2}z^{-u-k}\,{}_1F_1(-k; -2u-2k; -z).$$

From this equation and Corollary 1, it is an easy computation to obtain the identity

$$z^r M_{u,-u-k-\frac{1}{2}}(z) = \sum_{n=0}^{k+s} \binom{k+s}{n} \frac{\Gamma(2u-2s+2n-2r+2)}{\Gamma(-s+2u-2r+n+k+2)} \times {}_3F_2(-k, -s-k+n, +s-2u-n-k+2r-1; -2u-2k, -s-k; 1) M_{u-s-r,-u+s+r-n-\frac{1}{2}}(z), \qquad k, s, \pm r = 0, 1, 2, \cdots, \tag{5.12}$$

expressing the function $z^r M_{u,-u-k-\frac{1}{2}}(z)$ as a linear combination of Whittaker functions $M_{u+s,-u-s-n-\frac{1}{2}}(z)$. [Compare this expansion with Eq. (3.2) of II.]

Let k be a nonnegative integer.

Lemma 2:

$$(P^3)^k f_m^{(u)} = \sum_{l=0}^{u-m+k} (2\zeta)^k \binom{u-m+k}{l} \times \frac{\Gamma(2m+2l+2)}{\Gamma(u+m+k+l+2)} \frac{\Gamma(-2m-2l)}{\Gamma(-2u)} \times {}_3F_2(m-u, -k+m-u+l, -u-m-k+l-1; -2u, -k+m-u; 1)(-1)^{u-m+l} f_m^{(m+l)}.$$

Lemma 3:

$$(P^+)^k f_m^{(u)} = \sum_{l=\max(u-m-2k,0)}^{u-m} (2\zeta)^k \binom{u-m}{l} \times \frac{\Gamma(2m+2k+2l+2)}{\Gamma(u+m+2k+l+2)} \times \frac{\Gamma(-2m-2k-2l)}{\Gamma(-u-m-l)} \frac{(2k)!}{(m-u+2k+l)!} \times (-1)^{u-m+k+l} f_{m+k}^{(m+k+l)}.$$

Lemma 4:

$$(P^-)^k f_m^{(u)} = \sum_{l=u-m}^{u-m+2k} \frac{\Gamma(2m-2k+2l+2)}{\Gamma(u+m+l+2)} \times \frac{\Gamma(-2m+2k-2l)}{\Gamma(-u-m)} \frac{\Gamma(-2m+2k-l)}{\Gamma(-u-m+2k-l)} \times \frac{(2k)!\,(-1)^{u-m-l}}{(u-m+2k-l)!(l-u+m)!} (2\zeta)^k f_{m-k}^{(m-k+l)}.$$

Proof: For our model these results can be obtained easily from expression (5.12). Since the lemmas are valid for the model, they must be true for the abstract representation $R_3'(u_0, \zeta)$.

According to Schafke[10] (Chap. 8), the functions (5.7) form an analytic basis for $R_3'(u_0, \zeta)$. Thus, this Lie-algebra representation can be extended to a local group representation of T_6. The matrix elements $\{v, n\,|\mathbf{w}, g|\,u, m\}$ can be defined by formulas analogous to (4.9) and satisfy the addition theorems (4.10). In particular, the matrix elements are completely determined by Lie algebra relations (5.1)–(5.4). We now compute the most important of these matrix elements.

Since Eqs. (5.1) and (2.1) are formally identical, the matrix elements $\{v, n|\,\mathbf{0}, g\,|u, m\}$ are given by (3.8).

[10] F. W. Schafke, *Einfuhrung in die Theorie der Speziellen Funktion der Mathematischen Physik* (Springer-Verlag, Berlin, 1963).

The elements of the form $\{v, n \mid \mathbf{w}, \mathbf{e} \mid u, m\}$ can be computed directly from Lemmas 2–4. In particular,

$$\begin{aligned}\{v, n|\, 0, 0, \gamma; \mathbf{e}\,|u, m\} &= \delta_{m,n} \sum_k \frac{(2\zeta\gamma)^k}{k!} \binom{u-m+k}{v-m} \\ &\times \frac{\Gamma(2v+2)\Gamma(-2v)}{\Gamma(u+v+k+2)\Gamma(-2u)} \\ &\times {}_3F_2(m-u, -k+v-u, -k+v-u \\ &\quad -2m-1; -2u, -k+m-u; 1)(-1)^{v-u}. \end{aligned} \tag{5.13}$$

If $m = u$, this simplifies to

$$\begin{aligned}\{v, n|\, 0, 0, \gamma; \mathbf{e}\,|u, u\} &= \delta_{u,n} \frac{(2v+1)\Gamma(2u+1)(-1)^{v-u}}{(v-u)!(2\zeta\gamma)^{u+1}} I_{2v+1}(\sqrt{8\gamma\zeta}) \\ &\qquad \text{if} \quad v-u \geq 0, \\ &= 0 \qquad \text{if} \quad u-v > 0. \end{aligned} \tag{5.14}$$

Here $I_\lambda(z)$ is a modified Bessel function.[9] (5.13) and (5.14) are entire functions of $\zeta\gamma$:

$$\begin{aligned}\{v, n|\, \alpha, 0, 0; \mathbf{e}\,|u, m\} &= \frac{(2\alpha\zeta)^{n-m}}{(n-m)!} \binom{u-m}{v-n} \\ &\times \frac{\Gamma(2v+2)\Gamma(-2v)(2n-2m)!\,(-1)^{v-u}}{\Gamma(u+v+n-m+2)} \\ &\quad \times \Gamma(-u-v+n-m)(v-u+n-m)! \\ &\qquad \text{if} \quad n-m \geq |u-v|, \\ &= 0 \quad \text{otherwise}; \end{aligned} \tag{5.15}$$

$$\begin{aligned}\{v, n|\, 0, \beta, 0; \mathbf{e}\,|u, m\} &= \frac{(-2\beta\zeta)^{m-n}}{(m-n)!} \frac{\Gamma(2v+2)}{\Gamma(u+v+m-n+2)} \\ &\times \frac{\Gamma(-2v)\Gamma(-v-n)}{\Gamma(-u-m)\Gamma(-u-v+m-n)} \\ &\times \frac{(2m-2n)!\,(-1)^{v-u}}{(u-v+m-n)!\,(v-u+m-n)!} \\ &\qquad \text{if} \quad m-n \geq |u-v|, \\ &= 0 \quad \text{otherwise}. \end{aligned} \tag{5.16}$$

The operators $\mathbf{T}(\mathbf{w}, g)$ defining the multiplier representation induced by the Lie derivatives (5.6) take the form

$$[\mathbf{T}(\alpha, \beta, \gamma; \mathbf{e})f](z, t) = \exp\left[\frac{2\zeta}{z}(\alpha t - \beta t^{-1} + \gamma)\right] \cdot f(z, t), \tag{5.17}$$

$$\begin{aligned}[\mathbf{T}(\mathbf{0}, g)f](z, t) &= \exp\left[\frac{z}{2}\left(\frac{bt}{d+bt} - \frac{c}{at+c}\right)\right] \\ &\times f\left(\frac{zt}{(at+c)(d+bt)}, \frac{at+c}{d+bt}\right), \\ &\qquad |c/at| < 1, \quad |bt/d| < 1, \\ g &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2). \end{aligned} \tag{5.18}$$

By construction, the basis functions of our model must satisfy the identity

$$[\mathbf{T}(\mathbf{w}, g)f_m^{(u)}](z, t) = \sum_{v,n} \{v, n|\, \mathbf{w}, g\,|u, m\} f_n^{(v)}(z, t). \tag{5.19}$$

We examine some special cases of this identity. If $\mathbf{w} = \mathbf{0}$, (5.19) simplifies to

$$\begin{aligned}&\exp\left[\frac{z}{2}\left(\frac{bt}{d+bt} - \frac{c}{at+c}\right)\right] \\ &\quad \times \left(1 + \frac{c}{at}\right)^m \left(1 + \frac{bt}{d}\right)^{-m} (1+bc)^{-m} \\ &\quad \times M_{m,-m-k-\frac{1}{2}}\left(\frac{zt}{(at+c)(d+bt)}\right) \\ &= \sum_{l=0}^{\infty} \frac{d^l a^k b^{k-l}}{l!}\, k!\, \frac{F(-l, -2m-k; k-l+1; bc/ad)}{\Gamma(k-l+1)} \\ &\quad \times M_{m+k-l,-m-k-\frac{1}{2}}(z) t^{k-l}, \\ &\qquad ad - bc = 1, \quad |c/at| < 1, \quad |bt/d| < 1. \end{aligned} \tag{5.20}$$

If $b = 0$, (5.20) reduces to

$$\begin{aligned}&e^{-(zc/2)/(1+c)} M_{m,-m-k-1/2}\left(\frac{z}{1+c}\right)(1+c)^m \\ &\qquad = \sum_{l=0}^{\infty} \binom{2m+k}{l} M_{m-l,-m-k-1/2}(z) c^l, \\ &\qquad\qquad |c| < 1, \quad k = 0, 1, 2, \cdots. \end{aligned} \tag{5.21}$$

In particular, if $k = 0$, this last expression yields the generating function

$$\begin{aligned}&z^{-m} \exp\left(\frac{z(1-c)}{2(1+c)}\right)(1+c)^{2m} \\ &\qquad = \sum_{l=0}^{\infty} \binom{2m}{l} M_{m-l,-m-\frac{1}{2}}(z) c^l. \end{aligned}$$

If $c = 0$, Eq. (5.20) reduces to

$$\begin{aligned}&\exp\left[\frac{zb}{2(1+b)}\right](1+b)^{-m} M_{m,-m-k-\frac{1}{2}}\left(\frac{z}{1+b}\right) \\ &\qquad = \sum_{l=0}^{k} \binom{k}{l} M_{m+l,-m-k-\frac{1}{2}}(z) b^l, \quad |b| < 1. \end{aligned} \tag{5.22}$$

Another interesting special case of (5.19) can be

obtained by setting $g = \mathbf{e}$, $\alpha = \beta = 0$, and $u = m$:

$$\exp\left(\frac{a}{z} + \frac{z}{2}\right) = \sum_{l=0}^{\infty} \frac{\Gamma(2u + 2l + 2)}{l!} a^{-u-\frac{1}{2}} I_{2u+2l+1}(2a^{\frac{1}{2}}) \times z^u M_{u,-u-l-\frac{1}{2}}(z). \quad (5.23)$$

6. WEISNER'S METHOD

In this section we will be concerned exclusively with the differential operators (3.11) which provide a realization of $\mathfrak{C}_6$ in three variables. So far, these operators have been used to construct identities for special functions which are simultaneous eigenfunctions of J^3, $C_{1,0}$, and $\mathbf{P} \cdot \mathbf{P}$. Furthermore, our identities have been valid only for group elements in a sufficiently small neighborhood of $\{\mathbf{0}, \mathbf{e}\}$. However, we can follow a method introduced by Weisner[6] and use the operators (3.11) to derive identities for special functions in which the above restrictions are lifted. We make the following observations. If $f(r, z, t)$ is a solution of the equation $\mathbf{P} \cdot \mathbf{P} f = -\omega^2 f$, i.e.,

$$\left[\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{t^2}{r^2}\frac{\partial^2}{\partial t^2} - \frac{2t}{r^2}\frac{\partial}{\partial t} + \frac{(1 - z^2)}{r^2}\frac{\partial^2}{\partial z^2} - \frac{2zt}{r^2}\frac{\partial^2}{\partial t \partial z} - \frac{2z}{r^2}\frac{\partial}{\partial z} + \omega^2\right] f(r, z, t) = 0, \quad (6.1)$$

then the function $\mathbf{T}(\mathbf{w}, g)f$, formally defined by Eqs. (3.14) and (3.15), is also a solution of (6.1). This remark is true whenever the formal expression for $\mathbf{T}(\mathbf{w}, g)f$ can be interpreted as an analytic function in (r, z, t) and is a consequence of the fact that the differential operators (3.11) commute with $\mathbf{P} \cdot \mathbf{P}$. In addition, if f is a solution of the equation

$$Lf = (x_1 J^+ + x_2 J^- + x_3 J^3 + y_1 P^+ + y_2 P^- + y_3 P^3) f = \lambda f$$

for complex constants x_i, y_i, λ, then $\mathbf{T}(\mathbf{w}, g)f = f'$ is a solution of an equation of the same form $L'f' = \lambda f'$ where $L' = \mathbf{T}(\mathbf{w}, g) L \mathbf{T}^{-1}(\mathbf{w}, g)$, i.e.,

$$\begin{aligned}
x_1' &= a^2 x_1 - b^2 x_2 + ab x_3, \\
x_2' &= -c^2 x_1 + d^2 x_2 - c\, d x_3, \\
x_3' &= 2ac x_1 - 2b\, d x_2 + (1 + 2bc) x_3, \\
y_1' &= a^2 y_1 - b^2 y_2 + ab y_3 \\
&\quad + \alpha[-ac x_1 + 2b\, d x_2 - (1 + 2bc) x_3] \\
&\quad + \gamma[a^2 x_1 - b^2 x_2 + ab x_3], \\
y_2' &= -c^2 y_1 + d^2 y_2 - c\, d y_3 \\
&\quad + \beta[2ac x_1 - 2b\, d x_2 + (1 + 2bc) x_3] \\
&\quad + \gamma[c^2 x_1 - d^2 x_2 + c\, d x_3], \\
y_3' &= 2ac y_1 - 2b\, d y_2 + (1 + 2bc) y_3 \\
&\quad + \alpha[-2c^2 x_1 + 2d^2 x_2 - 2c\, d x_3] \\
&\quad + \beta[-2a^2 x_1 + 2b^2 x_2 - 2ab x_3].
\end{aligned} \quad (6.2)$$

As an application of these remarks, consider a simultaneous eigenfunction of the commuting operators $\mathbf{P} \cdot \mathbf{P}$, P^3, J^3:

$$\mathbf{P} \cdot \mathbf{P} f = -f, P^3 f = \lambda f, J^3 f = m f, \lambda, m \in \mathbb{C}. \quad (6.3)$$

A straightforward computation shows that the solutions of (6.3) are of the form

$$f(r, z, t) = [t(z^2 - 1)^{\frac{1}{2}}(\lambda^2 - 1)^{\frac{1}{2}}]^m \times e^{\lambda r z} I_{\pm m}(r(z^2 - 1)^{\frac{1}{2}}(\lambda^2 - 1)^{\frac{1}{2}}).$$

Choosing the I_m solution, we note the validity of the expansion

$$f(r, z, t) = r^{-\frac{1}{2}} \sum_{k=0}^{\infty} a_k(\lambda) I_{m+k+\frac{1}{2}}(r) C_k^{m+\frac{1}{2}}(z) t^m, \quad (6.4)$$

giving f as a sum of simultaneous eigenfunctions of the operators $\mathbf{P} \cdot \mathbf{P}$, $C_{1,0}$, J^3.[10] It remains only to compute $a_k(\lambda)$. Since f is symmetric in z and λ, $a_k(\lambda) = b_k C_k^{m+\frac{1}{2}}(\lambda)$. Furthermore, if $\lambda = 1$, then

$$f(r, z, t) = \left(\frac{rt}{2}\right)^m \frac{e^{rz}}{\Gamma(m + 1)},$$

which has the well-known expansion

$$\left(\frac{r}{2}\right)^m e^{rz} = \sum_{k=0}^{\infty} \left(\frac{2}{r}\right)^{\frac{1}{2}} \Gamma(m + \tfrac{1}{2})(m + k + \tfrac{1}{2}) \times I_{m+k+\frac{1}{2}}(r) C_k^{m+\frac{1}{2}}(z)$$

(see I, Corollary 7, or Ref. 9). This last expansion enables us to compute the coefficients $a_k(\lambda)$ with the result

$$[(z^2 - 1)(\lambda^2 - 1)]^{-m/2} e^{r\lambda z} I_m(r[(z^2 - 1)(\lambda^2 - 1)]^{\frac{1}{2}}) = \frac{2^{2m+1}}{(2\pi r)^{\frac{1}{2}}} [\Gamma(m + \tfrac{1}{2})]^2 \sum_{k=0}^{\infty} \frac{k!\,(m + k + \frac{1}{2})}{\Gamma(2m + k + 1)} \times I_{m+k+\frac{1}{2}}(r) C_k^{m+\frac{1}{2}}(z) C_k^{m+\frac{1}{2}}(\lambda) \quad (6.5)$$

convergent for all z, $\lambda \in \mathbb{C}$.

Similarly, it is easy to show that

$$j(r, z, t) = \left(\frac{t}{\lambda(z^2 - 1)^{\frac{1}{2}}}\right)^m e^{\lambda z r} I_m[\lambda r(z^2 - 1)^{\frac{1}{2}}]$$

is a solution of the equations

$$\mathbf{P} \cdot \mathbf{P} j = 0, \quad P^3 j = \lambda j, \quad J^3 j = m j. \quad (6.6)$$

There exists an expansion of the form

$$j(r, z, t) = \sum_{k=0}^{\infty} a_k(\lambda) r^{m+k} C_k^{m+\frac{1}{2}}(z) t^m, \quad (6.7)$$

expressing j as a sum of simultaneous eigenfunctions of $\mathbf{P} \cdot \mathbf{P}$, $C_{1,0}$, and J^3. The constants $a_k(\lambda)$ can be evaluated by setting $z = 1$ on both sides of Eq. (6.7). The result is

$$\Gamma(m + 1)[r(z^2 - 1)^{\frac{1}{2}}]^{-m} e^{rz} I_m[r(z^2 - 1)^{\frac{1}{2}}] = \sum_{k=0}^{\infty} \frac{\Gamma(2m + 1)}{\Gamma(2m + k + 1)} C_k^{m+\frac{1}{2}}(z) r^k, \quad (6.8)$$

convergent for all r, $z \in \mathbb{C}$.[9]

As a final example we consider

$$h(r, z, t) = r^{-\frac{1}{2}} I_{u+\frac{1}{2}}(r) C_k^{u-k+\frac{1}{2}}(z) t^{u-k}, \quad u \in \mathcal{C},$$

$2u$ not an integer, $k = 0, 1, 2, \cdots$. h is a solution of the simultaneous equations

$$\mathbf{P} \cdot \mathbf{P}h = -h, \; C_{1,0}h = u(u+1)h, \; J^3 h = (u-k)h. \tag{6.9}$$

Note that the function

$$h' = \mathbf{T}(0, \tfrac{1}{2}, 0; \mathbf{e})h = \left(r^2 + \frac{(1-z^2)r}{t}\right)^{(-u+k-\frac{1}{2})/2} \times I_{u+\frac{1}{2}}\left[\left(r^2 + \frac{(1-z^2)r}{t}\right)^{\frac{1}{2}}\right] C_k^{u-k+\frac{1}{2}} \times \left[rz\left(r^2 + \frac{(1-z^2)r}{t}\right)^{-\frac{1}{2}}\right](1+rt)^{u-k}, \quad 0 < |t| < |r^{-1}|, \tag{6.10}$$

can be expanded in a Laurent series in t. Thus, the following expansion is valid:

$$h'(r, z, t) = \sum_{n=-\infty}^{\infty} \sum_{s=\max([-2n],0)}^{\infty} a_{n,s} s!\, \Gamma(n + \tfrac{1}{2}) \times r^{-\frac{1}{2}} I_{n+s+\frac{1}{2}}(r) C_s^{n+\frac{1}{2}}(z)(2t)^n. \tag{6.11}$$

We now determine the constants $a_{n,s}$. According to expression (6.2), $h' = \mathbf{T}(0, \frac{1}{2}, 0; \mathbf{e})h$ satisfies the equation $(J^3 + \frac{1}{2}P^-)h' = (u-k)h!$. This implies the recursion relation

$$2(u-k-n)a_{n,s} = \frac{-a_{n+1,s-2}}{(2n+2s-1)} + \frac{(2n+s+1)(2n+s+2)}{(2n+2s+3)} \times a_{n+1,s} \tag{6.12}$$

for the coefficients $a_{n,s}$. On the other hand, if $z = 1$, (6.11) reduces to a power series in t:

$$h'(r, 1, t) = \binom{2u-k}{k} r^{-u+k-\frac{1}{2}} I_{u+\frac{1}{2}}(r)(1+rt)^{u-k}, \quad |rt| < 1. \tag{6.13}$$

By comparing (6.13) with the well-known expansion

$$\left(\frac{r}{2}\right)^{-u+k} I_{u+\frac{1}{2}}(r) = \sum_{l=0}^{\infty} \frac{\Gamma(k+l+\frac{1}{2})\Gamma(-u+k+l)(k+2l+\frac{1}{2})}{l!\,\Gamma(-u+k)\Gamma(u+l+\frac{3}{2})} \times I_{k+2l+\frac{1}{2}}(r), \quad k = 0, 1, 2, \cdots,$$

[see I, Eq. (5.10), or Ref. 9, p. 129], we find

$$a_{n,s} = \frac{(2)^{-u+k+2n}}{(\pi)^{\frac{1}{2}}} \binom{2u-k}{k} \binom{-u+n+k+l-1}{l} \times \frac{\Gamma(u-k+1)\Gamma(n+k+l+\frac{1}{2})}{\Gamma(u-k-n+1)\Gamma(u+l+\frac{3}{2})} \times \frac{(n+k+2l+\frac{1}{2})}{(2n+k+2l)!}$$
$$\text{if } s = k+2l, \quad k, l = 0, 1, 2, \cdots,$$
$$= 0 \quad \text{if } s \neq k+2l, \quad k, l = 0, 1, 2, \cdots. \tag{6.14}$$

Thus, we have computed $a_{n,s}$ for $n \geq 0$. However, formulas (6.14) make sense for all integers n such that $2n + s \geq 0$, and they satisfy the recursion relations (6.12) even when n is negative. Therefore, formulas (6.14), defined for all integers n and nonnegative integers s such that $2n + s \geq 0$, are the solution to our problem:

$$\frac{k!\,\Gamma(u-k+\frac{1}{2})}{\Gamma(2u-k+1)}\left(r^2 + \frac{(1-z^2)r}{t}\right)^{(-u+k-\frac{1}{2})/2} I_{u+\frac{1}{2}}\left[\left(r^2 + \frac{(1-z^2)r}{t}\right)^{\frac{1}{2}}\right] C_k^{u-k+\frac{1}{2}}\left[rz\left(r^2 + \frac{(1-z^2)r}{t}\right)^{\frac{1}{2}}\right](1+rt)^{u-k}$$
$$= \sum_{n=-\infty}^{\infty} \sum_{l=\max([-n-k/2],0)}^{\infty} 8^{-u+k+n} \binom{-u+n+k+l-1}{s} \frac{\Gamma(n+k+l+\frac{1}{2})\Gamma(n+\frac{1}{2})(k+2l)!\,(n+k+2l+\frac{1}{2})r^{-\frac{1}{2}}}{(2n+k+2l)!\,\Gamma(u+l+\frac{3}{2})\Gamma(u-k-n+1)} \times I_{n+k+2l+\frac{1}{2}}(r) C_{k+2l}^{n+\frac{1}{2}}(z) t^n, \quad 0 < |t| < |r^{-1}|, \; u \in \mathcal{C}, \; 2u \text{ not an integer}, \; k = 0, 1, 2, \cdots.$$

The above examples should suffice to indicate the scope of Weisner's approach—though many other results could be obtained using the same method.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Natural Boundary and Initial Conditions from a Modification of Hamilton's Principle

H. F. TIERSTEN*
Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey

(Received 29 November 1967)

In a Hamiltonian variational formulation of a field theory, certain boundary conditions arise naturally whereas others arise as constraints on the admissible variations, and no initial conditions arise naturally. When the Hamiltonian formulation is used to obtain approximate solutions to boundary-value problems, the approximating functions need not satisfy the natural conditions but must satisfy the constraint conditions. Although, in many instances, it is desirable for the approximating functions to satisfy the constraint—and even the natural—conditions, in other instances it is imperative that the approximating functions do not satisfy certain constraint conditions. A procedure is introduced for transforming Hamilton's principle so that the initial conditions and all conditions at boundaries and internal surfaces of discontinuity arise naturally and no constraint conditions are required. The transformation is effected by modifying the principle slightly, using Lagrange multipliers in the classical manner, and adding an appropriate initial-value term to the Lagrangian. A particularly useful approximation technique is applied to a problem with an internal surface of discontinuity, and it is shown that the transformed principle can be used whereas the usual form of Hamilton's principle cannot. It is noted that the transformed principle has an important advantage over the method of least squares.

1. INTRODUCTION

The calculus of variations has proven to be a useful tool for obtaining approximate solutions to boundary-value problems for which exact solutions cannot be found. There are in existence a wide variety[1-3] of techniques for actually making the approximation, all of which lie within the variational formulation of the field theory. Classically, a variational formulation can usually be regarded as Hamilton's principle[4] for a particular field. There are many other variational formalisms, including, for example, the method of complementary energy[5] in linear elasticity and Reissner's variational formulation of linear elasticity.[6] However, this discussion will be confined to those variational formulations that can be regarded as Hamilton's principle in the classical sense. From this author's point of view, the crucial distinction between Hamilton's principle and those of others is that in Hamilton's principle the Lagrange density is expressed in terms of the minimum number of dependent field variables and their derivatives; whereas in the others, the Lagrange density is expressed in terms of a larger number of field variables. In Hamilton's principle, the variations of each of the field variables are independent (unconstrained) within the domain, i.e., excluding the boundary; and therefore, each variation yields an independent differential equation. Moreover, the variation of each of the field variables is constrained to vanish at t and t_0 throughout the domain and the boundary. Furthermore, if, on any portion of the boundary, a particular field variable is prescribed, its variation is constrained to vanish on that portion of the boundary for all time. If, in addition to the aforementioned boundary, the domain contains a surface of discontinuity, the variations of each of the field variables are unconstrained in the regions on either side of the discontinuity, but are constrained to be continuous across the surface. The removal of the constraints on the variations of the field variables on the portions of the boundary where they are prescribed and across an internal surface of discontinuity, for all time, and, at t and t_0, throughout the domain and the boundary, in Hamilton's principle is precisely the subject of this paper. The removal of these constraints on the variations of the field variables can be very important in obtaining approximate solutions to boundary- or initial-value problems because, if the constraints are imposed on the variations, the approximating functions *must* satisfy these constraints; but if no constraints are imposed on the variations, the approximating functions need not satisfy any constraints. Although in many instances it is desirable for the approximating functions to satisfy the constraints so that the approximation will be more accurate, in other instances it is imperative that the approximating functions do not

* Present address: Department of Mechanics, Rensselaer Polytechnic Institute, Troy, N.Y. 12181.

[1] L. V. Kantorovich and V. I. Krylov, *Approximate Methods of Higher Analysis* (Interscience Publ. Inc., New York, and P. Noordhoff Ltd., Groningen, The Netherlands, transl. by C. D. Benster from 3rd Russian ed., 1964), Chap. IV, Sec. 2 and 3.

[2] R. D. Mindlin, Quart. Appl. Math. **19**, 51 (1961).

[3] M. Onoe, J. Acoust. Soc. Am. **30**, 1159 (1958).

[4] C. Lanczos, *The Variational Principles of Mechanics* (University of Toronto Press, Toronto, 1949), Chap. V, Sec. 1.

[5] I. Sokolnikoff, *Mathematical Theory of Elasticity* (McGraw-Hill Book Company, New York, 1956), 2nd ed. Chap., 7, Sec. 108.

[6] E. Reissner, *On Variational Principles in Elasticity* (*Proceedings of the Eighth Symposium on Applied Mathematics, Calculus of Variations and Its Applications*) (McGraw-Hill Book Company, New York, 1958), pp. 1–6.

satisfy certain constraints required by the classical form of Hamilton's principle. A particularly important situation arising in connection with a boundary-value problem, in which an internal surface of discontinuity is present and in which functions constrained in the classical Hamiltonian manner *cannot* be used, is discussed in the last section of this paper. The problem treated is an eigenvalue problem in which the initial conditions may be left out of account. Although the variational formulation permits a similar approximate treatment of initial-value problems, no such problem is discussed. It is interesting to note that the removal of the constraints on the variations at t and t_0 applies in the case of particle mechanics also. In that case the modified Lagrangian yields the differential equations and the *initial conditions*.

The technique for removing the boundary constraints and the time constraint at t required by the Hamiltonian form of the variations is the classical one of adding to the Lagrangian, each constraint as a zero times a Lagrange multiplier, so that the variations of the field variables can be treated as free (unconstrained). The only difference between this situation and the usual one is that the constraints treated here by the Lagrangian technique are boundary constraints, while those usually treated by the Lagrangian technique are either domain or isoperimetric constraints. In this situation the Lagrangian multiplier is either a function on a surface for all time or a function over a domain at a fixed time, instead of a function over a domain, for all time, as it usually is in the absence of isoperimetric constraints. In any case, the Lagrangian multiplier can be varied freely.[7] The technique for removing the constraints at t_0 consists of simply introducing initial prescribed (inertial) terms in exactly the same way that prescribed boundary conditions are introduced. Additional variational terms are introduced at t_0 in order to account for the initial conditions on the field variables.

The resultant Lagrangian has been given for the scalar Helmholtz equation by Morse and Feshbach,[8] for the biharmonic equation by Hildebrand,[9] for the electromagnetic equations as applied to waveguides and resonators by Berk,[10] and for a special case in linear piezoelectricity by Eer Nisse[11] and Eer Nisse and Holland.[12] In each instance the modified Lagrangian is simply presented without a convincing explanation of why the new Lagrangian is permissible, i.e., without either the explicit use of Lagrange multipliers or any clear discussion. However, the demonstration of validity in the case of Refs. 8 and 9 is rather straightforward. Moreover, all the modified Lagrangians that have been presented are applicable only when field variables are *prescribed* on portions of the boundary and none are applicable to time-dependent systems. The additional modifications required for either the important case of an internal surface of discontinuity or for time-dependent systems have not been presented. In all the aforementioned cases,[8–12] specifically when there is no surface of discontinuity, and even when there is a surface of discontinuity, the proof of the validity of the Lagrangians presented can be provided without the use of Lagrange multipliers by using the appropriate version of what might well be called "Reissner's unconstrained variational technique." [6] Furthermore, for the scalar systems presented in Refs. 8 and 9, the proof can be provided without the use of either Lagrange multipliers or Reissner's technique, simply by employing the definition of the variation. This same procedure can be used to obtain a transformed version of Hamilton's principle, which yields initial conditions naturally, without the use of either Lagrange multipliers or Reissner's technique.

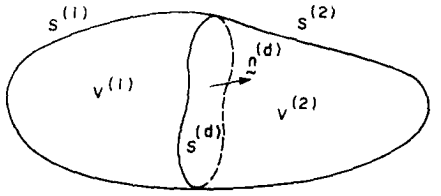


FIG. 1. Diagram of a bounded region containing an internal surface of discontinuity.

2. CLASSICAL FORM OF HAMILTON'S PRINCIPLE

In this section we briefly present Hamilton's principle so that we may clearly note the consequences when we transform it in the next section. Consider the diagram shown in Fig. 1. The diagram contains a region (possibly a body) which is divided in two by a surface of discontinuity $S^{(d)}$. The remaining boundary of region 1 is labeled $S^{(1)}$ and of region 2, $S^{(2)}$. Let the behavior of this body be governed by a vector[13] field

[7] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publ. Inc., New York, 1953), Vol. I, Chap. IV, Sec. 9.2.

[8] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, New York, 1953), part II, p. 1132.

[9] F. B. Hildebrand, *Methods of Applied Mathematics* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1965), 2nd ed., p. 219, Problem 100.

[10] A. D. Berk, Inst. Radio Engrs., Trans. Antennas Propagation **44**, 104 (1956).

[11] E. P. Eer Nisse, IEEE Trans. Sonics Ultrasonics **14**, 153 (1967).

[12] E. P. Eer Nisse and R. Holland, Proc. IEEE Letters **55**, 1524 (1967).

[13] Cartesian tensor notation is employed throughout; the summation convention for repeated tensor indices is employed as is the dot notation for differentiation with respect to time, and a comma followed by an index denotes differentiation with respect to a space coordinate.

$\varphi_k^{(m)}(x_j, t)$; $k, j = 1, 2, 3$; $m = 1, 2$. Let the Lagrange density in regions 1 and 2, respectively, be given by[14]

$$\mathfrak{L}^{(m)} = \tfrac{1}{2}\rho^{(m)}\dot{\varphi}_k^{(m)}\dot{\varphi}_k^{(m)} - \mathfrak{U}^{(m)}(\varphi_k^{(m)}, \varphi_{k,l}^{(m)}, x_j), \quad m = 1, 2. \tag{2.1}$$

In Eq. (2.1), $\mathfrak{L}^{(m)}$ has continuous first and second derivatives with respect to all its arguments in $V^{(m)}$, and $\varphi_k^{(m)}$ is twice continuously differentiable with respect to all independent variables in $V^{(m)}$ and attains prescribed values $\bar{\varphi}_k^{(m)}$ on portions of $S^{(m)}$, which we denote by $S_C^{(m)}$. On the remaining portions of the boundaries $S^{(m)}$, the $\varphi_k^{(m)}$ are unknown, but the quantities $n_l\partial\mathfrak{U}^{(m)}/\partial(\varphi_{k,l}^{(m)}) \equiv n_l G_{lk}^{(m)} \equiv F_k^{(m)}$ are prescribed and are given by

$$F_k^{(m)} \equiv n_l\partial\mathfrak{U}^{(m)}/\partial(\varphi_{k,l}^{(m)}) \equiv n_l G_{lk}^{(m)} = \bar{F}_k^{(m)}, \quad m = 1, 2, \tag{2.2}$$

in which the $\bar{F}_k^{(m)}$ are prescribed functions of the x_j and t, and n_l denotes the components of the outward-directed unit normal. The portions of the boundaries $S^{(m)}$ on which (2.2) holds are denoted by $S_N^{(m)}$. When the $\varphi_k^{(m)}$ are prescribed, we have what are termed constraint conditions, and when the $F_k^{(m)}$ are prescribed, natural conditions. At this point it is clear that Hamilton's principle may be written in the form

$$\delta\int_{t_0}^{t} dt \sum_{m=1}^{2}\left[\int_{V^{(m)}} \mathfrak{L}^{(m)}\, dV + \int_{S_N{}^{(m)}} \bar{F}_k^{(m)}\varphi_k^{(m)}\, dS\right] = 0, \tag{2.3}$$

where the $\delta\varphi_k^{(m)}$ are constrained to vanish everywhere at t and t_0 and for all time on $S_C^{(1)}$ and $S_C^{(2)}$, and to be continuous across $S^{(d)}$. Taking the variations,[15] utilizing the fact that the variation operation commutes with differentiation, integrating by parts with respect to time, and employing the divergence theorem and the constraints on the variations, we obtain

$$\begin{aligned}\int_{t_0}^{t} dt \sum_{m=1}^{2}\Bigg[\int_{V^{(m)}}&(-\partial\mathfrak{U}^{(m)}/\partial\varphi_k^{(m)} \\ &- \rho^{(m)}\ddot{\varphi}_k^{(m)} + G_{lk,l}^{(m)})\delta\varphi_k^{(m)}\, dV \\ &+ \int_{S_N{}^{(m)}}(-n_l^{(m)}G_{lk}^{(m)} + \bar{F}_k^{(m)})\delta\varphi_k^{(m)}\, dS\Bigg] \\ &- \int_{t_0}^{t} dt \int_{S^{(d)}} n_l^{(d)}(G_{lk}^{(1)} - G_{lk}^{(2)})\delta\varphi_k^{(1)}\, dS = 0,\end{aligned} \tag{2.4}$$

where $n_l^{(d)}$ denotes the components of the unit normal to the surface of discontinuity $S^{(d)}$ directed from $V^{(1)}$ to $V^{(2)}$. Since the volumetric and surface variations $\delta\varphi_k^{(m)}$ appearing in (2.4) are arbitrary, we have the Euler differential equations

$$-\partial\mathfrak{U}^{(m)}/\partial\varphi_k^{(m)} - \rho^{(m)}\ddot{\varphi}_k^{(m)} + G_{lk,l}^{(m)} = 0, \quad m = 1, 2, \tag{2.5}$$

the boundary conditions on $S_N^{(m)}$

$$-n_l^{(m)}G_{lk}^{(m)} + \bar{F}_k^{(m)} = 0, \quad m = 1, 2, \tag{2.6}$$

and the discontinuity (jump) condition across $S^{(d)}$

$$n_l^{(d)}(G_{lk}^{(1)} - G_{lk}^{(2)}) = 0. \tag{2.7}$$

In addition, by virtue of the constraints on the variations, we have the boundary conditions on $S_C^{(m)}$

$$\varphi_k^{(m)} - \bar{\varphi}_k^{(m)} = 0, \quad m = 1, 2, \tag{2.8}$$

and the continuity conditions on $S^{(d)}$

$$\varphi_k^{(1)} - \varphi_k^{(2)} = 0. \tag{2.9}$$

Thus, this Hamiltonian variational principle (2.3), constrained in the conventional manner, yields the differential equations (2.5), the boundary conditions (2.6) and (2.8), the jump conditions (2.7), and the continuity conditions (2.9) of the field in which we are interested, but does not yield all the initial conditions.

As usual, Eqs. (2.3) or (2.4) can be used to obtain approximate solutions to boundary-value problems, provided that initial conditions may be left out of account by a variety of techniques, such as the Rayleigh–Ritz procedure,[1] Kantorovich's procedure,[1] and many others. However, any approximating solution *must* satisfy the constraint conditions (2.8) and (2.9) in accordance with the variational principle; but the approximating solution need not satisfy the differential Eqs. (2.5) or the natural conditions (2.6) and (2.7). This point is discussed very thoroughly by Kantorovich and Krylov[16] and Collatz.[17] Clearly, the formulation is completely inadequate for obtaining an approximate solution to any problem in which initial conditions must be considered, since the approximating solutions *must* satisfy the constraints on the variations at t and t_0 and the values of the field variables at t are *unknown*. Moreover, there are no conditions on the *known* initial time derivatives of the field variables. These inadequacies of the classical form of Hamilton's principle have been noted by Gurtin[18] in a somewhat similar context. In connection with purely boundary-value problems, it should be noted that, although the requirement of satisfying (2.8) and (2.9) is frequently not restrictive and even desirable (for obtaining accuracy) in practice, at

[14] The extension to more than one field and to spaces of higher (different) dimensionality is evident, as is the extension to more than one surface of discontinuity.

[15] Reference 4, Chap. II, Secs. 8–11.

[16] Reference 1, pp. 258–260, 272–273, and 279–281.

[17] L. Collatz, *The Numerical Treatment of Differential Equations*, translated by P. G. Williams (Springer-Verlag, Berlin, 1960), 2nd ed., pp. 202–207 and 213–216.

[18] M. E. Gurtin, Arch. Ratl. Mech. Anal. **16**, 34 (1964).

times it is excessively restrictive and for certain types of procedures, which are discussed in Sec. 4, it is so stringent as to be prohibitive. In the next section we will transform the Hamiltonian variational principle, with the aid of Lagrange multipliers, in order that *all* conditions, including those at boundaries, at surfaces of discontinuity, and at the initial time, appear as natural conditions and there are no constraint conditions which an approximating solution must satisfy.

3. THE TRANSFORMED PRINCIPLE

In Sec. 2, the variations $\delta\varphi_k^{(m)}$ are constrained to vanish on those portions of the boundaries where the $\varphi_k^{(m)}$ are prescribed, to be continuous across the surface of discontinuity $S^{(d)}$, and to vanish everywhere at t. This situation is analogous to the usual one in the calculus of variations where holonomic[19] constraints are used to reduce the number of independent variations. Instead of eliminating a particular field variable for each holonomic constraint, it is common practice to add to the Lagrangian each constraint as a zero times a Lagrange multiplier,[7,20] and then to treat all variations as free. As stated in the Introduction, the constraint on the variations $\delta\varphi_k^{(m)}$ at t_0 can be eliminated by introducing initial prescribed terms in the Lagrangian in the same way that prescribed surface terms are introduced; additional variational terms must be introduced at t_0 in order to take account of the initial conditions on the field variables. Thus, in accordance with the above, we introduce Lagrangian undetermined multipliers $\lambda_k^{(m)}$, $\lambda_k^{(d)}$, and $\lambda_k^{(t)}$; then we dot $\lambda_k^{(m)}$ into (2.8) and integrate over $S_C^{(m)}$ and t; $\lambda_k^{(d)}$ into (2.9) and integrate over $S^{(d)}$ and t; and finally $\lambda_k^{(t)}$ into $\delta\varphi_k^{(m)}$ and integrate over $V^{(m)}$. And then we add the three integrals to the left-hand side of (2.3) while introducing the aforementioned terms at t_0 to obtain, in place of (2.3),

$$\begin{aligned}
&\delta\int_{t_0}^{t} dt \sum_{m=1}^{2}\Bigg[\int_{V^{(m)}} \mathfrak{L}^{(m)}\,dV + \int_{S_N^{(m)}} \bar{F}_k^{(m)}\varphi_k^{(m)}\,dS \\
&\quad + \int_{Sc^{(m)}} \lambda_k^{(m)}(\varphi_k^{(m)} - \bar{\varphi}_k^{(m)})\,dS\Bigg] \\
&\quad + \delta\int_{t_0}^{t} dt \int_{S^{(d)}} \lambda_k^{(d)}(\varphi_k^{(2)} - \varphi_k^{(1)})\,dS \\
&\quad + \sum_{m=1}^{2}\int_{V^{(m)}} [\lambda_k^{(t)}\delta q_k^{(m)}(t) + \bar{P}_k^{(m)}\delta q_k^{(m)}(t_0) \\
&\quad + (q_k^{(m)}(t_0) - \bar{q}_k^{(m)}(t_0))\rho^{(m)}\delta\dot{q}_k^{(m)}(t_0)]\,dV = 0, \quad (3.1)
\end{aligned}$$

where, now, the $\delta q_k^{(m)}$ are unconstrained on $S_C^{(m)}$ and $S^{(d)}$ as well as on $S_N^{(m)}$ and in $V^{(m)}$ at all times, and at t and t_0 everywhere, and the $\lambda_k^{(m)}$, $\lambda_k^{(d)}$, and $\lambda_k^{(t)}$ are to be varied freely,[7] and $\bar{P}_k^{(m)}(t_0)$ denotes the initial value of $\rho^{(m)}\dot{\varphi}_k^{(m)}$. In place of (2.4) we now have

$$\begin{aligned}
&\int_{t_0}^{t} dt \sum_{m=1}^{2}\Bigg[\int_{V^{(m)}} (-\partial\mathfrak{U}^{(m)}/\partial\varphi_k^{(m)} \\
&\quad - \rho^{(m)}\ddot{\varphi}_k^{(m)} + G_{lk,l}^{(m)})\delta\varphi_k^{(m)}\,dV \\
&\quad + \int_{S_N^{(m)}} (-n_l^{(m)}G_{lk}^{(m)} + \bar{F}_k^{(m)})\delta\varphi_k^{(m)}\,dS \\
&\quad + \int_{Sc^{(m)}} [(-n_l^{(m)}G_{lk}^{(m)} + \lambda_k^{(m)})\delta\varphi_k^{(m)} \\
&\quad + \delta\lambda_k^{(m)}(\varphi_k^{(m)} - \bar{\varphi}_k^{(m)})]\,dS\Bigg] \\
&\quad + \int_{t_0}^{t} dt \int_{S^{(d)}} [-(n_l^{(d)}G_{lk}^{(1)} + \lambda_k^{(d)})\delta\varphi_k^{(1)} \\
&\quad + (n_l^{(d)}G_{lk}^{(2)} + \lambda_k^{(d)})\delta\varphi_k^{(2)} + \delta\lambda_k^{(d)}(\varphi_k^{(2)} - \varphi_k^{(1)})]\,dS \\
&\quad + \sum_{m=1}^{2}\int_{V^{(m)}} [(\rho\dot{\varphi}_k^{(m)}(t) + \lambda_k^{(t)})\delta\varphi_k^{m}(t) \\
&\quad + (\bar{P}_k^{(m)} - \rho\dot{\varphi}_k^{(m)}(t_0))\delta\varphi_k^{(m)}(t_0) \\
&\quad + (\varphi_k^{(m)}(t_0) - \bar{\varphi}_k^{(m)}(t_0))\rho^{(m)}\delta\dot{\varphi}_k^{(m)}(t_0)]\,dV = 0. \quad (3.2)
\end{aligned}$$

Since the volumetric variations $\delta\varphi_k^{(m)}$ are arbitrary, we have (2.5) in $V^{(m)}$; and since all the surface variations $\delta q_k^{(m)}$, $\delta\lambda_k^{(m)}$, and $\delta\lambda_k^{(d)}$ are now free, we have the boundary conditions (2.6) on $S_N^{(m)}$, the conditions

$$-n_l^{(m)}G_{lk}^{(m)} + \lambda_k^{(m)} = 0, \quad (3.3)$$

$$\varphi_k^{(m)} - \bar{\varphi}_k^{(m)} = 0 \quad (3.4)$$

on $S_C^{(m)}$, and the conditions

$$n_l^{(d)}G_{lk}^{(1)} + \lambda_k^{(d)} = 0, \quad (3.5)$$

$$n_l^{(d)}G_{lk}^{(2)} + \lambda_k^{(d)} = 0, \quad (3.6)$$

$$\varphi_k^{(2)} - \varphi_k^{(1)} = 0 \quad (3.7)$$

across $S^{(d)}$. Since the variations $\delta q_k^{(m)}$ are arbitrary at t and t_0, we have

$$\lambda_k^{(t)} + \rho\dot{\varphi}_k^{(m)}(t) = 0, \quad (3.8)$$

$$\bar{P}_k^{(m)} - \rho^{(m)}\dot{\varphi}_k^{(m)}(t_0) = 0. \quad (3.9)$$

Since the $\delta\dot{q}_k^{(m)}(t_0)$ are independent of the $\delta\varphi_k^{(m)}(t_0)$, we have

$$\varphi_k^{(m)}(t_0) - \bar{\varphi}_k^{(m)}(t_0) = 0, \quad (3.10)$$

all in $V^{(m)}$. The subtraction of (3.6) from (3.5) yields (2.7). Equation (3.4) is identical with (2.8) and (3.7) with (2.9); but (3.4) and (3.7) appear as natural conditions, whereas (2.8) and (2.9) are constraint conditions. The Lagrange multipliers $\lambda_k^{(m)}$ may now be obtained from (3.3), $\lambda_k^{(d)}$ from either (3.5) or (3.6), and $\lambda_k^{(t)}$ from (3.8). Moreover, it is clear that the initial conditions may be read off from (3.9) and (3.10). Thus

[19] Reference 4, Chap. I, Sec. 6.
[20] Reference 4, Chap. II, Sec. 12.

it is clear that the variational principle (3.1) with unconstrained variations yields the Euler equations (2.5), the boundary conditions (2.6) and (2.8), and the conditions across $S^{(d)}$ (2.7) and (2.9), as does the variational principle (2.3) with constrained variations. It is also clear that the unconstrained variational principle (3.1) yields the initial conditions (3.9) and (3.10), whereas the constrained variational principle (2.3) does not.

As already mentioned, the unconstrained variational principle (3.1) has an advantage over the constrained variational principle (2.3) for obtaining approximate solutions to boundary-value problems, since, with (3.1), an approximating solution need satisfy no conditions. In order to find the most appropriate forms of (3.1) and (3.2) to be used in obtaining approximate solutions, add (3.5) and (3.6) to obtain

$$\lambda_k^{(d)} = -\tfrac{1}{2}n_l^{(d)}(G_{lk}^{(1)} + G_{lk}^{(2)}), \tag{3.11}$$

and substitute from (3.3), (3.8), and (3.11) into (3.1) and (3.2) to obtain, respectively,

$$\begin{aligned}
&\delta\int_{t_0}^{t} dt \sum_{m=1}^{2}\Big[\int_{V^{(m)}} \mathfrak{L}^{(m)}\, dV + \int_{S_N{}^{(m)}} \bar{F}_k^{(m)}\varphi_k^{(m)}\, dS \\
&\quad + \int_{S_C{}^{(m)}} n_l^{(m)}G_{lk}^{(m)}(\varphi_k^{(m)} - \bar{\varphi}_k^{(m)})\, dS\Big] \\
&\quad - \delta\int_{t_0}^{t} dt \int_{S^{(d)}} \tfrac{1}{2}n_l^{(d)}(G_{lk}^{(1)} + G_{lk}^{(2)})(\varphi_k^{(2)} - \varphi_k^{(1)})\, dS \\
&\quad + \sum_{m=1}^{2}\int_{V^{(m)}}[-\rho^{(m)}\dot{\varphi}_k^{(m)}\delta\varphi_k^{(m)}(t) + \bar{P}_k^{(m)}\delta\varphi_k^{(m)}(t_0) \\
&\quad + (\varphi_k^{(m)}(t_0) - \bar{\varphi}_k^{(m)}(t_0))\rho^{(m)}\delta\dot{\varphi}_k^{(m)}(t_0)]\, dV = 0, \qquad (3.12)
\end{aligned}$$

$$\begin{aligned}
&\int_{t_0}^{t} dt \sum_{m=1}^{2}\Big[\int_{V^{(m)}}(-\partial\mathfrak{L}^{(m)}/\partial\varphi_k^{(m)} - \rho^{(m)}\ddot{\varphi}_k^{(m)} \\
&\quad + G_{lk,l}^{(m)})\delta\varphi_k^{(m)}\, dV + \int_{S_N{}^{(m)}}(-n_l^{(m)}G_{lk}^{(m)} \\
&\quad + \bar{F}_k^{(m)})\delta\varphi_k^{(m)}\, dS + \int_{S_C{}^{(m)}} n_l^{(m)}(\varphi_k^{(m)} - \bar{\varphi}_k^{(m)})\delta G_{lk}^{(m)}\, dS\Big] \\
&\quad - \tfrac{1}{2}\int_{t_0}^{t} dt \int_{S^{(d)}} n_l^{(d)}[(G_{lk}^{(1)} - G_{lk}^{(2)})(\delta\varphi_k^{(1)} + \delta\varphi_k^{(2)}) \\
&\quad + (\varphi_k^{(2)} - \varphi_k^{(1)})(\delta G_{lk}^{(1)} + \delta G_{lk}^{(2)})]\, dS \\
&\quad + \sum_{m=1}^{2}\int_{V^{(m)}}[(\bar{P}_k^{(m)} - \rho\dot{\varphi}_k^{(m)}(t_0))\delta\varphi_k^{(m)}(t_0) \\
&\quad + (\varphi_k^{(m)}(t_0) - \bar{\varphi}_k^{(m)}(t_0))\rho^{(m)}\delta\dot{\varphi}_k^{(m)}(t_0)]\, dV = 0. \qquad (3.13)
\end{aligned}$$

The integrals over $S_C^{(m)}$ in (3.12) and (3.13) are precisely the forms presented in Refs. 8–12, but without the explicit use of Lagrange multipliers or any reference to Reissner's[6] technique or any meaningful discussion. The integrals over $S^{(d)}$ and over $V^{(m)}$ at t and t_0 in (3.12) and (3.13) do not appear in any of the references. Clearly, the expression (3.11), and consequently (3.12) and (3.13), is not unique, because of the difference of (3.5) and (3.6). However, the forms chosen are obviously the most straightforward and give equal weight to each adjacent region.

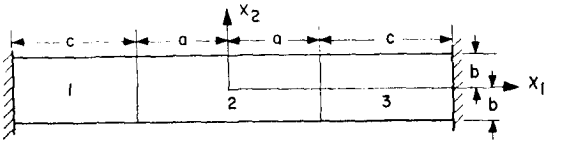


FIG. 2. Diagram of a symmetric bounded elastic plate with internal surfaces of discontinuity.

4. APPLICATION TO A BOUNDARY-VALUE PROBLEM

In this section we consider a technique of solution of a particular boundary-value problem in which the transformed principle in Sec. 3 can be used and the classical principle in Sec. 2 cannot. For clarity, simplicity, and definiteness, we confine ourselves to a prototype problem in isotropic, linear elasticity,[21] in which case the $\mathfrak{L}^{(m)}$ are of the form

$$\begin{aligned}
\mathfrak{L}^{(m)} = &-\tfrac{1}{2}[\lambda^{(m)}u_{i,i}^{(m)}u_{j,j}^{(m)} \\
&+ \mu^{(m)}(u_{i,j}^{(m)}u_{i,j}^{(m)} + u_{i,j}^{(m)}u_{j,i}^{(m)}) - \rho^{(m)}\dot{u}_j^{(m)}\dot{u}_j^{(m)}].
\end{aligned} \tag{4.1}$$

In Eq. (4.1), the $\varphi_k^{(m)}$ have been replaced by the $u_k^{(m)}$, which represent the components of mechanical displacement in region m, $\lambda^{(m)}$ and $\mu^{(m)}$ are the Lamè constants in region m, and $\rho^{(m)}$ is the mass density in region m. Consider the bounded plate shown in Fig. 2, in which region 1 is identical with region 3 while region 2 is different, and in which the length out of the paper is infinite and may be left out of account. The length out of the paper may also be infinitesimal, in which case the $\lambda^{(m)}$ in (4.1) must be converted to the plane stress value[22] $2\mu^{(m)}\lambda^{(m)}/(\lambda^{(m)} + 2\mu^{(m)})$. Let the upper and lower surfaces be traction free ($\bar{F}_k^{(1)} = \bar{F}_k^{(2)} = \bar{F}_k^{(3)} = 0$) and the left and right surfaces be displacement free ($\bar{u}_k^{(1)} = \bar{u}_k^{(3)} = 0$).

Suppose we are interested in determining the eigenfrequencies and mode shapes corresponding to the extensional solutions that are symmetric (and/or antisymmetric) about $x_1 = 0$ so that initial conditions may be left out of account. Although there are many procedures for obtaining an approximate solution to such a problem, a particularly interesting, useful, and

[21] A. E. H. Love, *A Treatise on the Mathematical Theory of Elasticity* (Cambridge University Press, Cambridge, England, 1927), 4th ed. (reprinted by Dover Publications, Inc., New York, 1944), Secs. 69 and 115.

[22] Reference 21, Sec. 146.

enlightening technique, suggested by Mindlin[23] and employed by Onoe[3] for the case of all traction-free boundaries and no internal surface of discontinuity, consists of expanding the solution of the bounded plate in an appropriate number of the exact solutions for the infinite plate with traction free surfaces[24] (Rayleigh modes) and satisfying the remaining conditions on the traction-free end edges approximately by means of what remains of (2.4). However, the technique of expanding in the Rayleigh modes cannot be used with (2.4) when the l and r edges are displacement free and/or there is an internal surface of discontinuity because the approximating functions *must* satisfy (2.8) and/or (2.9) in order to be used in (2.4); that is not possible when expanding in the Rayleigh modes. Nevertheless, the technique of expanding in the Rayleigh modes can be used with (3.13) when the l and r edges are displacement-free and/or there is an internal surface of discontinuity—or any other combination of conditions for that matter—because the approximating functions need not satisfy *any* conditions in order to be used in (3.13).

Explicitly, the Rayleigh solution for standing waves, which are symmetric about the center plane, in an infinite, isotropic plate may be written[25]

$$\begin{aligned} u_1 &= [\xi B \cos \alpha x_2 + \beta C \cos \beta x_2] \cos \xi x_1 \cos \omega t, \\ u_2 &= [-\alpha B \sin \alpha x_2 + \xi C \sin \beta x_2] \sin \xi x_1 \cos \omega t, \\ u_3 &= 0, \end{aligned} \tag{4.2}$$

so that

$$\begin{aligned} T_{11} &= -\mu[(\beta^2 + \xi^2 - 2\alpha^2)B \cos \alpha x_2 \\ &\qquad + 2\beta\xi C \cos \beta x_2] \sin \xi x_1 \cos \omega t, \\ T_{22} &= \mu[(\xi^2 - \beta^2)B \cos \alpha x_2 \\ &\qquad + 2\beta\xi C \cos \beta x_2] \sin \xi x_1 \cos \omega t, \\ T_{33} &= -\lambda(\alpha^2 + \xi^2)B \cos \alpha x_2 \sin \xi x_1 \cos \omega t, \\ T_{12} &= \mu[-2\xi\alpha B \sin \alpha x_2 \\ &\qquad + (\xi^2 - \beta^2)C \sin \beta x_2] \cos \xi x_1 \cos \omega t, \\ T_{23} &= T_{13} = 0, \end{aligned} \tag{4.3}$$

where the G_{ik} of Secs. 2 and 3 have been replaced by the T_{ik}, and where the differential equations are satisfied if

$$\alpha^2 = \frac{\omega^2}{V_1^2} - \xi^2, \quad \beta^2 = \frac{\omega^2}{V_2^2} - \xi^2, \tag{4.4}$$

and V_1 and V_2 are the velocities of dilatational and equivoluminal waves, respectively. The boundary conditions $T_{21} = T_{22} = T_{23} = 0$ on the upper and lower surfaces of the plate are satisfied if[26]

$$\tan \beta b / \tan \alpha b = -4\xi^2\alpha\beta/(\xi^2 - \beta^2)^2, \tag{4.5}$$

$$B = -2\xi\beta \cos \beta b, \quad C = (\xi^2 - \beta^2) \cos \alpha b. \tag{4.6}$$

Equation (4.5) is the Rayleigh frequency equation for waves that are symmetric about the center plane of the plate. Equations (4.4) and (4.5) determine dispersion curves $\omega = \omega(\xi)$, which are now well known.[27] For any given real, positive ω, there are a denumerably infinite number of branches—some are real, others imaginary, and still others complex. Nevertheless, in many frequency ranges one can conclude that a good approximation can be obtained by omitting all branches except a rather small number.[28] In fact, for the symmetric modes, below a certain frequency (edge mode) the number of branches that must be included is only one,[3,29] and, for the antisymmetric modes, below a certain frequency the number of branches that must be included is only two.[30] Although the actual number of branches required for obtaining an accurate solution is very important, it is not particularly important in this paper because we are interested only in giving an example of the use of (3.13) under circumstances in which (2.4) cannot be used. Consequently, we simply consider the situation where only the one lowest, real, symmetric branch is needed. The extension to more than one branch is trivial in principle, although cumbersome in practice.

Before proceeding further, it is important to note that a solution, symmetric about the center plane of the plate and linearly independent of the one in (4.2) and (4.3), can be obtained from that one simply by interchanging $\sin \xi x_1$ and $\cos \xi x_1$ and, in the quantities in square brackets, replacing ξ by $-\xi$. Equations (4.4) and (4.5) remain unchanged and ξ should be replaced by $-\xi$ in (4.6). Since we are interested only in solutions that are symmetric about $x_1 = 0$, this latter solution will be the only one needed in region 2, but both linearly independent solutions associated with the one branch will be required in region 1. Thus, it is clear that, for each additional real or imaginary branch that must be included, two linearly independent solutions occur when symmetry with respect

[23] R. D. Mindlin, *Investigations in the Mathematical Theory of Vibrations of Anisotropic Bodies*, CU-4-56-SC-64687-CE, Final Report, U.S. Army Signal Corps Engineering Lab., Fort Monmouth, N.J. (1956) (unpublished).

[24] Lord Rayleigh (J. W. Strutt), Proc. London Math. Soc. **20**, 225 (1889).

[25] R. D. Mindlin, *An Introduction to the Mathematical Theory of Vibrations of Elastic Plates*, DA-36-039-SC-56772, U.S. Army Signal Corps Engineering Lab., Fort Monmouth, N.J. (1955) (unpublished), Sec. 2.06.

[26] Ref. 25, Eqs. (2.115) and (2.116).

[27] R. D. Mindlin, *Proceedings of the First Symposium on Naval Structural Mechanics* (Pergamon Press, Inc., New York, 1960), Sec. 12, pp. 214–219.

[28] Reference 25, Chaps. 4 and 5; Ref. 27, Secs. 16 and 17.

[29] D. C. Gazis and R. D. Mindlin, J. Appl. Mech. **27**, 541 (1960).

[30] H. Deresiewicz and R. D. Mindlin, J. Appl. Mech. **22**, 86 (1955).

to x_1 does not exist and only one when symmetry exists. When the branches are complex, each linearly independent solution is composed of a pair of complex conjugate branches so that the resultant solution is real.[29]

At this stage it should be noted that, since the differential equations and boundary conditions on the upper and lower surfaces of the plate are satisfied in each region and the solution is symmetric (or antisymmetric) about $x_1 = 0$ and the initial conditions may be left out of account, all that remains of Eq. (3.13) is

$$+\left[\int_{-b}^{b} u_k^{(1)}\delta T_{1k}^{(1)}\,dx_2\right]_{x_1=a+c} - \frac{1}{2}\left[\int_{-b}^{b}[(T_{1k}^{(1)} - T_{1k}^{(2)})(\delta u_k^{(1)} + \delta u_k^{(2)}) + (u_k^{(2)} - u_k^{(1)})(\delta T_{1k}^{(1)} + \delta T_{1k}^{(2)})]\,dx_2\right]_{x_1=a} = 0, \tag{4.7}$$

where the first integral is evaluated at the edge $x_1 = a + c$ and the second at the internal surface of discontinuity $x_1 = a$. It should also be noted that those terms that still remain in (3.13) and appear in (4.7) are precisely the ones that do not exist in (2.4). The first integral in Eq. (4.7), which is associated with a displacement-free edge, is the type of term that appears in Refs. 8–12, but the second integral in Eq. (4.7), which is associated with an internal surface of discontinuity, does not appear in any of those references.

The solution for the bounded plate may now be written:

$$\begin{aligned}
u_1^{(1)} &= (L_1^{(1)}A^{(1)}\cos\xi^{(1)}(x_1 - a) + M_1^{(1)}D^{(1)}\sin\xi^{(1)}(x_1 - a))\cos\omega t,\\
u_2^{(1)} &= (L_2^{(1)}A^{(1)}\sin\xi^{(1)}(x_1 - a) + M_2^{(1)}D^{(1)}\cos\xi^{(1)}(x_1 - a))\cos\omega t,\\
u_1^{(2)} &= M_1^{(2)}D^{(2)}\sin\xi^{(2)}x_1\cos\omega t,\\
u_2^{(2)} &= M_2^{(2)}D^{(2)}\cos\xi^{(2)}x_2\cos\omega t,\\
u_3^{(1)} &= u_3^{(2)} = 0,
\end{aligned} \tag{4.8}$$

where the $L_k^{(1)}$ are the terms appearing in square brackets in (4.2) times the shear modulus in region 1, $\mu^{(1)}$, and the $M_k^{(m)}$ are the equivalent terms that would appear in place of the $L_k^{(m)}$ in the aforementioned linearly independent standing-wave solution for the infinite plate. The $T_{1k}^{(m)}$ may readily be obtained from (4.8) in the same way that (4.3) were obtained from (4.2), i.e., by means of

$$T_{ij}^{(m)} = \lambda^{(m)}u_{k,k}^{(m)}\delta_{ij} + \mu^{(m)}(u_{i,j}^{(m)} + u_{j,i}^{(m)}). \tag{4.9}$$

Thus, at this stage, given an ω, $\xi^{(m)}$ can be determined from the appropriate dispersion curve, which was calculated[27] from (4.4) and (4.5), after which the $\alpha^{(m)}$ and $\beta^{(m)}$ may be found from (4.4), and then the $B^{(m)}$ and $C^{(m)}$ from (4.6), so that the $L_k^{(1)}$ and $M_k^{(m)}$ are known. In the usual way, all quantities are now substituted in (4.7), in which the independent variations are $\delta A^{(1)}$, $\delta D^{(1)}$, and $\delta D^{(2)}$, and the integrations through the thickness $2b$ are performed at the appropriate values of x_1. Since the variations $\delta A^{(1)}$, $\delta D^{(1)}$, and $\delta D^{(2)}$ are arbitrary, the coefficient of each must vanish, thereby yielding three homogeneous, linear, algebraic equations in the three constants $A^{(1)}$, $D^{(1)}$, and $D^{(2)}$. The vanishing of the determinant of the coefficients of $A^{(1)}$, $D^{(1)}$, and $D^{(2)}$ in the three equations yields the geometric ratios for the selected eigenfrequency. There is no point in carrying the solution any further, since we have no intention of presenting any results because we are interested only in giving a definite example of a specific situation in which (3.13) can be used and (2.4) cannot. Since this has been done, we proceed no further.

In closing, it should be noted that the method of least squares[31] could have been applied to the situation presented in this section in place of (3.13). However, when the method of least squares is applied in the case of a surface of discontinuity and in many other cases,[31] there exists a dimensional problem in that the different conditions are not naturally dimensionally compatible. All the procedures for making them dimensionally compatible introduce a degree of arbitrariness into the approximation. This undesirable feature is automatically avoided by using the transformed variational principle (3.13) in place of the method of least squares in treating such a problem.

ACKNOWLEDGMENTS

This investigation was motivated by a preprint of a paper by Dr. E. P. Eer Nisse[11] and subsequent discussions with Dr. D. Amos, both of Sandia Laboratory. The author wishes to thank Dr. R. Holland, also of Sandia Laboratory, for carefully reading and commenting on the manuscript.

[31] A. Sommerfeld, *Partial Differential Equations in Physics*, translated by E. G. Straus (Academic Press Inc., New York, 1949), Sec. 6C.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Quantum-Ergodic Problem

GARRISON SPOSITO
Department of Physics, Sonoma State College, Rohnert Park, California

(Received 2 October 1967)

The Gibbs approach to quantum dynamics is justified for nonrelativistic systems. It is shown that infinite-time and microcanonical ensemble averages may be equated wherever the former exists and whenever a maximal set of constants of the motion can be determined, provided that the ensemble average is over states whose images under the maximal set are virtually stationary. The relevance of the demonstration for quantum-statistical mechanics is discussed in the light of the ensemble formalism and the classical ergodic problem.

INTRODUCTION

It is generally accepted that the Gibbs approach to Hamiltonian dynamics has been justified to a large degree by the fundamental theorems of von Neumann[1] and Birkhoff.[2] These propositions, when taken in a dynamical sense, assert the existence (in the mean of order two or almost everywhere, respectively) of the infinite-time average of any property of a physical system expressible as a Lebesgue-integrable function on phase space. Moreover, if the infinite-time average is independent of almost every initial state (point in phase space) of the system, it may be equated with a Gibbs microcanonical ensemble average. These statements admit of an interesting but unfortunate reciprocity as regards their significance in mathematics and physics: The first is difficult to prove but, in an operational sense, is simple to verify experimentally, while the second is easy to show but hard to apply. The epigram is unfortunate because it is the second statement which is of greater relevance to dynamics. It was with this problem in mind that Lewis[3] has shown recently that the infinite-time average need not be independent of nearly every initial state in order that it may be put equal to a microcanonical average, provided that the meaning of the latter is generalized somewhat. To be specific, if a maximal set of constants of the motion is known, in that every infinite-time average can be expressed as a Borel-measurable function of it, then the time average may be equated with a microcanonical average, wherein the average is taken over those points in phase space whose images under the maximal set (not just the Hamiltonian) are stationary. This represents a natural extension of the microcanonical ensemble average and permits the difficult task of ascertaining the independence of the time average from initial states to be replaced by the possibly less arduous problem of determining a maximal set of constants of the motion.

The picture is not quite so clear as this if one looks into the vindication of Gibbs's approach as extended to quantum dynamics. von Neumann's argument,[4] for example, has an unsavory probabilistic air about it since it invokes the random phase assumption. More recently, Klein[5] and Emch[6] have dealt with the problem rigorously by pointing out that von Neumann's *classical* mean-ergodic theorem applies to any separable Hilbert space and so may be adapted to that comprising the dynamical observables for a physical system. However, their arguments are limited in the same way as is von Neumann's theorem and, therefore, cannot produce a general physical criterion for equating infinite-time and ensemble averages whenever constants of the motion exist which are not stationary on every eigensubspace of the energy operator.

This paper represents an attempt to remedy the present situation. A precise formulation of the nonrelativistic quantum-ergodic problem is developed which employs, in contradistinction to the work of Klein and Emch, the notion of *state* in Lebesgue space, rather than *observable*, as primitive. In this way it is not difficult to illustrate the significance of the formalism for quantum-statistical mechanics and to outline its relationship with the ergodic problem in Hamiltonian dynamics. The point of view taken here thus differs in an essential way from that of the recent, very interesting attempts to formulate quantum-statistical mechanics as a problem in the theory of Banach algebras.[7] In the C^*-algebra approach, one discusses only functionals on sets of observables for an

[1] J. von Neumann, Proc. Natl. Acad. Sci. U.S. **18**, 70 (1932).
[2] G D. Birkhoff, Proc. Natl. Acad. Sci. U.S. **17**, 650 and 656 (1931).
[3] R. M. Lewis, Arch. Ratl. Mech. Anal. **5**, 355 (1960).
[4] J. von Neumann, Z. Physik **57**, 30 (1929).
[5] M. J. Klein, Phys. Rev. **87**, 111 (1952).
[6] G. G. Emch, in *Lectures in Theoretical Physics* (University of Colorado Press, Boulder 1966), Vol. VIII-A, p. 65. See also G. G. Emch and C. Favre, "Coarse Graining in Liouville Space and Ergodicity," preprint, University of Geneva (1965).
[7] See, for example, D. Kastler and D. W. Robinson, Commun. Math. Phys. **3**, 151 (1966); and R. Haag, N. M. Hugenholtz, and M. Winnik, Commun. Math. Phys. **5**, 215 (1967).

infinite system, thereby obviating the existence of recurrences. On the other hand, it has become apparent[7,8] that this approach has the bad side effect of invalidating the use of compact density operators. It will be observed that this phenomenon causes no difficulty in what follows, when the limit of infinite volume is considered.

THE QUANTUM-ERGODIC THEOREM

The ergodic problem is seen in the clearest light when discussed in the language of measure theory.[9] The physical motivation for casting the argument into this form should become clear when the formalism is brought to bear upon the subject of quantum ensembles in the next section.

Lemma: Let $\mathcal{S}$ be the product space generated by the empty square $\varnothing \times \varnothing$ and the class of all rectangles in $\mathfrak{L}_2 \times \mathfrak{L}_2$. Let $\mathcal{M}$ be the class of all subsets of $\mathcal{S}$. Then $(\mathcal{S}, \mathcal{M})$ is a measurable space.

Proof: If $\mathbf{L}$ is the class consisting of $\varnothing$ (the empty set) and all the subsets of $\mathfrak{L}_2$, then $\mathbf{L}$ is certainly closed under complementation and the taking of countable unions. Thus $\mathbf{L}$ is a σ algebra. It follows by definition that $(\mathfrak{L}_2', \mathbf{L})$, where $\mathfrak{L}_2'$ is $\mathfrak{L}_2$ adjoined to $\varnothing$, is a measurable space. Moreover, if $\mathcal{M} = \mathbf{L} \times \mathbf{L}$, then $(\mathfrak{L}_2' \times \mathfrak{L}_2', \mathbf{L} \times \mathbf{L}) = (\mathcal{S}, \mathcal{M})$ is a measurable space because there exist for every $\{\psi, \varphi\} \in \mathfrak{L}_2' \times \mathfrak{L}_2'$ the sets $\mathbf{A}$ and $\mathbf{B}$ such that $\psi \in \mathbf{A} \in \mathbf{L}$ and $\varphi \in \mathbf{B} \in \mathbf{L}$—which means $\{\psi, \varphi\} \in \mathbf{A} \times \mathbf{B} \in \mathbf{L} \times \mathbf{L}$.

Definition: The elements of the nonempty ordered pair $\{\psi, \varphi\} \in \mathcal{S}$ are *equivalent* ($\psi \sim \phi$ in symbols) if $\|\psi\| = \|\varphi\| = |(\psi, \varphi)| = 1$. An equivalence class in $\mathcal{S}$ is therefore a square whose side is a unit ray.

Lemma: Let μ be a function from the elements of $\mathcal{M}$ into the set of extended nonnegative integers, such that

$$\mu(\mathbf{M} \times \mathbf{N}) = \begin{cases} \text{cardinality of the collection of linearly} \\ \quad \text{independent equivalence classes in} \\ \quad \mathbf{M} \times \mathbf{N} \text{ if this collection is finite} \\ +\infty \text{ if the collection is infinite} \end{cases}$$

for any rectangle $\mathbf{M} \times \mathbf{N} \in \mathcal{M}$. Then $(\mathcal{S}, \mathcal{M}, \mu)$ is a measure space.

Proof: Evidently, $\mu(\mathbf{M} \times \mathbf{N}) \geq 0$ for any $\mathbf{M} \times \mathbf{N} \in \mathcal{M}$ and $\mu(\varnothing \times \varnothing) = 0$. If $\{\mathbf{M}_k\}$ is a sequence of disjoint rectangles, each of which contains at most a countably infinite number of equivalence classes, then the rule for adding integers requires that

$$\mu\left(\bigcup_{k=1}^{\infty} M_k\right) = \sum_{k=1}^{\infty} \mu(M_k).$$

If one of the rectangles in the sequence contains an uncountable number of equivalence classes, then the above equality is perforce satisfied. It follows by definition that μ is a measure function on all of $\mathcal{S}$ and, by the first lemma, that $(\mathcal{S}, \mathcal{M}, \mu)$ is a measure space.

Lemma: Let $\{\mathbf{T}_t\}$ be a single-parameter group of one-to-one, unitary transformations of $\mathfrak{L}_2 \times \mathfrak{L}_2$ onto itself which are continuous in $\mathbf{R}^1$. Then $(\mathcal{S}, \mathcal{M}, \mu, \{\mathbf{T}_t\})$ is a measure-preserving space.

Proof: For any $\mathbf{M} \in \mathcal{M}$ and fixed $\mathbf{t} \in \mathbf{R}^1$:

(a) $\mathbf{T}_t^{-1}$ exists and $\mathbf{T}_t^{-1} M \in \mathcal{M}$ because $\{\mathbf{T}_t\}$ is a group and $\mathbf{T}_t$ maps $\mathfrak{L}_2 \times \mathfrak{L}_2$ onto itself.

(b) $\mu(\mathbf{T}_t^{-1}\mathbf{M}) = \mu(\mathbf{M})$ because $\mathbf{T}_t$ preserves the scalar product on $\mathfrak{L}_2 \times \mathfrak{L}_2$.

For any $\mathbf{M} \in \mathcal{M}$ and any $\mathbf{t}$ in the Borel sets of $\mathbf{R}^1$, $\{\mathbf{T}_t\}$ is a measurable transformation from $(\mathbf{R}^1 \times \mathcal{S}, \mathcal{B} \times \mathcal{M}, \nu_{\mathfrak{B}} \times \mu)$ into $(\mathcal{S}, \mathcal{M}, \mu)$, where $\mathcal{B}$ is the class of Borel sets in $\mathbf{R}^1$ and $\nu_{\mathfrak{B}}$ is Borel measure. This is so because $\{\mathbf{T}_t\}$ is a continuous transformation and because of statement (a) above.

Since $\mathbf{T}_t$ is measurable and $\mu(\mathbf{T}_t^{-1}\mathbf{M}) = \mu(\mathbf{M})$, $(\mathcal{S}, \mathcal{M}, \mu)$ is by definition a measure-preserving space.

Definition: A *Hilbert function* is a composite, elementary, complex-valued function on the orthonormal elements of $\mathcal{S}$. It is expressed by

$$(f_t(\psi) \circ H)(\varphi) = \sum_{\{k,k'\}} f_{kk'}^t \chi(M_k \times M_{k'}), \tag{1}$$

where $\mathbf{t} \in \mathbf{R}^1$,

$$\chi(M_k \times M_{k'}) = \begin{cases} 1 & \{\psi, \varphi\} \in M_k \times M_{k'}, \\ 0 & \text{otherwise}, \end{cases}$$

$$M_k \times M_{k'} = \{\{\psi, \varphi\} \ni (\psi, H\varphi) = (\psi_{tn}^{jk}, H\psi_{tn'}^{j'k'}) = f_{kk'}^t\}, \tag{2}$$

and the sum is over all $\{\mathbf{k}, \mathbf{k}'\}$ corresponding to distinct $f_{kk'}^t$. Here $\mathbf{H}$ is a bounded, self-adjoint, linear operator mapping $\mathfrak{L}_2$ into itself and ψ_{tn}^j is any one of the countably many orthonormal vectors in $\mathfrak{L}_2$ which are solutions of the Schrödinger equation

$$i\hbar \frac{\partial \psi_{tn}^j}{\partial t} = \mathfrak{H}\psi_{tn}^j \quad (j = 1, \cdots, m_n; n = 1, \cdots, \infty). \tag{3}$$

In Eq. (2) it is understood that unless $(\mathbf{f}_t(\psi) \circ \mathbf{H})(\varphi)$ is identically zero, $\psi_{tn}^j \sim \psi \in \mathbf{M}_k$, $\psi_{tn'}^{j'} \sim \varphi \in M_{k'}$, and the $\mathbf{M}_k$ are not necessarily disjoint sets of orthonormal eigenvectors. Moreover, $\bigcup_{k=1}^{\infty} \mathbf{M}_k$ is equivalent to the set of all orthonormal solutions of Eq. (3).

[8] The problem has been discussed carefully by G. G. Emch, J. Math. Phys. 7, 1413 (1966).

[9] Those unacquainted with the basic notions of measure theory will find them presented very clearly in P. R. Halmos, *Measure Theory* (D. Van Nostrand Co., Inc., Princeton, N.J., 1950).

Theorem: The infinite-time average

$$(\bar{f}_0(\psi) \circ H)(\varphi) = \lim_{T\to\infty} \frac{1}{T} \int_0^T U_t(f_0(\psi) \circ H)(\varphi)\, dt \quad (4)$$

exists almost everywhere on $(\mathcal{S}, \mathcal{M}, \mu)$ and is invariant under the group of continuous transformations $\{\mathbf{U}_t\}$ generated by $\mathcal{H}$ in Eq. (3). Moreover, if $\mathbf{S} \subset \mathcal{S}$ is a subset of finite μ measure, then

$$\int_S (\bar{f}_0(\psi) \circ H)(\varphi)\, d\mu = \int_S (f_0(\psi) \circ H)(\varphi)\, d\mu \quad (5)$$

almost everywhere (a.e.).

Proof: It is easy to deduce from Eq. (3) that, if $\mathcal{H}$ is time-independent,

$$\psi^j_{tn} = \exp(-i\mathcal{H}t/\hbar)\psi^j_n \equiv U_t\psi^j_n,$$

where $\hbar$ is Dirac's constant and ψ^j_n is one of the orthonormal solutions of the eigenvalue problem

$$\mathcal{H}\psi^j_n = E_n\psi^j_n \quad (j = 1, \cdots, m_n;\ n = 1, \cdots, \infty),$$

$\mathbf{E}_n$ being a real number. Thus

$$U_t(f_0(\psi) \circ H)(\varphi) = (f_0(U_t\psi) \circ H)(U_t\varphi) = (f_t(\psi) \circ H)(\varphi) \quad (6)$$

and, because $\{\mathbf{U}_t\}$ is measure-preserving, each value (at fixed $\{\psi, \varphi\}$) of the time-dependent Hilbert function is of the form

$$(\psi^j_n, H\psi^{j'}_{n'}) \exp[-i(E_{n'} - E_n)t/\hbar],$$

which is strictly periodic in $\mathbf{t}$. The measure $\int \mathbf{dt}$ may be suitably completed so that the integral in Eq. (4) is a Lebesgue integral. In the present case the integral is evaluated easily, with the result

$$(\bar{f}_0(\psi) \circ H)(\varphi) = \sum_{\{k,k'\}} \bar{f}^0_{kk'}\chi(\bar{M}_k \times \bar{M}_{k'}),$$

where

$$\bar{M}_k \times \bar{M}_{k'} = \{\{\psi, \varphi\} \ni (\psi, H\varphi) = (\psi^{jk}_n, H\psi^{j'k'}_n) = \bar{f}^0_{kk'}\}. \quad (7)$$

Equation (6) makes sense only if $\{\psi, \varphi\}$ is a member of an orthonormal rectangle; the ordered pairs which are not have μ measure zero. It should also be noted that

$$M_k \times M_{k'} \supset \bar{M}_k \times \bar{M}_{k'}.$$

But

$$\mu(M_k \times M_{k'}) = \mu(\bar{M}_k \times \bar{M}_{k'}). \quad (8)$$

The time-averaged Hilbert function is invariant under $\{\mathbf{U}_t\}$ because of Eq. (6) and because the ψ^j_n are eigenvectors of $\mathcal{H}$.

Now suppose $\mathbf{S}$ is a subset of $\mathcal{S}$ such that $\{\mathbf{U}_t\}$ is measure-preserving and $\mu(\mathbf{S}) < \infty$. Then

$$\int_S (\bar{f}_0(\psi) \circ H)(\varphi)\, d\mu \equiv \sum_{\{k,k'\}} \bar{f}^0_{kk'}\mu(S \cap \bar{M}_k \times \bar{M}_{k'})$$

and

$$\int_S (f_0(\psi) \circ H)(\varphi)\, d\mu \equiv \sum_{\{k,k'\}} f^0_{kk'}\mu(S \cap M_k \times M_{k'}).$$

Because μ assigns zero measure to every argument of $(f_0(\psi) \circ \mathbf{H})(\varphi)$ that is not an equivalent-ordered pair, and because of Eqs. (7) and (8) and the finiteness of $\mu(\mathbf{S})$,

$$\int_S (\bar{f}_0(\psi) \circ H)(\varphi)\, d\mu = \int_S (f_0(\psi) \circ H)(\varphi)\, d\mu, \quad (5)$$

except on sets of μ measure zero.

QUANTUM-STATISTICAL MECHANICS

The principal result of the foregoing section, insofar as equilibrium quantum-statistical mechanics is concerned, is Eq. (5). Its application to the Gibbs microcanonical ensemble becomes evident if, following Lewis,[3] a *maximal set of invariant Hilbert functions* is defined.

$$\mathcal{C} \equiv (\mathcal{C}(\psi) \circ 0_{\mathcal{C}})(\varphi) = \{C_1, C_2, \cdots, C_N\}$$

is a maximal set of invariant Hilbert functions if each of the $\mathbf{C}_j \equiv (\mathbf{C}_j(\psi) \circ \mathbf{0}_{C_j})\ (\varphi)(j = 1, \cdots, \mathbf{N})$ is invariant under $\{\mathbf{U}_t\}$, as defined in the previous section, and if every measurable-invariant Hilbert function $(\mathbf{g}(\psi) \circ \mathbf{0}_g)(\varphi)$ is a measurable function of $(\mathcal{C}(\psi) \circ \mathbf{0}_{\mathcal{C}})(\varphi)$ almost everywhere on $(\mathcal{S}, \mathcal{M}, \mu)$:

$$(g(\psi) \circ 0_g)(\varphi) = G[(\mathcal{C}(\psi) \circ 0_{\mathcal{C}})(\varphi)] \quad \text{a.e.}$$

If $(\mathcal{S}, \mathcal{M}, \mu, \{\mathbf{U}_t\})$ is a measure-preserving space and $\mathcal{C}$ is a maximal set of invariant Hilbert functions on $(\mathcal{S}, \mathcal{M}, \mu)$, then $(\mathcal{S}, \mathcal{M}, \mu, \{\mathbf{U}_t\}, \mathcal{C})$ is called a *complete space.*

Now suppose that a maximal set of invariant Hilbert functions has been determined for some physical system. Then, for all infinite-time averages,

$$(\bar{f}_0(\psi) \circ H)(\varphi) = F[(\mathcal{C}(\psi) \circ 0_{\mathcal{C}})(\varphi)] \quad \text{a.e.}, \quad (9)$$

where $\mathbf{H}$ is a dynamical property of the system. If some region of the "state space" $(\mathcal{S}, \mathcal{M}, \mu)$ is measure-preserving,

$$\int_S F[(\mathcal{C}(\psi) \circ 0_{\mathcal{C}})(\varphi)]\, d\mu = \int_S (f_0(\psi) \circ H)(\varphi)\, d\mu, \quad (10)$$

according to Eqs. (5) and (9). In particular, let

$$S = \{\{\psi, \varphi\} \ni |(\mathcal{C}(\psi) \circ 0_{\mathcal{C}})(\varphi) - \mathcal{K}| \leq \delta\mathcal{C}\}, \quad (11)$$

where $\mathcal{K} \in \mathbf{R}^N$,

$$|(\mathcal{C}(\psi) \circ 0_{\mathcal{C}})(\varphi) - \mathcal{K}| = \{|(C_j(\psi) \circ 0_{C_j})(\varphi) - K_j|\},$$

$\mathbf{K}_j \in \mathbf{R}^1$, and $\delta\mathcal{C} = \{\delta\mathbf{C}_j\}$ is a very small interval in $\mathbf{R}^N$ whose (Borel) measure is determined by uncertainty relations. Because $(\mathcal{C}(\psi) \circ \mathbf{0}_{\mathcal{C}})(\varphi)$ is essentially stationary on $\mathbf{S}$,

$$F[(\mathcal{C}(\psi) \circ 0_{\mathcal{C}})(\varphi)] = \text{const} \equiv \bar{H}(\mathcal{C}^{-1}(\mathcal{K})). \quad (12)$$

It follows from Eqs. (9)–(12) that

$$\bar{H}(S) = \int_S (f_0(\psi) \circ H)(\varphi)\, d\mu/\mu(S),$$

or, in the notation of the previous section,

$$\bar{H}(\mathcal{C}^{-1}(\mathcal{K})) = \frac{\sum_{\{k,k'\}} f^0_{kk'}\mu(S \cap M_k \times M_{k'})}{\mu(\{\psi, \varphi\} \ni |(\mathcal{C}(\psi) \circ 0_{\mathcal{C}})(\varphi) - \mathcal{K}| \leq \delta\mathcal{C})}. \quad (13)$$

Equation (13) is a generalization of the Gibbs microcanonical average.

The foregoing analysis provides a justification of the Gibbs approach to quantum dynamics in the following sense. If a maximal set of constants of the motion is known for a chosen system, and if a measure-preserving set in the "state-space" $(\mathcal{S}, \mathcal{M}, \mu)$ can be found upon which the maximal set is very nearly stationary in value, then the infinite-time average of any property of the system is equal to its (generalized) microcanonical average over the determined set. It is to be noted that

$$\frac{\mu(S \cap M_k \times M_{k'})}{\mu(\{\psi, \varphi\} \ni |(\mathcal{C}(\psi) \circ 0_{\mathcal{C}})(\varphi) - \mathcal{K}| \leq \delta\mathcal{C})}$$

is the "statistical weight" of the contribution of $f^0_{kk'}$ to the microcanonical average of $\mathbf{H}$. The measure μ itself is analogous to the Liouville measure (Lebesgue measure on phase space) encountered in the Gibbs–Hamilton dynamics in that, if $\mathbf{L}$ is a subspace of $\mathfrak{L}'_2$, $\mu(\mathbf{L} \times \mathbf{L})$ is the dimension of $\mathbf{L}$.

Generalized canonical and grand canonical distributions may be derived by considering physical systems in *weak interaction*, in the sense of Lewis.[3] The generalization of these distributions amounts to replacing the conventional canonical partition function by

$$Q(\beta) = \sum_{\{k,k'\}} [\exp(-\beta\mathcal{C})]_{kk'}\mu(S \cap P_k \times P_{k'}),$$

where $\beta = \{\beta_1, \cdots, \beta_N\}$ is an arbitrary vector in $\mathbf{R}^N$,

$$\beta\mathcal{C} = \sum_{j=1}^{N} \beta_j C_j,$$

$\mathbf{S}$ is a measure-preserving subspace of $\mathcal{S}$, and

$$P_k \times P_{k'} = \{\{\psi, \varphi\} \ni (\psi, \exp(-\beta\mathcal{C})\varphi) = [\exp(-\beta\mathcal{C})]_{kk'}\}.$$

The arguments leading to this result are quite analogous to those put forth for the classical case in Lewis' paper, which may be consulted for details.

THE CLASSICAL LIMIT

In general, there is no classical limit of Eq. (13). A classical limit will exist only if $\mathbf{H}$ and the maximal set $(\mathcal{C}(\psi) \circ \mathbf{0}_{\mathcal{C}})(\varphi)$ possess classical analogs. If this condition is met, the facts that $\mathbf{S}$ is a product of invariant subspaces and that unitary transformations on $\mathbf{S}$ are measure-preserving may be used to rewrite Eq. (13) in the "expanded" form

$$\sum_{\{k,k'\}} f^0_{kk'}\mu(S \cap M_k \times M_{k'}) = \sum_{\{\mathbf{P}\}} (\varnothing_{\mathbf{P}}, H\varnothing_{\mathbf{P}}), \quad (14)$$

where $\varnothing_P$ is an eigenvector (in $\mathfrak{L}_2$) of the kinetic energy and $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_N\} \equiv \{\mathbf{p}_k\}$, $\mathbf{p}_i$ being the momentum of the ith of $\mathbf{N}$ particles making up the system. [In writing down Eq. (14), the invariance property of the trace under changes of basis is invoked.] The expression may be simplified a little by noting that a sum over each of the $\mathbf{P}$ is the same as a sum over the $\mathbf{p}_i$ individually, provided the latter sum is corrected by dividing it through with $\mathbf{N}!$—the number of ways of permuting the $\mathbf{p}_i$ in a given $\mathbf{P}$. Thus

$$\sum_{\{k,k'\}} f^0_{kk'}\mu(S \cap M_k \times M_{k'}) = \frac{1}{N!} \sum_{\{\mathbf{p}_k\}} (\varnothing_{\mathbf{P}}, H\varnothing_{\mathbf{P}}). \quad (15)$$

The classical limit is obtained by writing

$$\sum_{\mathbf{p}_i} \to \frac{V}{h^3} \int d\mathbf{p}_i$$

(in the sense of Lebesgue) into Eq. (15) and by replacing $\varnothing_P$ by its finite-volume, unsymmetrized counterpart:

$$\sum_{\{k,k'\}} f^0_{kk'}\mu(S \cap M_k \times M_{k'}) \to \frac{(2\pi\hbar)^{-3N}}{N!} \int \exp\left[-i\sum_{j=1}^{N} \mathbf{p}_j \cdot \mathbf{r}_j/\hbar\right] H \times \exp\left[i\sum_{j=1}^{N} \mathbf{p}_j \cdot \mathbf{r}_j/\hbar\right] d\nu, \quad (16)$$

where ν is the Liouville measure and the domain of (Lebesgue) integration is understood to be the set of all points in $\mathbf{R}^{2N}$ whose images under the maximal set of *classical* constants of the motion are, in a limiting sense, stationary. Equation (16) is brought into the desired form by making use of the following proposition. Let $\mathbf{0}$ be observable and possess a classical analog $\mathbf{0}_C$. Then

$$0_C = \exp\left[-i\sum_{j=1}^{N} \mathbf{p}_j \cdot \mathbf{r}_j/\hbar\right] 0 \exp\left[i\sum_{j=1}^{N} \mathbf{p}_j \cdot \mathbf{r}_j/\hbar\right], \quad (17)$$

where $\mathbf{N}$ is the number of particles in the system described by $\mathbf{0}$. Equation (17) is an immediate consequence of the fact that observables with classical analogs depend only upon fixed parameters, the coordinates, and derivatives with respect to the coordinates. Equation (17), when applied to Eq. (16), yields

$$\sum_{\{k,k'\}} f^0_{kk'}\mu(S \cap M_k \times M_{k'}) \to \frac{(2\pi\hbar)^{-3N}}{N!}\int H_C\, dv.$$

Once it has been noted that

$$\mu(S) = \sum_{m,m'\in M_S}(\psi_m, \psi_{m'}) = \sum_{m,m'\in M_S}(\psi_m, I\psi_{m'}),$$

$$M_S = \{\{m,m'\} \ni \{\psi_m, \psi_{m'}\} \in S;\ \psi_m \sim \psi_{m'} \sim \psi^j_n\},$$

where $\mathbf{I}$ is the identity operator, the denominator in Eq. (13) can be transformed according to the argument just given. Equation (13) then becomes, in the classical limit,

$$\bar{H}(\mathcal{C}^{-1}(\mathcal{K})) \to \lim_{\delta\mathcal{C}_C\to 0} \frac{\int H_C\, dv}{\nu(\{\mathbf{p},\mathbf{r}\} \ni |\mathcal{C}_C\{\mathbf{p},\mathbf{r}\} - \mathcal{K}_C| \le \delta\mathcal{C}_C)}, \tag{18}$$

where $\mathcal{C}_C\{\mathbf{p},\mathbf{r}\}$ is the classical analog of $(\mathcal{C}(\psi)\circ\mathbf{0}_{\mathcal{C}})(\varphi)$. The quantity on the right-hand side of Eq. (18) is Lewis'[3] generalization of the Gibbs microcanonical average.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Iterative Solutions by Means of Trial Operators

RUBIN GOLDSTEIN
Brookhaven National Laboratory, Upton, New York

(Received 28 September 1967)

The problem of finding a weighted average of an unknown solution to an inhomogeneous equation is examined. An analytic approximation technique is developed in terms of an iterative series involving a trial operator. By choosing the operator so that successive terms in the series vanish, one obtains a solution which has characteristics similar to variational solutions to the problem. The iterative approach has the added features of giving error estimates, the sign of the error, a testing ground for the quality of classes of trial operators or functions, and a possible means of determining upper and lower bounds to the exact result. Several examples are given for both self-adjoint and non-self-adjoint systems. It is shown that the trial operator approach can give useful analytic approximations, with results which may be superior to variational calculations.

Consider the class of problems characterized by the equation

$$H\psi = s, \tag{1}$$

where H and s are a given operator and a given source, respectively, and ψ is the unknown solution. Assuming that Eq. (1) is not amenable to exact solution, one must resort to approximate methods. A conventional approximation technique to equation solving is by iteration. Equation (1) may be put into a form which is convenient for iteration, by formally dividing H into two parts:

$$H = H_0 + H_1. \tag{2}$$

The hope is that the operator H_0 is a good representation of H. That is to say, H_0 is assumed to be the major part of H, while H_1 is treated as a perturbing operator. Using Eq. (2) and assuming that H_0 is nonsingular, Eq. (1) may be rewritten as

$$\psi = H_0^{-1}s - H_0^{-1}H_1\psi, \tag{3}$$

which is now in a form that is convenient for iteration.

A first approximation to ψ may be obtained by neglecting the term in Eq. (3) which involves the perturbing operator H_1:

$$\psi^{(1)} = H_0^{-1}s. \tag{4}$$

Iterating once on this first-order solution, a second-order approximation is obtained,

$$\psi^{(2)} = H_0^{-1}s - H_0^{-1}H_1H_0^{-1}s, \tag{5}$$

and this process may be continued.

Suppose one is interested in a weighted average of the unknown solution

$$I = (w, \psi), \tag{6}$$

where w is an arbitrary weight function and the scalar product (f, g) denotes integration of the two functions f and g over the entire domain of interest. A first approximation to I, the integral of interest, may be obtained from Eqs. (4) and (6):

$$I^{(1)} = (w, H_0^{-1}s). \tag{7}$$

where $\mathbf{N}$ is the number of particles in the system described by $\mathbf{0}$. Equation (17) is an immediate consequence of the fact that observables with classical analogs depend only upon fixed parameters, the coordinates, and derivatives with respect to the coordinates. Equation (17), when applied to Eq. (16), yields

$$\sum_{\{k,k'\}} f^0_{kk'}\mu(S \cap M_k \times M_{k'}) \to \frac{(2\pi\hbar)^{-3N}}{N!}\int H_C\, dv.$$

Once it has been noted that

$$\mu(S) = \sum_{m,m'\in M_S}(\psi_m, \psi_{m'}) = \sum_{m,m'\in M_S}(\psi_m, I\psi_{m'}),$$

$$M_S = \{\{m,m'\} \ni \{\psi_m, \psi_{m'}\} \in S;\ \psi_m \sim \psi_{m'} \sim \psi^j_n\},$$

where $\mathbf{I}$ is the identity operator, the denominator in Eq. (13) can be transformed according to the argument just given. Equation (13) then becomes, in the classical limit,

$$\bar{H}(\mathcal{C}^{-1}(\mathcal{K})) \to \lim_{\delta\mathcal{C}_C\to 0} \frac{\int H_C\, dv}{\nu(\{\mathbf{p},\mathbf{r}\} \ni |\mathcal{C}_C\{\mathbf{p},\mathbf{r}\} - \mathcal{K}_C| \le \delta\mathcal{C}_C)}, \tag{18}$$

where $\mathcal{C}_C\{\mathbf{p},\mathbf{r}\}$ is the classical analog of $(\mathcal{C}(\psi)\circ\mathbf{0}_{\mathcal{C}})(\varphi)$. The quantity on the right-hand side of Eq. (18) is Lewis'[3] generalization of the Gibbs microcanonical average.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Iterative Solutions by Means of Trial Operators

RUBIN GOLDSTEIN
Brookhaven National Laboratory, Upton, New York

(Received 28 September 1967)

The problem of finding a weighted average of an unknown solution to an inhomogeneous equation is examined. An analytic approximation technique is developed in terms of an iterative series involving a trial operator. By choosing the operator so that successive terms in the series vanish, one obtains a solution which has characteristics similar to variational solutions to the problem. The iterative approach has the added features of giving error estimates, the sign of the error, a testing ground for the quality of classes of trial operators or functions, and a possible means of determining upper and lower bounds to the exact result. Several examples are given for both self-adjoint and non-self-adjoint systems. It is shown that the trial operator approach can give useful analytic approximations, with results which may be superior to variational calculations.

Consider the class of problems characterized by the equation

$$H\psi = s, \tag{1}$$

where H and s are a given operator and a given source, respectively, and ψ is the unknown solution. Assuming that Eq. (1) is not amenable to exact solution, one must resort to approximate methods. A conventional approximation technique to equation solving is by iteration. Equation (1) may be put into a form which is convenient for iteration, by formally dividing H into two parts:

$$H = H_0 + H_1. \tag{2}$$

The hope is that the operator H_0 is a good representation of H. That is to say, H_0 is assumed to be the major part of H, while H_1 is treated as a perturbing operator. Using Eq. (2) and assuming that H_0 is nonsingular, Eq. (1) may be rewritten as

$$\psi = H_0^{-1}s - H_0^{-1}H_1\psi, \tag{3}$$

which is now in a form that is convenient for iteration.

A first approximation to ψ may be obtained by neglecting the term in Eq. (3) which involves the perturbing operator H_1:

$$\psi^{(1)} = H_0^{-1}s. \tag{4}$$

Iterating once on this first-order solution, a second-order approximation is obtained,

$$\psi^{(2)} = H_0^{-1}s - H_0^{-1}H_1H_0^{-1}s, \tag{5}$$

and this process may be continued.

Suppose one is interested in a weighted average of the unknown solution

$$I = (w, \psi), \tag{6}$$

where w is an arbitrary weight function and the scalar product (f, g) denotes integration of the two functions f and g over the entire domain of interest. A first approximation to I, the integral of interest, may be obtained from Eqs. (4) and (6):

$$I^{(1)} = (w, H_0^{-1}s). \tag{7}$$

To second order, from Eqs. (5) and (6) we have

$$I^{(2)} = (w, H_0^{-1}s) - (w, H_0^{-1}H_1H_0^{-1}s). \tag{8}$$

The general result is

$$I^{(n+1)} = I^{(n)} + (-1)^n R_n, \qquad I^{(0)} = 0, \tag{9}$$

where

$$R_n = (w, [H_0^{-1}H_1]^n H_0^{-1}s), \tag{10}$$

for $n = 0, 1, 2, \cdots$.

Note that the "remainder terms" R_n are of nth order in the perturbation $H_0^{-1}H_1$. Assuming that the iteration procedure given by Eq. (9) converges, one may write the formal solution for the integral I as the Neumann series

$$I = \sum_{n=0}^{\infty}(-1)^n R_n. \tag{11}$$

(Note that, by definition, $R_0 = I^{(1)}$.) How well the iteration procedure (9) or the series solution (11) converges, will depend, essentially, upon how well H_0 represents H.

The remaining problem then is the determination of H_0. Clearly, one good criterion for finding the "best" H_0, would be to reduce the perturbation H_1 as much as possible (in some sense); for example, one might require that the norm $\|H_0^{-1}H_1\| \ll 1$. This choice of H_0 and H_1 should certainly give a rapidly convergent solution. Finding this best H_0 is, of course, another problem.

Suppose, however, one is able to make only a limited number of iterations due to practicality. This may be due to time or economic considerations, or perhaps the desire for an analytic solution causes this limitation. We propose to treat H_0 as a trial operator, determining it by a scheme which is especially useful when one is limited to a small number of iterations. Furthermore, we give here a specific procedure for the determination of H_0.

Suppose we formally choose H_0 such that $R_1 = 0$. It follows from Eqs. (9–11) that

$$I^{(2)} = I^{(1)} \tag{12}$$

and

$$I = I^{(1)} + R_2 - R_3 + \cdots.$$

This choice of H_0 thus eliminates terms of first order in H_1, and the quantity of interest I is calculated accurate to second order. In this sense, this procedure should give results which are equivalent to a variational formulation of the problem which renders a functional stationary to first order. Furthermore, if we require that

$$R_1 = R_2 = \cdots = R_n = 0, \tag{13}$$

then, at least formally, the integral I will be calculated accurate to order $(n + 1)$. This should be equivalent to higher-order variational principles.

A general expression for the remainders R_n, which involves only H_0 and known quantities, may be obtained from Eq. (10). This is

$$R_n = (w, [H_0^{-1}H]^n H_0^{-1}s) - (w, H_0^{-1}s) - \sum_{m=1}^{n-1} \frac{n!}{m!(n-m)!} R_m. \tag{14}$$

Equation (14) gives a direct relation between H_0 and the remainders R_n. If we require that all the lower-order remainders vanish ($R_m = 0$, for $m < n$), then the nth-order remainder is given by

$$R_n = (w, [H_0^{-1}H]^n H_0^{-1}s) - (w, H_0^{-1}s), \tag{15}$$

for $n = 1, 2, 3, \cdots$.

We still require a concrete method for finding an operator H_0 which satisfies the set of conditions (13). One approach is to introduce free parameters into the trial operator and then fix the parameters by requiring that the conditions (13) be satisfied. If one wishes, the trial operator may be determined in terms of a trial function ψ_t. For example, let ψ_t be the solution to the fundamental equation (1), with H replaced by H_0:

$$H_0\psi_t = s \quad \text{or} \quad H_0^{-1}s = \psi_t. \tag{16}$$

Equation (16) does not yet determine H_0 uniquely. How specific one must be about H_0 depends upon the type of calculation one is performing. In some cases, for example, if we require that H_0 be self-adjoint, then only the combination $H_0^{-1}s$ may appear. In these cases it is sufficient that H_0 satisfy Eq. (16) and nothing additional is required. In other cases, an explicit expression for H_0^{-1} is needed, and then a particularly simple choice, for example, is to choose H_0^{-1} as a multiplication operator. That is to say,

$$H_0^{-1} = \psi_t/s. \tag{17}$$

By the use of trial functions with free parameters one also has an indirect means of introducing parameters into the trial operator:

$$H_0^{-1}(\kappa\lambda\mu\cdots)s = \psi(\kappa\lambda\mu\cdots). \tag{18}$$

If we have an n-parameter trial operator, it may be fixed by requiring Eq. (13) to hold, and then I is estimated to order $(n + 1)$.

We will now apply to some general examples the formal theory which has been developed. At the same time we will compare the iterative approach involving trial operators given here with variational formulations of the problems. Consider first a self-adjoint

operator H. The functional

$$F(\psi_t) = 2(s, \psi_t) - (\psi_t, H\psi_t) \tag{19}$$

has the property of being stationary about the solution ψ to Eq. (1), with the stationary value of the functional being $I = (s, \psi)$. If we use as the trial function

$$\psi_t = Af(x), \tag{20}$$

where A is a free arbitrary constant and $f(x)$ is a given fixed function, then the functional (19) becomes (note that everything is assumed to be real here)

$$F(A) = 2A(s,f) - A^2(f, Hf).$$

The amplitude A which renders this functional stationary is that satisfying the condition $\partial F/\partial A = 0$. Putting this value of A back into the functional in order to obtain the stationary value, as the variational estimate to I we obtain

$$F = (s,f)^2/(f, Hf). \tag{21}$$

The iterative solution to the problem is obtained by requiring that $R_1 = 0$. In this example, the weight function w is equal to s. Putting this into Eq. (15) for $n = 1$, we have

$$R_1 = (s, H_0^{-1}H\ H_0^{-1}s) - (s, H_0^{-1}s). \tag{22}$$

The trial operator may be chosen as

$$H_0^{-1}s = Af. \tag{23}$$

This choice is consistent with Eq. (20) for comparison purposes. For this problem, the only requirement additional to (23) that need be made on H_0 is that it be self-adjoint (and not necessarily multiplicative). In this case, R_1 may be written as

$$R_1 = (H_0^{-1}s, H\ H_0^{-1}s) - (s, H_0^{-1}s), \tag{24}$$

and H_0 now appears only in the combination $H_0^{-1}s$. After applying Eq. (23) to (24), the latter becomes

$$R_1 = A^2(f, Hf) - A(s,f).$$

For R_1 to be zero, a nontrivial A must be given by

$$A = (s,f)/(f\ Hf). \tag{24'}$$

This gives, for the iterative solution to the problem,

$$I^{(2)} = I^{(1)} = (s, H_0^{-1}s) = A(s,f) = (s,f)^2/(f, Hf). \tag{25}$$

For this case, the iterative result (25) is identical to the variational result (21).

A second parameter α may be introduced into the trial function by putting it into the function $f(x)$; in other words, $\psi_t = Af_\alpha(x)$. The two methods again give the estimate

$$(s,f_\alpha)^2/(f_\alpha, Hf_\alpha) \tag{26}$$

for the weighted average I, but the values of α will differ in general. The variational estimate is determined from the conditions

$$\frac{\partial F}{\partial A} = \frac{\partial F}{\partial \alpha} = 0, \tag{27}$$

whereas the iterative result comes from the solution to

$$R_1(A, \alpha) = R_2(A, \alpha) = 0. \tag{28}$$

Whether one estimate will be better than the other depends upon the particular problem. For example, if the problem is such that I is an extremum for Eq. (26), then the variational result will be superior, because Eq. (27) is an extremum condition. However, if the problem is not an extremum, then the two approximations are on a par with each other and it is difficult, in general, to choose one over the other. The class of problems which are not self-adjoint is an example of the latter situation.

A specific example in which the iterative result is superior to the variational result is given in the theory of resonance absorption of neutrons. This example will be discussed in detail later.

We are trying to indicate that the iterative approach involving trial operators is useful for obtaining analytic solutions and is particularly convenient when a limited number of iterations are required. The latter condition exists not only because higher-order iterations may become increasingly complex, but also because it is not always clear that the series (11) converges. However, we have no intention of summing the series, but only to obtain an "asymptotic type of approximation" in the sense that we have described.

It is conceivable that with a particularly poor choice of the trial operator, even though R_1 is chosen to be zero, the next term in the series R_2 may be large. But this same situation exists in variational calculations. One can render a functional stationary to first order, but, by choosing a particularly poor trial function, the second-order terms may be large, thus producing a poor estimate of the stationary value.

We do not claim to have a method which is generally superior to variational formulations, but only claim that this approach is a comparable approximation technique. However, the iterative approach does have the additional advantage of containing a great deal of other information. If the solution $R_1(H_0) = 0$ gives a reasonably convergent series, then the remainder term R_2 contains much information

about the approximation. This includes:

(1) If R_2 is small with respect to $R_0 = I^{(1)}$, this is an indication of the reasonableness of the approach, and R_2 gives one an order of accuracy estimate of the approximation $I = I^{(1)}$.

(2) The sign of R_2 indicates whether the approximation is an overestimate or an underestimate of the exact result. Again this depends upon the series approximation being good, since we are neglecting other remainder terms.

(3) By putting different classes of trial operators into R_2, one has a means of testing their quality with respect to the particular calculation. Hence, without knowing the exact result, one has a direct means of comparing different classes of operators or functions.

(4) By choosing forms for the trial operator or function so that one obtains both positive and negative values for R_2, one may be able to bound the exact result from above and from below. This would be an accomplishment for non-self-adjoint problems, since such bounds are especially difficult to obtain by other means.

We now give some examples which we hope will establish the utility of the technique. Mathematical rigor is not intended. We simply wish to indicate that when good trial operators are used, not only does one obtain a good approximation to the exact answer, but additional valuable information may be obtained from the iterative series.

Consider the self-adjoint problem

$$H\psi(x) = s, \quad -a \leq x \leq a, \quad \psi(\pm a) = 0, \quad (29)$$

where H is the one-dimensional operator

$$H = 1 - \frac{1}{\kappa^2}\frac{d^2}{dx^2}.$$

This system of equations may be used to represent the diffusion-theory problem as treated in Ref. 1. If we choose the trial operator as

$$H_0^{-1}s = Af$$

for arbitrary A, then the solution for $R_1 = 0$ is given by Eq. (25). This is identical to the variational solution of the problem. If s is spatially independent, then the weighted average $I = (s, \psi)$ is related to the spatially averaged flux ($\bar{\psi} = I/2as$). The approximate solution is given by Eq. (25) or by

$$\bar{\psi} = (s/2a)(1,f)^2/(f, Hf). \quad (30)$$

For any given shape function $f(x)$, the approximation to the spatial average $\bar{\psi}$ is determined by evaluating the scalar products in (30). There are a number of possible trial shapes. For example,

$$f_1 = a^2 - x^2 \quad \text{or} \quad f_2 = \cos(\pi/2a)x. \quad (31)$$

Both of these functions are symmetric in x and satisfy the homogeneous boundary conditions in (29). Which trial function shall we use?

We can make use of the R_2 term in the iterative series to test the quality of the trial functions without reference to the exact answer. If we choose H_0 as a multiplication operator ($H_0^{-1} = Af/s$), then we have from Eq. (15) that

$$R_2(H_0) = \frac{A^3}{s}(f, H[fHf]) - A(s,f). \quad (32)$$

From the equation for $R_1(H_0) = 0$ [i.e., Eq. (24′)], the amplitude A is given by

$$A = (s,f)/(f, Hf). \quad (33)$$

Applying Eq. (33) to Eq. (32), we have that

$$R_2(f) = \left\{\frac{(1,f)[f, H(fHf)]}{(f, Hf)^2} - 1\right\} I^{(1)}(f). \quad (34)$$

If each of the trial shapes (31) are now inserted into (34), the results are

$$R_2(f_1) = \frac{1}{14}\frac{(\kappa a)^4}{(\kappa a)^4 + 5(\kappa a)^2 + \frac{25}{4}} I^{(1)}(f_1),$$

which varies from 0 to 7.14% of $I^{(1)}$ as κa varies from zero to infinity, and

$$R_2(f_2) = (0.081)I^{(1)}(f_2),$$

which is 8.1% of $I^{(1)}$ for all κa.

We immediately learn the following from this calculation:

(1) The results are accurate to the order of 10%. This is not to be taken rigorously, because we are using only the first remainder in the series and we have not evaluated the other terms, R_n for $n > 2$. We are using this R_2 calculation as an indicator more than anything else. The smallness of R_2 suggests the reasonableness of the approximation.

(2) Because R_2 is positive in both cases, both trial shapes yield what is apparently an underestimate to the exact result.

(3) Since for all κa, $R_2(f_1)/I^{(1)}(f_1) < R_2(f_2)/I^{(1)}(f_2)$, f_1 appears to be better than f_2 as a trial shape.

All of this additional information about the approximation was obtained from R_2 without much labor. We now compare the approximate solutions (30) for each f, with the exact result.

[1] R. Goldstein, J. Math. Phys. **8**, 473 (1967).

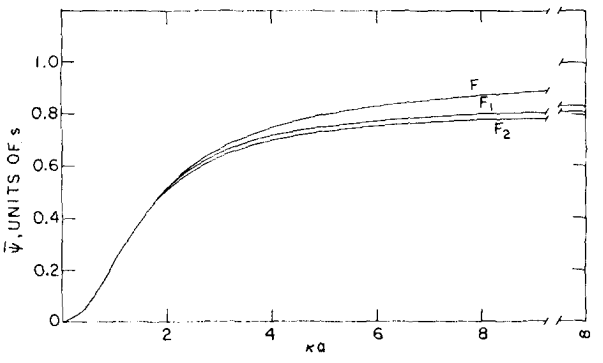


FIG. 1. The exact solution (F) and the iterative estimates (F_1 and F_2) for the average flux ($\bar{\psi}$) as a function of κa.

Let $F = s^{-1}\bar{\psi}$. The exact solution to Eq. (29) for F is

$$F = 1 - (\kappa a)^{-1} \tanh \kappa a. \tag{35}$$

The approximate solutions are

$$F_1 = \frac{(2a)^{-1}(1, f_1)^2}{(f_1, Hf_1)} = \frac{0.811(\kappa a)^2}{(\kappa a)^2 + 2.467} \tag{36}$$

and

$$F_2 = \frac{(2a)^{-1}(1, f_2)^2}{(f_2, Hf_2)} = \frac{0.833(\kappa a)^2}{(\kappa a)^2 + 2.500}. \tag{37}$$

Equations (35)–(37) are plotted in Fig. 1 as a function of κa. From the figure, we note, indeed, that both approximations F_1 and F_2 are reasonable, both are underestimates to the exact result, and F_1 does give the better estimate. The actual errors vary from 0 to 16%, so that even the errors are reasonably predicted by R_2.

We chose a particularly simple case to illustrate the utility of the method. Actually, the operator used here is positive-definite, so that we had a maximum principle.[1] Thus we had other means of obtaining some of the above information. However, in the next example, involving a non-self-adjoint operator, we do not have an extremum and the information obtained from R_2 is even more revealing.

We consider the problem of resonance absorption of neutrons in a homogeneous mixture. The equations here are not self-adjoint, in general, and they may be put into the form of Eq. (1) by writing[2]

$$H = s + \sigma - K\sigma_s,$$

where K is a slowing-down integral operator of the Volterra type and the other terms are known cross sections:

$$K = \int_{u-\Delta}^{u} du' \frac{e^{-(u-u')}}{1-\alpha}.$$

[2] R. Goldstein and E. R. Cohen, Nucl. Sci. Eng. **13**, 132 (1962); R. Goldstein, "Intermediate Resonance Absorption" in *Reactor Physics in the Resonance and Thermal Regions* (The M.I.T. Press, Cambridge, Mass., 1966), Vol. II, p. 37.

An appropriate trial function for the flux per unit lethargy ψ is given by

$$\psi_\lambda = \frac{s + \lambda\sigma_p}{s + \sigma_a + \lambda\sigma_s}. \tag{38}$$

This trial function has the property of being normalized to unity above the resonance and of yielding the corresponding wide- or narrow-resonance solutions when λ is set equal to zero or unity, respectively. The problem is to find the value of the parameter λ for each particular resonance, and then the corresponding resonance integral. For the variational solution to the problem one can construct a functional analogous to Eq. (19), which has the resonance integral as its stationary value. However, since the problem is not self-adjoint, one has to examine the equation adjoint to Eq. (1) and use correspondingly appropriate adjoint trial functions.[2]

The iterative solution to the problem may be obtained by choosing the trial operator as

$$H_0^{-1}s = \psi_\lambda = (s + \lambda\sigma_p)/(s + \sigma_a + \lambda\sigma_s). \tag{39}$$

When H_0^{-1} is needed explicitly, we take it as a multiplication operator, which means that we need not concern ourselves with adjoint trial functions or operators for this iterative approach. The iterative solution for λ is obtained from the equation $R_1(\lambda) = 0$ and is given in the notation of Ref. 2 by

$$\lambda = 1 - X_{\lambda\lambda}. \tag{40}$$

The solution (40) has the appropriate behavior in the limiting narrow- and wide-resonance extremes. Once the parameter λ is known, the resonance integral may be found directly.

The R_2 calculation is more complicated. Assuming that $R_1 = 0$, we obtain after some manipulation (in the notation of Ref. 2, where I^* stands for the "infinite-dilution resonance integral")

$$\begin{aligned} R_2(\lambda) = \frac{I^*}{\beta_\lambda}\Bigg\{ & \frac{3}{8}\left(\frac{s+\sigma_p}{s}\right)^2\left[\left(\frac{\beta_1}{\beta_\lambda}\right)^4 + \frac{2}{3}\left(\frac{\beta_1}{\beta_\lambda}\right)^2 + 1\right] \\ & - \frac{\sigma_p(s+\sigma_p)}{s^2}\left[\frac{1}{2}\frac{\beta_1^2+\beta_\lambda^2}{\beta_\lambda^2}\left(1 + \frac{2A}{\Delta}\tan^{-1}\frac{\Delta}{2}\right)\right. \\ & \left. + \frac{\beta_1^2-\beta_\lambda^2}{\beta_\lambda^2}\left(\frac{A}{\Delta^2+4} + \frac{1}{\Delta}\tan^{-1}\frac{\Delta}{2}\right) + 1\right] \\ & + \left(\frac{\sigma_p}{s}\right)^2\left[1 + \frac{A}{\Delta^2}\left(2\Delta\tan^{-1}\Delta - \Delta\tan^{-1}\frac{\Delta}{2}\right)\right. \\ & - \log(1+\Delta^2) + 2\log\left(1+\frac{\Delta^2}{4}\right) \\ & \left. + \frac{A^2}{\pi\Delta^2}(2J(\Delta,0) - J(\Delta,-\Delta) - J(0,0))\right] - 1\Bigg\}, \end{aligned} \tag{41}$$

where

$$A = \frac{\beta_\infty^2 - \beta_\lambda^2}{\beta_\lambda^2}, \quad \Delta = \frac{\delta}{\beta_\lambda},$$

and

$$J(\Delta, -\Delta) = \int_{-\infty}^{\infty} \frac{dx}{1 + x^2} \tan^{-1}(x - \Delta) \tan^{-1}(x + \Delta).$$

The J integrals are easily evaluated numerically [$J(0, 0) = \pi^3/12$].

As a specific numerical example, we consider the 192 eV resonance of U238 in a 1:1 atomic-ratio mixture with hydrogen. This resonance is chosen because it is poorly represented by the narrow- or wide-resonance approximations and thus constitutes a good intermediate case.[2]

If we call λ_0 the solution to Eq. (40), then the resonance integral is given by

$$I = I^*/\beta_{\lambda_0},$$

which is 0.165 b for the above example. Note that it is not necessary to evaluate R_2 in order to obtain this result.

If we do evaluate R_2 for this case, we find from Eq. (41) that

$$R_2(\lambda_0) = -0.010 \text{ b}.$$

This calculation reveals two things: The accuracy is of the order of 6% and the approximation yields an overestimate to the exact result.

Consider now another possible trial operator

$$H_0^{-1}s = \psi_\mu = (1 - \mu)f_0 + \mu f_1, \tag{42}$$

where

$$f_\lambda = \frac{s + \lambda\sigma_p}{s + \sigma_a + \lambda\sigma_s} \quad \text{for} \quad \lambda = 0 \quad \text{or} \quad 1.$$

This trial form has properties similar to that of Eq. (39). Equations (39) and (42) are identical for the limits $\mu = \lambda = 0$ and $\mu = \lambda = 1$, but they differ for the intermediate cases.

If we call μ_0 the solution to $R_1(\mu) = 0$, then

$$I = I^*/\beta_{\mu_0} = 0.159 \text{ b}$$

and

$$R_2(\mu_0) = +0.003 \text{ b}.$$

This approximation thus apparently yields an underestimate of the exact result, to an accuracy of the order of 2%.

Assuming the two approximations are reasonable, we have succeeded, therefore, in bounding the exact stationary value (the resonance integral) to this non-self-adjoint problem:

$$0.159 \text{ b} < I < 0.165 \text{ b}. \tag{43}$$

A numerical evaluation of the integral equation[3] gives for the "exact" resonance integral

$$I_{\text{exact}} = 0.161 \text{ b}.$$

Not only does the exact result fall nicely between the bounds of Eq. (43), but it is slightly closer to the lower bound, as predicted by the R_2's.

The corresponding variational calculations for this problem[2] yield for the resonance integral $I = 0.173$ b. The iterative results thus give approximately a 5% improvement over the variational estimates.

We see, therefore, that excellent results may be obtained by the trial operator approach. We do not claim that these isolated examples imply the universality of the method, but we hope to have established the usefulness of the approach.

ACKNOWLEDGMENTS

This work was performed under the auspices of the USAEC. Some of the material in this paper was presented to a meeting of the American Nuclear Society in Mexico City, May, 1967.

[3] G. C. Pomraning and M. W. Dyos (private communication).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Reduction of the Bethe–Salpeter Equation for a Zero-Mass System

R. F. KEAM
Department of Physics, University of Auckland, Auckland, New Zealand

(Received 29 November 1967)

The ladder-approximation Bethe–Salpeter equation for a bound spin-$\frac{1}{2}$ fermion–antifermion system of zero total mass is reduced in the general case to coupled radial equations for general radial potential functions. Symmetries of the radial equations are also demonstrated.

I. INTRODUCTION

The angular analysis of the ladder-approximation Bethe–Salpeter equation for a bound spin-$\frac{1}{2}$ fermion–antifermion system has been investigated by Günther,[1] Kummer,[2] and Delbourgo *et al.*[3] In the last paper, the Bethe–Salpeter equation is completely reduced to a set of coupled radial equations for a special class of representations of the symmetry group N_{4P} of the equation. We extend this by performing a reduction for all classes of representation and by considering general radial potential functions. A feature of the approach is that only ordinary three-dimensional Clebsch–Gordan coefficients appear.

II. MOMENTUM OPERATORS

With a spin-$\frac{1}{2}$ field operator $\psi(x)$, the momentum operators P_μ, $M_{\mu\nu}$ satisfy the commutation relations (for details of conventions and notation, see the Appendix)

$$[P_\mu, \psi(x)] = i\partial_\mu\psi(x),$$
$$[M_{\mu\nu}, \psi(x)] = \{i(x_\mu\partial_\nu - x_\nu\partial_\mu) - \tfrac{1}{2}\sigma_{\mu\nu}\}\psi(x). \quad (1)$$

Taking the Hermitian conjugate of (1), and inserting γ_4 where appropriate, we have for the adjoint field $\bar{\psi} = \psi^\dagger\gamma_4$

$$[P_\mu, \bar{\psi}(x)] = i\partial_\mu\bar{\psi}(x),$$
$$[M_{\mu\nu}, \bar{\psi}(x)] = \bar{\psi}\{i(x_\mu\overleftarrow{\partial}_\nu - x_\nu\overleftarrow{\partial}_\mu) + \tfrac{1}{2}\sigma_{\mu\nu}\}. \quad (2)$$

Let $|B\rangle$ be a spin-$\frac{1}{2}$ fermion–antifermion state in the Heisenberg picture and consider the two-body wavefunction or amplitude defined as

$$\chi(x_1, x_2) = \langle 0|\, T\{\psi(x_1)\bar{\psi}(x_2)\}\,|B\rangle, \quad (3)$$

where T is the Wick chronological product operator. [We follow the usual practice of regarding χ as a 4×4 matrix whose (μ, ν) element is

$$\chi_{\mu\nu} = \langle 0|\, T\{\psi_\mu(x_1)\bar{\psi}_\nu(x_2)\}\,|B\rangle.]$$

We wish to identify the operators $\mathcal{P}_\mu$, $\mathcal{M}_{\mu\nu}$, which, acting in χ, are equivalent, respectively, to P_μ, $M_{\mu\nu}$ acting on $|B\rangle$, i.e.,

$$\mathcal{P}_\mu\chi = \langle 0|\, T\{\psi(x_1)\bar{\psi}(x_2)\}P_\mu\,|B\rangle,$$
$$\mathcal{M}_{\mu\nu}\chi = \langle 0|\, T\{\psi(x_1)\bar{\psi}(x_2)\}M_{\mu\nu}\,|B\rangle. \quad (4)$$

Using the commutation relations (1) and (2), we find indeed

$$\mathcal{P}_\mu\chi = -i\Delta_\mu\chi,$$
$$\mathcal{M}_{\mu\nu}\chi = -i\{(X_\mu\Delta_\nu - X_\nu\Delta_\mu) + (x_\mu\partial_\nu - x_\nu\partial_\mu)\}\chi + \tfrac{1}{2}[\sigma_{\mu\nu}, \chi], \quad (5)$$

where $x = x_1 - x_2$ and $X = \mu_a x_1 + \mu_b x_2$ for arbitrary μ_a, μ_b satisfying $\mu_a + \mu_b = 1$ and where ∂, Δ are derivatives with respect to x, X, respectively. In the center-of-mass system we have

$$\chi = e^{-iET}f(x), \quad (6)$$

where E is the total energy of state $|B\rangle$ in that system, and $X_4 = iT$.

For a system so tightly bound that E vanishes,[4] the expression for $\mathcal{M}_{\mu\nu}$ simplifies to

$$\mathcal{M}_{\mu\nu} = -i(x_\mu\partial_\nu - x_\nu\partial_\mu) + \tfrac{1}{2}[\sigma_{\mu\nu}, \ ]. \quad (7)$$

III. BETHE–SALPETER EQUATION

For a suitable range of E, Wick[5] showed that the Bethe–Salpeter equation may be analytically continued in the relative coordinate x_4 (and its conjugate momentum p_4) between real and imaginary axes. We assume such a range for E in the following, taking x_4 and p_4 real.

Separating χ according to (6), in ladder approximation, the Bethe–Salpeter equation in the Euclidean relative-coordinate space becomes

$$(\gamma\cdot\partial - i\mu_a E\gamma_4 + m_a)f(x)(\gamma\cdot\overleftarrow{\partial} + i\mu_b E\gamma_4 + m_b) = -\lambda\mathcal{U}f(x),$$

or briefly

$$\mathcal{B}_E f(x) = -\lambda\mathcal{U}f(x), \quad (8)$$

[1] M. Günther, J. Math. Phys. **5**, 188 (1964).
[2] W. Kummer, Nuovo Cimento **31**, 219 (1964).
[3] R. Delbourgo, A. Salam, and J. Strathdee, Nuovo Cimento **50**, 193 (1967).
[4] We ignore, in common with other authors, the fact that a center-of-mass system does not exist for zero-mass states.
[5] G. C. Wick, Phys. Rev. **96**, 1124 (1954).

where the linear operator $\mathcal{U}$ describes the interaction between the fermion of mass m_a and the antifermion of mass m_b.

The most general form of potential we shall consider is given by

$$\mathcal{U}f = V_S(R)f + V_V(R)\gamma_\mu f\gamma_\mu + \tfrac{1}{2}V_T(R)\sigma_{\mu\nu}f\sigma_{\mu\nu} + V_A(R)i\gamma_5\gamma_\mu f i\gamma_5\gamma_\mu + V_P(R)\gamma_5 f\gamma_5. \quad (9)$$

The set of 16 linear operators $\{\hat{\gamma}_A\}$, defined by

$$\hat{\gamma}_A f(x) = \gamma_A f(x)\gamma_A \quad \text{(not summed)}, \quad (10)$$

forms an Abelian group whose simultaneous eigenfunctions are γ_B.

If we define

$$\hat{\Gamma}_i = \sum_{A\in i}\hat{\gamma}_A, \quad (11)$$

summed over all γ_A in Γ_i for $i = 1, 2, 3, 4$, or 5, we may write

$$\mathcal{U} = \sum_i V_i(R)\hat{\Gamma}_i. \quad (12)$$

Now f may be expressed as

$$f = \sum_{i=1}^{5} f_i\Gamma_i, \quad (13)$$

where for each i

$$f_i\Gamma_i = \sum_{A\in i} f_A\gamma_A \quad \text{(no summation in } i\text{, l.h.s.)}, \quad (14)$$

so that

$$\mathcal{U}f = \sum_{ij} V_i(R)f_j\hat{\Gamma}_i\Gamma_j = \sum_{ij} c_{ij}V_i(R)f_j\Gamma_j, \quad (15)$$

where

$$\hat{\Gamma}_i\Gamma_j = c_{ij}\Gamma_j. \quad (16)$$

The numerical values of the coefficients c_{ij} are given in Table I. For $E = 0$, we find

$$[\mathcal{B}_0, \hat{\Gamma}_2 - \hat{\Gamma}_3]f = -2(m_a - m_b)\hat{\Gamma}_2(f\gamma\cdot\overleftarrow{\partial} - \gamma\cdot\partial f),$$
$$[\mathcal{B}_0, \hat{\Gamma}_5 - \hat{\Gamma}_4]f = -2(m_a - m_b)\hat{\Gamma}_5(f\gamma\cdot\overleftarrow{\partial} - \gamma\cdot\partial f), \quad (17)$$

so that, when $m_a = m_b$, there are two operators that commute with $\mathcal{B}_0$. They are, however, algebraically dependent so we select one of them

$$\mathcal{R} = \hat{\Gamma}_5 - \hat{\Gamma}_4 \quad (18)$$

TABLE I. Coefficients c_{ij}: $\hat{\Gamma}_i\Gamma_j = c_{ij}\Gamma_j$.

	Γ_1	Γ_2	Γ_3	Γ_4	Γ_5
$\hat{\Gamma}_1$	1	1	1	1	1
$\hat{\Gamma}_2$	4	−2	0	2	−4
$\hat{\Gamma}_3$	6	0	−2	0	6
$\hat{\Gamma}_4$	4	2	0	−2	−4
$\hat{\Gamma}_5$	1	−1	1	−1	1

for further consideration. Its eigenfunctions γ_A belong to eigenvalues -3 (for $\gamma_A \in \Gamma_1, \Gamma_2$), 1 (for $\gamma_A \in \Gamma_3, \Gamma_4$), and 5 (for $\gamma_A \in \Gamma_5$).[6] One may regard its commuting as being responsible for the well-known decoupling of S–V, T–A, and P sectors when $m_a = m_b$.

Clearly

$$[\mathcal{R}, \mathcal{U}] = 0, \quad (19)$$

and one may readily show that

$$[\mathcal{R}, \mathcal{M}_{\mu\nu}] = 0. \quad (20)$$

Since

$$[\mathcal{M}_{\mu\nu}, \mathcal{B}_0] = [\mathcal{M}_{\mu\nu}, \mathcal{U}] = 0, \quad (21)$$

we may require solutions of (8) for $E = 0$ to be simultaneous eigenfunctions of a maximal commuting set of the $\mathcal{M}_{\mu\nu}$, and of $\mathcal{R}$ as well when $m_a = m_b$.

Define, for $i, j, k \in 1, 2, 3$,

$$J_i = \tfrac{1}{2}\epsilon_{ijk}\mathcal{M}_{jk}, \quad Q_i = \tfrac{1}{2}(\mathcal{M}_{i4} - \mathcal{M}_{4i}) \quad (22)$$

or, since $\mathcal{M}_{\nu\mu} = -\mathcal{M}_{\mu\nu}$,

$$J_i = \mathcal{M}_{jk} = -\mathcal{M}_{kj}, \qquad Q_i = \mathcal{M}_{i4} = -\mathcal{M}_{4i}, \quad (23)$$

where i, j, k cycle 1, 2, 3.

Intercommutation relations between the $M_{\mu\nu}$ imply

$$[J_i, J_j] = i\epsilon_{ijk}J_k, \quad [Q_i, Q_j] = i\epsilon_{ijk}J_k, \quad [J_i, Q_j] = i\epsilon_{ijk}Q_k = [Q_i, J_j]. \quad (24)$$

Also,

$$\alpha = \tfrac{1}{2}\mathcal{M}_{\mu\nu}\mathcal{M}_{\mu\nu} = \mathbf{J}^2 + \mathbf{Q}^2,$$
$$\beta = \tfrac{1}{4}\mathcal{M}^D_{\mu\nu}\mathcal{M}_{\mu\nu} = \tfrac{1}{8}\epsilon_{\mu\nu\pi\rho}\mathcal{M}_{\pi\rho}\mathcal{M}_{\mu\nu} = \mathbf{J}\cdot\mathbf{Q} \quad (25)$$

commute with all $\mathcal{M}_{\mu\nu}$, and hence with J_i and Q_j.

A maximal commuting set is then given by

$$\text{Set I} \qquad \{\alpha, \beta, \mathbf{J}^2, J_3\}, \quad (26)$$

where we recognize $\mathbf{J}$ as the ordinary angular-momentum operator. Indeed, for spin-zero particles, the operators corresponding to $\mathcal{M}_{\mu\nu}$ are obtained from (7) by omitting the $\frac{1}{2}[\sigma_{\mu\nu}, \ ]$ term, and in terms of spherical polar coordinates (R, ψ, θ, ϕ),

$$\Box = \frac{\partial^2}{\partial R^2} + \frac{3}{R}\frac{\partial}{\partial R} - \frac{\alpha}{R^2},$$
$$\alpha = -\frac{\partial^2}{\partial\psi^2} - 2\cot\psi\frac{\partial}{\partial\psi} + \frac{\mathbf{J}^2}{\sin^2\psi},$$
$$\mathbf{J}^2 = -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{J_3^2}{\sin^2\theta}, \quad (27)$$
$$J_3 = -i\frac{\partial}{\partial\phi},$$

[6] Since $\mathcal{R}$ has three eigenvalues, it satisfies a cubic equation so that there exist no further algebraically dependent but linearly independent operators commuting with $\mathcal{B}_0$ than the above, and the trivial $\hat{\Gamma}_1$.

so that the four-dimensional surface harmonic[7]

$$\sin^l \psi C_{n-l}^{l+1}(\cos\psi) Y_{lm}(\theta, \phi)$$

is simply a simultaneous eigenfunction of α, $\mathbf{J}^2$, and J_3 belonging, respectively, to eigenvalues $n(n+2)$, $l(l+1)$, and m. (In the spin-zero case $\beta \equiv 0$.)

IV. BISPHERICAL BASIS

Let us exploit the homomorphism $N_{4I} \sim N_{3I} \times N_{3I}$ by considering the alternative maximal commuting set:

$$\text{Set II} \qquad \{(\mathbf{J}^+)^2, J_3^+, (\mathbf{J}^-)^2, J_3^-\}, \tag{28}$$

where [8]

$$\mathbf{J}^\pm = \tfrac{1}{2}(\mathbf{J} \pm \mathbf{Q}). \tag{29}$$

The $\mathbf{J}^\pm$ behave as a commuting pair of angular momenta, and we may write

$$\mathbf{J}^\pm = \mathbf{L}^\pm + \mathbf{S}^\pm, \tag{30}$$

where

$$\begin{aligned} L_i^\pm &= -\tfrac{1}{2}i\{(x_j\partial_k - x_k\partial_j) \pm (x_i\partial_4 - x_4\partial_i)\}, \\ S^\pm &= \tfrac{1}{4}[\sigma_{jk} \pm \sigma_{i4}, \quad]. \end{aligned} \tag{31}$$

All commutators between $L_i^\pm$ and $S_j^\pm$ vanish, other than $[L_i^+, L_j^+]$, $[L_i^-, L_j^-]$, $[S_i^+, S_j^+]$, and $[S_i^-, S_j^-]$.

Combining eigenfunctions of $(\mathbf{L}^+)^2$, L_3^+, $(\mathbf{S}^+)^2$, S_3^+ in the familiar fashion with Clebsch–Gordan coefficients, we can therefore build eigenfunctions of $(\mathbf{J}^+)^2$ and J_3^+, and simultaneously we can build eigenfunctions of $(\mathbf{J}^-)^2$, J_3^- from eigenfunctions of $(\mathbf{L}^-)^2$, L_3^-, $(\mathbf{S}^-)^2$, and S_3^-.

First, consider simultaneous eigenfunctions of $(\mathbf{L}^\pm)^2$, $L_3^\pm$. Spherical polar coordinates are no longer appropriate. Instead, consider "bispherical" coordinates (R, ν, ω, ϕ) defined by

$$\begin{aligned} x_1 &= R \sin\nu \cos\phi, \\ x_2 &= R \sin\nu \sin\phi, \\ x_3 &= R \cos\nu \cos\omega, \\ x_4 &= R \cos\nu \sin\omega, \end{aligned} \tag{32}$$

and with ranges: $R \in [0, \infty)$, $\nu \in [0, \tfrac{1}{2}\pi]$, $\phi \in [0, 2\pi)$, and $\omega \in [0, 2\pi)$. R and ϕ coincide with the spherical polar coordinates represented by these respective symbols. Since $(\mathbf{L}^+)^2 \equiv (\mathbf{L}^-)^2$, we denote these by $\mathbf{L}^2$. One finds that

$$\mathbf{L}^2 = -\left\{\frac{1}{4}\frac{\partial^2}{\partial\nu^2} + \tfrac{1}{2}\cot 2\nu \frac{\partial}{\partial\nu} - \frac{1}{\sin^2 2\nu}[(L_3^+)^2 + (L_3^-)^2] - \frac{2\cos 2\nu}{\sin^2 2\nu} L_3^+ L_3^-\right\}, \tag{33}$$

$$L_3^\pm = -\tfrac{1}{2}i\left(\frac{\partial}{\partial\phi} \pm \frac{\partial}{\partial\omega}\right). \tag{34}$$

Normalized, properly phased, simultaneous eigenfunctions are given in terms of Jacobi polynomials by

$$\begin{aligned} &Z_{lm^+m^-}(\nu, \omega, \phi) \\ &\quad = (-1)^{\frac{1}{2}(m+|m|)} N \sin^{|m|}\nu \\ &\qquad \times \cos^{|m'|}\nu P_{l-\frac{1}{2}(|m|+|m'|)}^{(|m|,|m'|)}(\cos 2\nu) e^{im\phi} e^{im'\omega} \\ &\quad = (-1)^{\frac{1}{2}(m+|m|)} N g_{lm^+m^-}(\nu) e^{im\phi} e^{im'\omega} \\ &\quad = (-1)^{\frac{1}{2}(m+|m|)} N H_{lm^+m^-}(\nu, \omega, \phi), \end{aligned} \tag{35}$$

where $l = 0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$; and given l, the $m^\pm$ are in the set $-l, -l+1, \cdots, l$; $m = m^+ + m^-$, $m' = m^+ - m^-$. The normalization constant

$$N = \left\{\frac{2l+1}{2\pi^2} \cdot \frac{(l+m^-)!(l-m^-)!}{(l+m^+)!(l-m^+)!}\right\}^{\frac{1}{2}}, \quad \text{for } |m^-| \geq |m^+|,$$

or

$$= \left\{\frac{2l+1}{2\pi^2} \cdot \frac{(l+m^+)!(l-m^+)!}{(l+m^-)!(l-m^-)!}\right\}^{2}, \quad \text{for } |m^+| \geq |m^-|. \tag{36}$$

Next, consider simultaneous eigenfunctions of $(\mathbf{S}^\pm)^2$ and $S_3^\pm$. If the quantum numbers corresponding to $(\mathbf{S}^\pm)^2$ are $s^\pm$, one finds that for $\gamma_A \in (\Gamma_1, \Gamma_5)$, $s^\pm = 0$ while for $\gamma_A \in (\Gamma_2, \Gamma_4)$, $s^\pm = \frac{1}{2}$. One also finds that the three linearly independent matrices of type $(1-\gamma_5)\Gamma_3$ have $s^+ = 1$ and $s^- = 0$, while the matrices of type $(1+\gamma_5)\Gamma_3$ have $s^+ = 0$, $s^- = 1$. Selecting now eigenfunctions of $S_3^\pm$ corresponding to quantum numbers $m_s^\pm$, we have, after defining

$$\begin{aligned} \sigma_1^\pm &= 2^{-\frac{1}{2}}(\gamma_1 \pm i\gamma_2), \\ \sigma_2^\pm &= 2^{-\frac{1}{2}}(\gamma_4 \pm i\gamma_3), \end{aligned} \tag{37}$$

the results displayed in Table II. We may regard the

TABLE II. Dirac-space eigenfunctions.

	m_s^+	m_s^-	Eigenfunctions (not normalized)	
$s^+ = s^- = 0$	0	0	I	γ_5
$s^+ = s^- = \frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$-\sigma_1^+$,	$-i\gamma_5\sigma_1^+$
	$-\frac{1}{2}$	$-\frac{1}{2}$	σ_1^-,	$i\gamma_5\sigma_1^-$
	$\frac{1}{2}$	$-\frac{1}{2}$	$i\sigma_2^-$,	$i \cdot i\gamma_5\sigma_2^-$
	$-\frac{1}{2}$	$\frac{1}{2}$	$-i\sigma_2^+$,	$-i \cdot i\gamma_5\sigma_2^+$
$s^+ = 1, s^- = 0$	1	0	$\sigma_1^+\sigma_2^-$	
	0	0	$-2^{-\frac{1}{2}}i(\sigma_1^+\sigma_1^- - \sigma_2^+\sigma_2^-)$	
	-1	0	$-\sigma_1^-\sigma_2^+$	
$s^+ = 0, s^- = 1$	0	1	$\sigma_1^+\sigma_2^+$	
	0	0	$2^{-\frac{1}{2}}i(\sigma_1^+\sigma_1^- - \sigma_2^-\sigma_2^+)$	
	0	-1	$-\sigma_1^-\sigma_2^-$	

[7] C. Schwartz, Phys. Rev. 137, B717 (1965).

[8] At first sight, the complete reflection operator $\mathcal{C}$, with $\mathcal{C}f(x) = \gamma_5 f(-x)\gamma_5$, would seem to extend Set II. In four-dimensional Euclidean space, it is, however, not independent of the rotation operators and one finds $\mathcal{C} = \exp(2\pi i J_i^\pm) = \exp(-2\pi i J_i^\pm)$.

matrices shown as a bispherical basis for the Dirac space of 4 × 4 matrices.

Since normalization of the Dirac-space-matrix eigenfunctions is achieved by requiring Tr $(A^+A) = 1$, each eigenfunction in Table II must be multiplied by a normalization factor, $N = \frac{1}{2}$.

Denoting $\sigma^{\pm}_{1,2}$ by σ_μ, with $\mu = \begin{pmatrix} + & + & - & - \\ 1 & 2 & 1 & 2 \end{pmatrix}$, analogous to the property $[\gamma_\mu, \gamma_\nu]_+ = 2\delta_{\mu\nu}$, we have $[\sigma_\mu, \sigma_\nu]_+ = 2\delta'_{\mu\nu}$, where $\delta'_{\mu\nu} = 1$ when μ, ν have the same numerical index and opposite sign index, and $\delta'_{\mu\nu} = 0$ otherwise.

V. REDUCTION OF THE EQUATION

Consider now the equal mass case $m_a = m_b = m$. In the S–V sector, since $s^\pm = 0$ for the S term, we must have $j^+ = j^-$ $(= j$, say). Distinct radial functions will exist for each l value for each type (S or V) of term. For the S term we have simply $l_S = j$, but for the V terms, since $s^\pm = \frac{1}{2}$, we have $l_V = j \pm \frac{1}{2}$ for $j \neq 0$. Written in full, the solution must therefore be of the form

$$
\begin{aligned}
&f_{jm^+m^-}(R, \nu, \omega, \phi) \\
&= f_S(R) Z_{jm^+m^-\frac{1}{2}I} \\
&\quad + f_{V1}(R) \sum_{m_s^+ m_s^-} (j + \tfrac{1}{2}\,\tfrac{1}{2}\, m_l^+ m_s^+ \,|\, jm^+) \\
&\quad \times (j + \tfrac{1}{2}\,\tfrac{1}{2}\, m_l^- m_s^- \,|\, jm^-) Z_{j+\frac{1}{2}m_l^+ m_l^-} \Lambda_{m_s^+ m_s^-} \\
&\quad + f_{V2}(R) \sum_{m_s^+ m_s^-} (j - \tfrac{1}{2}\,\tfrac{1}{2}\, m_l^+ m_s^+ \,|\, jm^+) \\
&\quad \times (j - \tfrac{1}{2}\,\tfrac{1}{2}\, m_l^- m_s^- \,|\, jm^-) Z_{j-\frac{1}{2}m_l^+ m_l^-} \Lambda_{m_s^+ m_s^-}, \qquad (38)
\end{aligned}
$$

where $(l\, s\, m_l\, m_s \,|\, jm)$ are Clebsch–Gordan coefficients and $\Lambda_{m_s^+ m_s^-}$ are Γ_2-type eigenfunctions from Table II, normalized. In order to perform the reduction, the Bethe–Salpeter equation must now be written in terms of bispherical coordinates and matrices, and this is done in the Appendix.

Define

$$
D^\pm_\alpha = \frac{\partial}{\partial R} \pm \frac{2(j + \alpha)}{R}. \qquad (39)
$$

We note that

$$
D^+_{\alpha+1} D^-_{\alpha-\frac{1}{2}} = D^-_{\alpha-1} D^+_{\alpha+\frac{1}{2}}, \qquad (40)
$$

and that if ψ_{j+a} is an eigenfunction of $\mathbf{L}^2$ with quantum number $j + a$, then

$$
\Box \psi_{j+a} = D^-_{a-\frac{1}{2}} D^+_{a+1} \psi_{j+a} = D^+_{a+\frac{3}{2}} D^-_a \psi_{j+a}. \qquad (41)
$$

If we define

$$
\bar{V}_i = \sum_{j=1}^{5} c_{ji} V_j(R), \qquad (42)
$$

the radial equations obtained after substituting $f_{jm^+m^-}$ into the Bethe–Salpeter equation are

$$
\begin{Bmatrix} D^+_{\frac{3}{2}} D^-_0 \\ D^-_{-\frac{1}{2}} D^+_1 \end{Bmatrix} s + 2m(D^+_{\frac{3}{2}} v_1 + D^-_{-\frac{1}{2}} v_2) = -(m^2 + \lambda \bar{V}_1) s, \qquad (43a)
$$

$$
D^-_0 \{D^+_{\frac{3}{2}} v_1 + 2(j + 1)(D^-_{-\frac{1}{2}} v_2 + ms)\} = -(2j + 1)(m^2 + \lambda \bar{V}_2) v_1, \qquad (43b)
$$

$$
D^+_1 \{-D^-_{-\frac{1}{2}} v_2 + 2j(D^+_{\frac{3}{2}} v_1 + ms)\} = -(2j + 1)(m^2 + \lambda \bar{V}_2) v_2, \qquad (43c)
$$

where

$$
s = (2j + 1)^{\frac{1}{2}} f_S, \quad v_1 = -(j + 1)^{\frac{1}{2}} f_{V1}, \quad v_2 = j^{\frac{1}{2}} f_{V2}. \qquad (44)
$$

In Eq. (43a), the operator pairs $D^+_{\frac{3}{2}} D^-_0$, $D^-_{-\frac{1}{2}} D^+_1$ are to be regarded as alternatives according to Eq. (40).

When $j = 0$, l_V has only the value $\frac{1}{2}$, then instead of Eqs. (43) we have

$$
\begin{aligned}
D^+_{\frac{3}{2}}(D^-_0 s + 2m v_1) &= -(m^2 + \lambda \bar{V}_1) s, \\
D^-_0(D^+_{\frac{3}{2}} v_1 + 2ms) &= -(m^2 + \lambda \bar{V}_2) v_1. \qquad (45)
\end{aligned}
$$

In the T–A sector, if the l quantum numbers for the $(1 + \gamma_5)\Gamma_3$, $(1 - \gamma_5)\Gamma_3$, and Γ_4 terms are respectively l_{T^+}, l_{T^-}, and l_A, then j^+ and j^- must simultaneously satisfy a combination of the following possibilities:

$$
\begin{array}{llll}
j^+ = l_{T^+}, & l_{T^-}, l_{T^-} \pm 1, & l_A \pm \tfrac{1}{2}, \\
j^- = l_{T^+}, l_{T^+} \pm 1, & l_{T^-}, & l_A \pm \tfrac{1}{2},
\end{array}
$$

where the signs in the last column are not correlated. There are therefore three possibilities:

(A) $j^+ = j^- = j$ $(j \neq 0)$, $l_{T^+} = l_{T^-} = j$, $l_A = j \pm \frac{1}{2}$;

(B) $j^+ = j^- - 1 = j$, $l_{T^+} = j$, $l_{T^-} = j + 1$, $l_A = j + \frac{1}{2}$;

(C) $j^+ - 1 = j^- = j$, $l_{T^+} = j + 1$, $l_{T^-} = j$, $l_A = j + \frac{1}{2}$.

Case A: There are four radial functions $f_{T1}(R)$, $f_{T2}(R)$, $f_{A1}(R)$, and $f_{A2}(R)$ corresponding respectively to terms of type $(1 - \gamma_5)\Gamma_3$, $(1 + \gamma_5)\Gamma_3$, Γ_4 with $l_A = j + \frac{1}{2}$, and Γ_4 with $l_A = j - \frac{1}{2}$. The coupled equations obtained after substitution into the Bethe–Salpeter equation are

$$
\begin{Bmatrix} D^-_{-\frac{1}{2}} D^+_1 \\ D^+_{\frac{3}{2}} D^-_0 \end{Bmatrix} w + 2m(D^+_{\frac{3}{2}} a_1 + D^-_{-\frac{1}{2}} a_2) = -(m^2 + \lambda \bar{V}_3) w, \qquad (46a)
$$

$$
D^-_0 \{-D^+_{\frac{3}{2}} a_1 + 2j(D^-_{-\frac{1}{2}} a_2 + mw)\} = -(2j + 1)(m^2 + \lambda \bar{V}_4) a_1, \qquad (46b)
$$

$$
D^+_1 \{D^-_{-\frac{1}{2}} a_2 + 2(j + 1)(D^+_{\frac{3}{2}} a_1 + mw)\} = -(2j + 1)(m^2 + \lambda \bar{V}_4) a_2, \qquad (46c)
$$

$$
\{D^+_{\frac{3}{2}} D^-_0 - (m^2 + \lambda \bar{V}_3)\} \bar{w} = 0, \qquad (47)
$$

where

$$a_1 = j^{\frac{1}{2}} f_{A1}, \quad a_2 = (j+1)^{\frac{1}{2}} f_{A2},$$
$$w = 2^{-\frac{1}{2}}(2j+1)^{\frac{1}{2}}(f_{T1} + f_{T2}), \tag{48}$$
$$\bar{w} = 2^{-\frac{1}{2}}(2j+1)^{\frac{1}{2}}(f_{T1} - f_{T2}).$$

Case B: The three radial functions f_{T1}, f_{T2}, and f_A give the coupled equations

$$\begin{pmatrix} D_2^+ D_{\frac{1}{2}}^- \\ D_0^- D_{\frac{3}{2}}^+ \end{pmatrix} a + m(D_2^+ t_1 + D_0^- t_2) = -(m^2 + \lambda \bar{V}_4) a, \tag{49a}$$

$$D_{\frac{1}{2}}^-(D_0^- t_2 + 2ma) = -(m^2 + \lambda \bar{V}_3) t_1, \tag{49b}$$

$$D_{\frac{3}{2}}^+(D_2^+ t_1 + 2ma) = -(m^2 + \lambda \bar{V}_3) t_2, \tag{49c}$$

where

$$a = f_A, \quad t_1 = 2^{\frac{1}{2}} f_{T1}, \quad t_2 = 2^{\frac{1}{2}} f_{T2}. \tag{50}$$

Case C: There are three radial functions as for Case B satisfying

$$\begin{pmatrix} D_2^+ D_{\frac{1}{2}}^- \\ D_0^- D_{\frac{3}{2}}^+ \end{pmatrix} a + m(D_0^- t_1 + D_2^+ t_2) = -(m^2 + \lambda \bar{V}_4) a \tag{51a}$$

$$D_{\frac{3}{2}}^+(D_2^+ t_2 + 2ma) = -(m^2 + \lambda \bar{V}_3) t_1, \tag{51b}$$

$$D_{\frac{1}{2}}^-(D_0^- t_1 + 2ma) = -(m^2 + \lambda \bar{V}_3) t_2. \tag{51c}$$

In Case A when $j = 0$, there are no tensor terms, and $l_A = \frac{1}{2}$ only. Instead of (46), we therefore have the single radial equation

$$\{D_0^- D_{\frac{3}{2}}^+ - (m^2 + \lambda \bar{V}_4)\} a_1 = 0. \tag{52}$$

In the P sector, since $s^{\pm} = 0$, we have $j^+ = j^- = l_P = j$ and (with $p = f_P$) the equation

$$\{D_{\frac{3}{2}}^+ D_0^- - (m^2 + \lambda \bar{V}_5)\} p = 0. \tag{53}$$

Equations (43), (46), and (47) are the analog, for general potentials, of the momentum-space equations obtained by Delbourgo *et al.*[3] A complete solution $f_{j^+ j^- m^+ m^-}(R, \nu, \omega, \phi)$ of the Bethe–Salpeter equation, involving angular variables, can be obtained from any radial solution by appropriate combination with Z functions and Dirac matrices [see, e.g., Eq. (38)].

To recover the ordinary angular-momentum content has only to note that

$$\mathbf{J} = \mathbf{J}^+ + \mathbf{J}^-,$$
$$m = m^+ + m^-,$$

so that

$$f_{j^+ j^- J m} = \sum_{m^+ m^-} (j^+ j^- m^+ m^- \,|\, J m) f_{j^+ j^- m^+ m^-}, \tag{54}$$

where once again three-dimensional Clebsch–Gordan coefficients only appear.

VI. SYMMETRIES

Consider the substitution $j \to -(j+1)$. Under the simultaneous substitutions $s \to s$, $v_1 \to v_2$, and $v_2 \to v_1$ Eqs. (43) are invariant. Under the simultaneous substitutions $w \to w$, $a_1 \to a_2$, and $a_2 \to a_1$ the equations (46) are invariant. Similarly (47) and (53) are separately invariant.

Under either the substitution $j \to -(j+2)$ or the substitution $a \to a$, $t_1 \to t_2$, and $t_2 \to t_1$, the Eqs. (49) and (51) interchange; and under the combination of these substitutions the equations are separately invariant.

The quasisymmetry between equations in the S–V and T–A sectors noted by Delbourgo *et al.*[3] is apparent also in the coordinate-space sets (43) and (46).

When the interaction potential contains only scalar and pseudoscalar contributions, Eqs. (47) and (53) are identical.

VII. UNEQUAL-MASS CASE

When $m_a \neq m_b$, except for the case $j = 0$, terms of all type (S, V, T, A, P) are coupled. Since for the S and P terms $s^{\pm} = 0$, we have $j^+ = j^- = j$. There are eight distinct radial functions corresponding to the quantum numbers $l_S = l_P = l_{T^+} = l_{T^-} = j$, $l_V = j \pm \frac{1}{2}$, $l_A = j \pm \frac{1}{2}$.

When $j = 0$, there are no tensor terms and the V, A quantum numbers are restricted to $l_V = l_A = \frac{1}{2}$. In this case the radial equations reduce to the coupled pairs

$$D_{\frac{1}{2}}^+\{D_0^- s + (m_a + m_b) v_1\} = -(m_a m_b + \lambda \bar{V}_1) s,$$
$$D_0^-\{D_{\frac{1}{2}}^+ v_1 + (m_a + m_b) s\} = -(m_a m_b + \lambda \bar{V}_2) v_1, \tag{55}$$

and

$$D_0^-\{D_{\frac{1}{2}}^+ a_1 + i(m_a - m_b) p\} = (m_a m_b + \lambda \bar{V}_4) a_1,$$
$$D_{\frac{1}{2}}^+\{D_0^- p - i(m_a - m_b) a_1\} = (m_a m_b + \lambda \bar{V}_5) p. \tag{56}$$

Under the simultaneous substitutions $m_b \to -m_b$, $\lambda \to \pm\lambda$, $a_1 \leftrightarrow v_1$, and $ip \leftrightarrow s$, the two sets interchange where the positive sign with λ applies when the potential contains only vector and axial-vector interactions, and the negative sign when there are only scalar, tensor, and pseudoscalar interactions. Indeed, since the ladder-approximation Bethe–Salpeter equation (8) is invariant under $m_b \to -m_b$, $f \to f\gamma_5$, and $\lambda \to \pm\lambda$ with the same restrictions on the potential for the alternative signs of λ as above, we see that the Delbourgo *et al.* pseudosymmetry is the remnant of a true symmetry that exists when particle and antiparticle masses are unequal.

ACKNOWLEDGMENT

I should like to thank Professor R. H. Dalitz for his interest in this work.

APPENDIX

We take $\hbar = c = 1$. Greek letter indices have the range 1, 2, 3, 4 with $x_4 = it$. ∂_μ is the derivative with respect to x_μ. We take γ_μ to be Hermitian and satisfying $\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2\delta_{\mu\nu}I$. As a basis for 4×4 matrices we take the set of sixteen Hermitian matrices γ_A in the sets:

$$\begin{aligned}
\Gamma_1 &= I,\\
\Gamma_2 &= \gamma_1, \gamma_2, \gamma_3, \gamma_4,\\
\Gamma_3 &= \sigma_{23}, \sigma_{31}, \sigma_{12}; \sigma_{14}, \sigma_{24}, \sigma_{34},\\
\Gamma_4 &= i\gamma_5\gamma_1, i\gamma_5\gamma_2, i\gamma_5\gamma_3, i\gamma_5\gamma_4,\\
\Gamma_5 &= \gamma_5,
\end{aligned}$$

where

$$\sigma_{\mu\nu} = -\tfrac{1}{2}i(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu) = -\sigma_{\nu\mu}, \quad \gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4.$$

Cast in bispherical form, the Bethe–Salpeter equation (8) for $E = 0$ reads

$$(\Lambda \cdot \Delta + m_a)f(x)(\Lambda \cdot \overleftarrow{\Delta} + m_b) = -\lambda \mathfrak{V} f(x),$$

where

$$\begin{aligned}
\Lambda \cdot \Delta &\equiv \Lambda_1^+\Delta_1^+ + \Lambda_1^-\Delta_1^- + \Lambda_2^+\Delta_2^+ + \Lambda_2^-\Delta_2^-\\
&= \tfrac{1}{2}(-\sigma_1^+\Delta_1^+ + \sigma_1^-\Delta_1^- - i\sigma_2^+\Delta_2^+ + i\sigma_2^-\Delta_2^-)\\
&= \gamma \cdot \partial,
\end{aligned}$$

so that

$$\Delta_1^\pm = 2^{\frac{1}{2}} \frac{e^{\mp i\phi}}{R \sin \nu}\left[i\frac{\partial}{\partial\phi} \mp \sin\nu\left(\cos\nu \frac{\partial}{\partial\nu} + \sin\nu\, R\frac{\partial}{\partial R}\right)\right],$$

$$\Delta_2^\pm = 2^{\frac{1}{2}} \frac{e^{\pm i\omega}}{R\cos\nu}\left[\pm i\frac{\partial}{\partial\omega} - \cos\nu \times \left(\sin\nu\frac{\partial}{\partial\nu} - \cos\nu\, R\frac{\partial}{\partial R}\right)\right].$$

If $f(x)$ is expanded as

$$f = S \cdot \tfrac{1}{2}I + V \cdot \Lambda + \mathbf{T}^+ \cdot \mathbf{\Sigma}^+ + \mathbf{T}^- \cdot \mathbf{\Sigma}^- + A \cdot i\gamma_5\Lambda + P \cdot \tfrac{1}{2}\gamma_5,$$

where

$$\mathbf{T}^\pm \cdot \mathbf{\Sigma}^\pm = T_1^\pm\Sigma_1^\pm + 2^{\frac{1}{2}}T_0^\pm\Sigma_0^\pm + T_{-1}^\pm\Sigma_{-1}^\pm,$$

we obtain in the equal mass case, $m_a = m_b = m$, the equations

$$\Box S - m\Delta' \cdot V = -(m^2 + \lambda\bar{V}_1)S,$$

$$\begin{aligned}
m\Delta_{1,2}^\pm S - \Box V_{1,2}^\pm - \tfrac{1}{2}\Delta_{1,2}^\pm(\Delta' \cdot V) &= -(m^2 + \lambda\bar{V}_2)V_{1,2}^\pm\\
&\quad - (\Delta_1^+\Delta_1^- + \Delta_2^+\Delta_2^-)T_0^\mp + (\Delta_1^\pm\Delta_2^+T_{\mp1}^\mp + \Delta_1^\mp\Delta_2^-T_{\pm1}^\mp)\\
&\quad - m(\Delta_1^-A_1^+ - \Delta_1^+A_1^- \pm \Delta_2^-A_2^+ \mp \Delta_2^+A_2^-)\\
&= -2(m^2 + \lambda\bar{V}_3)T_0^\pm,
\end{aligned}$$

$$\begin{aligned}
&\tfrac{1}{2}\{(\Delta_1^+)^2T_{-1}^\mp + (\Delta_2^\mp)^2T_1^\mp - 2\Delta_1^+\Delta_2^\mp T_0^\mp\}\\
&\qquad - m(\Delta_2^\mp A_1^+ - \Delta_1^+A_2^\mp)\\
&\qquad = -(m^2 + \lambda\bar{V}_3)T_1^\pm,
\end{aligned}$$

$$\begin{aligned}
&\tfrac{1}{2}\{(\Delta_1^-)^2T_1^\mp + (\Delta_2^\pm)^2T_{-1}^\mp - 2\Delta_1^-\Delta_2^\pm T_0^\mp\}\\
&\qquad + m(\Delta_2^\pm A_1^- - \Delta_1^-A_2^\pm)\\
&\qquad = -(m^2 + \lambda\bar{V}_3)T_{-1}^\pm,
\end{aligned}$$

$$\begin{aligned}
&m\{\pm\Delta_1^\pm(T_0^+ + T_0^-) \mp (\Delta_2^-T_{\pm1}^\mp + \Delta_2^+T_{\pm1}^\pm)\}\\
&\qquad + \Box A_1^\pm + \tfrac{1}{2}\Delta_1^\pm(\Delta' \cdot A)\\
&\qquad = -(m^2 + \lambda\bar{V}_4)A_1^\pm,
\end{aligned}$$

$$\begin{aligned}
&m\{\mp\Delta_2^\pm(T_0^+ - T_0^-) + (\Delta_1^+T_{-1}^\pm - \Delta_1^-T_1^\mp)\}\\
&\qquad + \Box A_2^\pm + \tfrac{1}{2}\Delta_2^\pm(\Delta' \cdot A)\\
&\qquad = -(m^2 + \lambda\bar{V}_4)A_2^\mp,
\end{aligned}$$

$$\Box P = (m^2 + \lambda\bar{V}_5)P,$$

$$\Delta' \cdot X = \Delta_1^+X_1^- + \Delta_1^-X_1^+ - \Delta_2^+X_2^- - \Delta_2^-X_2^+.$$

Of importance are the following formulas, valid for m and m' positive:

$$\begin{aligned}
&2^{-\frac{1}{2}}(2l+1)\Delta_1^+H_{lm^+m^-}\\
&\qquad = -(l + m^-)H_{l-\frac{1}{2}m^+-\frac{1}{2}m^--\frac{1}{2}}D_1^+\\
&\qquad\quad + (l - m^+ + 1)H_{l+\frac{1}{2}m^+-\frac{1}{2}m^--\frac{1}{2}}D_0^-,
\end{aligned}$$

$$\begin{aligned}
&2^{-\frac{1}{2}}(2l+1)\Delta_1^-H_{lm^+m^-}\\
&\qquad = -(l - m^-)H_{l-\frac{1}{2}m^++\frac{1}{2}m^-+\frac{1}{2}}D_1^+\\
&\qquad\quad + (l + m^+ + 1)H_{l+\frac{1}{2}m^++\frac{1}{2}m^-+\frac{1}{2}}D_0^-,
\end{aligned}$$

$$\begin{aligned}
&2^{-\frac{1}{2}}(2l+1)\Delta_2^+H_{lm^+m^-}\\
&\qquad = (l + m^-)H_{l-\frac{1}{2}m^++\frac{1}{2}m^--\frac{1}{2}}D_1^+\\
&\qquad\quad + (l + m^+ + 1)H_{l+\frac{1}{2}m^++\frac{1}{2}m^--\frac{1}{2}}D_0^-,
\end{aligned}$$

$$\begin{aligned}
&2^{-\frac{1}{2}}(2l+1)\Delta_2^-H_{lm^+m^-}\\
&\qquad = (l - m^-)H_{l-\frac{1}{2}m^+-\frac{1}{2}m^-+\frac{1}{2}}D_1^+\\
&\qquad\quad + (l - m^+ + 1)H_{l+\frac{1}{2}m^+-\frac{1}{2}m^-+\frac{1}{2}}D_0^-.
\end{aligned}$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Background Dependence of Local States*

A. L. LICHT
U.S. Naval Ordnance Laboratory, Silver Spring, Maryland

(Received 29 March 1967)

We consider states prepared by a series of actions performed in a finite space–time region in the presence of a background state described by a density matrix. It is shown that there will always be some observable whose expectation value in such a state will depend to some extent on the background. There will, however, be a certain set of observables whose expectation values will be independent of the background. We obtain a characterization of such sets of observables.

1. INTRODUCTION

The state set up in any experiment consists of a part, the local state, which is under the control of the experimenter, and also of a part not under his control, the background state.[1] The local state is set up by a series of actions performed in a finite space–time region. The background state describes the rest of the universe. The expectation values of observables may depend on both the background and the local state.

It is generally assumed that an experimenter can perform a measurement or a finite series of measurements that sets up an ideal state in which the background has no influence on expectation values.[2] Such a state we say is *background-independent*. We show here that such an ideal state cannot be set up by measurements made in finite space–time regions. In every experiment there will be some observable whose expectation value will depend to some extent on the background state.

For a given local state, however, there will be a certain set of observables whose expectation values will be determined by the local state independently of the background. We say that such observables are *fixed* by the local state. We give here a complete characterization of such sets of observables for a wide class of local states.

In Secs. 2–4 we consider mainly those "pure-selective"[3] local states prepared by measuring the value one for a projection in a finite space–time region. The more general case is reduced to this special case in Sec. 5.

In Sec. 2 we show that a pure-selective state is background-independent if and only if it is formed by measuring a one-dimensional projection. A heuristic discussion is given to show that every projection measurable in a finite space–time region is infinite-dimensional. This result is proven rigorously in Appendix A. We conclude that no local pure-selective state is background-independent.

In Sec. 3 we discuss the notions of observables *determined* and *fixed* by a local state. Propositions 3 to 6 are used to characterize the set of all observables fixed by a pure-selective state.

In Sec. 4 we consider in greater detail the set of projections fixed by a pure-selective state. The notion of *linked* projections is introduced.

In Sec. 5 we consider the more general case of local states consisting of finite sequences of pure-selective and nonselective local states. We show in Proposition 9 that no such state can be background-independent. In Proposition 10 we show that the sets of local observables fixed by such general states can be considered in terms of the sets fixed by certain purely selective states.

In the following we will use some of the concepts and notation of Ref. 1. In particular, we assume that to each observable there corresponds an operator on a Hilbert space $\mathcal{H}$. The operators corresponding to observables measurable in a space–time region α generate a weakly closed ring $R(\alpha)$. We will occasionally speak of an operator A being measurable in α, by which we mean $A \in R(\alpha)$.

2. DIMENSIONALITY

Suppose a local state S is prepared in the presence of a background state T. The total state is then denoted by TS. The expectation value of an observable A in TS we denote by $E(TS, A)$.[4]

We regard both T and TS as being described by density matrices ρ_T, ρ_{TS}, respectively. Then

$$E(TS, A) = \mathrm{Tr}\,(\rho_{TS}A). \tag{2.1}$$

We assume that the vacuum state exists and is a possible background. It has been shown[5] that any

* Part of this work was done in 1966 at the Institute for Advanced Study, Princeton, New Jersey.

[1] A. L. Licht, J. Math. Phys. 7, 1656 (1966).

[2] R. M. F. Houtappel, H. Van Dam, and E. P. Wigner, Rev. Mod. Phys. 37, 595 (1965). See Postulate (a) on p. 611.

[3] Reference 1, Secs. 3A and 4B.

[4] This is the "mathematical" expectation value determined by a density matrix as in Eq. (2.1). It may differ from the physical expectation value if A is measured prior to the preparation of S. See Ref. 1, Sec. 3C; Ref. 2, Eq. (4.7); and also S. Watanabe, Rev. Mod. Phys. 27, 179 (1955).

[5] Reference 1, Theorem 4.

density matrix state can be approximated in a certain sense by a series of measurements performed in the presence of the vacuum. We will therefore accept any density matrix state as a possible background.

We will consider in particular the purely selective local state $S = \dot{p}$ prepared by measuring the value one for a projection p. This state can always be prepared in a background T if $E(T, p) \neq 0$. It is well known that[6]

$$\rho_{T\dot{p}} = p\rho_T p(\mathrm{Tr}\ \rho_T p)^{-1}. \tag{2.2}$$

In the ideal state of ordinary quantum mechanics the expectation value $E(T\dot{p}, A)$ will be independent of T for all observables A. We show in Proposition 1 that this is possible if and only if p projects onto a single vector state, i.e., if $p = |\varphi)(\varphi|$ for some unit vector state φ in the Hilbert space $\mathcal{H}$.

All actual experiments take place in finite regions of space–time. Suppose p is a projection measurable in the finite region α. According to Proposition 2, p must project onto an infinite-dimensional subspace of $\mathcal{H}$. We conclude that the ideal state cannot be prepared by a finite experimenter.

Proposition 1: Let p be a projection, T any background state such that $T\dot{p}$ is a possible state. $E(T\dot{p}, A)$ is independent of T for all operators A if and only if $\dim [p\mathcal{H}] = 1$.

Proof: Sufficiency: If $p = |\varphi)(\varphi|$, for a unit vector φ, then

$$\begin{aligned} E(T\dot{p}, A) &= \mathrm{Tr}\,(\rho_T |\varphi)(\varphi| A |\varphi)(\varphi|)/\mathrm{Tr}\,(\rho_T |\varphi)(\varphi|) \\ &= (\varphi| A |\varphi), \end{aligned}$$

independent of T for all A.

Necessity: Let $\{\varphi_n\}$, $n = 1, 2, \cdots, d = \dim [p\mathcal{H}]$, be an orthonormal basis for $p\mathcal{H}$. Let

$$\rho_{T_i} = |\varphi_i)(\varphi_i|, \quad A_j = |\varphi_j)(\varphi_j|.$$

Then

$$\begin{aligned} E(T_i\dot{p}, A_j) &= (\varphi_i, pA_jp\varphi_i)(\varphi_i, p\varphi_i)^{-1} \\ &= |(\varphi_i, \varphi_j)|^2 \\ &= \delta_{ij}. \end{aligned}$$

By hypothesis this is independent of i, which is consistent only with $d = 1$.

Proposition 2: Let p be a projection measurable in a finite space–time region α. Then $\dim [p\mathcal{H}] = \infty$.

Proof: This was first proved by Guenin and Misra[7] under the assumption that the local rings of observables are factors. We give a proof that does not require this factor hypothesis in Appendix A. Here we will only show the plausibility of the proposition by considering two examples drawn from nonrelativistic quantum mechanics.

Consider first a single nonrelativistic point particle. The Hilbert space $\mathcal{H}_1$ for this particle is spanned by the functions $\psi(\mathbf{x})$, L^2 integrable over Euclidean 3-space. Let p_v denote the projection corresponding to finding the particle in the volume V. Clearly

$$p_v\psi(\mathbf{x}) = \chi_v(\mathbf{x})\psi(\mathbf{x}),$$

where $\chi_v(\mathbf{x})$ is the characteristic function for the volume V. For any V there exists an infinite sequence of disjoint subvolumes $\{V_n, n = 1, 2, \cdots\}$. The unit vectors

$$\psi_n(\mathbf{x}) = V_n^{-\frac{1}{2}}\chi_n(\mathbf{x})$$

are all orthogonal eigenfunctions of p_v with eigenvalue 1. They span an infinite-dimensional subspace of $p_v\mathcal{H}_1$, which therefore must itself be of infinite dimension.

Consider now the nonrelativistic quantum mechanics of any number of point Bose particles. A vector state Ψ in this model is a sequence of symmetric functions,

$$\Psi = \{\psi_n(\mathbf{x}_1 \cdots \mathbf{x}_n), n = 1, 2, \cdots\},$$

such that the norm

$$\|\Psi\|^2 = \sum_{n=1}^{\infty} d^3x_1 \cdots d^3x_n\, |\psi_n(\mathbf{x}_1 \cdots \mathbf{x}_n)|^2$$

is finite. It is an element of the Hilbert space

$$\mathcal{H} = \mathcal{H}_1 \oplus (\mathcal{H}_1 \otimes \mathcal{H}_1) \oplus (\mathcal{H}_1 \otimes \mathcal{H}_1 \otimes \mathcal{H}_1) \oplus \cdots.$$

Suppose one particle is observed in the volume V. Strictly speaking, this says that some one particle is in V and all other particles are not in V. This observation corresponds to the operator

$$\begin{aligned} \bar{p}_v = p_v &+ p_v \otimes (1 - p_v) + (1 - p_v) \otimes p_v \\ &+ p_v \otimes (1 - p_v) \otimes (1 - p_v) + \cdots. \end{aligned}$$

The subspace $\bar{p}_v\mathcal{H}_1$ includes $p_v\mathcal{H}_1$, which is infinite-dimensional by the previous argument. It also includes tensor products of $p_v\mathcal{H}_1$ with $(1 - p_v)\mathcal{H}_1$, which are again infinite-dimensional. Thus

$$\dim [\bar{p}_v\mathcal{H}_1] = \infty.$$

We conclude from these two examples that a projection p measurable in a finite region α is likely to be infinite for two reasons. First, the event corresponding to the eigenvalue 1 of p could occur within α in an infinite number of disjoint ways. We see this

[6] G. Ludwig, *Die Grundlagen der Quantenmechanik* (Springer-Verlag, Berlin, 1954), Chap. II, Sec. 3. Also see Ref. 1, Theorem 1.

[7] M. Guenin and B. Misra, Nuovo Cimento **30**, 1272 (1963), Corollary to Theorem A.

in the first example, where the particle in V could be in any of an infinite number of disjoint subvolumes. Secondly, an infinite number of disjoint events could occur outside α, as in the second example, where particles could be distributed in an infinite number of ways outside V.

3. FIXED OBSERVABLES

Suppose the operator A is such that $E(TS, A)$ is independent of T, as T runs over all backgrounds such that TS is possible. Then we say that S *determines* A and write

$$S \to A. \tag{3.1}$$

If A is self-adjoint, then a sequence of independent trials that gives the mean value $E(TS, A)$ also gives as higher moments the expectation values $E(TS, A^n)$, for $n = 2, 3, \cdots$.

If A is not self-adjoint, we would measure $E(TS, A)$ by combining the mean values for the self-adjoint operators $A + A^\dagger$, $iA - iA^\dagger$. Suppose we measure independently all the self-adjoint operators $e^{i\theta}A + e^{-i\theta}A^\dagger$. By suitable combinations of the higher moments we could determine the expectation values $E(TS, B)$, for all operators B in the ring $\mathcal{R}_f(A)$ of finite polynomials in A, $A^\dagger$ and the unit operator.

It is clear from Eq. (2.1) that $S \to A$ implies $S \to A^\dagger$, and of course always $S \to 1$. However, there may be other elements of $\mathcal{R}_f(A)$ not determined by S. The measurement of A is then not completely independent of T. If, however, $S \to B$ for every $B \in \mathcal{R}_f(A)$, then we say that S *fixes* A and we write

$$S \Rightarrow A. \tag{3.2}$$

Let $F(S)$ denote the set of observables fixed by the state S. In this section we will derive a complete characterization of $F(S)$ in the case when $S = \dot{p}$. We will show in Sec. 5 that for a wide class of local states the set $F(S)$ can be considered in terms of this special case.

Proposition 3: Let p be a projection, and let A be some operator. Then $\dot{p} \to A$ if and only if there is a scalar a such that

$$pAp = ap. \tag{3.3}$$

Proof: Sufficiency: If $T\dot{p}$ is a possible state, then

$$E(T\dot{p}, A) = E(T, pAp)[E(T, P)]^{-1} = a, \quad \text{independently of } T.$$

This number a is the expectation value for A, given a measured value for p equal to 1.

Necessity: Let $\rho_T = |\psi)(\psi|$, with $(\psi, p\psi) \neq 0$. $E(T\dot{p}, A) = (\psi, pAp\psi)(\psi, p\psi)^{-1} = a$, say, independent of ψ by hypothesis. Therefore

$$(\psi, (pAp - ap)\psi) = 0,$$

for any ψ. This implies, by polarization,[8] that

$$pAp = ap.$$

Remarks: In the particular case when A is a projection, $0 \leq a \leq 1$. We then call a "the probability of A given p" and write

$$a = P(p \mid A).$$

Let $\mathcal{N}$ be some set of operators. Suppose $\dot{p}$ determines each element of $\mathcal{N}$, that is,

$$\dot{p} \to \mathcal{N}.$$

Let φ be some vector such that $p\varphi \neq 0$. By Proposition 3, for all $B \in \mathcal{N}$,

$$pBp = (\varphi, pBp\varphi)(\varphi, p\varphi)^{-1}p.$$

This equation is weakly continuous.[9] We can therefore extend it to the weak closure $\overline{\mathcal{N}}$ of $\mathcal{N}$. Since $\overline{\mathcal{N}} \supset \mathcal{N}$, we see that:

Corollary 1: $\dot{p} \to \mathcal{N}$ if and only if $\dot{p} \to \overline{\mathcal{N}}$.

Let $\mathcal{R}(A)$ denote the weak closure of the ring $\mathcal{R}_f(A)$. It is a von Neumann algebra.[10] We also see that:

Corollary 2: $\dot{p} \Rightarrow A$ if and only if $\dot{p} \to \mathcal{R}(A)$.

Proposition 4: Let $\mathcal{M}$ be some von Neumann algebra, and let p be some nonzero projection in $\mathcal{M}$. Then $\dot{p} \to \mathcal{M}$ if and only if p is minimal for $\mathcal{M}$.[11]

Proof: Necessity: Let r be some projection in $\mathcal{M}$, $0 < r \leq p$. By Proposition 3 there exists a scalar ρ such that

$$prp = \rho p.$$

But $prp = r$, and we must have $r = p$. Thus p is minimal.

Sufficiency: Let A be some positive operator in $\mathcal{M}$. Consider the positive operator

$$B = pAp.$$

Since $Bp = pB = B$, there is a spectral resolution of

[8] F. Riesz and B. Sz-Nagy, *Leçons D'analyse Fonctionelle*, (Akadémiai Kiadó, Budapest, 1953), 2nd ed., p. 227, Sec. 92, Eq. (2).

[9] J. Dixmier, *Les algebres d'operateurs dans l'espace Hilbertien* (Gauthier-Villars, Paris, 1957), p. 33. Referred to in the text as Dixmier.

[10] Reference 9, Chap. I, Sec. 3.4, p. 44, Theorem 2, Corollary I.

[11] Reference 9, Chap. I, Sec. 8.2, p. 122, Def. 2.

p, $F(\lambda)$ in $\mathcal{M}$,[12] such that $F(\lambda) \leq p$, $F(\|B\|) = p$, and

$$B = \int_0^{\|B\|} \lambda \, dF(\lambda).$$

The projection p is minimal; therefore $F(\lambda) = 0$ for $\lambda < \|B\|$ and

$$B = \|B\| \, p.$$

Any operator $A \in \mathcal{M}$ may be written as a sum of positive operators; therefore $pAp = \alpha p$ for some scalar α. By Proposition 3, $p \to \mathcal{M}$.

Proposition 5: Let $\mathcal{N}$ be a von Neumann algebra fixed by p. Let $\mathcal{M}$ be the von Neumann algebra generated by p and $\mathcal{N}$. Then (a) $p \to \mathcal{M}$ and p is minimal for $\mathcal{M}$; (b) $\mathcal{M}$ is a direct sum $\mathcal{M} = \mathcal{M}_s + \mathcal{M}_p$, with $\mathcal{M}_s p = 0$, $p \in \mathcal{M}_p$; (c) $\mathcal{M}_p$ is isomorphic to $\mathfrak{L}(\mathcal{K})$, the ring of all bounded operators on some Hilbert space $\mathcal{K}$.

Proof: Part (a): Let $\mathcal{T}$ denote the ring of finite polynomials in p and $\mathcal{N}$. The ring $\mathcal{M}$ is weakly generated by $\mathcal{T}$. By repeated application of Proposition 3, we see that $p \to \mathcal{T}$. By Corollary 2 to Proposition 3, we see that $p \to \mathcal{M}$. Proposition 4 then shows that p is minimal for $\mathcal{M}$.

Part (b): Let C be the central support of p.[13] Let $\mathcal{M}_s = (1 - C)\mathcal{M}$, $\mathcal{M}_p = C\mathcal{M}$, and Part (b) is immediate.

Part (c): This is essentially Example 4 of Dixmier (Ref. 9, Chap. I, Sec. 8, p. 126). For completeness we give the proof in Appendix B.

Remark: Let Φ denote the isomorphism that takes $\mathcal{M}_p$ onto $\mathfrak{L}(\mathcal{K})$. Clearly Φ takes minimal projections onto minimal projections. The minimal projections of $\mathfrak{L}(\mathcal{K})$ are just the one-dimensional projections. We can find an orthonormal basis for $\mathcal{K}$, $\{\varphi_n, n = 1, 2, \cdots\}$[14,15] such that

$$\Phi(p) = |\varphi_1)(\varphi_1|.$$

The operators $|\varphi_n)(\varphi_m|$ span $\mathfrak{L}(\mathcal{K})$. Define

$$\Phi^{-1}(|\varphi_n)(\varphi_m|) = V_{nm}.$$

Clearly

$$\begin{aligned} V_{11} &= p, \\ V_{nm}V_{kl} &= \delta_{m,k}V_{nl}, \\ V_{nm}^{\dagger} &= V_{mn}. \end{aligned} \tag{3.4}$$

Any operator $B \in \mathcal{M}_p$ can be written as

$$B = \sum c_{nm}V_{nm},$$

with scalars c_{nm} determined by

$$V_{nn}BV_{mm} = c_{nm}V_{nm}.$$

A set of operators $\{V_{nm}\}$ that satisfy Eq. (3.4) will be called a set of *matrix units*.[16]

This leads to the following complete characterization of all operators fixed by a given projection p.

Proposition 6: $\dot{p} \Rightarrow A$ if and only if $A = A_s + A_p$, where $A_sA_p = A_pA_s = pA_s = A_sp = 0$, and there exists a set of matrix units V_{nm} as in Eq. (3.4) and scalars c_{nm} such that

$$A_p = \sum c_{nm}V_{nm}.$$

Proof: Necessity: Take $\mathcal{N} = \mathcal{R}(A)$ and apply Proposition 5. Let A_s to be the part of A in $\mathcal{M}_s$, A_p that part in $\mathcal{M}_p$, and the result follows by the above.

Sufficiency: The orthogonality conditions imply that $\mathcal{R}(A)$ is the direct sum $\mathcal{R}(A_s) + \mathcal{R}(A_p)$, with $p\mathcal{R}(A_s) = 0$. The weakly closed ring generated by the V_{nm}, $\mathcal{R}(\{V_{nm}\})$, is clearly fixed by $\dot{p}$ and contains $\mathcal{R}(A_p)$. Thus $\dot{p} \to \mathcal{R}(A_p)$, and therefore $\dot{p} \Rightarrow A$.

4. FIXED PROJECTIONS

An important subset of $F(\dot{p})$ consists of the projections fixed by $\dot{p}$. In this section we will construct explicitly the decomposition of such projections according to Proposition 6.

Proposition 7: Suppose $\dot{p} \to q$. Let $a = P(p \mid q)$. Then, as in Proposition 6, $q = q_s + q_p$ where (a) $q_p = qpqa^{-1}$; (b) $\dot{q}_p \to p$ and $\dot{p} \to q_p$; (c) $q_p = aV_{11} + [a(1 - a)]^{\frac{1}{2}}(V_{12} + V_{21}) + (1 - a)V_{22}$, where the operators

$$V_{11} = p, \quad V_{22} = (1 - p)q_p(1 - p)(1 - a)^{-1},$$
$$V_{21} = (1 - p)q_pp[a(1 - a)]^{-\frac{1}{2}}$$

are matrix units.

Proof: Part (a): The ring $\mathcal{R}(q)$ is the set of all operators of the form $\alpha 1 + \beta q$, for all scalars α, β. Therefore the relation $\dot{p} \to q$ is equivalent to $\dot{p} \Rightarrow q$. Let $\mathcal{M}$ denote the ring generated by p and $\mathcal{R}(q)$, and let C be the central support of p in $\mathcal{M}$. According to Proposition 6,

$$q = q_s + q_p, \quad \text{with} \quad q_s = (1 - C)q,$$
$$q_p = Cq.$$

[12] Consider the operator B as restricted to $p\mathcal{K}$, and apply the spectral resolution theorem of Ref. 8, Sec. 107, p. 272, and Proposition 2 of Ref. 9, Chap. I, Sec. 1.2, p. 3.

[13] Reference 9, Chap. I, Sec. 1.3, p. 7. The central support of p is the least central projection greater than p.

[14] In general $\mathcal{K}$ does not have to be separable. However, we assume the over-all Hilbert space to be separable. This forces $\mathcal{M}_p$ to be countably decomposable [Ref. 9, Chap. I, Sec. 1]. The set $\{V_{nn}\}$ must then be countable, and $\mathcal{K}$ therefore separable.

[15] Reference 9, Chap. I, Sec. 8, Example 2, p. 126.

[16] I. Kaplansky, Ann. Math. 56, 460 (1952).

By Proposition 3, $pqp = ap$. Consider the operator

$$e = qpqa^{-1}.$$

This is a projection, as

$$e^2 = qpqqpqa^{-2} = q(pqp)qa^{-2} = e.$$

Also

$$e = qCpqa^{-1} = q_p pq_p a^{-1} = Ce.$$

Thus $e < C$ and $< q_p$. The operator $e' = q_p - e$ is clearly a projection $< q_p$ and $< C$. Since

$$\begin{aligned} e'p &= C(q - qpqa^{-1})p \\ &= 0, \end{aligned}$$

we see that e' commutes with both p and q. It is therefore in the center of $\mathcal{M}C$. In Proposition 5 it is shown that $\mathcal{M}_C$ is a factor. Therefore $e' = C$ or 0. It is orthogonal to p; therefore $e' \neq C$. Thus $e' = 0$ and

$$q_p = qpqa^{-1}.$$

Part (b): By Part (a),

$$\begin{aligned} pq_p p &= p(qpq)pa^{-1} = (pqp)(pqp)a^{-1} \\ &= ap; \end{aligned}$$

thus $\dot{p} \to q_p$,

$$\begin{aligned} q_p pq_p &= (qpq)p(qpq)a^{-2} = q(pqp)(pqp)qa^{-2} \\ &= aq_p, \end{aligned}$$

and $\dot{q}_p \to p$.

Part (c) follows immediately from Parts (a) and (b).

Remark: If two projections p, q are such that both $\dot{p} \to q$ and $\dot{q} \to p$, then we say that both p and q are *linked* and write $p \leftrightarrow q$. In Proposition 7 we see that every projection fixed by p breaks up into a part orthogonal to p and a part linked to p. For linked projections the following interesting reciprocity relation holds:

Proposition 8: If $p \leftrightarrow q$, then $P(p \mid q) = P(q \mid p)$.

Proof: By Proposition 3, the respective probabilities are the constants a, b in the equations

$$pqp = ap, \tag{4.1}$$

$$qpq = bq. \tag{4.2}$$

If $a = 0$, it follows that q and p are orthogonal and $b = 0 = a$. Suppose now that $a \neq 0$. Consider the operator $W = qpa^{-\frac{1}{2}}$. By Eq. (4.1)

$$W^\dagger W = p;$$

therefore W is a partial isometry.[17] The product $WW^\dagger$ must therefore be a projection. It is $ba^{-1}q$, by Eq. (4.2), and this is a projection only if $b = a$.

5. LOCAL STATES

In the previous sections we have investigated those local states that can be prepared by the measurement of a projection. We will consider in this section a wider class of local states.

In the following, the von Neumann algebra of observables based on a region α will be denoted $R(\alpha)$.[18] The spacelike complement of α will be denoted by α'. The ring $R(\alpha')$ will be assumed to be in the commutant of $R(\alpha)$, i.e.,

$$R(\alpha') \subset R'(\alpha).$$

The commutant of $R(\alpha')$, $R'(\alpha')$ then includes $R(\alpha)$.

There are two main types of local states, selective and nonselective.[1] In a pure-selective state $\dot{p}$, a projection p is measured in a finite space–time region α and those independent trials are selected in which the measured value is one. A selective state takes the background density matrix ρ_T into[6]

$$\rho_{T\dot{p}} = p\rho_T p(\mathrm{Tr}\ \rho_T p)^{-1}. \tag{5.1}$$

In a nonselective state $\hat{S}$, some action is performed in the region α, but no selection of trials is made. The density matrix ρ_T is taken into[19]

$$\rho_{T\hat{S}} = \sum_{n=1}^{\infty} A_n \rho_T A_n^\dagger, \tag{5.2}$$

where $\{A_n, n = 1, 2, \cdots\}$ is a sequence of operators in $R'(\alpha')$ such that

$$\sum_n A_n^\dagger A_n = 1. \tag{5.3}$$

We will consider in this section local states S_α set up in the region α by preparing a finite sequence of selective and nonselective states. For example, suppose $S_\alpha = \dot{p}\hat{S}$. Applying in turn Eqs. (5.2), (5.1), and (5.3), we find

$$\begin{aligned} \rho_{T\dot{p}\hat{S}} &= \sum_n A_n \rho_{T\dot{p}} A_n^\dagger \\ &= \sum_n A_n p\rho_T p A_n^\dagger (\mathrm{Tr}\ \rho_T p)^{-1} \\ &= \sum_n A_n p\rho_T p A_n^\dagger \left(\mathrm{Tr}\ \rho_T \sum_m p A_m^\dagger A_m p\right)^{-1}. \end{aligned}$$

By repeated application of Eqs. (5.1)–(5.3), it can be shown that for any such state S_α there is a sequence of

[17] M. A. Naimark, *Normed Rings*, translated from the first Russian edition by L.F. Boron (P. Noordhoff, Ltd., Groningen, The Netherlands, 1959), Chap. I, Sec. 5.14, p. 112.

[18] R. Haag, *Colloque internationale sur les problemes mathematique de la theorie quantique des champs, Lille, 1957* (Centre National de la Recherche Scientifique, Paris, 1959); H. Araki, Lecture Notes, University of Zurich (1961); H. Araki, Progr. Theoret. Phys. (Kyoto) **32**, 844 (1964); H. J. Borchers, Lecture Notes, Princeton University (1966).

[19] Reference 1, Theorem 5.

operators B_n in $R'(\alpha')$ such that

$$\rho_{\mathrm{TS}_\alpha} = \sum_{n=1}^{\infty} B_n \rho_T B_n^\dagger \Big(\mathrm{Tr}\ \rho_T \sum_n B_n^\dagger B_n \Big)^{-1}, \qquad (5.4)$$

where $\sum_n B_n^\dagger B_n$ is a bounded operator.

We will show in Proposition 9 that no states S_α of the above form can fix all operators. We will then investigate in Proposition 10 the local operators fixed by such a state.

Proposition 9: Let S_α be a local state preparable in the finite region α and specified as in Eq. (5.4) by the sequence $\{B_n \in R'(\alpha'), n = 1, 2, \cdots\}$. Then S_α cannot fix all operators.

Proof: Let A be some operator. If S_α determines A, then the expectation value

$$E(\mathrm{TS}_\alpha, A) = \mathrm{Tr} \Big(\rho_T \sum_n B_n^\dagger A B_n \Big) \Big(\mathrm{Tr}\ \rho_T \sum_n B_n^\dagger B_n \Big)^{-1}$$

is independent of T. Just as in Proposition 3, this is equivalent to

$$\sum_n B_n^\dagger A B_n = a \sum_n B_n^\dagger B_n, \qquad (5.5)$$

for some scalar a. Suppose S_α fixes all operators. Then in particular it fixes all $A \in R(\alpha')$. For such A, Eq. (5.5) implies that

$$\sum_n B_n^\dagger B_n (A - a1) = 0. \qquad (5.6)$$

If $R(\alpha')$ were a factor,[20] this would imply that

$$A = a1,$$

for all $A \in R(\alpha')$, and $R(\alpha')$ would be the trivial ring of scalars, which we assume is not the case. Thus we need only consider the case when $R(\alpha')$ is not a factor. There then exists a projection G in the center $\mathfrak{Z}$ of $R(\alpha')$ such that[21]

$$GA = \alpha G \qquad (5.7)$$

and

$$G \sum B_n^\dagger B_n = \sum B_n^\dagger B_n. \qquad (5.8)$$

Equation (5.8) implies that

$$GB_n = B_n, \quad \text{for all } n, \qquad (5.9)$$

and we can take G to be the smallest central projection such that Eq. (5.9) holds. Thus Eq. (5.7) holds for all $A \in R(\alpha')$ and G fixed. Let Ω denote the vacuum vector. Then[22]

$$[R(\alpha')\Omega] = \mathcal{H}. \qquad (5.10)$$

By Eq. (5.7),

$$GA\Omega = (\Omega, GA\Omega)(\Omega, G\Omega)^{-1} G\Omega.$$

Let $G\Omega\, \|G\Omega\|^{-1} = \Phi$. Then $GA\Omega = (\Phi, A\Omega)\Phi$. By Eq. (5.10) the vectors $A\Omega$ are dense in $\mathcal{H}$. Therefore

$$G = |\Phi)(\Phi|,$$

and thus G is a finite-dimensional projection. But $G \in \mathfrak{Z} \subset R(\alpha')$ and, as shown in Appendix A, $\dim [G\mathcal{H}] = \infty$, a contradiction.

Definition: Let β be some finite space–time region. Let $F_\beta(S_\alpha)$ denote the set of all operators A_β in the local $R(\beta)$ such that

$$S_\alpha \Rightarrow A_\beta.$$

The following proposition shows that this set can be completely characterized by Proposition 6.

Proposition 10: Let γ be some finite region spacelike relative to $\alpha \cup \beta$. Let $\delta = \alpha \cup \gamma$. There exists a projection $q \in R'(\delta')$ such that

$$F_\beta(S_\alpha) = R(\beta) \cap F(q).$$

Proof: Let $A_0 \in F_\beta(S_\alpha)$, then $S_\alpha \Rightarrow A_0$ and Eq. (5.5) holds for all A in $\mathfrak{R}(A_0)$. Let $H = \sum_n B_n^\dagger B_n$.

It is known that $R(\gamma)$ is of infinite type.[23,24] In Appendix A we show that therefore there exists an infinite sequence of orthogonal projections, summing to 1, and each equivalent to 1 mod $R(\gamma)$. There exist then partial isometries $W_n \in R(\gamma)$ such that

$$\begin{aligned} W_n^\dagger W_n &= \delta_{n,m} 1, \\ \sum_n W_n W_n^\dagger &= 1. \end{aligned} \qquad (5.11)$$

Consider the operator

$$B = \sum_n B_n W_n.$$

Since γ is spacelike relative to α,

$$\begin{aligned} B^\dagger B &= \sum_{nm} B_n^\dagger W_n^\dagger W_m B_m \\ &= \sum_n B_n^\dagger B_n = H. \end{aligned}$$

Thus $\|B^\dagger B\| = \|H\|$, and B must be a bounded operator, with $\|B\| = \|H\|^{\frac{1}{2}}$. The operators W_n are in $R(\gamma)$ and the B_n are in $R'(\alpha')$. Thus $B \in \{R'(\alpha'), R(\gamma)\}''$. The region $\delta = \alpha \cup \gamma$ includes both α and γ. It follows that $\delta' \subset \alpha'$. By Assumption (A1) of Appendix A,

$$R(\delta) \supset R(\gamma)$$

[20] J. von Neumann and F. J. Murray, Ann. Math. **37**, 116 (1936), Corollary to Theorem III.
[21] Reference 9, Chap. I, Sec. 2, Example 6.
[22] H. Reeh and S. Schlieder, Nuovo Cimento **22**, 1051 (1961).
[23] Reference 7, Theorem A.
[24] R. V. Kadison, J. Math. Phys. **4**, 1511 (1963).

and $R(\delta') \subset R(\alpha')$, which implies

$$R'(\delta') \supset R'(\alpha').$$

Since also $R'(\delta') \supset R(\delta)$, we have

$$R'(\delta') \ni B.$$

We claim that Eq. (5.5) is equivalent to

$$B^\dagger AB = aH. \tag{5.12}$$

For, since γ is also spacelike relative to β,

$$B^\dagger AB = \sum_{nm} B_n^\dagger A W_n^\dagger W_m B_m$$
$$= \sum_n B_n^\dagger A B_n.$$

By the polar-decomposition theorem,[25] there exists a partial isometry $W \in R'(\delta')$ such that

$$B = WH^{\frac{1}{2}},$$

where $W^\dagger W = E$ projects onto the range of H and $WW^\dagger = q$ projects onto the range of B.

Let $F(\lambda)$ denote the spectral resolution of H[26]:

$$H = \int_0^{\|H\|} \lambda \, dF(\lambda).$$

Note that $E = 1 - F(0)$. For $\eta > 0$, consider the operators

$$G_\eta = \int_\eta^{\|H\|} \lambda^{-\frac{1}{2}} \, dF(\lambda).$$

We have

$$HG_\eta^2 = G_\eta^2 H = G_\eta H G_\eta = 1 - F(\eta).$$

The upper strong continuity of the spectral resolution $F(\lambda)$ implies that

$$E = \text{strong limit}_{\eta \to 0} \, HG_\eta^2$$
$$= \text{strong limit}_{\eta \to 0} \, H^{\frac{1}{2}} G_\eta.$$

From Eq. (5.12) we get

$$G_\eta H^{\frac{1}{2}} W^\dagger A W H^{\frac{1}{2}} G_\mu = a G_\eta H G_\mu.$$

Taking the strong limits as first η and then $\mu \to 0$, this becomes

$$W^\dagger A W = aE.$$

Premultiplying by W and postmultiplying by $W^\dagger$ yields

$$qAq = aq. \tag{5.13}$$

Equation (5.5) thus implies Eq. (5.13). It is actually equivalent to Eq. (5.13), as may be seen by premultiplying Eq. (5.13) by $H^{\frac{1}{2}} W^\dagger$ and postmultiplying by $WH^{\frac{1}{2}}$.

Thus $S_\alpha \to A$ if and only if Eq. (5.13) holds. By Proposition 3, this equation is, however, exactly the condition that $\dot{q} \to A$. The operator A is an arbitrary element of $\mathcal{R}(A_0)$, and thus $\dot{q} \Rightarrow A_0$ if and only if $S_\alpha \Rightarrow A_0$. Since A_0 is an arbitrary element of $F_\beta(S_\alpha)$, we see that

$$F_\beta(S_\alpha) = R(\beta) \cap F(\dot{q}).$$

6. DISCUSSION

We have seen that a local state composed of a finite sequence of selective and nonselective local states will not fix all operators. According to Proposition 10 it will fix in each region just those operators that are also fixed by a pure-selective state.

Propositions 5 and 6 serve to characterize all operators A that are fixed by a pure-selective state $\dot{q}$. Essentially, A must be such that q is a minimal projection for the von Neumann algebra generated by A and q.

It is an open question at present whether a local state of the above type exists that will fix all operators measurable in a finite space–time region α. This would require the existence of a projection q measurable in some larger region β that was minimal for the ring $\{R(\alpha), q\}''$.

ACKNOWLEDGMENT

The author wishes to express his gratitude for the hospitality extended to him by the late J. R. Oppenheimer at the Institute for Advanced Study.

APPENDIX A

Let α be any region such that its spacelike complement α' contains an open subregion. Let p be any projection in the local ring $R(\alpha)$. We will prove here that $\dim [p\mathcal{H}] = \infty$.

We assume the standard axioms for local rings.[18] In particular, we assume:

Isotony: $\alpha \subset \beta$ implies that $R(\alpha) \subset R(\beta)$, (A1)

Locality: $R(\alpha') \subset R'(\alpha)$, (A2)

and we assume the existence of a common vector Ω, cyclic and separating for all the local rings[22]

$$[R(\alpha)\Omega] = \mathcal{H}. \tag{A3}$$

We do not assume that these rings are factors.

By Kadison's Lemma (2),[23,24] Assumptions (A1) and (A3) imply that the rings $R(\alpha)$ are not of finite type. The same lemma can be seen to show that no direct summand of a local ring $R(\alpha)$ can be of finite type. Therefore by Ref. 9 (Chap. I, Sec. 6.7, Proposition 8, p. 97), the rings $R(\alpha)$ must be properly infinite.[27]

[25] J. von Neumann, Ann. Math. 33, 294 (1932); Ref. 9, Appendix III and p. 5.

[26] Reference 8, Sec. 107.

[27] "Proprement infini" in Ref. 8. A properly infinite factor would be either type I_∞, II_∞, or III_∞.

For the regions α considered here, there always exist open regions β, γ, δ such that

$$\beta \supset \alpha, \quad \gamma \subset \beta \cap \alpha', \quad \text{and} \quad \delta \subset \beta'.$$

We will show that any projection $p \in R(\alpha)$ is properly infinite in $R(\beta)$.[28] The result will then follow. For then[29]

$$p = \sum_{n=0}^{\infty} p_n,$$

the p_n all orthogonal and nonzero. Since

$$p\mathcal{H} = \sum \oplus p_n\mathcal{H}, \quad \dim [p\mathcal{H}] = \infty.$$

Following Borchers,[30] we consider the projection

$$q\mathcal{H} = [R(\beta)p\Omega].$$

Borchers shows that q is the central support of p in $R(\beta)$ and also that $P \sim q \bmod R(\beta)$. By Ref. 9, (Chap. I, Sec. 6.7, Proposition 7, p. 97), the projection q must be properly infinite in $R(\beta)$. Therefore p must be properly infinite in $R(\beta)$.

The identity operator is in each $R(\alpha)$. Therefore the identity is properly infinite in each $R(\alpha)$.[28] Applying Ref. 29, we see that the identity can be written as a sum over an infinite sequence of nonzero orthogonal projections, each equivalent to 1 mod $R(\alpha)$.

[28] Reference 9, Chap. III, Sec. 2.1, p. 241, Definition 1.

[29] Reference 9, Chap. III, Sec. 8.6, Theorem 1, p. 319, Corollary 2.

[30] H. J. Borchers, "A Remark on a Theorem of B. Misra," preprint, Institut für Theoretische Physik der Universität Göttingen (1967), Theorem III. 3.

APPENDIX B

We give here the proof of Part (c) of Proposition 5.

Part (c): The projection C is minimal for the center $\mathfrak{Z}$ of $\mathcal{M}$. For if not, there exists $r \in \mathfrak{Z}$ or $r < C$. By Part (a) and Proposition 3, there exists a scalar ρ such that $prp = \rho p$. But $r \in \mathfrak{Z}$; therefore $prp = rp$ and $rp = \rho p$. This is possible only if $\rho = 0$ or 1. In the first case the central projection $C - r < C$ is greater than p. In the second case $r < C$ is greater than p. In either case we have a contradiction to the definition of C.

Thus C is minimal for $\mathfrak{Z}$. The center of $\mathcal{M}_p$ is $\mathfrak{Z}C$ and must therefore consist of just the scalar multiples of C. The ring $\mathcal{M}_p$ is isomorphic to $\mathcal{M}_p$ restricted to $C\mathcal{H}$, $\mathcal{M}_p|_C$.[31] The center of $\mathcal{M}_p|_C$ is then the scalar multiples of the identity, and $\mathcal{M}_p|_C$ is a factor. It is clear that $p|_C$ is minimal for $\mathcal{M}_p|_C$. By Dixmier (Ref. 9, Chap. I, Sec. 8, Theorem I, Corollary 3, p. 124) $\mathcal{M}_p|_C$ is isomorphic to $\mathfrak{L}(\mathcal{K})$, for some Hilbert space $\mathcal{K}$. Therefore $\mathcal{M}_p$ is isomorphic to $\mathfrak{L}(\mathcal{K})$.

[31] Reference 9, Chap. I, Sec. 2.1, Proposition 2, p. 19.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Coherent Superposition of Local States*

A. L. LICHT
U.S. Naval Ordnance Laboratory, Silver Spring, Maryland

(Received 29 March 1967)

The notion of coherent superposition is generalized so as to apply to local states.

1. INTRODUCTION

A unit vector state $|\chi)$ can always be formally constructed by superimposing two other vector states,

$$|\chi) = [a\,|\psi) + b\,|\varphi)]\;\|a\,|\psi) + b\,|\varphi)\|^{-1}.$$

In practice $|\chi)$ would be set up by measuring the value one for the projection $p_\chi = |\chi)(\chi|$.

This process of coherent superposition applies only to vector states. The projections onto single vector states are however not observable in finite regions of space–time.[1,2] A finite observer therefore cannot construct a vector state from an arbitrary background, but only from a background that is already a vector state. The coherent superposition of vector states is then strongly background dependent.

The coherent superposition of local states is, however, a common laboratory practice, in which the influence of the background state is generally felt to be negligible. This laboratory superposition cannot therefore correspond exactly to the simple addition of vector states.

In the following we will derive a generalized definition of coherent superposition that does apply to local states. In Sec. 2 we investigate a typical example of coherent superposition in a finite laboratory. This leads to a general definition in terms of sets of linked projections[3] in Sec. 3. Proposition 1 in Sec. 4 gives the mathematical structure of these sets. Proposition 2 in Sec. 5 shows the relationship of the general coherent superposition to the usual addition of vector states.

Notation: The results and notation of Refs. 2 and 4 will be used throughout. To each observable there corresponds an operation on a Hilbert space $\mathcal{H}$. The measurement of the value one for a projection p produces a state denoted by $\dot{p}$. If this measurement takes place in the presence of a background state T, described by a density matrix ρ_T, the total state we denote by $T\dot{p}$. It is described by the density matrix

$$\rho_{T\dot{p}} = p\rho_T p/\mathrm{Tr}\,(\rho_T p).$$

An observable A is said to be *determined* by the state $\dot{p}$,

$$\dot{p} \to A,$$

if the expectation value of A in the state $T\dot{p}$ is independent of T. It is said to be *fixed*[5] by $\dot{p}$,

$$\dot{p} \Rightarrow A,$$

if every operator B in the von Neumann algebra $\mathfrak{R}(A)$ generated by A is determined by $\dot{p}$.

The relations $\dot{p} \to q$ and $\dot{p} \Rightarrow q$ are equivalent if q is a projection. The expectation value of q in a state $T\dot{p}$ then depends entirely on p. We call it the probability of q given p, and write $P(p\,|\,q)$. If two projections p, q are such that $\dot{p} \to q$ and $\dot{q} \to p$, then we say that they are *linked*,[3] and write

$$p \leftrightarrow q.$$

2. ELECTRON SPIN

Consider the apparatus shown in Fig. 1. Enclosed in an evacuated chamber are an electron gun G, an electron counter C, and several devices for measuring spin $X^{1,2}_{\uparrow,\downarrow}, Z_{\uparrow,\downarrow}, Y_{\uparrow,\downarrow}$. Electrons are emitted by G in a well-collimated beam directed along the y axis. The counter C records each electron without absorbing it or causing it to deviate appreciably from its path. The counters $Z_\uparrow$, $Z_\downarrow$ when turned on record the passage of an electron with spin up, down, respectively, relative to the z axis. The counters $X^{1,2}_{\uparrow,\downarrow}, Y_{\uparrow,\downarrow}$ do the same for spin relative to the x and y axes. The x spin may be measured either immediately before the z-spin measurement in X^1, or immediately afterward in X^2. We reserve the right to introduce other counters into the chamber without changing the diagram.

Let c denote the projection corresponding to the observation of an electron by C. Let $x^i_\uparrow, x^i_\downarrow$, $i = 1, 2$ be the projections corresponding to finding the x spin either up or down in counters $X^{1,2}$. Let $z_\uparrow, z_\downarrow, y_\uparrow, y_\downarrow$ be the corresponding projections for the z and y spin.

* Part of this work was done in 1966 at the Institute for Advanced Study, Princeton, New Jersey.

[1] M. Guenin and B. Misra, Nuovo Cimento **30**, 1272 (1963), corollary to Theorem A.

[2] A. L. Licht, J. Math. Phys. 9, 1468 (1968), Proposition 2. This reference will be referred to in the text as BD.

[3] Reference 2, Sec. 4.

[4] A. L. Licht, J. Math. Phys. **7**, 1656 (1966).

[5] Reference 2, Sec. 3.

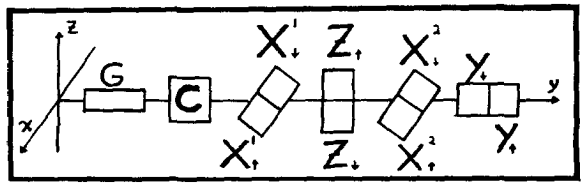


FIG. 1. Electron spin apparatus; G is an electron gun, C an electron counter; $X^{1}_{\uparrow,\downarrow}, X^{2}_{\uparrow,\downarrow}, Z_{\uparrow,\downarrow}, Y_{\uparrow,\downarrow}$ are counters that record electrons with spin up, down, relative to the x, z, y axes.

The whole apparatus is of finite size, and is assumed to be in existence for only a finite time, so these projections are all measurable in some finite space–time region α. They are then infinite projections, according to BD Proposition 2.

We assume ideal collimation and sensitivity. This is, of course, unobtainable in practice. There is always a small probability that an electron will diverge from the beam between counters and be absorbed by the chamber wall. A counter may fail to record an electron that passes through it. Spurious electrons may be emitted by the walls. We consider here the ideal case, and assume that such possibilities can be ignored.

If the z- and y-spin counters are turned off, and the x-spin counters are turned on, in this ideal apparatus X^1 will click if and only if the $X^2_\uparrow$ counter clicks. This implies[6] that

$$x^1_\uparrow = x^2_\uparrow = x_\uparrow, \quad \text{say.}$$

Similarly we must have

$$x^1_\downarrow = x^2_\downarrow = x_\downarrow .$$

If $X_\uparrow$ clicks, then $X_\downarrow$ does not and vice versa. The same holds for the pairs $Z_\uparrow, Z_\downarrow$ and $Y_\uparrow, Y_\downarrow$. This implies that the corresponding projections are orthogonal,

$$x_\uparrow x_\downarrow = y_\uparrow y_\downarrow = z_\uparrow z_\downarrow = 0.$$

Every electron counted by C will be recorded by the X, Y, or Z counters if they are turned on, and every electron recorded by a spin counter must have been counted by C. This is equivalent to

$$c = x_\uparrow + x_\downarrow = y_\uparrow + y_\downarrow = z_\uparrow + z_\downarrow .$$

Let T denote the background state. The condition of the universe outside the chamber, the temperature of the chamber walls, the stray magnetic field within the chamber, etc., are all specified by T. Suppose the counter $Z_\uparrow$ clicks. The state is then $T\dot{z}_\uparrow$. We would then expect $X^2_\uparrow$ or $X^2_\downarrow$ to click, each with probability $\frac{1}{2}$.

[6] Reference 2, Sec. 4.

Stray magnetic fields in the chamber could cause the electron spin to flip between Z and X^2. The probability of $X^2_\uparrow$ clicking would then be different from $\frac{1}{2}$ by an amount depending on T. In a magnetically well-shielded chamber the probability will be very close to $\frac{1}{2}$ for a wide range of T. In the ideal chamber it will be exactly $\frac{1}{2}$, independently of T. Then $z_\uparrow$ fixes $x_\downarrow, x_\uparrow$,

$$\dot{z}_\uparrow \rightarrow x_\downarrow, x_\uparrow ,$$

with $P(z_\uparrow \mid x_\downarrow) = P(z_\uparrow \mid x_\downarrow) = \frac{1}{2}$.

If counter $X^1_\uparrow$ or $X^1_\downarrow$ were to click, then in the ideal chamber we expect that $z_\uparrow$ will click with probability $\frac{1}{2}$, independently of T. Thus

$$\dot{x}_\downarrow, \dot{x}_\uparrow \rightarrow z_\uparrow ,$$

and $z_\uparrow$ is linked to $x_\uparrow$ and $x_\downarrow$.

We expect in fact that all pairs of projections in the set $S = \{z_{\uparrow\downarrow}, x_{\uparrow\downarrow}, y_{\uparrow\downarrow}\}$ are linked with probabilities $\frac{1}{2}$ or 0.

Consider the operators formed from products of projections in S, $x_\uparrow y_\uparrow$, $y_\uparrow x_\uparrow$, etc. We expect also that in this ideal apparatus the expectation values of all these operators should be fixed by each projection in S.

It is customary to regard the state $T\dot{z}_\uparrow$ as a coherent superposition of the states $T\dot{x}_\uparrow$ and $T\dot{x}_\downarrow$, and also as a coherent superposition of $T\dot{x}_\uparrow$ and $T\dot{y}_\uparrow$. By BD Proposition 2, these are not likely to be vector states, so that this type of superposition is not just the addition of vector states. We give below the mathematical structure of this type of superposition. We will show first that $\{x_\uparrow, x_\downarrow\}, \{x_\uparrow, y_\downarrow\}$ are examples of "coherent sets of projections." We will then give a general definition of the coherent superposition of such coherent sets, and we will show that $z_\uparrow$ fits this definition.

Remarks: In this ideal apparatus the two different counters $X^1_\uparrow, X^1_\uparrow$ correspond to the same operator $x_\uparrow$. They measure $x_\uparrow$ in different places and at different times. The states $T\dot{x}^1_\uparrow$ and $T\dot{x}^2_\uparrow$ are therefore very different.

The state $T\dot{x}^1_\uparrow$ is prepared before the measurement of $z_\uparrow$. The expectation value $E(T\dot{x}^1_\uparrow, z_\uparrow)$ is therefore a *prediction.*[7] We expect it in this apparatus to equal $\frac{1}{2}$ independently of T.

The state $T\dot{x}^2_\uparrow$ is prepared after the measurement of $z_\uparrow$. The expectation value of $z_\uparrow$ in this state, denoted by $M(T\dot{x}^2_\uparrow z\, z_\uparrow)$, is therefore a *retrodiction.*[7] It can be

[7] See Ref. 6 and also S. Watanabe, Rev. Mod. Phys. 27, 179 (1955).

shown[8] to equal $E(T\dot{c}, z_\uparrow)$, and will not be independent of T.

The state $T\dot{x}_\uparrow$ is therefore not well defined. It is not completely specified until the place where $x_\uparrow$ is measured is given. We will however use $T\dot{x}_\uparrow$ as an abbreviation for either $T\dot{x}^1_\uparrow$ or $T\dot{x}^2_\uparrow$.

A retrodiction is only in certain very special cases independent of the background.[9] It follows that the relation $\dot{p} \to q$ is physically meaningful only if q can be measured after the measurement of p. We will assume this to be always possible for the projections considered here.

3. COHERENCE

Let $Q = \{q_i, i = 1, 2, \cdots\}$ be a sequence of projections. Let $\mathfrak{W}(Q)$ denote the weakly closed * ring generated by Q. The projection $q = \cup_i q_i$ is the biggest projection contained in $\mathfrak{W}(Q)$, and is not necessarily the identity.

Each operator A in $\mathfrak{W}(Q)$ we interpret as a measure of the relative coherence of the states $\dot{q}_i$. We will say that Q is a *coherent set* if all such measures are background independent, that is:

Definition 1: Q is a *coherent set* if for all $q_i \in Q$ and all $A \in \mathfrak{W}(X)$,

$$\dot{q}_i \to A.$$

Since $\mathfrak{W}(Q)$ contains the q_i's themselves, it follows from this that the q_i's must be pairwise linked.

Example 1: A collection Q_1 of one-dimensional projections $q_i = |\varphi_i)(\varphi_i|$ is automatically a coherent set. The ring $\mathfrak{W}(Q_1)$ consists of operators A of the form

$$A = \sum \alpha_{ij} |\varphi_i)(\varphi_j| \, (\varphi_i \mid \varphi_j),$$

for appropriate scalars α_{ij}.

Example 2: $Q_2 = \{x_\uparrow, x_\downarrow\}$ is a coherent set. The projections $x_\uparrow$, $x_\downarrow$ are orthogonal, therefore trivially linked.

$$\mathfrak{W}(Q_2) = \{\alpha x_\uparrow + \beta x_\downarrow, \text{ for all scalars } \alpha, \beta\}$$

Note that $x_\uparrow \cup x_\downarrow = x_\uparrow + x_\downarrow = c$.

Example 3: Consider the set $Q_3 = \{x_\uparrow, y_\uparrow\}$. In our ideal apparatus we must have $x_\uparrow \leftrightarrow y_\uparrow$. By BD Proposition 3, we then have

$$\begin{aligned} x_\uparrow y_\uparrow x_\uparrow &= \tfrac{1}{2} x_\uparrow, \\ y_\uparrow x_\uparrow y_\uparrow &= \tfrac{1}{2} y_\uparrow. \end{aligned} \tag{3.1}$$

It then follows that

$$\mathfrak{W}(Q_3) = \{a x_\uparrow + \beta y_\uparrow + \gamma x_\uparrow y_\uparrow + \delta y_\uparrow x_\uparrow, \quad \text{for all scalars } \alpha, \beta, \gamma, \delta\}.$$

Equations (3.1) imply that $Q_3 \to \mathfrak{W}(Q_3)$, and therefore Q_3 is a coherent set. Indeed, it is clear that any pair of linked projections forms a coherent set.

Note that $x_\uparrow \cup y_\uparrow = c$. For, let $e = x_\uparrow \cup y_\uparrow$. Since $c > x_\uparrow$ and $c > y_\uparrow$, we have $c \geq e$. Now $e > x_\uparrow$ and $e > y_\uparrow$. Therefore

$$\begin{aligned} e y_\uparrow &= y_\uparrow, \\ ecy_\uparrow &= cy_\uparrow, \\ e(x_\uparrow + x_\downarrow)y_\uparrow &= (x_\uparrow + x_\downarrow)y_\uparrow \\ &= x_\uparrow y_\uparrow + e x_\downarrow y_\uparrow. \end{aligned}$$

Thus

$$e x_\downarrow y_\uparrow = x_\downarrow y_\uparrow. \tag{3.2}$$

We assume that $x_\downarrow \leftrightarrow y_\uparrow$ with probability $\frac{1}{2}$. By BD Proposition 3, $x_\downarrow y_\uparrow x_\downarrow = \frac{1}{2} x_\downarrow$. Applying this to Eq. (3.2) yields

$$e x_\downarrow = x_\downarrow,$$

and $e > x_\downarrow$. We have $e > x_\uparrow$ and $e > x_\downarrow$, therefore $e \geq x_\uparrow + x_\downarrow = c$. Thus $e = c$.

Remark: Consider the projection $z_\uparrow$. We have seen that $z_\uparrow$ is linked to each projection in both Q_2 and Q_3 with probability $\frac{1}{2}$. Therefore it neither equals nor is orthogonal to any such projection. Also

$$z_\uparrow < c = x_\uparrow \cup x_\downarrow = x_\uparrow \cup y_\uparrow,$$

and we expect that in both cases

$$z_\uparrow \to \mathfrak{W}(Q).$$

This suggests the following definition:

Definition 2: Suppose that $Q = \{q_i, i = 1, 2, \cdots\}$ is a coherent set of projections. Let p be some projection. We will say that p is a coherent superposition over Q, if:

(1) $p \leftrightarrow q_i$, for all i;
(2) $\dot{p} \to \mathfrak{W}(Q)$;
(3) $P(p \mid q_i) \neq 0, 1$ for all i;
(4) $p \leq \cup_i q_i$.

These conditions (1) to (4) can be interpreted as follows. Condition (1) states that the relationship

[8] From Ref. 6, Eq. (2.14),
$M(T\dot{x}^2_\uparrow, z_\uparrow) = M(T, z_\uparrow x^2_\uparrow z_\uparrow)[M(T, z_\uparrow x_\uparrow z_\uparrow) + M(T, z_\downarrow x^2_\downarrow z_\downarrow)]^{-1}$
$= \frac{1}{2} M(T, z_\uparrow)[\frac{1}{2} M(T, z_\uparrow) + \frac{1}{2} M(T, z_\downarrow)]^{-1} = M(T, z_\downarrow)[M(T, c)]^{-1}$
$= M(T, cz\ c)[M(T, c)]^{-1} = M(T\dot{c}, z_\downarrow) = E(T\dot{c}, z_\downarrow)$.

[9] Reference 6, Sec. 3.

between p and q_i is background-independent. Condition (2) insures that the relative coherence of the $\dot{q}_i$'s is determined in the state $\dot{p}$, independently of the background. Condition (3) states that p is a coherent superposition over no subset of Q, and Condition (4) states that p is a coherent superposition over no set Q' greater than Q.

In Examples 2 and 3, $z_\uparrow$ is clearly a coherent superposition in this sense over $\{x_\uparrow, x_\downarrow\}$ or over $\{x_\uparrow, y_\uparrow\}$.

In Example 1, let p_ψ be the projection onto the vector state $\psi = \sum a_i \varphi_i$. The state $\dot{p}_\psi$ is the vector state ψ. The projection p_ψ is a coherent superposition in this sense over the q_i provided that ψ neither equals nor is orthogonal to any of the φ_i's.

Remark: If p, Q are as in Definition 1, we will occasionally say that the state $\dot{p}$ is a coherent superposition of the states $\dot{q}_i$.

If Q consists of just one element $q_i = q$, then it is clear that Conditions (1), (2), and (4) are automatically satisfied by $p = q$. Indeed, BD Proposition 8 can be used to show that if $Q = \{q\}$ then these conditions are only satisfied by $p = q$. In this case we will relax Condition (3) and say that q is a coherent superposition over q.

4. MATHEMATICAL STRUCTURE

Let p be a coherent superposition over Q. The projection p may not be in the ring $\mathfrak{W}(Q)$. For example, if the q_i's are mutually orthogonal, then it is easily seen that no projection in $\mathfrak{W}(Q)$ can be a coherent superposition over Q. Consider however the weakly closed * ring $\mathfrak{U}$ generated by p and $\mathfrak{W}(Q)$. This ring does contain p. It is completely characterized by the following proposition.

Proposition 1: (a) The projections p and q_i are minimal for $\mathfrak{U}$; and (b) $\mathfrak{U}$ is isomorphic to $\mathfrak{L}(\mathcal{K})$, for some Hilbert space $\mathcal{K}$.

Proof: The projection q is the biggest projection in the rings $\mathfrak{W}(Q)$ and $\mathfrak{U}$. These rings are therefore isomorphic to their restrictions to $q\mathcal{K}$. The restricted rings $\mathfrak{W}_q(Q)$,$\mathfrak{U}_q$ are von Neumann algebras. We can therefore apply BD Proposition 5 with $\mathcal{N} = \mathfrak{W}_q(Q)$ and $\mathcal{M} = \mathfrak{U}_q$. We see that p is minimal for $\mathfrak{U}$ and that $C\mathfrak{U}$ is isomorphic to $\mathfrak{L}(\mathcal{K})$, for some Hilbert space $\mathcal{K}$, where C is the central support of p. It remains only to show that the q_i's are minimal for $\mathfrak{U}$, and that $C = q$.

By hypothesis, $p \leftrightarrow q_i$, and $P(p \mid q_i) = a_i > 0$. By BD Proposition 3,

$$q_i = q_i p q_i a_i^{-1}.$$

Let $A \in \mathfrak{U}$. The operator $B = q_i A q_i$ is in $\mathfrak{U}$, therefore by BD Proposition 4,

$$pBp = \beta p, \quad \text{for some scalar } \beta.$$

Now

$$\begin{aligned} q_i A q_i &= a_i^{-2} q_i p q_i A q_i p q_i = a_i^{-2} q_i p B p q_i \\ &= a_i^{-2} \beta q_i p q_i = a_i^{-1} \beta q_i. \end{aligned}$$

Therefore, each $q_i \to \mathfrak{U}$, and each q_i must be minimal for $\mathfrak{U}$ by BD Proposition 4.

By definition, C is the smallest central projection in $\mathfrak{U}$ such that

$$Cp = p.$$

Consider

$$\begin{aligned} Cq_i &= Cq_i p q_i a_i^{-1} = q_i C p q_i a_i^{-1} \\ &= q_i p q_i a_i^{-1} = q_i. \end{aligned}$$

Therefore $C > q_i$ for all i, which implies $C \geq q$. But $C \in \mathfrak{U}$, therefore $C \leq q$, and we must have $C = q$

5. LINEAR SUPERPOSITION

Let Φ denote the isomorphism taking $\mathfrak{U}$ onto $\mathfrak{L}(\mathcal{K})$. Under Φ, minimal projections must go into minimal projections. Therefore, for each projection r minimal in $\mathfrak{U}$, there exists a unit vector $\Psi(r) \in \mathcal{K}$ such that

$$\Phi(r) = |\Psi(r))(\Psi(r)|.$$

In particular, let $\Phi_i = \Psi(q_i)$. Since $\Phi(q) = 1$, the Φ_i's must span $\mathcal{K}$. Therefore there is a least set of scalars $\alpha^i(r)$ such that

$$\Psi(r) = \sum \alpha^i(r) \Phi_i. \tag{5.1}$$

We see from BD Proposition 4 that each such r is a coherent superposition in the sense of Definition 2 of some subset of the q_i's. This subset is clearly specified by just those $\alpha^i(r)$'s in Eq. (5.1) which are nonzero.

Let A be any operator in $\mathfrak{U}$. From BD Proposition 3 there exists a scalar α such that

$$rAr = \alpha r, \tag{5.2}$$

and for any T such that $T\dot{r}$ is a possible state,

$$E(T\dot{r}, A) = \alpha.$$

Under the map Φ, Eq. (5.2) becomes

$$|\Psi(r))(\Psi(r)|\, \Phi(A)\, |\Psi(r))(\Psi(r)| = \alpha\, |\Psi(r))(\Psi(r)|.$$

Thus

$$E(T\dot{r}, A) = (\Psi(r)\, |\Phi(A)|\, \Psi(r)).$$

We summarize these results in the following proposition:

Proposition 2: Any minimal projection r in $\mathfrak{U}$ is a coherent superposition in the sense of Definition 2 of a subset of Q. To each such projection r there corresponds a unit vector $\Psi(r)$ in $\mathcal{K}$ that is a coherent superposition in the usual sense of the vector $\Psi(q_i)$. No possible state of the form Tr can be distinguished from $\Psi(r)$ by the measurement of an operator in $\mathfrak{U}$.

The Eq. (5.1) expresses a certain linear relation between the states r and the q_i, in terms of the corresponding vectors in $\mathcal{K}$. It is interesting to note that it can also be expressed in terms of operators on $\mathcal{K}$. Let $|\theta)$ be any unit vector in $\mathcal{K}$. Equation (5.1) is equivalent to

$$|\Psi(r))(\theta| = \sum \alpha^i(r)\, |\Phi^i)(\theta|. \tag{5.3}$$

Denote

$$s = \Phi^{-1}(|\theta)(\theta|),$$
$$V_{rs} = \Phi^{-1}(|\Psi(r))(\theta|),$$
$$V_{is} = \Phi^{-1}(|\Phi^i)(\theta|).$$

The projection s is minimal in $\mathfrak{U}$. The operators V_{rs}, V_{is} are partial isometries in $\mathfrak{U}$ that take s into r, q_i, respectively; that is,

$$V_{rs}^\dagger V_{rs} = V_{is}^\dagger V_{is} = s,$$
$$V_{rs} V_{rs}^\dagger = r, \quad V_{is} V_{is}^\dagger = q_i.$$

Equation (5.3) is equivalent to

$$V_{rs} = \sum \alpha^i(r) V_{is}. \tag{5.4}$$

We see that in general coherent superposition, the linear superposition of vectors is replaced by the linear superposition of partial isometries.

6. DISCUSSION

The operators in the weakly closed $*$ ring $\mathfrak{W}(Q)$ generated by a set of projections $Q = \{q_i\}$ we interpret as measures of the relative coherence of the states $\dot{q}_i$. If all such measures are fixed by each $\dot{q}_i$, then we say in Definition 1 that Q is a coherent set.

A state $\dot{p}$ is defined in Definition 2 as a coherent superposition of the $\dot{q}_i$. Such a state $\dot{p}$ fixes each element of $\mathfrak{W}(Q)$. The projection p is linked to each q_i, it is not orthogonal to any q_i, and it is less than the projection $\cup_i q_i$. The projection p does not equal any q_i, except when Q contains just one element.

In Proposition 1 we have seen that such a projection p and the q_i together generate a Type I ring $\mathfrak{U}$ for which they are minimal. According to Proposition 2 any minimal projection r, say, in $\mathfrak{U}$ is a coherent superposition over some subset of Q.

The measurement of an operator in $\mathfrak{U}$ cannot distinguish the state $\dot{r}$ from a vector state in a certain auxiliary Hilbert space $\mathcal{K}$. This vector state is a linear superposition in the usual sense of the vector states corresponding to the q_i. This linear superposition, when expressed in terms of operators on the over-all Hilbert space $\mathcal{K}$, yields a linear relation between partial isometries that generate the projections r and q_i.

ACKNOWLEDGMENT

The author wishes to express his gratitude for the hospitality extended to him by the late J. R. Oppenheimer at the Institute for Advanced Study.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Elastic-Wave Propagation and the Factorization of Matrix-Valued Functions of Several Complex Variables

EDGAR A. KRAUT
Science Center of North American Rockwell Corporation, Thousand Oaks, California

(Received 5 December 1967)

The solution of systems of equations of the Wiener–Hopf type leads directly to the problem of factoring matrix-valued functions of one or more complex variables. The matrix-factorization problem is solved here for matrices that can be diagonalized by a similarity transformation and also have eigenvalues whose logs are analytic and quadratically integrable in a set of strips. As an application of the technique, Green's tensor for the elastic wave equation is factored.

1. INTRODUCTION

The local particle-displacement vector $\mathbf{U}(\mathbf{x}, t)$ in a perfectly elastic solid whose volume is V and whose boundary surface is S satisfies the time-dependent elastic wave equation

$$\mu\nabla^2\mathbf{U} + (\lambda + \mu)\nabla\nabla\cdot\mathbf{U} - \rho\frac{\partial^2\mathbf{U}}{\partial t^2} = -\mathbf{f}(\mathbf{x}, t), \quad (1.1)$$

where λ and μ are the Lamé parameters characterizing the elastic material, ρ is its density, and $\mathbf{f}(\mathbf{x}, t)$ represents the local body-force density acting at position $\mathbf{x}$ and at time t. If the body force is time-harmonic and of the form

$$\mathbf{f}(\mathbf{x}, t) = \mathbf{F}(\mathbf{x}, \omega)\exp(i\omega t), \quad (1.2)$$

then one seeks solutions of (1.1) of the form

$$\mathbf{U}(\mathbf{x}, t) = \mathbf{u}(\mathbf{x}, \omega)\exp(i\omega t), \quad (1.3)$$

where

$$\mu\nabla^2\mathbf{u} + (\lambda + \mu)\nabla\nabla\cdot\mathbf{u} + \rho\omega^2\mathbf{u} = -\mathbf{F}(\mathbf{x}, \omega). \quad (1.4)$$

For an aperiodic body force turned on at $t = 0$, it is convenient to assume that the initial-particle velocity and displacement vanish at $t = 0$. Then a one-sided Laplace transform in time applied to (1.1) yields

$$\mu\nabla^2\mathbf{u} + (\lambda + \mu)\nabla\nabla\cdot\mathbf{u} - \rho s^2\mathbf{u} = -\mathbf{F}(\mathbf{x}, s), \quad (1.5)$$

where

$$\mathbf{u}(x, s) = \int_0^\infty U(\mathbf{x}, t)\exp(-st)\,dt \quad (1.6)$$

and

$$\mathbf{F}(\mathbf{x}, s) = \int_0^\infty \mathbf{f}(x, t)\exp(-st)\,dt. \quad (1.7)$$

The reduced wave equations (1.4) and (1.5) have the integral solution[1]

$$u_i(\mathbf{x}) = \int_V G_{iJ}F_J\,dV + \int_S c_{jkpq}G_{iJ}(\partial u_p/\partial\xi_q)m_k\,dS + (\partial/\partial x_q)\int_S c_{jkpq}G_{ip}u_j m_k\,dS, \quad (1.8)$$

where

$$c_{ijpq} = \lambda\delta_{ij}\delta_{pq} + \mu(\delta_{ip}\delta_{jq} + \delta_{jp}\delta_{iq}). \quad (1.9)$$

δ_{ij} is the Kronecker delta, and the summation convention on repeated indices is understood with (i, j, p, q) running from 1 to 3. For the time-harmonic case (1.4) the infinite-medium Green's tensor G_{ij} appearing in (1.8) is given by[1]

$$G_{ij}(\mathbf{x} - \boldsymbol{\xi}, \omega) = \frac{1}{4\pi\rho}\left\{\frac{1}{\omega^2}\frac{\partial^2}{\partial x_i\partial x_j}\left\{\frac{\exp(-ik_s r)}{r} - \frac{\exp(-ik_p r)}{r}\right\} + v_s^{-2}\frac{\exp(-ik_s r)}{r}\delta_{ij}\right\}, \quad (1.10)$$

where

$$r = \{(x_1 - \xi_1)^2 + (x_2 - \xi_2)^2 + (x_3 - \xi_3)^2\}^{\frac{1}{2}},$$

$k_p = \omega/v_p$, $k_s = \omega/v_s$, $\mathrm{Im}\,(k_p) \leq 0$, $\mathrm{Im}\,(k_s) \leq 0$, and $\rho v_p^2 = \lambda + 2\mu$, $\rho v_s^2 = \mu$.

The corresponding form of the infinite-medium Green's tensor for the aperiodic case (1.5) is given by

$$G_{ij}(\mathbf{x} - \boldsymbol{\xi}, s) = \frac{1}{4\pi\rho}\left\{\frac{1}{s^2}\frac{\partial^2}{\partial x_i\partial x_j}\left\{\frac{\exp(-k_p r)}{r} - \frac{\exp(-k_s r)}{r}\right\} + v_s^{-2}\frac{\exp(-k_s r)}{r}\delta_{ij}\right\}, \quad (1.11)$$

where $k_p = s/v_p$ and $k_s = s/v_s$.

In what follows we shall be interested in obtaining certain factorizations of the double-bilateral Laplace transform of the Green's tensors (1.10) and (1.11). Let $\mathfrak{L}$ represent the double bilateral Laplace-transform operator:

$$\mathfrak{L} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp(-s_1x_1 - s_2x_2)\,dx_1\,dx_2, \quad (1.12)$$

where $s_1 = \sigma_1 + i\tau_1$ and $s_2 = \sigma_2 + i\tau_2$.

[1] A. T. De Hoop, Sc.D. thesis, Technische Hogeschool, Delft, 1958.

Operating on (1.10) with $\mathfrak{L}$ yields

$$\mathfrak{L}\{G_{ij}(\mathbf{x}, \omega)\} = (\rho\omega^2)^{-1}\{L^p_{ij}K_p(s_1, s_2)\exp(|x_3|/2K_p) + L^s_{ij}K_s(s_1, s_2)\exp(|x_3|/2K_s)\}, \quad (1.13)$$

where for purposes of computing $\mathfrak{L}\{G_{ij}\}$, ξ_1, ξ_2, and ξ_3 can be equated to zero in (1.10) without any loss of generality. In (1.13) the matrices L^p_{ij} and L^s_{ij} are given by

$$L^p_{ij} = \begin{Vmatrix} s_1^2 & s_1s_2 & s_1\,\mathrm{sgn}\,(x_3)/2K_p \\ s_1s_2 & s_2^2 & s_2\,\mathrm{sgn}\,(x_3)/2K_p \\ s_1\,\mathrm{sgn}\,(x_3)/2K_p & s_2\,\mathrm{sgn}\,(x_3)/2K_p & \tfrac{1}{4}K_p^{-2} \end{Vmatrix}, \quad (1.14)$$

$$L^s_{ij} = \begin{Vmatrix} (s_2^2 + \tfrac{1}{4}K_s^{-2}) & -s_1s_2 & -s_1\,\mathrm{sgn}\,(x_3)/2K_s \\ -s_1s_2 & (s_1^2 + \tfrac{1}{4}K_s^{-2}) & -s_2\,\mathrm{sgn}\,(x_3)/2K_s \\ -s_1\,\mathrm{sgn}\,(x_3)/2K_s & -s_2\,\mathrm{sgn}\,(x_3)/2K_s & (s_1^2 + s_2^2) \end{Vmatrix}, \quad (1.15)$$

where

$$K_p(s_1, s_2) = (i/2)(s_1^2 + s_2^2 + k_p^2)^{-\frac{1}{2}}, \quad (1.16)$$

$$K_s(s_1, s_2) = (i/2)(s_1^2 + s_2^2 + k_s^2)^{-\frac{1}{2}}, \quad (1.17)$$

$$\mathrm{sgn}\,(x_3) = \begin{cases} 1; & x_3 > 0 \\ 0; & x_3 = 0 \\ -1; & x_3 < 0. \end{cases} \quad (1.18)$$

The branches of the radicals in (1.16) and (1.17) are fixed by the conditions that

$$\mathrm{Im}\,\{s_1^2 + s_2^2 + k_p^2\}^{\frac{1}{2}} < 0 \quad (1.19)$$

and

$$\mathrm{Im}\,\{s_1^2 + s_2^2 + k_s^2\}^{\frac{1}{2}} < 0. \quad (1.20)$$

This guarantees that as $|x_3| \to \infty$

$$\mathfrak{L}\{G_{ij}\} \to 0.$$

The corresponding result for the bilateral Laplace transform of (1.11) when $\xi_1 = \xi_2 = \xi_3 = 0$ is

$$\mathfrak{L}\{G_{ij}(\mathbf{x}, s)\} = M^p_{ij}\frac{\exp(-s\gamma_p|x_3|)}{2\rho s\gamma_p} + M^s_{ij}\frac{\exp(-s\gamma_s|x_3|)}{2\rho s\gamma_s}, \quad (1.21)$$

where

$$M^p_{ij} = \begin{Vmatrix} \frac{s_1^2}{s^2} & \frac{s_1s_2}{s^2} & -\mathrm{sgn}\,(x_3)\frac{s_1\gamma_p}{s} \\ \frac{s_1s_2}{s^2} & \frac{s_2^2}{s^2} & -\mathrm{sgn}\,(x_3)\frac{s_2\gamma_p}{s} \\ -\mathrm{sgn}\,(x_3)\frac{s_1\gamma_p}{s} & -\mathrm{sgn}\,(x_3)\frac{s_2\gamma_p}{s} & \gamma_p^2 \end{Vmatrix}, \quad (1.22)$$

$$M^s_{ij} = \begin{Vmatrix} \left(\gamma_s^2 + \frac{s_2^2}{s^2}\right) & \frac{-s_1s_2}{s^2} & \mathrm{sgn}\,(x_3)\frac{s_1\gamma_s}{s} \\ \frac{-s_1s_2}{s^2} & \left(\gamma_s^2 + \frac{s_1^2}{s^2}\right) & \mathrm{sgn}\,(x_3)\frac{s_2\gamma_s}{s} \\ \mathrm{sgn}\,(x_3)\frac{s_1\gamma_s}{s} & \mathrm{sgn}\,(x_3)\frac{s_2\gamma_s}{s} & \left(\frac{s_1^2 + s_2^2}{s^2}\right) \end{Vmatrix}, \quad (1.23)$$

with

$$s\gamma_p = [k_p^2 - (s_1^2 + s_2^2)]^{\frac{1}{2}}, \quad (1.24)$$

$$s\gamma_s = [k_s^2 - (s_1^2 + s_2^2)]^{\frac{1}{2}}. \quad (1.25)$$

In order to ensure that $\mathfrak{L}\{G_{ij}\}$, as given by (1.21), vanishes as $|x_3| \to \infty$, we require that the branches of the radicals in (1.24) and (1.25) be chosen so that

$$\mathrm{Re}\,(s\gamma_p) > 0, \quad (1.26)$$

$$\mathrm{Re}\,(s\gamma_s) > 0. \quad (1.27)$$

2. FORMULATION OF THE FACTORIZATION PROBLEM

In another investigation[2] the writer showed that the analysis of diffraction of a plane elastic wave by a rigid right-angle wedge, occupying the first quadrant ($x_1 \geq 0, x_2 \geq 0$) of the $x_3 = 0$ plane in an unbounded elastic medium, could be reduced to the problem of solving a system of Wiener–Hopf equations in two complex variables. A solution of this system was suggested by analogy with the known solution for a system of Wiener–Hopf equations depending on only one complex variable.[3] In developing the analog solution, the problem arose of obtaining a product decomposition of the matrix $\mathfrak{L}(G_{ij})_{x_3=0}$ into matrix factors analytic in certain pairs of half-planes. It is to this problem that the present paper is addressed.

As preparation for the analysis which we are about to undertake, consider the problem of obtaining such a product decomposition for a scalar function of several complex variables $f(z_1, z_2, \cdots, z_k)$. Sufficient conditions for the existence of the decomposition have been stated by Bochner.

Bochner's Decomposition Theorem[4]

Let $\log f(z_1, z_2, \cdots, z_k)$ be analytic and of bounded L_2 norm in the tube $\gamma_j \leq \mathrm{Re}\,(z_j) \leq \delta_j, j = 1, 2, \cdots, k$, where the L_2 norm of $\log f(z_1, \cdots, z_k)$ is defined by

$$\|\log f\|_2 = \left\{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}|\log f(x_1 + iy_1, \cdots, x_k + iy_k)|^2 \times dy_1, \cdots, dy_k\right\}^{\frac{1}{2}}. \quad (2.1)$$

Then $\log f(z_1, \cdots, z_k)$ can be additively decomposed into the sum of 2^k functions $[\log f(z_1, \cdots, z_k)]_n$, $n = 1, 2, \cdots, 2^k$, so that

$$\log f(z_1, \cdots, z_k) = \sum_{n=1}^{2^k}[\log f(z_1, \cdots, z_k)]_n, \quad (2.2)$$

[2] E. A. Kraut, Bull. Seism. Soc. Am. **58**, 1083 (1968).
[3] B. Noble, *Methods Based on The Wiener–Hopf Technique* (Pergamon Press, Inc., New York, 1958), p. 157.
[4] S. Bochner, Am. J. Math. **59**, 732 (1937).

where for each n the function $[\log f(z_1, \cdots, z_k)]_n$ is analytic and bounded within a set of k half-planes containing the interior of the tube $\gamma_j \leq \mathrm{Re}\,(z_j) \leq \delta_j$, $j = 1, \cdots, k$. The functions $[\log f(z_1, \cdots, z_k)]_n$ are uniquely determined[5] and are representable by Cauchy integrals. For example, $[\log f(z_1, \cdots, z_k)]_1$ is a bounded analytic function in the set of k half-planes $\mathrm{Re}\,(z_j) > \gamma_j, j = 1, \cdots, k$ and is represented as

$$[\log f(z_1, \cdots, z_k)]_1 = \frac{1}{(2\pi i)^k} \int_{Z_1} \cdots \int_{Z_k} \frac{\log f(\zeta_1, \cdots, \zeta_k)\, d\zeta_1\, d\zeta_2 \cdots d\zeta_k}{(z_1 - \zeta_1)(z_2 - \zeta_2) \cdots (z_k - \zeta_k)}, \tag{2.3}$$

where Z_j is a rectilinear contour running from $\gamma_j - i\infty$ to $\gamma_j + i\infty$ in the ζ_j plane. The required product decomposition is given by

$$f(z_1, \cdots, z_k) = \prod_{n=1}^{2^k} \exp\,[\log f(z_1, \cdots, z_k)]_n. \tag{2.4}$$

Examples

Consider the problem of obtaining a product decomposition of the form (2.4) with $k = 2$ for the function

$$s\gamma = \{k^2 - (s_1^2 + s_2^2)\}^{\frac{1}{2}}, \tag{2.5}$$

where k is real. Bochner's theorem is not immediately applicable to (2.5). However, on differentiating (2.5) with respect to k, one finds

$$\frac{1}{\gamma}\frac{d\gamma}{dk} = \frac{k}{(s\gamma)^2}. \tag{2.6}$$

The right member of (2.6) is analytic in the tube $T(B)$ defined by

$$|\mathrm{Re}\,(s_j)| < b_j \quad (j = 1, 2), \tag{2.7}$$

$$b_j > 0, \quad |(b_1^2 + b_2^2)^{\frac{1}{2}}| < k, \tag{2.8}$$

and has a bounded L_2 norm in $T(B)$. Therefore, Bochner's theorem guarantees a unique additive decomposition for the right side of (2.6). In order to obtain this decomposition, we make use of the following result:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_0[k(x_1^2 + x_2^2)^{\frac{1}{2}}] \exp\,(-s_1x_1 - s_2x_2)\, dx_1\, dx_2 = 2\pi(s\gamma)^{-2} = \mathfrak{L}\{K_0(kr)\}, \tag{2.9}$$

where K_0 is a modified Bessel function of the second kind of order zero.

Proof: The function $K_0(kr)$ satisfies the reduced wave equation

$$(\nabla^2 - k^2)K_0(kr) = -2\pi\delta(x_1)\delta(x_2), \tag{2.10}$$

where δ is Dirac's delta function. On multiplying (2.10) by $\exp\,(-s_1x_1 - s_2x_2)$ and integrating by parts over all (x_1, x_2) space, the result follows immediately, provided s_1 and s_2 are restricted to be in the tube $T(B)$ of (2.7) and (2.8).

With the aid of (2.9), it follows that

$$\frac{1}{\gamma}\frac{d\gamma}{dk} = \frac{k}{2\pi}\,\mathfrak{L}\{K_0(kr)\}. \tag{2.11}$$

Now let $\mathfrak{L}$ be expressed as the sum of its restrictions to the four quadrants of the (x_1, x_2) plane, i.e.,

$$\mathfrak{L} = \sum_{n=1}^{4} \mathfrak{L}_n, \tag{2.12}$$

where $\mathfrak{L}_n$ is the operator

$$\mathfrak{L}_n = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx_1\, dx_2 H(\epsilon_{n1}x_1)H(\epsilon_{n2}x_2) \times \exp\,(-s_1x_1 - s_2x_2) \tag{2.13}$$

and H is Heaviside's unit step function with

$$\begin{aligned} \epsilon_{n1} &= 1 \quad (n = 1, 4), \\ \epsilon_{n1} &= -1 \quad (n = 2, 3), \\ \epsilon_{n2} &= 1 \quad (n = 1, 2), \\ \epsilon_{n2} &= -1 \quad (n = 3, 4). \end{aligned} \tag{2.14}$$

The convolution theorem for Laplace transforms can be used to show that the operators $\mathfrak{L}_n$ map $K_0(kr)$ into functions analytic in the pairs of half-planes (B, n) defined by

$$\begin{aligned} &(B, 1): \quad \mathrm{Re}\,(s_1) > -b_1, \quad \mathrm{Re}\,(s_2) > -b_2, \\ &(B, 2): \quad \mathrm{Re}\,(s_1) < b_1, \quad \mathrm{Re}\,(s_2) > -b_2, \\ &(B, 3): \quad \mathrm{Re}\,(s_1) < b_1, \quad \mathrm{Re}\,(s_2) < b_2, \\ &(B, 4): \quad \mathrm{Re}\,(s_1) > -b_1, \quad \mathrm{Re}\,(s_2) < b_2, \end{aligned} \tag{2.15}$$

with b_1, b_2 subject to (2.8).

Proof: From the convolution theorem, it follows that

$$\int_0^{\infty}\int_0^{\infty} f_1(x_1, x_2)f_2(x_1, x_2) \exp\,(-s_1x_1 - s_2x_2)\, dx_1\, dx_2 = (2\pi i)^{-2}\int_{c_1-i\infty}^{c_1+i\infty}\int_{c_2-i\infty}^{c_2+i\infty} g_1(z_1, z_2) \times g_2(s_1 - z_1, s_2 - z_2)\, dz_1\, dz_2, \tag{2.16}$$

where

$$g_i(z_1, z_2) = \int_0^{\infty}\int_0^{\infty} f_i(x_1, x_2) \exp\,(-x_1z_1 - x_2z_2)\, dx_1\, dx_2, \quad i = 1, 2, \tag{2.17}$$

and if f_2 is chosen so that $f_2(x_1, x_2) = H(x_1)H(x_2)$,

[5] E. Kraut, S. Busenberg, and W. Hall, Bull. Am. Math. Soc. **74**, 372 (1968).

then

$$g_2(s_1 - z_1, s_2 - z_2) = (s_1 - z_1)^{-1}(s_2 - z_2)^{-1}. \quad (2.18)$$

Therefore

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x_1)H(x_2)f_1(x_1, x_2)\exp(-s_1x_1 - s_2x_2)\,dx_1\,dx_2$$
$$= (2\pi i)^{-2}\int_{c_1-i\infty}^{c_1+i\infty}\int_{c_2-i\infty}^{c_2+i\infty}\frac{g_1(z_1, z_2)\,dz_1\,dz_2}{(s_1 - z_1)(s_2 - z_2)}$$
$$= g_1(s_1, s_2), \quad (2.19)$$

for $\mathrm{Re}\,(s_1) > c_1$, $\mathrm{Re}\,(s_2) > c_2$.

We conclude from this that the operator $\mathfrak{L}_1$ defined in (2.13) maps $f_1(x_1, x_2)$ into a function analytic in the pair of half-planes $\mathrm{Re}\,(s_1) > c_1$, $\mathrm{Re}\,(s_1) > c_2$. Now let $f_1(x_1, x_2) = K_0[k(x_1^2 + x_2^2)^{\frac{1}{2}}]$ and choose $c_1 = -b_1$, $c_2 = -b_2$; then $\mathfrak{L}_1$ maps K_0 into $(B, 1)$ and similar arguments suffice to show that $\mathfrak{L}_n$ maps $K_0(kr)$ into functions analytic in (B, n).

Equation (2.11) may be written as

$$\frac{d}{dk}\log(s\gamma/k) = \sum_{n=1}^{4}\left[\frac{k}{2\pi}\mathfrak{L}_n\{K_0(kr)\} - \frac{1}{4k}\right], \quad (2.20)$$

and, on integrating both sides of (2.20) with respect to k, one finds

$$s\gamma(k) = k\exp\sum_{n=1}^{4}\int_{\infty}^{k}\left[\frac{k}{2\pi}\mathfrak{L}_n\{K_0(kr)\} - \frac{1}{4k}\right], \quad (2.21)$$

and therefore the required factorization of (2.5) is given by

$$s\gamma(k) = k\prod_{n=1}^{4}\exp\int_{\infty}^{k}\left[\frac{k}{2\pi}\mathfrak{L}_n\{K_0(kr)\} - \frac{1}{4k}\right]dk. \quad (2.22)$$

For computational purposes, consider the integral

$$V(s_1, s_2) = \mathfrak{L}_1\{K_0(kr)\}$$
$$= \int_0^{\infty}\int_0^{\infty} K_0(kr)\exp(-s_1x_1 - s_2x_2)\,dx_1\,dx_2 \quad (2.23)$$

and observe that by symmetry

$$\mathfrak{L}_n\{K_0(kr)\} = V(\epsilon_{n1}s_1, \epsilon_{n2}s_2), \quad n = 1, 2, 3, 4, \quad (2.24)$$

where ϵ_{n1} and ϵ_{n2} are given in (2.14). To evaluate (2.23), let (2.10) be multiplied by $\exp(-s_1x_1 - s_2x_2)$ and integrated from 0 to ∞ with respect to x_1 and x_2 with the understanding that

$$\int_0^{\infty}\int_0^{\infty}\delta(x_1)\delta(x_2)\,dx_1\,dx_2 = \tfrac{1}{4}. \quad (2.25)$$

The result is

$$\mathfrak{L}_1\{K_0(kr)\} = (s_1^2 + s_2^2 - k^2)^{-1}$$
$$\times\left\{-\frac{\pi}{2} + s_1\int_0^{\infty}K_0(kx_2)\exp(-s_2x_2)\,dx_2\right.$$
$$\left. + s_2\int_0^{\infty}K_0(kx_1)\exp(-s_1x_1)\,dx_1\right\}, \quad (2.26)$$

where use has been made of the relation

$$\left\{\frac{\partial}{\partial x_1}K_0[(x_1^2 + x_2^2)^{\frac{1}{2}}]\right\}_{\substack{x_1=0\\x_2\neq 0}}$$
$$= \{K_0'[(x_1^2 + x_2^2)^{\frac{1}{2}}]x_1(x_1^2 + x_2^2)^{-\frac{1}{2}}\}_{\substack{x_1=0\\x_2\neq 0}} = 0. \quad (2.27)$$

The one-sided Laplace transforms remaining in (2.26) are well known and lead to the following expression for $V(s_1, s_2)$:

$$V(s_1, s_2) = \mathfrak{L}_1\{K_0(kr)\} = (s_1^2 + s_2^2 - k^2)^{-1}$$
$$\times\left\{-\frac{\pi}{2} + s_1\begin{cases}(s_2^2 - k^2)^{-\frac{1}{2}}\\ \quad\times\log\dfrac{[s_2 + (s_2^2 - k^2)^{\frac{1}{2}}]}{k}; & s_2 > k,\\ 2(k^2 - s_2^2)^{-\frac{1}{2}}\\ \quad\times\tan^{-1}\left(\dfrac{k - s_2}{k + s_2}\right); & s_2 < k,\end{cases}\right.$$
$$\left. + s_2\begin{cases}(s_1^2 - k^2)^{-\frac{1}{2}}\\ \quad\times\log\dfrac{[s_1 + (s_1^2 - k^2)^{\frac{1}{2}}]}{k}; & s_1 > k,\\ 2(k^2 - s_1^2)^{-\frac{1}{2}}\\ \quad\times\tan^{-1}\left(\dfrac{k - s_1}{k + s_1}\right); & s_1 < k.\end{cases}\right\} \quad (2.28)$$

The definite integrals

$$\int_{\infty}^{k}k\mathfrak{L}_n\{K_0(kr)\}\,dk$$

involved in (2.22) unfortunately are not expressible in terms of elementary functions and this complicates the further analysis of the factorization.

If, instead of (2.5), we wish to factor

$$K(s_1, s_2) = (i/2)(s_1^2 + s_2^2 + k^2)^{-\frac{1}{2}}, \quad (2.29)$$

the same technique gives[6]

$$K(s_1, s_2)$$
$$= (i/2k)\prod_{n=1}^{4}\exp\int_{\infty}^{k}\left[(-ik/4)\mathfrak{L}_n\{H_0^{(2)}(kr)\} + \frac{1}{4k}\right]dk, \quad (2.30)$$

where $H_0^{(2)}(kr)$ is a Hankel function of the second kind of order zero. In this case

$$\mathfrak{L}_n\{H_0^{(2)}(kr)\} = W(\epsilon_{n1}s_1, \epsilon_{n2}s_2), \quad (2.31)$$

where

$$W(s_1, s_2) = \frac{1}{s_1^2 + s_2^2 + k^2}$$
$$\times\left\{-i + \frac{s_1}{(s_2^2 + k^2)^{\frac{1}{2}}}\left[1 + \frac{2i}{\pi}\log\frac{s_2 + (s_2^2 + k^2)^{\frac{1}{2}}}{k}\right]\right.$$
$$\left. + \frac{s_2}{(s_1^2 + k^2)^{\frac{1}{2}}}\left(1 + \frac{2i}{\pi}\log\frac{s_1 + (s_1^2 + k^2)^{\frac{1}{2}}}{k}\right)\right\}, \quad (2.32)$$

and where $(s_i^2 + k^2)^{\frac{1}{2}} = k$ at $s_i = 0$, $i = 1, 2$.

[6] J. Radlow, Arch. Ratl. Mech. Anal. **8**, 139 (1961).

In the case of (2.29) it is understood that Im $(k) < 0$. The tube $T(B)$ on which (2.29) is analytic is defined as in (2.7) and (2.8) except that k, where it appears in (2.8), must be replaced by $|\text{Im}\,(k)|$. The factors in (2.30) are then, respectively, analytic in the pairs of half-planes (B, n) defined in (2.15).

Now consider the following matrix version of the factorization problem.

Problem: Given an $(N \times N)$ matrix-valued function $K_{ij}(z_1, \cdots, z_k)$ of k complex variables analytic in the tube $\gamma_j \leq \text{Re}\,(z_j) \leq \delta_j$, $j = 1, 2, \cdots, k$, obtain a product decomposition for K_{ij} of the form

$$K_{ij}(z_1, \cdots, z_k) = \prod_{n=1}^{2^k} M_{ij}^{(n)}(z_1, \cdots, z_k), \quad (2.33)$$

where for each n the matrix $M_{ij}^{(n)}(z_1, \cdots, z_k)$ is analytic in a certain set of k half-planes containing the interior of the tube $\gamma_j \leq \text{Re}\,(z_j) \leq \delta_j$, $j = 1, 2, \cdots, k$.

3. DISCUSSION OF THE MATRIX-FACTORIZATION PROBLEM

When $\log K_{ij}(z_1, \cdots, z_k)$ represents a matrix which is analytic in the tube $\gamma_j < \text{Re}\,(z_j) < \delta_j$, $j = 1, 2, \cdots, k$, and if the required Cauchy integrals exist, it is possible to additively decompose $\log K_{ij}$ in a manner analogous to (2.2), i.e.,

$$\log K_{ij}(z_1, \cdots, z_k) = \sum_{n=1}^{2^k} [\log K_{ij}(z_1, \cdots, z_k)]_n . \quad (3.1)$$

However, the required product decomposition corresponding to (2.4) can only be obtained if the matrices $[\log K_{ij}]_n$ commute with one another[7] and this usually is not the case. An exception to the rule is the instance in which $K_{ij}(z_1, \cdots, z_k)$ is diagonal. Then the matrices $[\log K_{ij}]_n$ must also be diagonal, and of course they will commute. We shall exclude this special case from further consideration.

The next simplest class of matrices to consider are those matrices which can be diagonalized by a similarity transformation. Suppose that $K_{ij}(z_1, \cdots, z_k)$ is such a matrix and let

$$\hat{\mathbf{K}}(z_1, \cdots, z_k) = \mathbf{S}^{-1}\mathbf{K}\mathbf{S} \quad (3.2)$$

be the representation of $\mathbf{K}$ in diagonal form. By our previous remarks, a product decomposition for $\hat{\mathbf{K}}(z_1, \cdots, z_k)$ of the form

$$\hat{\mathbf{K}}(z_1, \cdots, z_k) = \prod_{n=1}^{2^k} \hat{\mathbf{M}}^{(n)}(z_1, \cdots, z_k) \quad (3.3)$$

can be obtained when the eigenvalues of $\hat{\mathbf{K}}$ (and thus of $\mathbf{K}$) can be factorized using Bochner's theorem. A product representation for the original matrix $K_{ij}(z_1, \cdots, z_k)$ can then be immediately written as

$$\mathbf{K}(z_1, \cdots, z_k) = \prod_{n=1}^{2^k} (\mathbf{S}\hat{\mathbf{M}}^{(n)}\mathbf{S}^{-1}). \quad (3.4)$$

Commutation is not a problem here because the matrices $\hat{\mathbf{M}}^{(n)}$, being diagonal, commute with one another and commutation remains invariant under a similarity transformation. Another problem arises to plague us instead. In general, the matrix $(\mathbf{S}\hat{\mathbf{M}}^{(n)}\mathbf{S}^{-1})$ will not remain analytic and bounded in the same set of half-planes as $\hat{\mathbf{M}}^{(n)}(z_1, \cdots, z_k)$. The analyticity domains of the factors appearing in (3.3) unfortunately are not generally invariant under a similarity transformation because the matrices $\mathbf{S}(z_1, \cdots, z_k)$ and $\mathbf{S}^{-1}(z_1, \cdots, z_k)$ themselves depend on the k complex variables.

Consider the example mentioned in the beginning of Sec. 2. We have from (1.13)–(1.15)

$$\mathfrak{L}\{G_{ij}\}_{x_3=0} = K_{ij}(s_1, s_2) = (\rho\omega^2)^{-1}\begin{Bmatrix} (\lambda_1 + s_1^2\psi) & s_1 s_2 \psi & 0 \\ s_1 s_2 \psi & (\lambda_2 - s_1^2\psi) & 0 \\ 0 & 0 & \lambda_3 \end{Bmatrix} \quad (3.5)$$

for the time-harmonic elastic-wave Green's tensor, where

$$\lambda_1 = \{\tfrac{1}{4}K_s^{-2} + (s_1^2 + s_2^2)\}K_s, \quad (3.6)$$

$$\lambda_2 = \{\tfrac{1}{4}K_p^{-1}K_s^{-1} + (s_1^2 + s_2^2)\}K_p, \quad (3.7)$$

$$\lambda_3 = \{\tfrac{1}{4}K_p^{-1}K_s^{-1} + (s_1^2 + s_2^2)\}K_s, \quad (3.8)$$

$$\psi = (\lambda_2 - \lambda_1)/(s_1^2 + s_2^2) = K_p - K_s, \quad (3.9)$$

and

$$\mathfrak{L}\{G_{ij}\}_{x_3=0} = K_{ij}(s_1, s_2) = (2\rho s)^{-1}\begin{Bmatrix} \left(\lambda_1 - \frac{s_1^2}{s^2}\psi\right) & \frac{-s_1 s_2}{s^2}\psi & 0 \\ \frac{-s_1 s_2}{s^2}\psi & \left(\lambda_2 + \frac{s_1^2}{s^2}\psi\right) & 0 \\ 0 & 0 & \lambda_3 \end{Bmatrix}, \quad (3.10)$$

$$\lambda_1 = \{s^2\gamma_s^2 + (s_1^2 + s_2^2)\}(s^2\gamma_s)^{-1}, \quad (3.11)$$

$$\lambda_2 = \{s^2\gamma_p\gamma_s + (s_1^2 + s_2^2)\}(s^2\gamma_p)^{-1}, \quad (3.12)$$

$$\lambda_3 = \{s^2\gamma_p\gamma_s + (s_1^2 + s_2^2)\}(s^2\gamma_s)^{-1}, \quad (3.13)$$

$$\psi = s^2(\lambda_1 - \lambda_2)/(s_1^2 + s_2^2) = (\gamma_s^{-1} - \gamma_p^{-1}), \quad (3.14)$$

for the corresponding aperiodic case (1.21)–(1.23).

[7] G. H. Weiss and A. A. Maradudin, J. Math. Phys. 3, 771 (1962).

The functions λ_1, λ_2, and λ_3 in (3.5) and (3.10) are the eigenvalues of $K_{ij}(s_1, s_2)$ and the matrices (3.5) and (3.10) are both diagonalized by the same matrix

$$\mathbf{S} = (s_1^2 + s_2^2)^{-\frac{1}{2}} \begin{Vmatrix} -s_2 & s_1 & 0 \\ s_1 & s_2 & 0 \\ 0 & 0 & (s_1^2 + s_2^2)^{\frac{1}{2}} \end{Vmatrix}. \quad (3.15)$$

The columns of $\mathbf{S}$ form the eigenvectors of $K_{ij}(s_1, s_2)$ and $\mathbf{S}$ satisfies the relation $\mathbf{S} = \mathbf{S}^{-1}$.

In diagonal form $\mathbf{K}(s_1, s_2)$ becomes

$$\hat{\mathbf{K}}(s_1, s_2) = \begin{Vmatrix} \lambda_1(s_1, s_2) & 0 & 0 \\ 0 & \lambda_2(s_1, s_2) & 0 \\ 0 & 0 & \lambda_3(s_1, s_2) \end{Vmatrix} \quad (3.16)$$

and the corresponding factor matrices $\hat{\mathbf{M}}^{(n)}(s_1, s_2)$ appearing in (3.3) are given by

$$\hat{\mathbf{M}}^{(n)}(s_1, s_2) = \begin{Vmatrix} \lambda_1^{(n)}(s_1, s_2) & 0 & 0 \\ 0 & \lambda_2^{(n)}(s_1, s_2) & 0 \\ 0 & 0 & \lambda_3^{(n)}(s_1, s_2) \end{Vmatrix}, \quad (3.17)$$

with

$$\lambda_i(s_1, s_2) = \prod_{n=1}^{4} \lambda_i^{(n)}(s_1, s_2), \quad i = 1, 2, 3. \quad (3.18)$$

For a fixed n each of the three scalar factors $\lambda_i^{(n)}(s_1, s_2)$, $i = 1, 2, 3$, is analytic in the pair of half-planes (B, n) of (2.15). In the time-harmonic case (3.5), the k_p and k_s appearing in (1.16) and (1.17) are complex and b_1 and b_2 in (2.15) are chosen so that

$$|(b_1^2 + b_2^2)^{\frac{1}{2}}| < |\mathrm{Im}\,(k_p)|. \quad (3.19)$$

In the aperiodic case (3.10), it is assumed that k_p and k_s in (1.24) and (1.25) are real and that

$$|(b_1^2 + b_2^2)^{\frac{1}{2}}| < |k_p|. \quad (3.20)$$

The problem of factorizing the matrix $K_{ij}(s_1, s_2)$ has now been reduced to that of factorizing its eigenvalues as indicated in (3.18). Consider the time-harmonic case (3.6)–(3.8) first. The eigenvalue $\lambda_1(s_1, s_2)$ given by (3.6) is a special case, for on using (1.17) one finds

$$\lambda_1(s_1, s_2) = -k_s^2 K_s(s_1, s_2), \quad (3.21)$$

and this has already been factored in (2.30). The remaining eigenvalues (3.7) and (3.8) each involve the product of

$$N(s_1, s_2) = \{(4K_pK_s)^{-1} + (s_1^2 + s_2^2)\}, \quad (3.22)$$

with K_p or K_s. These latter functions factorize as shown in (2.30). To complete the factorization in (3.18) it is necessary to obtain a product decomposition for (3.22) of the type indicated in (2.4). For this purpose we shall obtain an additive decomposition of $\log N(s_1, s_2)$ with the aid of a double Cauchy integral. Using (1.16) and (1.17), $N(s_1, s_2)$ becomes

$$N(s_1, s_2) = (s_1^2 + s_2^2) - (s_1^2 + s_2^2 + k_p^2)^{\frac{1}{2}}(s_1^2 + s_2^2 + k_s^2)^{\frac{1}{2}}. \quad (3.23)$$

It follows from (3.23) that with the choice of branches made in (1.19) and (1.20) $N(s_1, s_2)$ is bounded and never vanishes.

Proof: To establish that $N(s_1, s_2)$ is bounded it suffices to observe that for large $|s_1|$ and $|s_2|$

$$\lim_{\substack{|s_1| \to \infty \\ |s_2| \to \infty}} N(s_1, s_2) = -\frac{(k_p^2 + k_s^2)}{2}\left\{1 + O\left(\frac{1}{s_1^2 + s_2^2}\right)\right\}. \quad (3.24)$$

Next note that $N(s_1, s_2)$ can vanish only if

$$(s_1^2 + s_2^2) = -\frac{k_p^2 k_s^2}{k_p^2 + k_s^2}. \quad (3.25)$$

On substituting (3.25) back into (3.23), one finds

$$N(s_1, s_2) = -\frac{k_p^2 k_s^2}{k_p^2 + k_s^2} - \left\{\frac{k_p^4}{k_p^2 + k_s^2}\right\}^{\frac{1}{2}}\left\{\frac{k_s^4}{k_p^2 + k_s^2}\right\}^{\frac{1}{2}}, \quad (3.26)$$

and in order for (3.26) to vanish the signatures of the two square roots which appear must differ. However, because of (1.19) and (1.20), this is not the case and consequently $N(s_1, s_2)$ never vanishes.

Let a parameter c be defined by

$$c = -2(k_p^2 + k_s^2)^{-1}. \quad (3.27)$$

Then, when $|s_1|$ or $|s_2|$ is sufficiently large, there is some positive real number A such that

$$0 < |\log cN(s_1, s_2)| < \frac{A}{|s_1^2 + s_2^2|}. \quad (3.28)$$

Furthermore, the analyticity and boundedness of $cN(s_1, s_2)$ in the tube $T(B)$ of (2.7) and (3.19) guarantees that $\log cN(s_1, s_2)$ is also analytic and bounded in $T(B)$. Therefore, the double Cauchy integrals

$$[\log cN(s_1, s_2)]_n = \frac{1}{(2\pi i)^2}\int_{-i\infty}^{i\infty}\int_{-i\infty}^{i\infty}\frac{\log cN(z_1, z_2)}{(s_1 - z_1)(s_2 - z_2)}\,dz_1\,dz_2 \quad (3.29)$$

converge and yield a complete additive decomposition

of $\log cN(s_1, s_2)$ of the form

$$\log cN(s_1, s_2) = \sum_{n=1}^{4} [\log cN(s_1, s_2)]_n, \quad (3.30)$$

where $[\log cN(s_1, s_2)]_n$ is analytic in the pair of half-planes (B, n) defined in (2.15) and (3.19).

The required product decomposition of $N(s_1, s_2)$ is then given by

$$N(s_1, s_2) = c^{-1} \prod_{n=1}^{4} \exp [\log cN(s_1, s_2)]_n. \quad (3.31)$$

In the time-harmonic case the factors appearing in the product decomposition (3.18) are given by

$$\lambda_1^{(n)} = k_s^{\frac{1}{2}} 2^{-\frac{1}{4}} \exp\left(\frac{3i\pi}{8}\right) \times \exp \int_{\infty}^{k_s} \left[(-ik/4)\mathfrak{L}_n\{H_0^{(2)}(kr)\} + \frac{1}{4k}\right] dk, \quad (3.32)$$

$$\lambda_2^{(n)} = (-2ik_p c)^{-\frac{1}{4}} \exp [\log cN(s_1, s_2)]_n \times \exp \int_{\infty}^{k_p} \left[(-ik/4)\mathfrak{L}_n\{H_0^{(2)}(kr)\} + \frac{1}{4k}\right] dk, \quad (3.33)$$

$$\lambda_3^{(n)} = (-2ik_s c)^{-\frac{1}{4}} \exp [\log cN(s_1, s_2)]_n \times \exp \int_{\infty}^{k_s} \left[(-ik/4)\mathfrak{L}_n\{H_0^{(2)}(kr)\} + \frac{1}{4k}\right] dk. \quad (3.34)$$

Using (3.15) and (3.4), the original matrix (3.5) is seen to have the product decomposition

$$K_{ij}(s_1, s_2) = \prod_{n=1}^{4} K_{ij}^{(n)}(s_1, s_2), \quad (3.35)$$

where

$$K_{ij}^{(n)} = (s_1^2 + s_2^2)^{-1} \times \begin{pmatrix} (\lambda_1^{(n)} s_2^2 + \lambda_2^{(n)} s_1^2) & s_1 s_2 (\lambda_2^{(n)} - \lambda_1^{(n)}) & 0 \\ s_1 s_2 (\lambda_2^{(n)} - \lambda_1^{(n)}) & (\lambda_1^{(n)} s_1^2 + \lambda_2^{(n)} s_2^2) & 0 \\ 0 & 0 & \lambda_3^{(n)}(s_1^2 + s_2^2) \end{pmatrix}. \quad (3.36)$$

The results for the aperiodic case (3.10)–(3.13) are quite similar. The factors appearing in the product decomposition (3.18) are now given by

$$\lambda_1^{(n)} = k_s^{\frac{1}{4}} \exp \int_{\infty}^{k_s} \left[\frac{-k}{2\pi} \mathfrak{L}_n\{K_0(kr)\} + \frac{1}{4k}\right] dk, \quad (3.37)$$

$$\lambda_2^{(n)} = (-2ck_p)^{-\frac{1}{4}} \exp [\log cN(s_1, s_2)]_n \times \exp \int_{\infty}^{k_p} \left[\frac{-k}{2\pi} \mathfrak{L}_n\{K_0(kr)\} + \frac{1}{4k}\right] dk, \quad (3.38)$$

$$\lambda_3^{(n)} = (-2ck_s)^{-\frac{1}{4}} \exp [\log cN(s_1, s_2)]_n \times \exp \int_{\infty}^{k_s} \left[\frac{-k}{2\pi} \mathfrak{L}_n\{K_0(kr)\} + \frac{1}{4k}\right] dk, \quad (3.39)$$

where in this case

$$N(s_1, s_2) = [\{k_s^2 - (s_1^2 + s_2^2)\}^{\frac{1}{2}} \times \{k_p^2 - (s_1^2 + s_2^2)\}^{\frac{1}{2}} + (s_1^2 + s_2^2)]. \quad (3.40)$$

The matrix (3.10) has a product decomposition (3.35) with factors of the form (3.36), however, with $\lambda_i^{(n)}$ $(i = 1, 2, 3)$ given by (3.37)–(3.39) instead of (3.32)–(3.34).

Each of the matrices $K_{ij}^{(n)}(s_1, s_2)$ in (3.36) would be analytic in a pair of half-planes (B, n) defined in (2.15), (3.19), and (3.20) if it were not for the singular scalar multiplier $(s_1^2 + s_2^2)^{-1}$ appearing in (3.36). The factorization which we have actually obtained has the form

$$K_{ij}(s_1, s_2) = f(s_1, s_2) \prod_{n=1}^{4} K_{ij}^{(n)}(s_1, s_2), \quad (3.41)$$

where the $K_{ij}^{(n)}(s_1, s_2)$ are analytic in the desired half-planes (B, n) and $f(s_1, s_2)$ is a singular scalar factor given by

$$f(s_1, s_2) = (s_1^2 + s_2^2)^{-4}. \quad (3.42)$$

The tube in which an expression like (2.29) is analytic disappears in the case of (3.42). This prohibits us from using Bochner's theorem (see Sec. 2) to factorize (3.42) and thereby prevents us from further simplifying (3.41). The source of this difficulty is clearly seen to be that the eigenvectors of $K_{ij}(s_1, s_2)$, i.e., the column vectors of (3.15), are singular on the manifold $(s_1^2 + s_2^2) = 0$, and therefore have no tube of analyticity, whereas the eigenvalues $\lambda_i(s_1, s_2)$ $(i = 1, 2, 3)$ of $K_{ij}(s_1, s_2)$ do share a common tube of analyticity.

ACKNOWLEDGMENTS

The author wishes to thank Dr. Guy W. Lehman and Dr. William Hall for their interest in this work and for a useful suggestion in connection with it.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Note on Zak's Method for Constructing Representations of Space Groups

L. T. KLAUDER, JR., AND J. G. GAY
Research Laboratories, General Motors Corporation, Warren, Michigan

(Received 8 January 1968)

Zak [J. Math. Phys. **1**, 165 (1960)] has given a method for constructing the irreducible representations of space groups which involves inducing the representations of the full group from those of an invariant subgroup. When a representation of the subgroup is self-conjugate, Zak's prescription for the induction is subject to a restriction which makes it inapplicable in some practical applications of the method. This paper presents a general prescription for carrying out the induction from self-conjugate representations.

When a space group is nonsymmorphic, the irreducible representations of its little groups at the Brillouin-zone edge cannot be constructed by the usual methods applicable to symmorphic groups.[1] Zak[2] has given a method whereby these little-group representations can be constructed. The method utilizes the fact that every space group contains an invariant subgroup of index two or three and consists of a procedure for inducing the irreducible representations of the full group from those of the invariant subgroup. If the invariant subgroup of the little group is symmorphic, its representations can be obtained in the usual way and the induction procedure can be applied to obtain the representations of the little group. If the subgroup is not symmorphic, one relies on the fact that it also contains an invariant subgroup of index two or three. The process is continued until a symmorphic subgroup is encountered. Then, by a series of inductions, the irreducible representations of the little group are constructed from those of its symmorphic subgroup.

However, Zak's induction procedure is not complete. When an irreducible representation of the invariant subgroup is self-conjugate, the prescription offered by Zak is applicable only in special circumstances which are not always met in practical applications.[3] Below we present a simple, general prescription for effecting induction from self-conjugate representations. This prescription provides, in conjunction with Zak's prescriptions for other cases, a complete induction procedure.

We begin by sketching the well-known[4] formal procedure for inducing associated irreducible representations of a group G from a self-conjugate irreducible representation of an invariant subgroup H of index two or three. Let a be any element of G not in H. Then, if H is of index two, let its coset be aH; if H is of index three, let its cosets be aH and $a^{-1}H$. Let r, s, t be typical elements of H, and let $\mathsf{D}(r)$ be the matrices of a self-conjugate irreducible representation of H. Because the representation is self-conjugate, there exists a matrix B such that

$$\mathsf{D}(ara^{-1}) = \mathsf{B}\mathsf{D}(r)\mathsf{B}^{-1}.$$

In the index-two case, $a^2 \in H$ and we have

$$\mathsf{D}(a^2)\mathsf{D}(r)\mathsf{D}^{-1}(a^2) = \mathsf{B}^2\mathsf{D}(r)(\mathsf{B}^{-1})^2$$

or

$$\mathsf{D}(r)\mathsf{D}^{-1}(a^2)\mathsf{B}^2 = \mathsf{D}^{-1}(a^2)\mathsf{B}^2\mathsf{D}(r), \quad \text{all} \quad r \in H,$$

so that by Schur's lemma

$$\mathsf{B}^2 = \lambda\mathsf{D}(a^2),$$

where λ is some constant. Introducing $\mathsf{A} = \lambda^{-\frac{1}{2}}\mathsf{B}$, we have $\mathsf{A}^2 = \mathsf{D}(a^2)$. This suggests that we can let the matrix A represent the element a and thereby obtain a representation of the group $G = H + aH$. Indeed, two associated representations of G may be induced from the representation of H. One is obtained by adding to $\mathsf{D}(r)$ the matrices

$$\mathsf{D}_1(ar) \equiv \mathsf{A}\mathsf{D}(r),$$

and the other by adding to $\mathsf{D}(r)$ the matrices

$$\mathsf{D}_2(ar) \equiv -\mathsf{A}\mathsf{D}(r).$$

The index-three case is analogous. Namely, because $\mathsf{D}(r)$ is self-conjugate, there exists a matrix A such that

$$\mathsf{D}(ara^{-1}) = \mathsf{A}\mathsf{D}(r)\mathsf{A}^{-1},$$

and such that

$$\mathsf{A}^3 = \mathsf{D}(a^3).$$

Then the representation $\mathsf{D}(r)$ of H induces three representations of $G = H + aH + a^{-1}H$. Each of these is obtained by adding to $\mathsf{D}(r)$ one of the following three sets of matrices:

$$\mathsf{D}_1(ar) \equiv \mathsf{A}\mathsf{D}(r), \qquad \mathsf{D}_1(a^{-1}r) \equiv \mathsf{A}^{-1}\mathsf{D}(r)$$

or

$$\mathsf{D}_2(ar) \equiv e^{2\pi i/3}\mathsf{A}\mathsf{D}(r), \qquad \mathsf{D}_2(a^{-1}r) \equiv e^{-2\pi i/3}\mathsf{A}^{-1}\mathsf{D}(r)$$

---

[1] G. F. Koster, Solid State Phys. **5**, 173 (1957).

[2] J. Zak, J. Math. Phys. **1**, 165 (1960).

[3] The representations of the space group D_{4h}^{14} (the "rutile" group) have been constructed by Zak's method [J. G. Gay, W. A. Albers, Jr., and F. J. Arlinghaus, J. Phys. Chem. Solids (to be published)]. It was necessary to use the prescription given in this note to obtain certain of the representations.

[4] H. Boerner, *Representations of Groups* (North-Holland Publishing Co., Amsterdam, 1959), pp. 95–101.

or

$$D_3(ar) \equiv e^{-2\pi i/3}AD(r), \quad D_3(a^{-1}r) \equiv e^{2\pi i/3}A^{-1}D(r).$$

With these formal results, all that is required to obtain the representations of G is a prescription for obtaining the matrix A. The prescription given by Zak is valid only when A turns out to be a scalar matrix. The prescription given below is applicable in general.

We consider the matrix

$$C(X) \equiv \sum_{r\in H} D(ara^{-1})XD(r^{-1}),$$

where X is an arbitrary matrix. We first observe that C does not vanish for all X. Namely, because $D(r)$ is self-conjugate, there exists a nonsingular matrix B such that

$$D(ara^{-1}) = BD(r)B^{-1},$$

and then

$$\begin{aligned} C(B) &= \sum_{r\in H} D(ara^{-1})BD(r^{-1}) \\ &= \sum BD(r)D(r^{-1}) \\ &= mB, \end{aligned}$$

where m is the order of H. Second, we observe that for arbitrary X

$$\begin{aligned} C(X)D(s) &= \sum_{r\in H} D(ara^{-1})XD(r^{-1})D(s) \\ &= \sum_{t\in H} D(asa^{-1}ata^{-1})XD(t^{-1}) \\ &= D(asa^{-1})C(X); \quad \text{for all } s \in H. \end{aligned}$$

We may thus infer from Schur's lemma that, as long as $C(X) \neq 0$, C(X) has an inverse and accomplishes the similarity transformation from $D(r)$ to $D(ara^{-1})$. Thus, A is equal to a constant times C(X), as long as $C(X) \neq 0$.

We thus have the prescription that, to obtain a matrix A to represent the coset representative element a, one first computes the matrix C(X) for a succession of simple matrices X until a nonzero result is obtained, and then one normalizes the result so that $A^2 = D(a^2)$ or $A^3 = D(a^3)$, according to whether the subgroup was of index two or three.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Diffraction Scattering and Vector-Meson Resonances*

D. ATKINSON
Lawrence Radiation Laboratory, University of California, Berkeley, California
AND
A. P. CONTOGOURIS†
Laboratoire de Physique Théorique et Hautes Energies, Orsay, Seine-et-Oise, France

(Received 1 May 1967)

It is shown that, under reasonable assumptions about inelasticity and asymptotic behavior, the usual diffraction picture combined with the N/D approach to π–π scattering leads to a singular integral equation. The authors' formalism is then used to show that, in conformity with the nearby-singularities philosophy, a constant left-hand discontinuity is by itself incapable of producing resonances in the GeV region. Next, a model for the creation of vector resonances, which combines a long-range force (defined by exchange of a cutoff vector meson) plus a short-range force (compatible with diffraction requirements), is introduced. The effect of the short-range force on the self-consistent (bootstrap) solutions is investigated in an approximate scheme. For a self-consistent solution with the correct ρ-meson mass, which is found to exist, inclusion of the short-range force is shown to decrease the self-consistent width by a factor of 2, which is nevertheless still greater than the experimental value.

1. INTRODUCTION

Most of the low-energy calculations on strongly interacting systems are based on the assumption that the scattering in the GeV region is determined by low-energy singularities and that the effects of the high-energy region are completely unimportant. This assumption is fully justified in pion–nucleon scattering, notably below 500 MeV, where detailed quantitative agreement between dispersion calculations and experiment has been found.[1] It is also justified in nucleon–nucleon scattering, where reasonable models account for all the important experimental features.[2]

* This work was done under the auspices of the United States Atomic Energy Commission.

† Postal address: Laboratoire de Physique Théorique et Hautes Energies, Faculté des Sciences, Orsay, France.

[1] J. Hamilton, "Dynamics of the π-N System," Fünfte Internationale Universitätswochen für Kernphysik, Schladming (1966).

[2] A. Scotti and D. Y. Wong, Phys. Rev. **138**, B145 (1965); H. G. Dosch and V. F. Müller, Nuovo Cimento **39**, 886 (1965).

or

$$D_3(ar) \equiv e^{-2\pi i/3}AD(r), \quad D_3(a^{-1}r) \equiv e^{2\pi i/3}A^{-1}D(r).$$

With these formal results, all that is required to obtain the representations of G is a prescription for obtaining the matrix A. The prescription given by Zak is valid only when A turns out to be a scalar matrix. The prescription given below is applicable in general.

We consider the matrix

$$C(X) \equiv \sum_{r\in H} D(ara^{-1})XD(r^{-1}),$$

where X is an arbitrary matrix. We first observe that C does not vanish for all X. Namely, because $D(r)$ is self-conjugate, there exists a nonsingular matrix B such that

$$D(ara^{-1}) = BD(r)B^{-1},$$

and then

$$\begin{aligned} C(B) &= \sum_{r\in H} D(ara^{-1})BD(r^{-1}) \\ &= \sum BD(r)D(r^{-1}) \\ &= mB, \end{aligned}$$

where m is the order of H. Second, we observe that for arbitrary X

$$\begin{aligned} C(X)D(s) &= \sum_{r\in H} D(ara^{-1})XD(r^{-1})D(s) \\ &= \sum_{t\in H} D(asa^{-1}ata^{-1})XD(t^{-1}) \\ &= D(asa^{-1})C(X); \quad \text{for all } s \in H. \end{aligned}$$

We may thus infer from Schur's lemma that, as long as $C(X) \neq 0$, C(X) has an inverse and accomplishes the similarity transformation from $D(r)$ to $D(ara^{-1})$. Thus, A is equal to a constant times C(X), as long as $C(X) \neq 0$.

We thus have the prescription that, to obtain a matrix A to represent the coset representative element a, one first computes the matrix C(X) for a succession of simple matrices X until a nonzero result is obtained, and then one normalizes the result so that $A^2 = D(a^2)$ or $A^3 = D(a^3)$, according to whether the subgroup was of index two or three.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Diffraction Scattering and Vector-Meson Resonances*

D. ATKINSON
Lawrence Radiation Laboratory, University of California, Berkeley, California
AND
A. P. CONTOGOURIS†
Laboratoire de Physique Théorique et Hautes Energies, Orsay, Seine-et-Oise, France

(Received 1 May 1967)

It is shown that, under reasonable assumptions about inelasticity and asymptotic behavior, the usual diffraction picture combined with the N/D approach to π–π scattering leads to a singular integral equation. The authors' formalism is then used to show that, in conformity with the nearby-singularities philosophy, a constant left-hand discontinuity is by itself incapable of producing resonances in the GeV region. Next, a model for the creation of vector resonances, which combines a long-range force (defined by exchange of a cutoff vector meson) plus a short-range force (compatible with diffraction requirements), is introduced. The effect of the short-range force on the self-consistent (bootstrap) solutions is investigated in an approximate scheme. For a self-consistent solution with the correct ρ-meson mass, which is found to exist, inclusion of the short-range force is shown to decrease the self-consistent width by a factor of 2, which is nevertheless still greater than the experimental value.

1. INTRODUCTION

Most of the low-energy calculations on strongly interacting systems are based on the assumption that the scattering in the GeV region is determined by low-energy singularities and that the effects of the high-energy region are completely unimportant. This assumption is fully justified in pion–nucleon scattering, notably below 500 MeV, where detailed quantitative agreement between dispersion calculations and experiment has been found.[1] It is also justified in nucleon–nucleon scattering, where reasonable models account for all the important experimental features.[2]

* This work was done under the auspices of the United States Atomic Energy Commission.

† Postal address: Laboratoire de Physique Théorique et Hautes Energies, Faculté des Sciences, Orsay, France.

[1] J. Hamilton, "Dynamics of the π-N System," Fünfte Internationale Universitätswochen für Kernphysik, Schladming (1966).

[2] A. Scotti and D. Y. Wong, Phys. Rev. **138**, B145 (1965); H. G. Dosch and V. F. Müller, Nuovo Cimento **39**, 886 (1965).

No similar agreement can be claimed for pion–pion scattering—in particular, for a self-consistent determination of the parameters of the ρ meson. For this system, it is quite possible that the effects of the high-energy region are less unimportant.[3,4] Furthermore, in a determination of the parameters of the ρ, the mass of which is rather high, the features of the amplitude in a region which starts at 2 or 3 GeV are, perhaps, of some importance.

The purpose of this work is to give a hint concerning the effect of the high-energy region on the features of the vector-meson resonances. For this, the experimentally well-established picture of diffraction scattering is combined with the usual N/D approach employed in self-consistent calculations of pion–pion scattering. This leads to a marginally singular integral equation[5] which can be solved by an application of methods developed by the authors elsewhere.[6]

In Sec. 2 it is shown that, under certain assumptions, both Regge behavior and the conventional diffraction picture (with nonshrinking forward peak) lead to marginally singular N/D equations. In Sec. 3 the possibility that the distant singularities (short-range forces) produce resonances in the low-energy region is studied separately. For this the basic integral equation formulated in Sec. 2 is applied to a model which consists of a constant left-hand discontinuity (and constant inelasticity). This model is compatible with the requirements of unitarity and of the diffraction picture, but completely neglects the structure of the nearby singularities; as a result it is known to be incapable of generating resonances in the GeV region. Section 4 contains the formulation of a more realistic model whose left-hand discontinuity combines a long-range part determined by vector exchange with a short-range part compatible with the requirements of diffraction scattering. Finally, in Sec. 5, the results of an approximate numerical calculation involving the model of Sec. 4 are presented and compared with the solutions of the conventional vector-meson bootstrap (without short-range part). The conclusion is that the short-range part tends to decrease significantly the coupling necessary to produce a resonance, affecting its width to a lesser extent, to reduce the self-consistent mass, and, for low cutoffs, to change the self-consistent width in the correct direction. In particular, the width of the self-consistent solution which corresponds to $m_\rho^2 \simeq 30\, m_\pi^2$ is reduced by a factor of 2.

In Appendix A, the model of Sec. 3 is reconsidered in the approximation of contracting to zero the gap between left- and right-hand cuts; this approximation has the advantage of providing explicit and relatively simple solutions. Again, it is concluded that a featureless left-hand discontinuity and inelasticity are incapable of producing acceptable resonances. Finally, in the contracted-gap case, certain features of the solutions and, in particular, the positions of the zeros of the denominator function are studied in Appendix B.

2. FORMULATION OF THE BASIC EQUATIONS

Consider the elastic scattering of two pseudoscalar particles of mass unity, and assume that the partial P-wave amplitude $A_1(\nu)$ admits the usual decomposition:

$$A_1(\nu) = N(\nu)/D(\nu); \tag{2.1}$$

ν is the square of the center-of-mass momentum. It is convenient to consider a once-subtracted representation for N and D, with the subtraction point at $\nu = 0$. Due to the usual threshold properties, $N(0) = 0$, so the equations are

$$N(\nu) = \frac{\nu}{\pi}\int_{-\infty}^{-\omega_L} d\nu' \frac{\sigma(-\nu')D(\nu')}{\nu'(\nu' - \nu)}, \tag{2.2}$$

$$D(\nu) = 1 - \frac{\nu}{\pi}\int_0^{\infty} d\nu' \frac{\rho(\nu')R_1(\nu')N(\nu')}{\nu'(\nu' - \nu)}. \tag{2.3}$$

Here $\sigma(-\nu)$ is the discontinuity along the left-hand cut $-\infty < \nu \leq -\omega_L$, $\rho(\nu) = [\nu/(\nu + 1)]^{\frac{1}{2}}$ is the usual phase-space factor, and $R_1(\nu)$ is the inelasticity of the P wave[7]; thus the unitarity condition reads

$$\text{Im}\, A_1(\nu) = \rho(\nu)R_1(\nu)\,|A_1(\nu)|^2, \quad 0 \leq \nu < \infty. \tag{2.4}$$

To derive the basic integral equation of the problem, one can substitute (2.2) into (2.3). Then the definitions

$$\nu = -\omega, \quad -D(\nu)/\nu = f(\omega) \tag{2.5}$$

give

$$f(\omega) = \frac{1}{\omega} + \frac{1}{\pi}\int_{\omega_L}^{\infty} d\omega' K(\omega, \omega')\sigma(\omega')f(\omega'), \tag{2.6}$$

where

$$K(\omega, \omega') = \frac{1}{\pi}\int_0^{\infty} dx \frac{\rho(x)R_1(x)}{(x + \omega)(x + \omega')}. \tag{2.7}$$

Next, assume that, for $\nu \to +\infty$, $A_1(\nu)$ becomes purely imaginary. Equation (2.4) implies

$$\rho(\nu)R_1(\nu) \sim [\text{Im}\, A_1(\nu)]^{-1}. \tag{2.8}$$

[3] H. Burkhardt, Nuovo Cimento **42**, 351 (1966).

[4] L. Van Hove, "Theoretical Problems in Strong Interactions at High Energies," CERN 65-22 (1965).

[5] That in pion–pion scattering the usual diffraction requirements lead to a singular integral equation has independently been concluded by D. H. Lyth, Phys. Rev. Letters **17**, 820 (1966).

[6] D. Atkinson and A. P. Contogouris, Nuovo Cimento **39**, 1082 (1965).

[7] G. F. Chew and S. Mandelstam, Phys. Rev. **119**, 467 (1960).

According to the Phragmen–Lindelof theorem,[8] if

$$A_1(\nu) \sim \nu^{\alpha_0}(\log \nu)^{\alpha_1}(\log\log \nu)^{\alpha_2}\cdots(\log\log\cdots\log \nu)^{\alpha_n} \quad \text{for} \quad \nu \to \infty,$$

and

$$A_1(\nu) \sim |\nu|^{\alpha_0'}(\log|\nu|)^{\alpha_1'} \times (\log\log|\nu|)^{\alpha_2'}\cdots(\log\log\cdots\log|\nu|)^{\alpha_n'} \quad \text{for} \quad \nu \to -\infty,$$

and if

$$|A_1(\nu)| < |ae^{(\pi-\epsilon)\nu}|, \quad \text{where} \quad a = \text{const},$$

for all complex ν and $\epsilon > 0$, then $\alpha_0 = \alpha_0'$, $\alpha_1 = \alpha_1', \cdots, \alpha_n = \alpha_n'$ and

$$\sigma(-\nu) \sim \operatorname{Im} A_1(\nu), \quad \text{for} \quad \nu \to \infty. \tag{2.9}$$

Lest it be felt that these conditions are too restrictive, one may prefer simply to assert (2.9). Any N/D system for which $\operatorname{Im} A_1(\nu)$ has different limits for $\nu \to \pm\infty$ would constitute a pathology lying outside the scope of this paper.

It is now shown that all the important models of high-energy elastic scattering imply an $\operatorname{Im} A_1(\nu)$, and hence a $\sigma(\nu)$, such that the kernel of (2.6) has an unbounded norm, so that (2.6) is a singular integral equation. For energies above a few GeV and momentum transfers $|t|^{\frac{1}{2}} \leqslant 1$ GeV/c, a good parametrization of the observed t dependence of the scattering amplitude is[4]

$$|A(\nu, t)|^2 = |A(\nu, 0)|^2 e^{tb(\nu)}.$$

It is assumed that, as $\nu \to \infty$, $A(\nu, t)$ becomes purely imaginary. Then projection onto the P wave gives[9]

$$A_1(\nu) = \frac{1}{4\nu}\int_{-4\nu}^{0} dt\, A(\nu, t) P_1\left(1 + \frac{t}{2\nu}\right)_{\nu\to\infty} \sim ib^{-1}(\nu)\sigma_{\text{tot}}(\nu); \tag{2.10}$$

$\sigma_{\text{tot}}(\nu)$ is the total cross section, which is taken to be asymptotically constant. As for $b(\nu)$, the width of the diffraction peak, two cases are of interest:

(i) $b(\nu) \sim \log \nu$, in accord with the hypothesis of asymptotic dominance by a Pomeranchuk–Regge trajectory of nonzero slope $[\alpha_P'(t = 0) \neq 0]$

(ii) $b(\nu) \sim$ const, corresponding to the conventional diffraction picture, or to a flat Pomeranchuk trajectory.[10]

[8] E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, New York, 1960).

[9] In deriving (2.10), it is assumed that the contribution to the integral from large $|t|$ can be neglected.

[10] See, e.g., L. Van Hove, Rapporteur's Report at XIIIth International Conference on High Energy Physics at Berkeley, CERN preprint (Th. 714), 1966.

Consider first Case (i), when for $\nu \to +\infty$:

$$\operatorname{Im} A_1(\nu) \sim (\log \nu)^{-1}\rho(\nu)R_1(\nu) \sim \log \nu. \tag{2.11}$$

The behavior of $K(\omega, \omega')$ for large ω, ω' is controlled by the large values of the integrand in (2.7). Thus,

$$K(\omega, \omega') \sim \int_0^\infty dx \frac{\log x}{(x + \omega)(x + \omega')}. \tag{2.12}$$

With the lower limit of integration taken at $x = 0$, this gives[11]

$$K(\omega, \omega') \sim \tfrac{1}{2}[(\log \omega')^2 - (\log \omega)^2]/(\omega' - \omega). \tag{2.13}$$

In view of (2.9), $\sigma(\omega) \sim (\log \omega)^{-1}$. Thus for large ω, ω' the kernel of (2.6) reduces to

$$K(\omega, \omega')\sigma(\omega') \sim \frac{1}{2}\frac{\log(\omega'/\omega)}{\omega' - \omega}\left(1 + \frac{\log \omega}{\log \omega'}\right). \tag{2.14}$$

The norm of this diverges (logarithmically) for large ω, ω', so that (2.6) is a marginally singular integral equation.[12]

Consider now Case (ii). The above considerations can easily be generalized to include the asymptotic behavior $\operatorname{Re} A_1(\nu) \sim$ const together with $\operatorname{Im} A_1(\nu) \sim$ const (for $\nu \to \infty$). In view of (2.4), (2.7) can be split as follows:

$$K(\omega, \omega') = \frac{1}{\pi}\frac{\operatorname{Im} A_1(\infty)}{|A_1(\infty)|^2}\cdot\frac{\log(\omega'/\omega)}{\omega' - \omega} + \frac{1}{\pi}\int_0^\infty dx \frac{\rho(x)R_1(x) - \operatorname{Im} A_1(\infty)\,|A_1(\infty)|^{-2}}{(x + \omega)(x + \omega')}. \tag{2.15}$$

Then use of (2.9) reduces (2.6) to the form

$$f(\omega) = \frac{1}{\omega} + \frac{\lambda}{\pi^2}\int_{\omega_L}^\infty d\omega' \frac{\log(\omega'/\omega)}{\omega' - \omega}f(\omega') + \frac{1}{\pi}\int_{\omega_L}^\infty d\omega' K_F(\omega, \omega')f(\omega'), \tag{2.16}$$

where

$$\lambda \equiv \left\{\frac{\operatorname{Im} A_1(\infty)}{|A_1(\infty)|}\right\}^2. \tag{2.17}$$

$K_F(\omega, \omega')$ of (2.16) contains the integral of (2.15) and the difference $\sigma(\omega) - \operatorname{Im} A_1(\infty)$. Thus, with suitable assumptions about the way $R_1(\nu)$ and $\operatorname{Im} A_1(\nu)$ approach their asymptotic limits, the norm of

[11] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Sums and Products* (Academic Press Inc., New York, 1965).

[12] Note that in (2.14) the first part of the kernel $\frac{1}{2}[\log(\omega'/\omega)]/(\omega' - \omega)$ has been treated in Ref. 6; and the second part leads to a singular equation which is reduced to the forms of Ref. 6 by a simple change of the unknown function $[f(\omega) \to \log\omega \cdot f(\omega)]$.

$K_F(\omega, \omega')$ is finite. Then (2.16) is a standard form solved and studied in Ref. 6.[13]

Furthermore, the definition (2.17) implies

$$0 \leq \lambda \leq 1. \tag{2.18}$$

Then it is known that solutions free of unwanted poles on the physical sheet of ν (ghost-free solutions) can be constructed by the N/D approach[6,14-16] (see also Appendix B).

The rest of this work is restricted to models of the type (ii), for which the methods developed in Ref. 6 are directly applicable.

3. INABILITY OF SHORT-RANGE FORCE TO GENERATE RESONANCES

As a first application of the foregoing formalism, consider a model with a constant left-hand discontinuity

$$\sigma(-\nu) = \operatorname{Im} A_1(\infty), \quad -\infty < \nu < -\omega_L, \tag{3.1}$$

and constant inelasticity

$$R_1(\nu) = R_1(\infty) = \operatorname{Im} A_1(\infty)\,|A_1(\infty)|^{-2}; \tag{3.2}$$

also, the approximation $\rho(\nu) = 1$ will be made. Then, it is asked whether resonances can be produced in the GeV region. This example, which neglects basic features of the low-energy part of the amplitude and of the long-range part of the potential, cannot be realistic for low-energy calculations. Nevertheless, it can give some idea of the relative importance of the distant singularities in the generation of the strong-interaction resonances.

Before consideration of the main problem, it is shown that a ghost-free amplitude can be constructed by direct application of the N/D equations. Substitution of (2.3) in (2.2) gives

$$N(\nu) = B(\nu) + \frac{\nu}{\pi} R_1(\infty) \int_0^\infty d\nu' \, \frac{B(\nu) - B(\nu')}{\nu - \nu'} \cdot \frac{N(\nu')}{\nu'}, \tag{3.3}$$

where

$$B(\nu) = \frac{\nu}{\pi} \int_{-\infty}^{-\omega_L} d\nu' \, \frac{\sigma(-\nu')}{\nu'(\nu' - \nu)}, \tag{3.4}$$

or, due to (3.1),

$$B(\nu) = [\operatorname{Im} A_1(\infty)/\pi] \log (1 + \nu/\omega_L). \tag{3.5}$$

In view of (3.2),

$$R_1(\infty) \operatorname{Im} A_1(\infty) \leq 1;$$

then Ref. 6 concludes that an iteration solution of (3.3) (Neumann–Liouville series expansion) exists. On the other hand, it is easy to see from (3.5) that, for any $\nu, \nu' \geq 0$,

$$B(\nu) \geq 0, \quad [B(\nu) - B(\nu')]/[\nu - \nu'] \geq 0. \tag{3.6}$$

Hence, all the terms in this iteration solution are positive and

$$N(\nu) > 0, \quad \nu > 0. \tag{3.7}$$

Then (2.3) implies that $-D(\nu)$ is a Herglotz function, i.e.,

$$\operatorname{Im} D(\nu) < 0, \quad \operatorname{Im} \nu > 0, \tag{3.8}$$

and that

$$D(\nu) > 0, \quad \nu < 0. \tag{3.9}$$

Thus $D(\nu)$ has no zeros on the first sheet of the complex ν plane. It is concluded that the solution of Eqs. (2.2) and (2.3) which admits a Neumann–Liouville expansion will be free of ghost poles.

With (3.1), (3.2), and the approximation $\rho(\nu) = 1$, $K_F(\omega, \omega')$ of (2.16) vanishes; then, with the simplification $\omega_L = 1$ (no loss of generality), Eq. (2.16) can be written

$$\tilde{f}(\omega) = \frac{1}{\omega} + \frac{\lambda}{\pi^2} \int_1^\infty d\omega' \, \frac{\log (\omega'/\omega)}{\omega' - \omega} \tilde{f}(\omega'); \tag{3.10}$$

λ is given by (2.17). A solution of this equation is

$$\tilde{f}(\omega) = \frac{1}{\omega} + \lambda \int_1^\infty d\omega' R(\omega, \omega'; \lambda) \frac{1}{\omega'}, \tag{3.11}$$

where $R(\omega, \omega'; \lambda)$ is a resolvent of (3.10). Reference 6 shows that there exists a unique resolvent with a branch point only at $\lambda = 1$:

$$R(\omega, \omega'; \lambda) = \frac{1}{2} \int_{-\infty}^\infty ds \, \frac{s \tanh (\pi s)}{\cosh^2 (\pi s) - \lambda} \times P_{-\frac{1}{2}+is}(2\omega - 1) P_{-\frac{1}{2}+is}(2\omega' - 1). \tag{3.12}$$

Thus, the extra requirement of analyticity at $\lambda = 0$ (and thus the existence of a Neumann–Liouville expansion) leads to the unique choice (3.12); in the next section this choice is further supported by certain continuity arguments.

Substitution of (3.12) into (3.11) and use of the identity[17]

$$P_{-\frac{1}{2}+is}(z) = \frac{\cosh \pi s}{\pi} \int_1^\infty \frac{dx}{x + z} P_{-\frac{1}{2}+is}(x), \quad z > -1, \tag{3.13}$$

gives the solution

$$\tilde{D}(-\omega) = \omega \tilde{f}(\omega) = 1 + \frac{\lambda \pi \omega}{2} \int_{-\infty}^\infty ds \, \frac{s \tanh (\pi s)}{\cosh^2 (\pi s) - \lambda} \times \frac{P_{-\frac{1}{2}+is}(2\omega - 1)}{\cosh (\pi s)}. \tag{3.14}$$

[13] Other possibilities, e.g., $b(\nu) \sim \nu^{-\frac{1}{2}}$, lead also to marginally singular integral equations of the type treated in Ref. 6.

[14] D. Atkinson and D. Morgan, Nuovo Cimento **41**, 559 (1966).

[15] D. Atkinson, J. Math. Phys. 7, 1607 (1966).

[16] A. P. Contogouris and A. Martin, Nuovo Cimento **49A**, 61 (1967).

[17] *Higher Transcendental Functions*, A. Erdelyi, Ed. (McGraw-Hill Book Co., New York, 1953), Vol. 2.

The foregoing conclusions on the absence of ghost zeros of this solution can be checked directly, at least for ν real and negative (according to Appendix B, this is a particularly relevant region). Figure 1 presents plots of (3.14) in the interval $1 < \omega < 700$ for various values of the parameter λ, subject to the condition (2.18); no ghost zeros of $D(-\omega)$ are indicated.

An analytic continuation of (3.14) to $\omega < 0$ is necessary in order to look for resonances. This is best accomplished by means of the identity[17]

$$P_l(-z) = e^{\mp il\pi}P_l(z) - \frac{2}{\pi}\sin(\pi l)Q_l(z)$$

($\pm$ according as $\operatorname{Im} z \gtrless 0$), which gives

$$\tilde{D}(\nu) = 1 - i\lambda\pi\nu\int_0^\infty ds\,\frac{s\tanh(\pi s)}{\cosh^2(\pi s) - \lambda}\,P_{-\frac{1}{2}+is}(2\nu+1) - \lambda\nu\int_{-\infty}^{\infty} ds\,\frac{s\tanh(\pi s)}{\cosh^2(\pi s) - \lambda}\,Q_{-\frac{1}{2}+is}(2\nu+1). \tag{3.15}$$

To simplify the first integrand, use has been made of the symmetry property[15]

$$P_{-\frac{1}{2}+is}(z) = P_{-\frac{1}{2}-is}(z). \tag{3.16}$$

For $\nu > 0$, because of (3.16), the contribution of the second term of (3.15) is purely imaginary. Hence the condition for a resonance at $\nu = \nu_R$ is

$$\operatorname{Re}\tilde{D}(\nu_R) = 1 - 2\lambda\nu_R\int_0^\infty ds\,\frac{s\tanh(\pi s)}{\cosh^2(\pi s) - \lambda} \times \operatorname{Re}Q_{-\frac{1}{2}+is}(2\nu_R+1) = 0. \tag{3.17}$$

The functions $\operatorname{Re}\tilde{D}(\nu)$, for $\lambda = 0.25$, 0.50, 0.75, 0.95, and 1, are plotted in Fig. 2. For $0 < \nu < 50$, there are no zeros of $\operatorname{Re}\tilde{D}(\nu)$; with the beginning of the left-hand cut defined by two-pion exchange ($\omega_L = m_\pi^2 = 1$), this region extends up to 2 GeV. Moreover, the weak dependence of $\operatorname{Re}\tilde{D}(\nu)$ on ν at large ν indicates that, probably, there are no zeros at all.

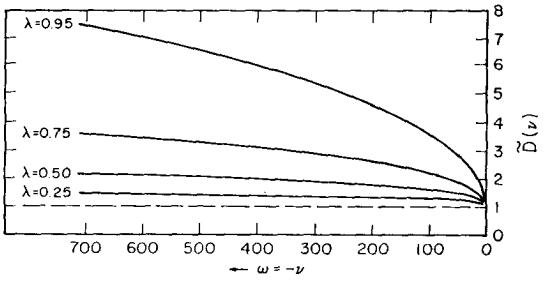


FIG. 1. The function $\tilde{D}(\nu)$, given by Eq. (3.14), for real negative ν.

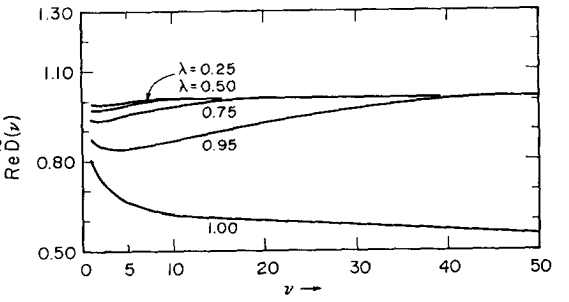


FIG. 2. The function $\operatorname{Re}\tilde{D}(\nu)$, given by Eq. (3.17), for real positive ν.

For λ close to unity ($\lambda \to 1-$), the majorizations of $R(\omega, \omega'; \lambda)$ given in Ref. 6 and the asymptotic properties of the Legendre functions for large argument indicate that, apart from logarithmic factors,

$$\operatorname{Im}\tilde{D}(\nu) \sim \nu^{\frac{1}{2}}, \quad \text{for} \quad \nu \to \infty. \tag{3.18}$$

Since a once-subtracted representation for $\tilde{D}(\nu)$ is used, this behavior is in agreement with the conclusions of Olesen and Squires.[18]

The conclusion of this section is that a constant left-hand discontinuity along with a constant inelasticity is incapable of generating strong-interaction resonances: the real part of the corresponding denominator function does not vanish at all, at least in the GeV region. This conclusion is further strengthened by the explicit solution of Appendix A (approximation $\omega_L = 0$).

In this model, a constant left-hand discontinuity and inelasticity can be considered as an abstraction representing the effects of the high-energy region. In this sense it can be said that for the generation of the known resonances, the high-energy effects are not primarily responsible; the resonances are generated by the long-range forces. This conclusion is, of course, hardly surprising. However, the model also shows that an asymptotically constant left-hand discontinuity, which is compatible with the present experimental information in the diffraction region, in no way contradicts the basic principles of dominance by nearby singularities.

On the other hand, although not primarily responsible, the short-range force may have a significant effect on certain features of the resonances. This question is taken up in the next two sections.

4. MODEL COMBINING LONG- AND SHORT-RANGE FORCES

The next application of the formalism of Sec. 2 is a more realistic model with a left-hand discontinuity, of which the nearby part is given by the exchange of an

[18] P. Olesen and E. J. Squires, Nuovo Cimento **39**, 956 (1965).

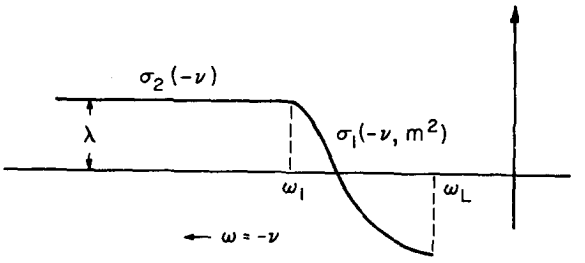


FIG. 3. One possible way to combine long- and short-range forces; however, the attraction is insufficient to generate resonances at $\nu \leq 50$.

elementary vector meson of mass m (long-range forces), and the distant part is constant (short-range force). Thus, for $-\omega_1 < \nu < -\omega_L = -m^2/4$,[19]

$$\sigma_1(-\nu, m^2) = \gamma\left(1 + \frac{m^2+4}{8\nu}\right)\left(1 + \frac{m^2}{2\nu}\right) \times \theta\left(-\nu - \frac{m^2}{4}\right)\theta(\omega_1 + \nu), \quad (4.1)$$

where γ is proportional to the $\pi\pi\rho$ coupling (the width $\Gamma_{\rho\to\pi\pi} \simeq 120$ MeV corresponds to $\gamma \simeq 3.8$); and for $\nu < -\omega_1$,

$$\sigma_2(-\nu) = \lambda\theta(-\nu - \omega_1), \quad \text{with} \quad 0 < \lambda < 1. \quad (4.2)$$

For simplicity, elastic unitarity [$R_1(\nu) = 1$] is assumed, the generalization to any asymptotically constant inelasticity being straightforward.

One way to combine (4.1) and (4.2) is indicated in Fig. 3: $\sigma_1(-\nu, m^2)$ rises on the left until it reaches the value λ. Correspondingly, ω_1 is the larger zero of the equation

$$\gamma[1 - (m^2 + 4/8\omega)][1 - (m^2/2\omega)] = \lambda. \quad (4.3)$$

Unfortunately, in this model, which has the advantage of not introducing additional parameters, calculations with $0 < \lambda < 1$ and $1 \leq \gamma \leq 50$ give no indication of zeros of Re $D(\nu)$ for energies up to $\nu = 50$. It can be said that the attractive part of the potential ($\equiv$ positive part of left-hand discontinuity) is not sufficiently strong to produce physical resonances. Notice that increase of γ strengthens the repulsion rather than the attraction.[20]

Thus one is led to a combination of (4.1) and (4.2) according to Fig. 4, where a sharp cutoff Λ is imposed on the vector meson contribution, so that

$$\omega_1 = \Lambda(\tfrac{1}{4}m^2 - 1). \quad (4.4)$$

It can be shown again that for all λ, $0 < \lambda < 1$, a ghost-free amplitude exists, because an iteration solution of (3.3) can be constructed; and $B(\nu)$, as defined in (3.4), has been found to satisfy (3.6), at least for $4 < m^2 < 50$ and $\Lambda \geq 10$.

The defect of this model is that it contains two free parameters, λ and Λ. In view of the smallness of the real part of the forward amplitudes observed in high-energy p–p and π–p scattering,[4] it is perhaps reasonable to assume that λ is close to unity, say $0.9 < \lambda < 1$. However, Λ remains in principle undetermined. Still, important information may be obtained by comparing the resulting solutions with those from a model with the same $\sigma_1(-\nu, m^2)$ (i.e., the same Λ) but with $\lambda = 0$. By keeping the same long-range part one may expect to get some information about the effect of the short-range force on various features of the amplitude.

This program is pursued in an approximate scheme defined as follows: With the left-hand discontinuity of Fig. 4, suppose that Eq. (2.16) is written in operator form:

$$f = f_0 + \lambda K_S \cdot f + K_F \cdot f, \quad (4.5)$$

where f_0 stands for $1/\omega$, K_S for the singular kernel $\pi^{-2}[\log(\omega'/\omega)/(\omega' - \omega)]$, etc. When the last part $(K_F \cdot f)$ is neglected, the solution of (4.5) is given in Sec. 3 and can be written

$$\tilde{f} = f_0 + \lambda R \cdot f_0; \quad (4.6)$$

R represents the resolvent (3.12). The approximate solution that will be used is

$$\hat{f} = f_0 + (\lambda K_S + K_F) \cdot \tilde{f}. \quad (4.7)$$

To compare this with the exact solution, note that (4.5) can be written

$$f = (1 + \lambda R) f_0 + (1 + \lambda R) \cdot K_F \cdot f. \quad (4.8)$$

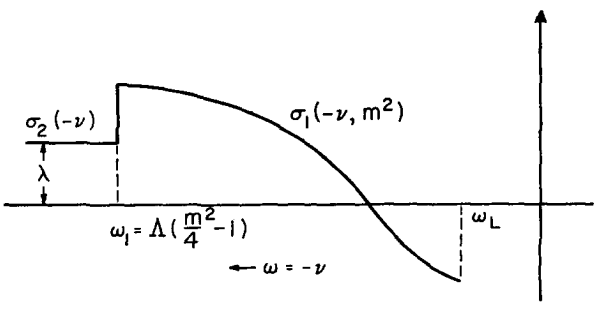


FIG. 4. Left-hand discontinuity in a realistic model combining long- and short-range forces.

[19] A. P. Contogouris and D. Atkinson, Nuovo Cimento **39**, 1102 (1965).

[20] The particular case $\gamma = \lambda(\omega_1 = \infty)$ corresponds to a left-hand discontinuity given entirely by exchange of an elementary vector meson. Here, for $0 < \gamma < 1$ no zeros of Re $D(\nu)$ are indicated in the region $\nu \leq 50$. For $\gamma > 1$, A. Bassetto and F. Paccanoni [Nuovo Cimento **44A**, 1139 (1966)] report the existence of a bootstrap solution free of arbitrary parameters, which in fact gives both the width and mass of the ρ in very good agreement with experiment. However, as has been stressed by those authors, the corresponding D function is expected to have unwanted zeros.

Reference 6 has shown that for $0 < \lambda < 1$, and with the resolvent (3.12), Eq. (4.8) is Fredholm; hence, for sufficiently small λ, an iteration solution exists. The first iteration gives

$$f_1 = (1 + \lambda R)f_0 + (1 + \lambda R) \cdot K_F \cdot \tilde{f}.$$

Use of the well-known identity[21]

$$K_S \cdot (1 + \lambda R) = R$$

shows that f_1 differs from $\hat{f}$ to terms of order $\lambda \, \|R \cdot K_F\| / \|K_F\|$.

Equations (2.15) and (2.16) show that, for $\lambda \to 0$, K_F tends to a finite limit, say $K_F^{(0)}$, so that (4.5) becomes

$$f = f_0 + K_F^{(0)} \cdot f \tag{4.9}$$

(a Fredholm equation). On the other hand, since R is the resolvent analytic at $\lambda = 0$, $\lim_{\lambda \to 0} \lambda R = 0$, and (4.8) reduces again to the form (4.9).

Suppose, however, that the calculation is carried out, not with R, but with another resolvent $R^{(1)}$, which contains a multiple of the homogeneous solution corresponding to $f = f_0 + \lambda K_S \cdot f$. It can be seen [Eq. (2.11) of Ref. 18] that, in the limit $\lambda \to 0$, $\tilde{f} \nrightarrow f_0$ and $\lambda R^{(1)} \nrightarrow 0$; hence in this case (4.9) is not reproduced. Clearly, on grounds of continuity, it is desirable that the limit $\lambda \to 0$ reproduce the situation that corresponds to simple exchange of a cutoff vector meson [Eq. (4.9)].

Note that in the limit $\lambda \to 0$, with $\tilde{f} \to f_0$, the approximate solution of (4.7) tends to $\hat{f} \to f_0 + K_F^{(0)} f_0$. Clearly, this is the first iteration of (4.9), usually called the "determinantal" solution[22]; most of the numerical results of bootstrap calculations have been obtained with this type of solution.

In terms of the solution $\tilde{D}(-\omega)$ of (3.14), Eq. (4.7) can be written as

$$\hat{D}(-\omega) = 1 + \frac{\omega}{\pi} \times \int_{\omega_L(m^2)}^{\infty} d\omega' \, \frac{K(\omega, \omega')\sigma(\omega', m^2)}{\omega'} \, \tilde{D}(-\omega'), \tag{4.10}$$

where $\sigma(\omega, m^2)$ is given by (4.1), (4.2) (or Fig. 4), and

$$K(\omega, \omega') = \frac{2}{\pi(\omega - \omega')} \left\{ \left(\frac{\omega}{\omega - 1} \right)^{\frac{1}{2}} \log \left[\omega^{\frac{1}{2}} + (\omega - 1)^{\frac{1}{2}} \right] - \left(\frac{\omega'}{\omega' - 1} \right)^{\frac{1}{2}} \log \left[\omega'^{\frac{1}{2}} + (\omega' - 1)^{\frac{1}{2}} \right] \right\}. \tag{4.11}$$

Again, to determine the resonances ($\nu > 0$), an analytic continuation to $\omega < 0$ is needed. For this, let

$$\omega = \cosh^2 (\tfrac{1}{2} i\pi + y) = -\sinh^2 y$$

and, for $\omega < 0$,

$$\sinh y = (-\omega)^{\frac{1}{2}}, \quad \text{i.e.,} \quad y = \log \left[(-\omega)^{\frac{1}{2}} + (-\omega + 1)^{\frac{1}{2}} \right].$$

Hence,

$$\left(\frac{\omega}{\omega - 1} \right)^{\frac{1}{2}} \log \left[\omega^{\frac{1}{2}} + (\omega - 1)^{\frac{1}{2}} \right] = (\tfrac{1}{2} i\pi + y) \tanh y = \left\{ \tfrac{1}{2} i\pi + \log \left[(-\omega)^{\frac{1}{2}} + (-\omega + 1)^{\frac{1}{2}} \right] \right\} \left(\frac{-\omega}{1 - \omega} \right)^{\frac{1}{2}}.$$

With this, one has

$$\operatorname{Re} \hat{D}(-\omega) = 1 + \frac{\omega}{\pi} \times \int_{\omega_L(m^2)}^{\infty} d\omega' \, \frac{\operatorname{Re} K(\omega, \omega')\sigma(\omega', m^2)}{\omega'} \, \tilde{D}(-\omega'), \tag{4.12}$$

where

$$\operatorname{Re} K(\omega, \omega') = \frac{2}{\pi(\omega - \omega')} \left\{ \left(\frac{-\omega}{1 - \omega} \right)^{\frac{1}{2}} \log \left[(-\omega)^{\frac{1}{2}} + (-\omega + 1)^{\frac{1}{2}} \right] - \left(\frac{\omega'}{\omega' - 1} \right)^{\frac{1}{2}} \log \left[\omega'^{\frac{1}{2}} + (\omega' - 1)^{\frac{1}{2}} \right] \right\}, \tag{4.13}$$

and

$$\operatorname{Im} \hat{D}(-\omega) = \frac{\omega}{\pi} \int_{\omega_L(m^2)}^{\infty} d\omega' \, \frac{\operatorname{Im} K(\omega, \omega')\sigma(\omega', m^2)}{\omega'} \, \tilde{D}(-\omega'), \tag{4.14}$$

where

$$\operatorname{Im} K(\omega, \omega') = \frac{1}{\omega - \omega'} \left(\frac{-\omega}{1 - \omega} \right)^{\frac{1}{2}}. \tag{4.15}$$

The condition, then, for a (narrow) resonance at $\nu = -\omega = \nu_R$ is $\operatorname{Re} \hat{D}(\nu_R) = 0$, and this results in an equation of the form

I $\quad \gamma = \gamma(\nu_R, m^2).$

The coupling necessary to produce a resonance at $\nu = \nu_R$ (input coupling) is determined by I. The self-consistent (bootstrap) solutions are defined by the condition

I′ $\quad m^2 = 4(\nu_R + 1),$

and by the relation of the width of the produced resonance (output width) to the coupling of the exchanged vector meson, which is

II $\quad \gamma = 6\pi \left(\frac{(1 + \nu)^{\frac{1}{2}}}{\nu^{\frac{3}{2}}} \frac{\operatorname{Im} \hat{D}(\nu)}{(d/d\nu) \operatorname{Re} \hat{D}(\nu)} \right)_{\nu = \nu_R}.$

[21] F. Smithies, *Integral Equations* (Cambridge University Press, New York, 1962).

[22] F. Zachariasen, Scottish Universities' Summer School Lecture Notes, Edinburgh, 1964.

This expression is approximate and is not a good estimate for the large widths which will prove necessary. Nevertheless, the main point, that the widths *are* reduced by the addition of a short-range force, is clearly valid.

5. NUMERICAL RESULTS AND DISCUSSION

Numerical calculations of the program of Sec. 4 have been carried out for several cutoff values in the range $5 \leq \Lambda \leq 80$. For each Λ, two cases were compared:

(a) $\lambda = 0$ (i.e., without short-range force);
(b) $\lambda = 0.95$ (i.e., with short-range force).

In Fig. 5, four curves are shown between γ, the coupling, and ν_R/m_π^2, the resonance position, for the cutoff $\Lambda = 10$. Curve (Ia) is a plot of equation (I), subject to equation (I′), that is to say, the relation between the input, or cross-channel coupling γ, and the mass of the produced resonance, this latter being constrained to be the same as the mass of the input resonance. Curve (Ib) is a similar plot, but this time with the short-range force added. Curve (IIa) is a plot of equation (II), so that γ is now the (reduced) width of the output resonance, ν_R/m_π^2 being, as before, the self-consistent resonance position. Curve (IIb) repeats this with the short-range force added. Finally, Fig. 6 is a similar graph for the cutoff $\Lambda = 40$; for other values of Λ the results remain qualitatively unchanged.

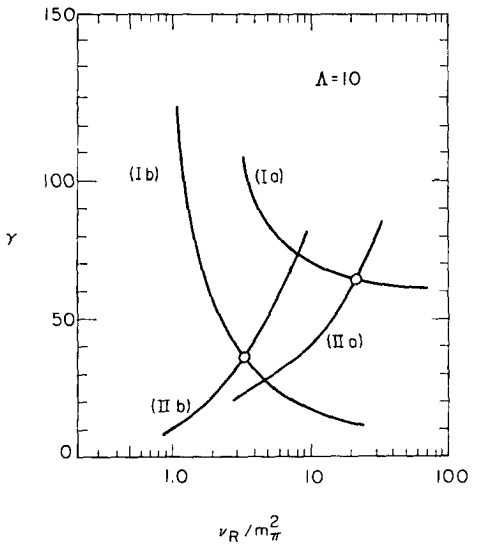


FIG. 5. Bootstrap solutions in the model of Fig. 4 for a cutoff $\Lambda = 10$. The curves I represent equation (I) of Sec. 4 subject to the condition (I′); the curves II represent equation (II). Case (a) corresponds to the absence of short-range force and Case (b) to the presence of short-range force.

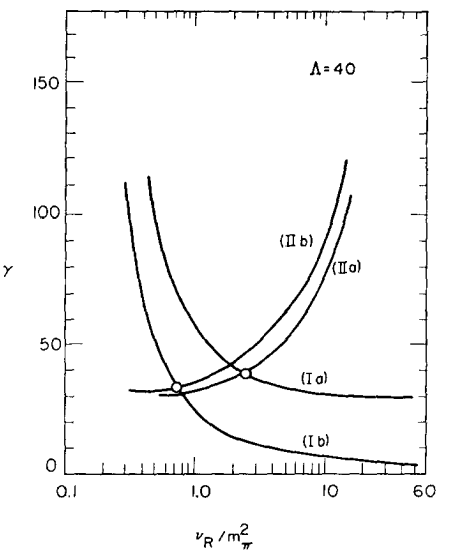


FIG. 6. The same as in Fig. 5 for $\Lambda = 40$.

Several observations can be made within this approximate model, on the basis of Figs. 5 and 6: In general, bootstrap solutions exist both without the short-range force [intersection of curves (Ia) and (IIa)] and with the short-range force [intersection of (Ib) and (IIb)].

For each case (a and b), the self-consistent mass and width as a function of Λ is presented in Fig. 7. For all Λ the addition of the short-range force decreases the self-consistent mass. For large Λ, the differences $\gamma_{(a)} - \gamma_{(b)}$ and $(m_\rho/m_\pi)^2_{(a)} - (m_\rho/m_\pi)^2_{(b)}$ are small, as they should be (most of the left-hand

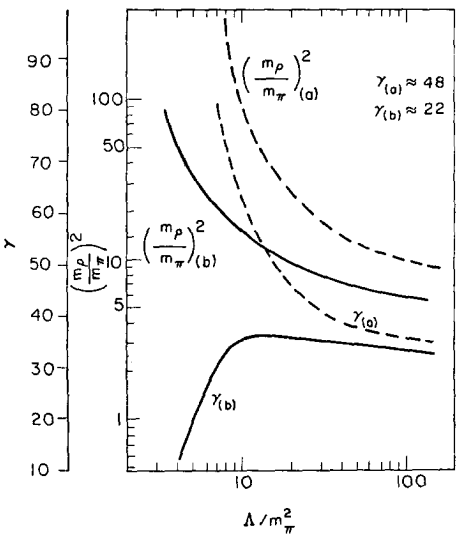


FIG. 7. The self-consistent mass ratio m_ρ/m_π and coupling γ as functions of the cutoff Λ. As always, Case (a) corresponds to the absence of short-range force and Case (b) to the presence of short-range force. Experimental values: $(m_\rho/m_\pi)^2 = 30$, $\gamma = 3.8$.

discontinuity being given by the ρ exchange); but as Λ decreases, these differences increase.

It is of particular interest that a bootstrap solution with the correct ρ-meson mass ($m_\rho^2 = 30$) does exist. Here the short-range attraction decreases the self-consistent width by a factor of 2 (from $\gamma_{(a)} \simeq 48$ to $\gamma_{(b)} \simeq 22$).

However, even with a short-range force, this width remains about six times as large as the experimental value. Thus, the conclusion is that the addition of the short-range attraction acts in the correct direction, but is not sufficient to explain the whole magnitude of the discrepancy.

(i) For a given resonance position ν_R/m_π^2, the addition of the short-range force significantly decreases the necessary input coupling γ [cf. curves (Ia) and (Ib)]. This is true for each cutoff and is an eminently reasonable state of affairs: if a short-range attraction is present, the long-range force needed to produce a resonance at a given position is reduced. Similarly, for a given γ, the resonance mass is decreased by the addition of the long-range force.

(ii) For a given ν_R, the short-range force increases the width of the output resonance, however [cf. curves (IIa) and (IIb)]. This is not surprising, for it is known that what is required to narrow the output resonance, for a given mass, is the addition of a long-range repulsion,[22] rather than a short-range attraction.

Of course, in this calculation many important contributions to the binding force have been omitted. For example, the exchange of two pions in relative S state could be significant, either if its contribution is strongly repulsive,[23] or if it is strongly attractive, with perhaps a resonance.[24] It is even possible that multiparticle exchange is important. Moreover, this simple model has neglected inelasticity. It is possible that a one-channel calculation would require a CDD pole, even if the correct inelasticity were used,[25,26] and that a dynamical calculation could only be done with good accuracy in a many-channel scheme. An $SU(3)$ model in which the $K\bar{K}$ channel was also incorporated suggested that this channel might not be too important[27]; but an $SU(6)$ model, in which the $\pi\omega$ channel also occurs, would, if it is to be believed, require a one-channel CDD pole, or equivalently a many-channel ND^{-1} system.[27] In this connection, it is interesting that a calculation by Fulco, Shaw, and Wong[28] of the ρ meson in the three-channel system ($\pi\pi$, $K\bar{K}$, $\pi\omega$), with a cutoff and no short-range force, gives, as usual, a resonance width that is too large, although the $K\bar{K}$ and $\pi\omega$ channels do assist in reducing the ρ width. It can be expected that the direction and order of magnitude of the effect of a short-range force will be the same in a more sophisticated model of this kind as it was in the work presented here. That is, we may expect a singular tail to assist materially in the narrowing of the π–π P-wave resonance.

From the mathematical point of view, the fact that the numerical calculations were done in an approximate scheme may be considered as unsatisfactory. It would certainly be of interest to repeat the whole program with the exact solution of (2.16) [or (4.5)]; and in view of the presented formalism and of the methods developed in Ref. 6, which reduce the singular to a Fredholm equation, this can be done in a straightforward manner.

ACKNOWLEDGMENTS

This work was started while the authors were at CERN, Geneva. One of them (D. A.) is indebted to Professor R. J. Eden for discussions. The other (A. P. C.) is grateful to Dr. V. Müller for helpful conversations and criticism, to Dr. J. D. Bessis, Dr. B. Diu, and Dr. A. Martin for discussions, and to Professor Ph. Meyer for the hospitality extended to him at Orsay. Finally, the authors are deeply indebted to G. Sheppey for performing the numerical calculations; for these the CERN CDC 6600 computer was used.

APPENDIX A

In this appendix, Eq. (3.10) is solved under the assumptions (3.1) and (3.2), plus the additional approximation involved in eliminating the gap between the left- and right-hand cuts ($\omega_L = 0$). This simplification, along with (3.1) and (3.2), can be characterized as a high-energy approximation. Such a situation is even more unrealistic than that of Sec. 3; however, insofar as one is concerned with the effects of the distant parts of the discontinuities on the resonance region, and because of the possibility of obtaining explicit and relatively simple solutions, its study is, perhaps, of some interest.

A disadvantage of this treatment is that, by replacing $\omega_L = 1$ by $\omega_L = 0$, the mass scale has been lost.

[23] G. F. Chew, Phys. Rev. **140**, B1427 (1965).
[24] C. Lovelace, R. M. Heinz, and A. Donnachie, Phys. Letters **22**, 332 (1966).
[25] E. J. Squires, Nuovo Cimento **34**, 1751 (1964).
[26] D. Atkinson, K. Dietz, and D. Morgan, Ann. Phys. (N.Y.) **37**, 77 (1966).
[27] D. Atkinson and M. B. Halpern, Phys. Rev. **150**, 1377 (1966).
[28] J. R. Fulco, G. L. Shaw, and D. Y. Wong, Phys. Rev. **137**, B1242 (1965).

Thus, it is necessary to subtract the N and D equations at some point $\nu = \nu_0 = -\omega_0$ (not the normal threshold), at which $D(\nu)$ can be normalized to unity; the value of ω_0 reintroduces a mass scale. With

$$R_1(\infty)N(-\omega_0) = a, \tag{A1}$$

the integral equation becomes

$$f(\omega) = \frac{1}{\omega - \omega_0} + \frac{a}{\pi}\frac{\log(\omega/\omega_0)}{\omega - \omega_0} + \frac{\lambda}{\pi^2}\int_0^\infty d\omega' \frac{\log(\omega'/\omega)}{\omega' - \omega} f(\omega'). \tag{A2}$$

In operator notation this can be written

$$f = a_1 + a_2 + (\lambda/\pi^2)Kf, \tag{A3}$$

where

$$a_1 \Rightarrow \frac{1}{\omega - \omega_0}, \quad a_2 \Rightarrow \frac{a}{\pi}\frac{\log(\omega/\omega_0)}{\omega - \omega_0}, \quad K \Rightarrow \frac{\log(\omega'/\omega)}{\omega' - \omega}.$$

Then, defining f_1, f_2 by

$$f_i = a_i + (\lambda/\pi^2)Kf_i \quad \text{for} \quad i = 1, 2, \tag{A4}$$

one has

$$f = f_1 + f_2. \tag{A5}$$

The functions f_i are solutions of (A4), which can be expressed in the form

$$f_i = a_i + \lambda R a_i, \tag{A6}$$

where R is a resolvent of the kernel K/π^2, satisfying

$$R = (1/\pi^2)K + (\lambda/\pi^2)K \cdot R. \tag{A7}$$

It was shown in Ref. 6 that R exists but is not unique: it has a two-parameter manifold. However, if the solution is required to have no singularity at $\lambda = 0$, and thus to admit a perturbation expansion in powers of λ, a unique resolvent is singled out, which has a branch point only at $\lambda = 1$ (see also Appendix B). This resolvent is

$$R(\omega, \omega'; \lambda) = \frac{1}{\pi[\lambda(1-\lambda)]^{\frac{1}{2}}} \cdot \frac{\sinh[S_0 \log(\omega'/\omega)]}{\omega' - \omega}, \tag{A8}$$

where

$$\lambda = \sin^2(\pi S_0) S_0 = (i\pi)^{-1} \log[(-\lambda)^{\frac{1}{2}} + (1-\lambda)^{\frac{1}{2}}],$$

with $0 < S_0 < \frac{1}{2}$ for $0 < \lambda < 1$.

One might solve the two equations (A4) by using the resolvent (A8). To find $f_1(\omega)$ it is necessary to evaluate (A6) for $i = 1$ by performing the integral explicitly. The result is

$$f_1(\omega) = \frac{1}{\omega - \omega_0}\cosh[S_0 \log(\omega/\omega_0)]. \tag{A9}$$

The equation for $f_2(\omega)$ is trivial, since

$$a_2(\omega) = (a/\pi)K(\omega, \omega_0), \tag{A10}$$

so that by comparing (A4) for $i = 2$ with (A7), one has immediately

$$f_2(\omega) = a\pi R(\omega, \omega_0; \lambda). \tag{A11}$$

Finally, by (A5), one has the solution

$$\begin{aligned} D(\nu) &= (\omega - \omega_0)f(\omega) \\ &= \cosh[S_0 \log(\omega/\omega_0)] \\ &\quad + B^{-1}\sinh[S_0 \log(\omega/\omega_0)], \end{aligned} \tag{A12}$$

where

$$B \equiv [\lambda(1-\lambda)]^{\frac{1}{2}}/a.$$

Substituting $\omega = -\nu$, one has, for ν real and positive ($\equiv$ physical region), the following real and imaginary parts:

$$\begin{aligned} \operatorname{Re} D(\nu) &= \cos(\pi S_0)\left[\cosh\left(S_0 \log\frac{\nu}{\omega_0}\right) + \frac{1}{B}\sinh\left(S_0 \log\frac{\nu}{\omega_0}\right)\right], \\ \operatorname{Im} D(\nu) &= \sin(\pi S_0)\left[\sinh\left(S_0 \log\frac{\nu}{\omega_0}\right) + \frac{1}{B}\cosh\left(S_0 \log\frac{\nu}{\omega_0}\right)\right]. \end{aligned} \tag{A13}$$

Now, the condition for a resonance at $\nu = \nu_R$ is $\operatorname{Re} D(\nu_R) = 0$, i.e.,

$$\tanh[S_0 \log(\nu_R/\omega_0)] = -B. \tag{A14}$$

As ν_R changes from zero to infinity, the left-hand side progresses monotonically from -1 to $+1$. Accordingly, if $|B| < 1$, there is one, and only one, solution of (A14), while if $|B| > 1$ there are no solutions. In the former case the coupling (or, equivalently, the width) of the resonance is

$$\begin{aligned} \frac{g^2}{4\pi} &\simeq \frac{3}{\nu_0}\left[\frac{\operatorname{Im} D(\nu)}{(d/d\nu)\operatorname{Re} D(\nu)}\right]_{\nu=\nu_0} \\ &= \frac{3}{S_0}\tan(\pi S_0) = \left(\frac{\lambda}{1-\lambda}\right)^{\frac{1}{2}}\frac{3\pi}{\arcsin\lambda^{\frac{1}{2}}}. \end{aligned} \tag{A15}$$

In the first equality of (A15), the phase-space factor $[(\nu_0 + 1)/\nu_0]^{\frac{1}{2}}$ has been replaced by unity. Notice that the subtraction point ω_0 and the subtraction constant $N(-\omega_0)$ (or equivalently B) have disappeared, and (A15) is a simple equation involving only λ and the width $g^2/4\pi$.

Hence, in the approximation $\omega_L = 0$, and provided that $|B| < 1$, a constant left-hand discontinuity is capable of producing a resonance at some point

$\nu = \nu_R$ satisfying (A14). However, one can see from (A15) that for all λ, $0 < \lambda < 1$, the corresponding width is exceedingly large ($g^2/4\pi > 3\pi$; compare, for example, with the experimental value for the width of $\rho \to 2\pi$: $g^2_{\rho\pi\pi}/4\pi \simeq 0.6$). Such a wide resonance can hardly produce any of the usual effects on the cross section of the corresponding process [and hardly justifies the use of formulas like (A15), which are meaningful only for narrow resonances]. Thus it cannot be considered acceptable.

To find solutions of the equation

$$g^2/4\pi = (3/S_0)\tan(\pi S_0)$$

corresponding to $g^2/4\pi \simeq 1$, one needs S_0 outside the interval $0 < S_0 < \frac{1}{2}$. As the first Riemann sheet of the λ plane corresponds to $-\frac{1}{2} < S_0 < \frac{1}{2}$, this means that one has to go to higher λ sheets. On the higher sheets, branch points exist at both $\lambda = 0$ and $\lambda = 1$ (see Appendix B); moreover, the continuation of a solution onto a higher sheet is not necessarily a solution of the original equation, because the integral in (A2) will no longer converge. The continued solution corresponds to a higher CDD class.[14] Thus, *in fine*, there is no resonance with acceptable width generated by a featureless, constant left-hand discontinuity.

APPENDIX B

The purpose of this appendix is twofold: First, to present the sheet structure in λ of the resolvent of Appendix A [Eq. (A8)]; second, to study in certain simple examples (corresponding to $\omega_L = 0$) the zeros of the denominator function. As the approximation $\omega_L = 0$ leads to relatively simple explicit solutions having a number of features in common with the exact ones, the conclusions are expected to provide useful insight.

To find the structure of (A8) it is convenient to map the infinity of Riemann sheets in λ onto the complex plane of another appropriate variable w defined by

$$\lambda = \cos^2(\pi w). \tag{B1}$$

Under this mapping,

$$R(\omega, \omega'; \lambda) = \frac{2}{\pi(\omega - \omega')} \cdot \frac{\sinh[(\frac{1}{2} - w)\log(\omega'/\omega)]}{\sin(2\pi w)}. \tag{B2}$$

This is a meromorphic function of w; its poles correspond to branch points in λ and appear at

$$w = m, \quad m = 0, \pm\tfrac{1}{2}, \pm 1, \pm\tfrac{3}{2}, \cdots. \tag{B3}$$

Thus, the various sheets of λ are mapped onto parallel strips of the w plane (Fig. 8).

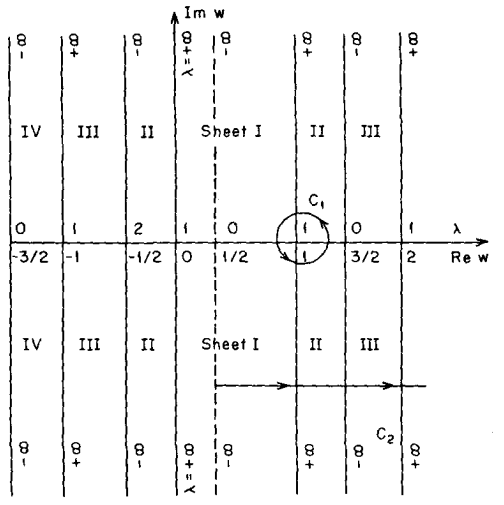


FIG. 8. Appropriate mapping [defined by (B1)] of the resolvent of Eq. (A8), which determines its structure in λ.

The point $\lambda = 0$ corresponds to $m = \pm\frac{1}{2}, \pm\frac{3}{2}, \cdots$. In particular, $w = \frac{1}{2}$ does not lead to a singularity of (B2); and as the first sheet of λ corresponds to $0 < \mathrm{Re}\, w < 1$, it is concluded that (B2) has no singularity at $\lambda = 0$ on the first sheet. However, on higher sheets $\lambda = 0$ is a branch point.

Sheet II is defined to be the sheet connected to sheet I across the cut $1 \leq \lambda < \infty$ and maps onto the strip $1 < \mathrm{Re}\, w < \frac{3}{2}$. Here there are branch points at both $\lambda = 0$ (cut $-\infty < \lambda < 0$) and $\lambda = 1$ (cut $1 < \lambda < \infty$), and this is true for all higher sheets. It follows that a double circuit around $\lambda = 1$ which does not enclose $\lambda = 0$ (C_1 of Fig. 8) brings one back onto sheet I; however, circuits enclosing $\lambda = 0$ and $\lambda = 1$ (C_2 of Fig. 8) lead into higher sheets. Notice the similarities with the sheet structure of the exact resolvent (for $\omega_L \neq 0$)[15]: as in that case, the branch points may be said to behave individually like square roots but together like a logarithm.

Consider now the zeros of $D(\nu)$ corresponding to the resolvent (B2), and take (for simplicity) $N(-\omega_0) = 0$. Then Eq. (A12) reduces to

$$D(-\omega) = \cosh[S_0 \log(\omega/\omega_0)]. \tag{B4}$$

Clearly, the physical sheet of ω is

$$-\pi < \arg\omega < \pi \tag{B5}$$

and corresponds to $0 < \arg\nu < 2\pi$.

Suppose now that λ varies over real values. At first $-\infty < \lambda < 0$ corresponds to $w = k + \frac{1}{2} + iv$, $k = 0, \pm 1, \pm 2, \cdots$, and $0 < v < \infty$ (or $-\infty < v < 0$). Hence $\lambda = -\sinh^2(\pi v)$, $S_0 = -iv$, and the zeros of (B4) appear at

$$\omega = \omega_0 \exp[(n + \tfrac{1}{2})(\pi/v)], \quad n = 0, \pm 1, \pm 2, \cdots. \tag{B6}$$

This gives two sets, each of an infinite number of zeros, lying on the positive real ω axis; for $0 < v < \infty$, the one set, which corresponds to $n = -1, -2, -3, \cdots$, accumulates at $\omega = 0$, and the other set ($n = 0, 1, 2, \cdots$) accumulates at $\omega = \infty$. In view of (B5), these zeros lie on the physical sheet of the complex ν plane (along the negative real ν axis).

Next, suppose that λ varies on its first sheet along $0 < \lambda < 1$. From Fig. 8 this corresponds to w = real, $0 < w < 1$. Now, the zeros of (B4) appear at

$$\omega = \omega_0 \exp\left[\frac{(2n+1)\pi}{2w-1}\right], \quad n = 0, \pm 1, \pm 2, \cdots.$$

For all n, and w in $0 < w < 1$,

$$|\arg \omega| > \pi.$$

Hence, none of these zeros lies on the physical sheet of ω (or ν).

Finally, let $\lambda > 1$. This corresponds to $w = k + iv$, $k = 0, \pm 1, \pm 2, \cdots$, and $0 < v < \infty$ (or $-\infty < v < 0$). Hence $\lambda = \cosh^2(\pi v)$ and the zeros of (B4) appear at

$$\omega = \omega_0 \exp\left[(2n+1)\pi \frac{2v - i}{(2v)^2 + 1}\right], \quad n = 0, \pm 1, \pm 2, \cdots.$$

For all v, $0 < v < \infty$, this relation gives at least one pair of complex zeros on the physical sheet of ω. Note that the zeros of each pair do *not* appear at complex conjugate positions; in this case the Riemann–Schwartz reflection symmetry is violated.

The conclusion is that for $\lambda < 0$ and $\lambda > 1$, the denominator function has zeros on the physical sheet of the complex ν plane which correspond to unwanted poles of the amplitude (ghosts). However, for $0 < \lambda < 1$ these zeros disappear from the physical sheet.

This conclusion can be further strengthened by similar analysis of a different resolvent. For example,

$$R^{(1)}(\omega, \omega'; \lambda) = \frac{1}{2\pi[\lambda(\lambda-1)]^{\frac{1}{2}}} \cdot \frac{\sin[q_0 \log(\omega'/\omega)]}{\omega' - \omega} \cdot \frac{\omega' + \omega}{(\omega'\omega)^{\frac{1}{2}}},$$

where

$$q_0 = \pi^{-1} \log[\lambda^{\frac{1}{2}} + (\lambda - 1)^{\frac{1}{2}}]$$

is one of the resolvents of (A2) having a branch point at $\lambda = 0$.[6] Here, for $-\infty < \lambda < 0$, $D(\nu)$ has at least one complex pair of zeros on the first sheet of ν (violating the Riemann–Schwartz symmetry); and for $\lambda > 1$ it has a double infinity of zeros along $-\infty < \nu < 0$. However, again for $0 < \lambda < 1$, no zeros appear on the physical sheet of ν.

In view of these examples and of more general theorems on the existence of ghost-free solutions of partial-wave dispersion relations,[16] one presumes that only the case $0 < \lambda < 1$ (considered in Secs. 2–5) can possibly lead to solutions of physical interest.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 9 SEPTEMBER 1968

Threshold-Behavior Requirements for P-Wave N/D Equations*

H. CORNILLE AND G. RUBINSTEIN
Institut de Physique Nucléaire, Division de Physique Théorique, 91 Orsay, France

(Received 17 July 1967)

For nonrelativistic P-wave N/D equations we consider the two different cases depending on whether the solutions of the subtracted equations, assuring automatically the right threshold-behavior conditions, satisfy the unsubtracted equations or not. In both cases we investigate the meaning of the corresponding interactions.

I. INTRODUCTION

Among different theoretical and practical difficulties connected with the use of partial-wave dispersion relations as well as N/D equations, a current one is the requirement of correct threshold behavior for the solutions.

When we consider the unsubtracted equation and assume for the left-hand-cut discontinuity some general condition, such that the solution exists, then in general this is not sufficient for the solution to exhibit the threshold-behavior requirements. Of course it can be argued that this comes mainly because we can determine only a part of the left-hand-cut discontinuity, and if the whole discontinuity were known this problem probably would not occur. Nevertheless it can be realized that these threshold-behavior conditions for the solutions of the unsubtracted equations are in fact intrinsic conditions for the left-hand-cut discontinuities.

Furthermore, it is usual to consider as "ansatz equations" some subtracted equation where the threshold behavior conditions are automatically satisfied. Whether or not such procedure is always allowed is not so clear. For instance, it is well known that in the N/D equations in general we modify the asymptotic D behavior. We have in fact two different cases. In the first case the solution of the "ansatz" subtracted equation is also a solution of the unsubtracted equation. This means that the above intrinsic conditions for the left-hand-cut discontinuity are satisfied. In the second case these intrinsic conditions are not satisfied and the solutions of the ansatz subtracted equations are not solutions of the unsubtracted equations. In that last case the following question arises: What is the meaning of the ansatz subtracted equation?

Consider for instance the nonrelativistic N/D equations; we know that the Yukawa-type family belongs to the first case. But we recall that dispersion relations leading to N/D have been proved only for Yukawa-type family. It follows that if we can show for the second case that our considered ansatz subtracted equation has a meaning at all, then we must necessarily find some kind of singular interactions. We recall that the Marchenko[1] inversion formalism is a powerful tool[2] for studying the meaning of the N/D equations in the $l = 0$ case. We shall use in this paper Blažek's[3] extension for $l \neq 0$ of the Marchenko formalism. In the following we shall give for $l \neq 0$ some general results concerning the inversion problem (mainly the dispersive formulation of the Jost solutions). However, we shall investigate explicitly whether or not we can attribute a real meaning for some ansatz subtracted equation only for the P-wave case.

II. SUBTRACTED AND UNSUBTRACTED EQUATIONS IN CONNECTION WITH THRESHOLD BEHAVIOR FOR $l \neq 0$

In order to reconstruct the $S_l(k)$ matrix from the left-hand-cut discontinuity, we consider the resulting integral equation of $f_l(k)/f_l(-k)$ [an approach equivalent to N_l/D_l, $f_l(k)$ being the Jost function $f_l(-k) = D_l(k^2)$] such that $S_l(k) = f_l(k)/f_l(-k)$.

Unsubtracted Equations

We assume in this section that the potentials are of the generalized Yukawa-type family "regular" at the origin

$$\lambda V(r) = \int_m^\infty e^{-\alpha r} \lambda C(\alpha)\, d\alpha \quad \text{and} \quad \int_m^\infty \frac{|C(\alpha)|}{\alpha^2}\, d\alpha < \infty.$$

In that case we have[4] the following resulting integral equation:

$$\begin{cases} f_l(k = -ix) = F_l(x), \\ F_l(x) = 1 + \int_{m/2}^\infty \mu \dfrac{\Delta_l(y)}{x + y} F_l(y), \end{cases} \tag{1}$$

* Paper presented at the 1967 International Conference on High Energy Physics at Heidelberg.

[1] Z. S. Agranovich and V. A. Marchenko, *The Inverse Problem of Scattering Theory* (Gordon and Breach, Science Publishers, Inc., New York, 1963).

[2] H. Cornille, J. Math. Phys. **8**, 2268 (1967); IPNO TH76 November 1966 (to be published in J. Math. Phys.).

[3] M. Blažek, Commun. Math. Phys. **3**, 282 (1966).

[4] E. J. Squires, *Strong Interactions and High Energy Physics* (Oliver and Boyd, London, 1961).

where we call $-\mu\Delta_l$ the discontinuity of the S_l matrix,

$$\mu\Delta_l(x) = -(2i\pi)^{-1}[S_l(\mu, ix + \epsilon) - S_l(\mu, ix - \epsilon)], \quad x > m/2. \quad (2)$$

For the Yukawa-type family (short-range type), the threshold-behavior conditions for the S_l matrix are satisfied[5] for any l,

$$|S_l(k) - 1| \underset{k\to 0}{\simeq} O(k^{2l+1}). \quad (3)$$

This means that the $F_l(x)$ solutions of (1) corresponding to the Yukawa-type family satisfy automatically

$$\left(\frac{\partial^{(2p-1)}}{\partial x^{2p-1}} F_l(x)\right)_{x=0} = 0, \quad p = 1, 2, \cdots, l, \quad (4a)$$

or equivalently

$$\sigma_l^p(\mu) = \int_{m/2}^{\infty} \frac{F_l(x)\Delta_l(x)}{x^{2p}}\, dx = 0, \quad p = 1, 2, \cdots, l. \quad (4b)$$

We emphasize that (4b) must be considered as an intrinsic condition for the discontinuity

$$\left(\int_{m/2}^{\infty} \frac{F_l(\mu_l)\Delta_l(x)}{x^{2p}}\, dx = 0\right)$$

and a boundary condition for (1). $\mu\Delta_l$, corresponding to Yukawa-type family, is such that (4b) is satisfied; then we can try to include these boundary conditions in (1). If (4a) and (4b) are satisfied, then

$$F_l(x) = F_l(0) + \left[F_l(x) - F_l(0) - \sum_{p=1}^{l} \frac{x^{2p-1}}{(2p-1)!}\left(\frac{\partial^{2p-1}}{\partial x^{2p-1}} F_l(x)\right)_{x=0}\right]$$

and we get, for $l \geq 1$,

$$F_l(x) = F_l(0) + \mu\int_{m/2}^{\infty} g_l(x, y)\Delta_l(y)F_l(y)\, dy,$$

$$g_l(x, y) = \frac{x^2}{y^{2l}}\left(\frac{x^{2l-1} - y^{2l-1}}{x^2 - y^2}\right). \quad (5)$$

B. Ansatz Subtracted Equations

Now we consider formally the following equation considered as ansatz:

$$\tilde{F}_l(x) = \tilde{F}_l(0) + \mu\int_{m/2}^{\infty} g_l(x, y)\Delta_l(y)\tilde{F}_l(y)\, dy \quad (6)$$

without any subsidiary condition and we put formally

$$\tilde{S}_l = \frac{\tilde{F}_l(x = ik)}{\tilde{F}_l(-x = -ik)}.$$

We define

$$\tilde{\sigma}_l^p(\mu) = \int_{m/2}^{\infty} \frac{\tilde{F}_l(y)\Delta_l(y)}{y^{2p}}\, dy \quad p = 1, 2, \cdots, l.$$

1. First, if

$$\tilde{\sigma}_l^p(\mu) = 0 \quad \text{for} \quad p = 1, 2, \cdots, l, \quad (4c)$$

it is easy to see from (6) that $\tilde{F}_l(x) = \text{const} \times F_l(x)$ where $F_l(x)$ is a solution of (1) satisfying (4b). (A normalization constant is not important for the ratio $\tilde{F}_l/\tilde{F}_l$ or F_l/F_l, so we are free to multiply both the numerator and the denominator by a constant.) Furthermore, if

$$\tilde{F}_l(0) - \int_{m/2}^{\infty} \frac{\tilde{F}_l(y)\Delta_l(y)}{y}\, dy = 1,$$

the constant is 1; $F_l = \tilde{F}_l$. So the conditions $\tilde{\sigma}_l^p = 0$ $(p = 1, \cdots, l)$ for the ansatz subtracted equation (6) are the conditions such that the solution $\tilde{F}_l$ is also solution of the unsubtracted equation (1) or equivalently these are the conditions such that F_l solutions of (1) satisfy the threshold conditions (4b). This means only that in this case the threshold-behavior requirements are intrinsic conditions for the discontinuity and do not depend on whether the equation considered is (1) or (6). In this case we note also that $\tilde{F}_l(x) \underset{x\to\infty}{\to} \text{const}$.

2. Second, for any solution of (6) we have automatically

$$\left(\frac{\partial^{2p-1}}{\partial x^{2p-1}} \tilde{F}_l(x)\right)_{x=0} = 0, \quad p = 1, 2, \cdots, l \quad (4d)$$

such that the right threshold behavior is always satisfied for $\tilde{S}_l$.

3. Third, we consider (6) where (4c) is not satisfied (at least one $\tilde{\sigma}_l^p \neq 0$). But of course (4d) is verified. In those cases where the solution of (1) is not a solution of (6) and where the solution of (1) does not satisfy (4) and (4b) the following question arises. Has the corresponding ansatz (6) a meaning? Or equivalently, can we give a real meaning to the interactions "$\mu\Delta(x)$" or to the $\tilde{S}_l$ matrix? We note that we know *a priori* in this case that even if we are able to analyze the interactions, we cannot find a member of the Yukawa family, so that we know *a priori* that we must find some kind of singularity. We emphasize that a different ansatz is possible in order to force the threshold behavior. We consider here one particular ansatz and try to understand the corresponding meaning. In fact we shall investigate here the P-wave case, but we shall give also some general results. In the following we consider "regular discontinuities" $\mu\Delta_l(x)$ of (1) such that the Fredholm-type solution of (1) exists.

We recall[2] that for $l = 0$ the Marchenko[1] inversion formalism gives the possibility to interpret the N/D equations. We shall use the extension of Marchenko

[5] With the modifications for exceptional cases; R. G. Newton, J. Math. Phys. **1**, 319 (1960).

formalism given by Blaźek[3] for $l \neq 0$ and try to see if for $l = 1$ the ansatz (6) has a meaning at all.

III. INVERSION FORMALISM

We do not want to give a complete analysis but only integral equations for the Jost solutions.

At the beginning we assume only that the potentials $V(r)$ satisfy sufficient conditions such that the following integral equation for the Jost solutions $f_l(k, r) \simeq i^l e^{-ikr}$ exists for $r > 0$:

$$f_l(k, r) = w_l(kr) - \int_r^\infty dr' g_l(k, r, r') V(r') f_l(k, r'), \quad (7)$$

$$w_l(\rho) = i^l e^{-i\rho} \sum_{n=0}^{l} \left(\frac{i}{-2\rho}\right)^n a_{l,n}, \quad a_{l,n} = \frac{(l+n)!}{n!(l-n)!}, \quad (8)$$

$$g_l(k, r, r') = i(-)^l (2k)^{-1} [w_l(kr) w_l(-kr') - w_l(-kr) w_l(kr')]. \quad (9)$$

We note that we do not assume anything for V when $r \to 0$, but for $r \to \infty$ we assume that V decreases exponentially or weaker conditions. Similarly, as for $l = 0$, we put formally

$$f_l(k, r) = w_l(k, r) + \int_r^\infty K_l(r, t) w_l(kt)\, dt. \quad (10)$$

In Appendix A, using the properties of spherical Hankel transform,[6] it is shown that if we substitute (10) into (7) we get

$$(d/dr)[K_l(r, t = r)] = -\tfrac{1}{2} V(r). \quad (11)$$

The method used gives also the possibility of obtaining an integral equation for $K_l(r, t)$ from $V(r)$ and consequently the possibility of finding sufficient conditions for $V(r)$ such that $K_l(r, t)$ exists (as has been done[1] for $l = 0$). For simplicity we shall not investigate this integral equation in this paper.

Now we assume that the potentials are of the Yukawa type. Blažek[3] has given the extension of the Marchenko equations for $K_l(r, t)$ in the case $l \neq 0$:

$$K_l(r, t) = \bar{\mathfrak{G}}_l(r, t) + \int_r^\infty K_l(r, u) \bar{\mathfrak{G}}_l^l(u, t)\, du, \quad (12)$$

where $\bar{\mathfrak{G}}_l$ is the scattering data[7]

$$\bar{\mathfrak{G}}_l(r, t) = \frac{(-1)^{l+1}}{2\pi} \times \int_{-\infty}^{+\infty} dk (1 - S_l(k)) h_l(-ikr) h_l(-ikt) \quad (13)$$

and[8] $i^l h_l(iz) = w_l(z)$.

For the Yukawa-type family, we can rotate the integration path in the upper-k complex plane (as for $l = 0$, see Ref. 2) and we get

$$\bar{\mathfrak{G}}_l(r, t) = (-)^l \mu \int_{m/2}^\infty \Delta_l(x) h_l(xr) h_l(xt)\, dx. \quad (14)$$

If we substitute (12) in (10), taking into account (14), then we get an integral equation for the Jost solutions

$$\begin{cases} F_l(x, r) = f_l(k = -ix, r)/i^l, \\ F_l(x, r) = h_l(xr) + \mu \int_{m/2}^\infty (-)^l \Delta_l(y) F_l(y, r) G_l(x, y, r)\, dy, \\ G_l(x, y, r) = \int_r^\infty h_l(xt) h_l(yt)\, dt. \end{cases} \quad (15)$$

For $l = 0$, (15) reduces to the previously found integral equations.[2]

For regular potentials like the Yukawa family, where the singularity of the solutions are given by the centrifugal potential, the Jost function $f_l(k)$ is

$$f_l(k) = \lim_{r \to 0} r^l k^l f_l(k, r)/(2l-1)!!. \quad (16)$$

So we define $\tilde{f}_l(k, r) = r^l k^l f_l(k, r)/(2l-1)!!$ and we get from (15)

$$\begin{aligned} \tilde{F}_l(x, r) &= \tilde{f}_l(k = -ix, r), \\ \tilde{F}_l(x, r) &= (xr)^l \frac{h_l(xr)}{(2l-1)!!} \\ &\quad + \mu \int_{m/2}^\infty (-)^l \Delta_l(y) \tilde{F}_l(y, r) \left(\frac{x}{y}\right)^l G_l(x, y, r). \end{aligned} \quad (17)$$

In conclusion, for the Yukawa family we know that $\tilde{F}_l(x, 0) = F_l(x)$, so that (17) must reduce to (1) when $r \to 0$.

Now we do not assume that $\mu\Delta_l(x)$ corresponds necessarily to a member of the Yukawa family, but we still consider a "regular discontinuity" for $\Delta(x)$ such that the Fredholm-type solutions of (1) exist [independently from the significance of this solution of (1)].

We get formally from (12), (14), and (15) a set of

[6] H. Cornille, Compt. Rend. **251**, 2135 (1960).

[7] Blažek has considered only the case where no bound states are present. But for the Yukawa family and $\mu\Delta(x)$ "sufficiently weak" we are in this case [see, for instance, $l = 0$ (Ref. 2)]. If bound states are present, we define the scattering data (similarly as for $l = 0$) as coming from two parts: the first one gives the contribution of the continuum as (13), the second supplementary part comes from these bound states. Finally, our fundamental equations (12) and (15) remain unchanged with this modified scattering data.

[8] Note that in the integrand of (13) the right threshold behavior for $S_l(k)$ is necessary in order to avoid a singularity when $k \to 0$.

results. First if we put

$$F_l(x, r) = \int_r^\infty h_l(xy)\phi_l(r, y)\, dy$$

and $\bar{\mathcal{C}}_l = \mu\mathcal{C}_l$, we get

$$\phi_l(r, y) = \frac{\delta(y - r)}{(2l - 1)!!} + \mu\int_r^\infty \mathcal{C}_l(y, t)\phi_l(r, t)\, dt, \quad (18a)$$

$$K_l(r, y) = \mu\mathcal{C}_l(r, y) + \mu\int_r^\infty \mathcal{C}_l(y, t)K_l(r, t)\, dt. \quad (18b)$$

We see that (18a) and (18b) have exactly the same kernel. This is in fact the main property giving the possibility of connecting the N/D formalism and the inversion formalism. For instance, for $r > 0$ the eigenvalues (μ values such that nontrivial solutions of the homogeneous equations exist) of (15) or (18) are the same. Furthermore, we remark that the kernel of (17) is $(x/y)^l$ times that of (15) such that the traces of both kernels are the same and consequently the Fredholm determinants for $r > 0$ of (15), (17), and (18) are also the same. Moreover, we shall show that, as for $l = 0$, this Fredholm determinant has a key role in the theory.

In Appendix B it is shown that the solution of (18b) is such that

$$K_l(r, r) = \frac{(d/dr)\mathfrak{D}_l(\mu, r)}{\mathfrak{D}_l(\mu, r)} \quad (19)$$

where $\mathfrak{D}_l(\mu, r)$ is the Fredholm determinant of (15), (17), and (18). Then the potential corresponding to the discontinuity[9] $\mu\Delta_l(x)$ is, from (11) and (19),

$$V(\mu, r) = -2\frac{d}{dr}\left(\frac{(d/dr)\mathfrak{D}_l(\mu, r)}{\mathfrak{D}_l(\mu, r)}\right), \quad (20)$$

as for[2] $l = 0$. We note also that as in the $l = 0$ case from (20) we can see that $V(\mu, r)$ are of short-range type and decrease exponentially.

On the one hand (similarly as for $l = 0$), if for some $r_0 > 0$, $\mathfrak{D}_l(\mu, r_0) = 0$, then V has second-order poles when $r \to r_0$, leading to a repulsive singular potential at r_0. If the multiplicity of the roots of $\mathfrak{D}_l(\mu, r_0) = 0$ is m, then

$$V(\mu, r) \underset{r\to r_0}{\simeq} 2m/(r - r_0)^2.$$

On the other hand, we note that when $r \to 0$ we have a great difference with the $l = 0$ case. The conditions about $\Delta_l(x)$, such that (1) is of Fredholm type, are not sufficient for $l \neq 0$ to assume that $\mathfrak{D}_l(\mu, 0)$ exists. This is due to the singular character of the kernel of (15), (17), and (18) becoming from

$$G_l(x, y, r) \simeq \frac{[(2l - 1)!!]^2}{x^l y^l r^{2l-1}}[1 + O(r)].$$

For instance, if we investigate $l = 1$, we find in general

$$\mathfrak{D}_{l=1}(\mu, r) \underset{r\to 0}{\simeq} \frac{\text{const}}{r}[1 + O(r)];$$

for $l = 2$,

$$\mathfrak{D}_{l=2}(\mu, r) \underset{r\to 0}{\simeq} \frac{\text{const}}{r^3}[1 + O(r)]$$

(these constants being functions of μ can vanish for special μ values). More generally, if

$$\mathfrak{D}_l(\mu, r) \underset{r\to 0}{\simeq} (\text{const}/r^n)[1 + O(r)],$$

we see that $V \underset{r\to 0}{\simeq} (-2n/r^2)$ leading to singular attractive potentials at the origin.

Of course, for a Yukawa-type family regular at the origin, the corresponding $\mathfrak{D}_l(\mu, r)$ cannot be singular at the origin.

First we note that (20) can be written

$$-\frac{1}{2}\int_r^\infty dx\int_x^\infty V(\mu, t)\, dt = \log \mathfrak{D}_l(\mu, r).$$

If we write V as a Laplace transform

$$V(\mu, r) = \int_m^\infty e^{-\alpha r}C(\alpha, \mu)\, d\alpha,$$

then

$$\frac{1}{2}\int_m^\infty e^{-\alpha r}\frac{C(\alpha, \mu)}{\alpha^2}\, d\alpha = -\log \mathfrak{D}_l(\mu, r).$$

As for $l = 0$, this last relation can be used in order to reconstruct the potential from the discontinuity.

Inversely, if we consider a Yukawa-type potential

$$\lambda V(r) = \lambda\int_m^\infty e^{-\alpha r}C(\alpha)\, d\alpha$$

and the corresponding discontinuities

$$\sum_p \lambda^p\Delta_l^p(x),$$

where Δ_l^p are the contributions coming from the nth Born approximation, then $\mathfrak{D}_l(\mu, r)$ becomes a function of λ:

$$\mathfrak{D}_l(\mu, r) \to d_l(\lambda, r)$$

and

$$\lambda\int_m^\infty e^{-\alpha r}\frac{C(\alpha)}{\alpha^2}\, d\alpha = -\log d_l(\lambda, r).$$

[9] Some care must be taken in order to understand these formulas. If we start from a given regular potential $V(r)$ and consider the corresponding set of left-hand-cut discontinuities Δ_l (which can be calculated from the Born expansions), it does not follow that there corresponds to the new set $(\mu\Delta_l)$, via (19), (20), an l-independent $V(\mu, r)$. For instance, when μ is varying, for $l = 0$ the only second-order poles which can appear for $V(\mu, r)$ correspond to repulsive singular potentials, whereas for $l \neq 0$ also singular attractive potentials can appear at the origin. In other words if we linearize the N/D equations with the same parameter $(\mu\Delta_l)$, then the corresponding potentials are, in general, l-dependent.

For the Yukawa-type family the Fredholm determinant of (15), (17), and (18) is well defined when $r = 0$.

In the following we consider only $l = 1$.

IV. CONNECTION BETWEEN INVERSION FORMALISM AND N/D EQUATIONS FOR P WAVE

In this section[10] we consider (17) for $l = 1$

$$\bar{F}(x, r) = e^{-xr}(1 + xr) - \int_{m/2}^{\infty} \mu\Delta(y) \times \bar{F}(y, r)\frac{x}{y}\left(\frac{1}{xyr} + \frac{1}{x + y}\right)e^{-(x+y)r}\,dy, \quad (17')$$

where

$$f(k, r) = \frac{\bar{F}(x = ik, r)}{kr} \underset{r\to\infty}{\simeq} ie^{-ikr}$$

is the Jost solution corresponding to

$$V(\mu, r) = -2\frac{d}{dr}\left(\frac{(d/dr)\mathfrak{D}(\mu, r)}{\mathfrak{D}(\mu, r)}\right),$$

$\mathfrak{D}(\mu, r)$ being the Fredholm determinant of (17'). We want to investigate in which cases we get from (17') either the unsubtracted equation

$$F(x) = 1 + \mu\int_{m/2}^{\infty} \frac{\Delta(y)F(y)}{x + y}\,dy \quad (1')$$

or the ansatz subtracted equation

$$\tilde{F}(x) = \tilde{F}(0) + \mu\int_{m/2}^{\infty} \frac{x^2\Delta(y)\tilde{F}(y)}{y^2(x + y)}\,dy. \quad (6')$$

We still assume that Δ is such that the Fredholm-type solution of (1') exists:

$$\int_{m/2}^{\infty} \frac{\Delta(x)}{x}\,dx < \infty,$$

$$\int_{m/2}^{\infty}\int_{m/2}^{\infty} \left(\frac{\Delta(y)}{x + y}\right)^2 dx\,dy < \infty. \quad (21)$$

The Fredholm-type solutions of (17'), (1'), and (6') can be written

$$\bar{F}(x, r) = e^{-xr}(1 + xr) + \frac{\mathcal{N}(x, \mu, r)}{\mathfrak{D}(\mu, r)}, \quad (22)$$

$$\tilde{F}(x) = \tilde{F}(0)\left[1 + \frac{\tilde{\mathcal{N}}(x, \mu)}{\mathfrak{D}(\mu)}\right],$$

$$F(x) = 1 + \frac{N(x, \mu)}{\mathfrak{D}(\mu)}, \quad (23)$$

$\mathfrak{D}(\mu)$ being both the Fredholm determinant of (1') and (6').

A simple example: the discontinuity replaced by one pole $\mu\Delta(x) = \mu\delta(x - b)$. First, from (1') we get

[10] We omit the index $l = 1$ in the following.

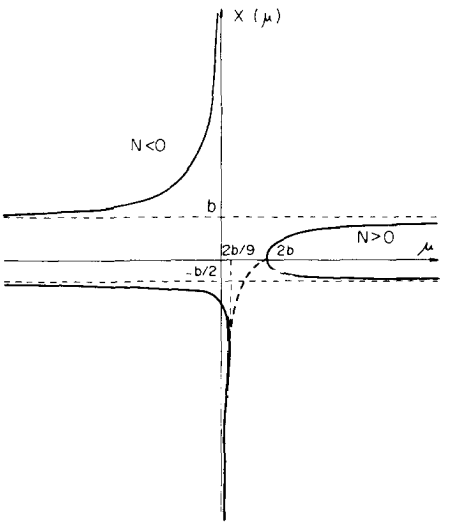


FIG. 1. $\mu\Delta = \mu\delta(x - b)$.

$F'(0)\mathfrak{D}(\mu) = -\mu/b^2$ such that $\mu = 0$ is the only value assuring the threshold behavior for (1'). This is a well-known result. It follows that (1') and (6') have no common solution. Second, from (17') and (6') we get

$$\frac{\bar{F}(x, r)}{r} \underset{r\to 0}{\simeq} \left(\frac{b^2}{\mu}\right)[1 + \mu g(x, b) - \mu g(b, b)],$$

$$g(x, b) = \frac{x^2}{b^2(x + b)},$$

$$\tilde{F}(x) = \frac{\tilde{F}(0)}{\mathfrak{D}(\mu)}[1 + \mu g(x, b) - \mu g(b, b)].$$

Then from inversion we get the solution of the ansatz equation (6') with $\tilde{F}(0) = -1/F'(0)$. We get also $\mathfrak{D}(\mu, r) = 1 + \mu[(b^2r)^{-1} + (2b)^{-1}]e^{-2br}$, so

$$V \underset{r\to 0}{\simeq} -2/r^2 \quad (\mu \neq 0).$$

But whereas in the case $\mu > 0$, V has no second-order pole for $r > 0$, on the contrary for $\mu < 0$, V has a second-order pole at $r_0 > 0$ such that $\mathfrak{D}(\mu, r_0) = 0$ and near this r_0, $V \simeq 2/(r - r_0)^2$ (we shall see that this type of repulsive potential leads to the presence of a ghost as in the $l = 0$ case).[2] In Fig. 1 we have represented the roots Re $x > 0$ of $\tilde{F}(x) = 0$ and the sign of the normalization N corresponding to $x > 0$ root (states in the physical sheet). For $\mu < 0$ we find always a ghost while for $\mu > 0$ when a state is present on the physical sheet ($\mu > 2b$) we have a true bound state.

From $S_1(k) = \tilde{F}(ik)/\tilde{F}(-ik)$ we get

$$tg\delta_{l=1} = \mu k^3/b^2(b^2 + k^2)\left(1 - \frac{\mu(b^2 + 3k^2)}{2b(k^2 + b^2)}\right),$$

and for $\mu > 0$ it is easy to verify that the usual Levinson theorem must be modified. We find $\delta(\infty) - \delta(0) = -\pi(n_{BS}) + \pi/2$ where $n_{BS} = 0, 1$ depending on whether a bound state is present or not. (We find also similar modification in the case $\mu < 0$ where a ghost appears.) This violation of the Levinson theorem with a supplementary term $\pi/2$ comes from the fact that $\tilde{F}(x) \xrightarrow[x\to\infty]{} x$ and not a constant as for the Yukawa family (note that $S_{l=1}(k) \xrightarrow[k\to\infty]{} -1$). We emphasize that for $\mu > 0$ the interaction is physically available[11,12] because the threshold behavior is verified by construction of the ansatz and because when a state is present on the physical sheet it is a true bound state.

If we use this left-hand-cut discontinuity of the $S_{l=1}$ matrix $-\mu\delta(x-b)$ in order to write a dispersion relation directly for the partial wave amplitude $[S_l(k)-1]/2ik$, it is easy to verify for $\mu > 0$ that the solution $\tilde{S}_1(k) = \tilde{F}_1(ik)/\tilde{F}_1(-ik)$ deduced from the ansatz (6′) satisfies this dispersion relation. In conclusion, the subtracted equation in this pole case is the only equation having a meaning in potential scattering but as a consequence the asymptotic behavior of the Jost function (or D in N/D) is modified: it is not 1 but x and further the corresponding V is $-2/r^2$ near the origin and a ghost is present when

$$V(r) \underset{r\to r_0}{\simeq} \frac{2}{(r-r_0)^2}, \quad r_0 > 0, \quad \mu < 0.$$

Two poles: $\mu\Delta(x) = \bar{\mu}_1 2b_1\delta(x-b_1) + \bar{\mu}_2 2b_2\delta(x-b_2)$; $b_1 > 0$, $b_2 > 0$. We get

$$\frac{\mathfrak{D}(\mu)F'(0)}{2} = \frac{\bar{\mu}_1}{b_1} + \frac{\bar{\mu}_2}{b_2} + \frac{\bar{\mu}_1\bar{\mu}_2(b_1-b_2)^2}{b_1b_2(b_1+b_2)}.$$

It is easy to verify that if $\mathfrak{D}(\mu)F'(0) = 0$, then the singular part of $\mathfrak{D}(\mu, r)$ vanishes so that we get a Yukawa-type family in general (in fact, a slight generalization as for $l = 0$, Ref. 2) and from (17′) it can be shown that $\bar{F}(x, 0)$ reduces to the solution of (1′). We see that in the $\bar{\mu}_1$, $\bar{\mu}_2$ plane only a curve corresponds to (1′) and this is the curve for the common solutions of both (1′) and (6′). If $F'(0)\mathfrak{D}(\mu) \neq 0$ then $\lim_{r\to 0} \bar{F}(x, r)/r$ is reduced to a solution of (6′) with $\tilde{F}(0) = -[F'(0)]^{-1}$ and $V \underset{r\to 0}{\simeq} -2/r^2$. In all cases the possibility of ghosts is still associated with second-order poles for $r > 0$ leading to marginally singular repulsive $V(\mu, r)$. In this case where the discontinuity is replaced by two poles, (1′) and (6′) have a meaning but they correspond to different domains in the $\bar{\mu}_1$, $\bar{\mu}_2$ plane. This example shall be treated in great details in a separate paper where we shall extend self-damping properties of the interactions from the S-wave case[13] to the P-wave one.[14]

General case: In Appendix C it is shown that

$$\mathfrak{D}(\mu, r) = \frac{1}{r}\bar{\mathfrak{D}}(\mu, r) + \bar{\bar{\mathfrak{D}}}(\mu, r) \underset{r\to 0}{\simeq} \frac{1}{r}\bar{\mathfrak{D}}(\mu, 0) + d(\mu) + O(r),$$

$$\mathcal{N}(x, \mu, r) = -\frac{e^{-xr}}{r}\bar{\mathfrak{D}}(\mu, r) + \bar{\bar{\mathcal{N}}}(x, \mu, r),$$

where, due to the condition (21), the limits $r \to 0$ of $\bar{\mathfrak{D}}$, $\bar{\bar{\mathfrak{D}}}$, and $\bar{\bar{\mathcal{N}}}$ exist. Then $\lim_{r\to 0}[\mathcal{N}(x, \mu, r) + \mathfrak{D}(\mu, r)]$ exists also. Further, it is shown in Appendix C, from the Fredholm-type solution of (1),

$$\bar{\mathfrak{D}}(\mu, 0) = -\mathfrak{D}(\mu)F'(0) \quad [F'(0) = ((d/dx)F(x))_{x=0}], \tag{24}$$

showing explicitly the connection between the threshold-behavior condition for (1′) and the singularity of $\mathfrak{D}(\mu, r)$ when $r \to 0$ [leading to $V(r) \underset{r\to 0}{\simeq} -2/r^2$].

In Appendix D it is shown that

$$\lim_{r\to 0}(\mathcal{N}(x, \mu, r) + \mathfrak{D}(\mu, r)) = \tilde{\mathcal{N}}(x, \mu) + \mathfrak{D}(\mu). \tag{D1}$$

If we apply these results to (22) we get

$$\frac{\bar{F}(x, r)}{r} \underset{r\to 0}{\simeq} \frac{\mathfrak{D}(\mu) + \tilde{\mathcal{N}}(x, \mu)}{-F'(0)\mathfrak{D}(\mu) + r\,d(\mu)}. \tag{25}$$

From (6′) it is easy to get

$$\tilde{F}(x) \xrightarrow[x\to\infty]{} x\int_{m/2}^{\infty} \frac{\Delta(y)\tilde{F}(y)}{y^2}\,dy.$$

In Appendix D it is shown that for the same $\mu\Delta(x)$ there exists a general relation between the corresponding solutions of (1′) and (6′):

$$\int_{m/2}^{\infty} \mu \frac{\Delta(y)F(y)}{y^2}\,dy = -F'(0) = \frac{1}{\tilde{F}(0)}\int_{m/2}^{\infty} \mu \frac{\Delta(y)\tilde{F}(y)}{y^2}\,dy. \tag{D2}$$

This shows explicitly that there exists a straightforward connection between threshold condition for the unsubtracted equation and asymptotic behavior for the subtracted one.

[11] We can verify in this case also that Martin's condition[12] about the sign of the contribution coming from the left-hand-cut discontinuity is satisfied.

[12] A. Martin, Nuovo Cimento **38**, 1326 (1965).

[13] H. Cornille, Nucl. Phys. **B3**, 655 (1967).

[14] H. Cornille and G. Rubinstein, Nuovo Cimento (to be published).

Coming back to (25) we see that we have two cases.

First case: $F'(0)\mathfrak{D}(\mu) = N'(0, \mu) \neq 0$.

This is the case where the solution of the ansatz (6') is not a solution of (1') and conversely. This is also the case where the solution of (1') does not give the right threshold behavior for the S matrix (we have an S-wave threshold behavior). In this case we see that $\mathfrak{D}(\mu, r) \underset{r\to 0}{\simeq} \text{const}/r$ and $V(\mu, r) \underset{r\to 0}{\simeq} -2/r^2$ such that the whole potential (V + centrifugal potential) simulates a regular potential like in the S-wave case when $r \to 0$. (But note that V decreases exponentially when $r \to \infty$ such that this whole potential goes like $2/r^2$ when $r \to \infty$.) From (25) we get

$$\lim_{r\to 0} \frac{\tilde{F}(x, r)}{r} = -\frac{1}{F'(0)}\left[1 + \frac{\tilde{\mathcal{N}}(x, \mu)}{\mathfrak{D}(\mu)}\right]. \quad (26)$$

We see that in this case the equation obtained from inversion formalism is the ansatz subtracted equation (6') with $\tilde{F}(0) = -1/F'(0)$ whereas (1') has no meaning (at least in potential scattering). But the interactions analyzed as potentials are not of the Yukawa-type family, regular at the origin, although they decrease exponentially. These potentials are marginally singular and attractive at the origin but can have poles of the second order for $r > 0$. We consider now μ values where no second-order poles for V appear at $r > 0$ [$\mathfrak{D}(\mu, r) \neq 0$ for $r > 0$], such that we have only the singularity $V \underset{r\to 0}{\simeq} -2/r^2$. If $\tilde{F}(x) = 0$ for some $x > 0$, then the corresponding state can be considered really as a bound state because the corresponding wave-function $\tilde{F}(x, r)$ is normalizable. Further, the threshold behavior is satisfied for the ratio $\tilde{F}(ik)/\tilde{F}(-ik)$ such that $S(k) = \tilde{F}(ik)/\tilde{F}(-ik)$ with its discontinuity $-\mu\Delta_t(x)$ really has a meaning. Nevertheless, such types of interactions lead to special features different from the usual Yukawa-type family.

From (D2) for $\tilde{F}(0) = -1/F'(0)$ we get

$$\tilde{F}(x) \underset{x\to\infty}{\longrightarrow} x \quad \text{and} \quad S(k) \underset{k\to\infty}{\longrightarrow} -1.$$

We see that the Jost functions (or D in N/D) do not go to constants when $x \to \infty$ (like regular Yukawa-type family) but on the contrary go to x. If we define $f_{l=1}(k) = \tilde{F}(ik)$ we can, as usual,[5] evaluate

$$\frac{1}{2i\pi}\int \frac{d}{dk} \log F_1(k)\, dk$$

in $\operatorname{Im} k < 0$ leading to a modified Levinson theorem (because $f(k) \underset{k\to\infty}{\longrightarrow} ik$, we have a supplementary contribution from the large semicircle in the lower half-plane). We get

$$\delta_{l=1}(0) - \delta_{l=1}(\infty) = \pi n_{\text{BS}} - \pi/2.$$

This type of Levinson modification in connection with asymptotic behavior of D not going to a constant is well known in the relativistic case, but we note that here we have $\pi/2$ and not a multiple of π. We want to show that the value $\pi/2$ is a nonrelativistic effect and can be well explained by Martin's theorem.[12] We consider relativistic partial wave amplitudes $a_l(k^2)$ for equal-mass particles (leading in the nonrelativistic limit to $k^{-1}e^{i\delta_l}\sin\delta_l$). We write a dispersion relation and if $\operatorname{Im} a_l(k^2) \underset{k^2\to\infty}{\longrightarrow} 0$ such that the left-hand-cut contribution $\underset{k^2\to\infty}{\longrightarrow} 0$ (the assumed properties of Δ_l in this paper leads also to this case), then[12] an unsubtracted relation holds and $\operatorname{Im} a_l(k^2) \underset{k^2\to\infty}{\longrightarrow} 0$. (Here we have found $\operatorname{Im} a_{l=1} \underset{k^2\to\infty}{\longrightarrow} k^{-1}$.) However, in the relativistic case the relativistic phase factor $\underset{k^2\to\infty}{\longrightarrow} 1$, then necessarily $\delta(\infty) = 0(\pi)$; whereas in the nonrelativistic case because of the present k^{-1} factor,

$$\delta_{l=1}(\infty) - \pi/2 = 0(\pi)$$

is allowed.

At the end we want to emphasize that these different possibilities for the asymptotic D behavior, $D \to$ const, as for a Yukawa family, or $D \to x$ as for the case considered in this section, do not come from the fact that we represent S_l by a quotient. If we require $D(x)$ (in N/D) or $\tilde{F}(x)$ going to constant when $x \to 0$, then these different asymptotic behaviors come entirely from different μ values of the discontinuity $\mu\Delta(x)$ or from different interactions.

Second case[15]: $F'(0)\mathfrak{D}(\mu) = N'(0, \mu) = 0$.

The corresponding solution is a solution of both (1') and (6'). The solution of (1') gives directly the right threshold behavior for the S matrix. From (25) we get

$$\lim_{r\to 0} \tilde{F}(x, r) = (\mathfrak{D}(\mu) + \tilde{\mathcal{N}}(x, \mu))/d(\mu).$$

In Appendix E it is shown that

$$F(0) = \mathfrak{D}(\mu)/d(\mu),$$

such that

$$\lim_{r\to 0} \tilde{F}(x, r) = F(0)\left[1 + \frac{\tilde{\mathcal{N}}(x, \mu)}{\mathfrak{D}(\mu)}\right].$$

We see that $\lim_{r\to 0} \tilde{F}(x, r)$ is a solution of (6') with $\tilde{F}(0) = F(0)$ because $F'(0) = 0$. It follows that $\lim_{r\to 0} \tilde{F}(x, r)$ is also a solution of the unsubtracted equation (1'). In this case $\mathfrak{D}(\mu, r)$ is not singular when $r = 0$ and (17') as well as (18b) are of the Fredholm type for $r > 0$ and $r = 0$.

[15] In fact, we do not consider $\mathfrak{D}(\mu) = 0$. If the Fredholm-type solution of (1) breaks down [$\mathfrak{D}(\mu) = 0$] and also $F'(0)\mathfrak{D}(\mu) = 0$, then a more detailed analysis is necessary.[14]

As in the $l = 0$ case, if $\mathfrak{D}(\mu, 0) = d(\mu) \neq 0$, then the potential is regular at the origin. This case includes the Yukawa family and, as for $l = 0$, not only these potentials. For instance, if $\mathfrak{D}(\mu, r) \xrightarrow[r \to r_0]{} 0$ $(r_0 \geq 0)$ with a root of multiplicity m, then V becomes repulsive as $2m/(r - r_0)^2$. We note also that in this case

$$\int_{m/2}^{\infty} \frac{\Delta(y)\tilde{F}(y)}{y^2}\, dy = 0,$$

such that F or $\tilde{F}$ goes to 1 when $x \to \infty$.

In conclusion to this section, concerning the two different cases considered above, we see that for $\mu\Delta_{l=1}(x)$ satisfying (21) the following are required: (a) right threshold behavior for the unsubtracted equation; (b) Fredholm character of the inversion equation for $r = 0$; (c) asymptotic behavior $D \to 1$ (in N/D); (d) no presence of marginally singular attractive potential like $-2/r^2$ at the origin; and (e) since the solution of the subtracted equation also is a solution of the unsubtracted one, they can be obtained by using the same intrinsic condition on the discontinuity.

V. GHOSTS AND BOUND STATES

For $l = 0$, we recall[2] the results [$\mathfrak{D}_{l=0}(\mu)$ being the Fredholm determinant of Ref.1]:

(i) When μ crosses some roots of $\mathfrak{D}_{l=0}(\mu) = 0$, then in general a real ghost appears at infinity in the physical sheet.

(ii) When μ crosses some roots of $\mathfrak{D}_{l=0}(-\mu) = 0$, then, in general, a bound state appears at the origin.

(iii) For a μ value such that both $\mathfrak{D}_{l=0}(\pm\mu) = 0$, a more detailed analysis is necessary.

We want to show that for the P wave and the subtracted equation (6'), the threshold-behavior conditions $N'(0, \mu) = 0$ and $\mathfrak{D}_{l=1}(\mu) = 0$ play the same role as $\mathfrak{D}_{l=0}(\mu) = 0$ and $\mathfrak{D}_{l=0}(-\mu) = 0$ in the S-wave case.

A. Bound States

When $N'(0, \mu) \neq 0$ from Sec. IV [Eq. (26)], then $\lim_{r \to 0} \tilde{F}(x, r)/r$ is a solution of (6') and

$$\tilde{F}(x) = -[N'(0, \mu)]^{-1}[\mathfrak{D}(\mu) + \tilde{N}(x, \mu)].$$

Because $\tilde{N}(0, \mu) \equiv 0$, we see that if $\mathfrak{D}(\mu) = 0$ then $\tilde{F}(0) = 0$ and a bound state appears at the origin [for instance, see Fig. 1 for the one-pole case $\mathfrak{D}(\mu) = 1 - \mu/2b$ and a bound state appears for $\mu = 2b$].

In conclusion, for the subtracted equation (6') bound states appear at $x = 0$ when $\mathfrak{D}(\mu) = 0$.

B. Ghosts

We consider states appearing at infinity in the physical sheet [roots of $\tilde{F}(x) = 0$ and $x \to \infty$]. We recall that if $N'(0, \mu) \neq 0$ then $\tilde{F}(x) \xrightarrow[x \to \infty]{} x$ such that no real ghost can appear at infinity (this does not mean that no ghost can be present at x finite, as can be seen with the one-pole case and $\mu < 0$; see Fig. 1). This result can be seen also if we subtract in (6') the equation for $x_0 > 0$. We get

$$\tilde{F}(x) = \tilde{F}(x_0) + \int_{m/2}^{\infty} \mu K_{x_0}(x, y)\tilde{F}(y)\, dy,$$

where

$$K_{x_0}(x, y) = \Delta(y)\left[\frac{x^2}{y^2(x + y)} - \frac{x_0^2}{y^2(x_0 + y)}\right].$$

We see that if for particular μ values $\tilde{F}(x_0) = 0$, then μ and $\tilde{F}(x)$ are eigenvalues and eigenfunctions of $K_{x_0}(x, y)$. This means that μ is a root of $\mathfrak{D}_{x_0}(\mu)$, the Fredholm determinant corresponding to $K_{x_0}(x, y)$. In general, the μ roots $\mathfrak{D}_{x_0}(\mu) = 0$ are functions of x_0, so we write them $\mu(x_0)$. In Appendix F we show that $\mathfrak{D}_{x_0}(\mu) \underset{x_0 \to \infty}{\simeq} -x_0[N'(0, \mu) + O(1/x_0)]$ and we see that roots $\tilde{F}(x) = 0$ can appear at infinity only if

$$F'(0)\mathfrak{D}(\mu) = N'(0, \mu) = 0.$$

For instance, in the one-pole case we see that for $\mu = 0$, which is the only value $N'(0, \mu) = 0$, a real ghost appears at infinity.

This proof works only if $\mu(x_0)$ is really a function of x_0. Assume now that there exist special μ values, roots of $\mathfrak{D}_{x_0}(\mu) = 0$, but independent of x_0. Then these μ values are also roots of $\mathfrak{D}_{x_0=0}(\mu) = 0$ corresponding to the kernel $K_{x_0=0}(x, y) = (x^2/y^2)(x + y)$. Therefore these μ values are roots of $\mathfrak{D}(\mu) = 0$ and for these values a more detailed analysis is necessary. For instance, there exist such values in the two-poles case.[14]

VI. CONCLUSION

In the first part of this paper, with the help of Blažek's extension of Marchenko's inversion formalism, we have established a dispersive integral equation for the Jost solutions for $l \neq 0$, where the kernel is proportional to the left-hand-cut discontinuity of the S_l matrix (the discontinuity being such that the N/D equations are of the Fredholm type). We have also obtained the integral equation giving the possibility of reconstructing the potential from this left-hand-cut discontinuity. From this last equation we get, for instance, that the potentials are short range (decrease exponentially), so we know that if we reconstruct the $S_l(k)$ matrix from the first equation (Jost-solutions integral equation), then the correct threshold behavior for $S_l(k)$ will be satisfied. So these two equations are sufficient tools in order to solve the problem of the threshold behavior for nonrelativistic N/D equations. Indeed, when we know

the two Jost solutions (asymptotically ingoing or outgoing wave), we can reconstruct the $S_l(k)$ matrix by taking the ratio of these two functions and investigating the limit of the radial coordinate going to zero. Further, knowing the behavior of the potential near the origin, we get the corresponding singularity of the Jost solution near the origin. Consequently, from the Jost solution, we can obtain the Jost function (or D in N/D) and see if the integral equation for the Jost function exists. The great difference with the $l = 0$ case is that, in general, the potentials (although having a Yukawa-type asymptotic behavior) are marginally singular attractive at the origin, so the behavior of the solutions is not given by the centrifugal potential alone (as for $l = 0$, we find also the possibility of second-order singularity leading to repulsive potentials). So if integral equations exist for the Jost functions, then the kernels in general will be modified at least in order to get the right threshold behavior for the S_l matrix. This is why we have *a priori* (without any justification and independently of inversion formalism) considered an ansatz subtracted equation by modifying the kernel of the unsubtracted one in such a way that the new Jost functions (or D in N/D) lead directly to a correct behavior for the S_l matrix.

In the second part of this paper we have restricted our study to the P-wave case and investigated whether or not such an ansatz subtracted equation has a meaning: more explicitly, if the ansatz can be deduced from integral equation for the Jost solution and what the features of the corresponding interactions are (we note that the same study can be made for $l > 1$ in order to see if our ansatz or another one has a meaning at all). For $l = 1$ we have found two cases following different values of the discontinuity. In the first case the unsubtracted equation satisfies the right threshold behavior. We find that the corresponding potentials include the Yukawa family and are in general not singular at the origin. The asymptotic value of D is 1.

In the second case the unsubtracted equation does not satisfy the correct threshold behavior: the ansatz subtracted equation is the equation having a meaning, the potentials $\simeq -2/r^2$ near the origin, D does not go to a constant at infinity, and the Levinson theorem must be modified. This case never corresponds to the Yukawa-type family which has been considered in order to prove the Mandelstam representation in potential scattering. Since these two cases correspond to different discontinuities of the S matrix, the correspondingly different asymptotic behavior of D cannot be attributed to any ambiguity coming from the use of the ratio N/D. (We fix the behavior of D as constant at the origin.) In this paper these two cases are distinguished by different behavior of the potentials near the origin. For $l = 0$ we recall that we obtained not only the Yukawa family but also marginally singular repulsive potentials. But for these badly behaved potentials we recall that there appear ghosts; so, requiring only available physical states in the physical sheet, we can reject these badly behaved potentials. For $l = 1$ and the above second case, the situation is different. For instance in this paper, for the discontinuity replaced by one pole, we have seen that always the potential is marginally singular attractive at the origin (except the trivial case where the residue of the pole vanishes); nevertheless values of this residue exist such that the present state on the physical sheet is a true bound state and consequently from spectrum requirements we cannot reject the corresponding interaction. In fact, in a further paper[14] we shall seek the possibility of a self-damping connected domain of physically available interactions for the subtracted equation considered here, where the discontinuity is replaced by a finite number of poles.

Finally, we recall that in potential scattering for the N/D equations the discontinuity is taken as input and the only thing we can do is to try to analyze the consequences corresponding to different input. It is why we think it can be interesting to investigate in the relativistic case, by taking into account crossing, whether or not such different features ($D \to$ const or not) connected with different discontinuities can be explained from the consideration of the other channels.

ACKNOWLEDGMENT

We thank Dr. A. Martin for his interest in this work.

APPENDIX A

We want to show that

$$K_l(r, r) = \frac{1}{2}\int_r^\infty V(r')\,dr'. \tag{A1}$$

We consider $\operatorname{Im} k < 0$. Substitution of (10) in (7) yields

$$\int_r^\infty K_l(r, t)w_l(kt)\,dt = I_1 + I_2,$$

$$I_1(k, r) = -\int_r^\infty dr' g_l(k, r, r')V(r')w_l(kr')$$

$$= \int_r^\infty K_l^{(1)}(r, t)w_l(kt)\,dt, \tag{A2}$$

$$I_2(k, r) = -\int_r^\infty dr' g_l(k, r, r')V(r') \times \int_{r'}^\infty dr'' K_l(r', r'')w_l(k, r'')$$

$$= \int_r^\infty K_l^{(2)}(r, t)w_l(kt)\,dt, \tag{A3}$$

where $K_l = K_l^{(1)} + K_l^{(2)}$. We want to find $K_l^{(i)}$ from (A2) and (A3). For $l = 0$, $w_0 = ie^{-ikt}$ and we have to consider[1] a Fourier or Laplace transform. But for $l \neq 0$ we recall that if we define the operator D_l such that

$$D_l(f(k)) = \left(\frac{1}{k}\frac{\partial}{\partial k}\right)^l (k^l f(k)),$$

then

$$D_l(w_l(kr)) = r^l e^{-ikr}$$

and we get

$$D_l(I^{(i)}(k, r)) = \int_r^\infty t^l K_l^{(i)}(r, t) e^{-ikt}\, dt. \qquad \text{(A4)}$$

Using D_l the problem is thus reduced to find a Fourier or Laplace transform. Using (8) for w_l we get

$$-g_l(k, r, r')w_l(k, r'') = e^{-ik(r-r'+r'')} \sum_0^{3l} \frac{\alpha_p(r, r', r'')}{k^{p+1}} - e^{-ik(r'-r+r'')} \sum_0^{3l} \frac{\alpha_p(-r, -r', r'')}{k^{p+1}},$$

where the α_p can be determined from (8) and are not singular for $r, r', r'' \neq 0$,

$$D_l(-g_l w_l) = e^{-ik(r-r'+r'')} \sum_0^{4l} \frac{\gamma_p(r, r', r'')}{k^{p+1}} - e^{-ik(r'-r+r'')} \sum_0^{4l} \frac{\gamma_p(-r, -r', r'')}{k^{p+1}},$$

where γ_p are still not singular for $r, r', r'' \neq 0$, $\gamma_0(r, r', r'') = -ir^l/2$. Now we use

$$k^{-p}e^{-ikr} = i^p \int_r^\infty e^{-ikt}(t - r)^{p-1}((p-1)!)^{-1}$$

and we get

$$D_l I_1 = \int_r^\infty V(r') \sum_0^{4l} \left\{ \int_r^\infty dt \bar{\gamma}_p(r, r', r'' = r')(t - r)^p - \int_r^\infty dt \bar{\gamma}_p(-r, -r', r'' = r')(t - 2r' + r)^p \right\} dt', \qquad \text{(A5)}$$

where $\bar{\gamma}_p = \gamma_p i^{p+1}/p!$

We change the order of integration in (A5) and comparing with (A4) we get

$$t^l K_l^{(1)}(r, t) = \sum_0^{4l} \left\{ (t - r)^p \int_r^\infty dr' \bar{\gamma}_p(r, r', r'' = r')V(r') + \int_r^{(t+r/2)} dr'(t - 2r' + r)^p \times \bar{\gamma}_p(-r, -r', r'' = r')V(r') \right\}. \qquad \text{(A6)}$$

With similar algebra from $D_l I_2$ we get

$$t^l K_l^{(2)}(r, t) = \sum_0^{4l} \left\{ \int_r^\infty dr' V(r') \int_{r'}^{t-r+r'} (t - r + r' - r'')^p \times \bar{\gamma}_p(r, r', r'')K_l(r', r'')\, dr'' - \int_r^{(t+r/2)} dr' V(r') \int_{r'}^{t+r-r'} (t - r' + r - r'')^p \times \bar{\gamma}_p(-r, -r', r'')K_l(r', r'')\, dr'' \right\}. \qquad \text{(A7)}$$

From (A6) and (A7) and $\bar{\gamma}_0(r, r', r'' = r') = \frac{1}{2}r'$ we see that (A1) follows.

APPENDIX B

We want to show that

$$K_l(r, r) = \frac{d}{dr}(\mathfrak{D}_l(\mu, r))/\mathfrak{D}_l(\mu, r), \qquad \text{(B1)}$$

where $\mathfrak{D}_l(\mu, r)$ is both the Fredholm denominator of the integral equations (15), (17), and (18). The Fredholm type of solution of $K_l(r, y)$, which is a solution of (18b), can be written, for $y = r$, as

$$\mathfrak{D}_l(\mu, r)K_l(r, r) = \mu\mathfrak{C}_l(r, r) + \mu^2 \int_r^\infty \mathcal{N}_l(\mu, r, t_1)\mathfrak{C}_l(t_1, r)\, dt_1, \qquad \text{(B2)}$$

where[16]

$$\mathfrak{D}(\mu, r) = 1 + \sum_1^\infty \frac{(-\mu)^n}{n!} \mathfrak{D}^{(n)}(r),$$

$$\mathcal{N}(\mu, r, t_1) = \sum_0^\infty \frac{(-\mu)^n}{n!} \mathcal{N}^{(n)}(r, t_1). \qquad \text{(B3)}$$

In order to prove (B1), due to (B2) and (B3), we have only to show that

$$\frac{1}{n+2}\frac{d}{dr}\mathfrak{D}^{(n+2)}(r) = -\mathfrak{C}(r, r)\mathfrak{D}^{(n+1)}(r) + \int_r^\infty (n+1)\mathcal{N}^{(n)}(r, t_1)\mathfrak{C}(t_1, r)\, dt_1. \qquad \text{(B4)}$$

From Fredholm's theory we get

$$\mathfrak{D}^{(n)}(r) = \int_r^\infty dt_1 \cdots \int_r^\infty dt_n E_n(t_1, t_2, \cdots, t_n),$$

$$\mathcal{N}^{(n)}(r, t_1) = \int_r^\infty dt_2 \cdots \int_r^\infty dt_{n+1} H_n(r, t_1, t_2, \cdots, t_{n+1}),$$

where

$$E_n(t_1, t_2, \cdots, t_n) = \begin{vmatrix} \mathfrak{C}(t_1, t_1) & \mathfrak{C}(t_1, t_2) & \cdots & \mathfrak{C}(t_1, t_n) \\ \mathfrak{C}(t_2, t_1) & \mathfrak{C}(t_2, t_2) & \cdots & \mathfrak{C}(t_2, t_n) \\ \cdots & \cdots & \cdots & \cdots \\ \mathfrak{C}(t_n, t_1) & \mathfrak{C}(t_n, t_2) & \cdots & \mathfrak{C}(t_n, t_n) \end{vmatrix},$$

[16] Because our proof is l-independent, we omit the index l in the following.

$$H_n(r_1, t_1, t_2, \cdots, t_{n+1}) = \begin{vmatrix} \mathcal{C}(r_1 t_1) & \mathcal{C}(r_1 r_2) & \cdots & \mathcal{C}(r_1 t_{n+1}) \\ \mathcal{C}(t_2, t_1) & \mathcal{C}(t_2, t_2) & \cdots & \mathcal{C}(t_2, t_{n+1}) \\ \cdots & \cdots & \cdots & \cdots \\ \mathcal{C}(t_{n+1}, t_1) & \mathcal{C}(t_{n+1}, t_2) & \cdots & \mathcal{C}(t_{n+1}, t_{n+1}) \end{vmatrix}.$$

We obtain

$$\begin{aligned} &\frac{1}{n+2}\frac{d}{dr}\mathfrak{D}^{(n+2)}(r) \\ &= \frac{-1}{n+2}\sum_{j=1}^{n+2}\int_r^\infty dt_1 \cdots \int_r^\infty dt_{j-1}\int_r^\infty dt_{j+1}\cdots\int_r^\infty dt_{n+2} \\ &\quad \times E_{n+2}(t_1, \cdots, t_{j-1}, r, t_{j+1}, \cdots, t_{n+2}) \\ &= -\int_r^\infty dt_1 \cdots \int_r^\infty dt_{n+1} E_{n+2}(r_1, t_1, \cdots, t_{n+1}), \end{aligned}$$

where we have used the symmetry

$$E_n(\cdots, t_i, \cdots, t_j, \cdots) = E_n(\cdots, t_j, \cdots, t_i, \cdots).$$

Now we develop E_{n+2} following the elements of the first column:

$$\begin{aligned} E_{n+2}(r_1, t_1, \cdots, t_{n+1}) &= \mathcal{C}(r, r)E_{n+1}(t_1, \cdots, t_{n+1}) \\ &\quad + \sum_{j=1}^{n+1}\mathcal{C}(t_j, r)(-)^j M_{n+1}^j(r_1, t_1, \cdots, t_j, \cdots, t_{n+1}), \end{aligned}$$

where

$$M_{n+1}^{j=1}(r_1, t_1, \cdots, t_{n+1}) = H_n(r_1, t_1, \cdots, t_{n+1})$$

and

$$M_{n+1}^j(r_1, t_1, \cdots, t_{n+1}) = \begin{vmatrix} \mathcal{C}(r_1, t_1) & \cdots & \mathcal{C}(r_1, t_{n+1}) \\ \mathcal{C}(t_1, t_1) & \cdots & \mathcal{C}(t_1, t_{n+1}) \\ \cdots & \cdots & \cdots \\ \mathcal{C}(t_{j-1}, t_1) & \cdots & \mathcal{C}(t_{j-1}, t_{n+1}) \\ \mathcal{C}(t_{j+1}, t_1) & \cdots & \mathcal{C}(t_{j+1}, t_{n+1}) \\ \cdots & \cdots & \cdots \\ \mathcal{C}(t_{n+1}, t_1) & \cdots & \mathcal{C}(t_{n+1}, t_{n+1}) \end{vmatrix}.$$

Similarly, as was done for the $l = 0$ case,[2] it is easy to get the identity

$$\begin{aligned} &\int_r^\infty dt_1 \cdots \int_r^\infty dt_{n+1}\mathcal{C}(t_j, r)M_{n+1}^j(r_1, t_1, \cdots, t_{n+1}) \\ &= \int_r^\infty dt_1 \cdots \int_r^\infty dt_{n+1}(-)^{j-1}\mathcal{C}(t_1, r)M_{n+1}^1(r_1, t_1, \cdots, t_{n+1}). \end{aligned}$$

Finally we get

$$\begin{aligned} \frac{1}{n+2}\frac{d}{dr}\mathfrak{D}^{(n+2)}(r) &= -\mathcal{C}(r, r)\int_r^\infty dt_1 \cdots \int_r^\infty dt_{n+1} \\ &\quad \times E_{n+1}(t_1, \cdots, t_{n+1}) + (n+1) \\ &\quad \times \int_r^\infty dt_1 \cdots \int_r^\infty dt_{n+1}\mathcal{C}(t_1, r)H_n(r_1, t_1, \cdots t_{n+1}), \end{aligned}$$

which is the relation (B4).

APPENDIX C

We want to show that the singular part of the numerator and the denominator in the Fredholm type of solution of (17) for $l = 1$ are linked to the threshold condition (4).

1. From the Fredholm type of solution of (1), the condition (4) can be written[17]

$$\begin{aligned} \mathfrak{D}(\mu)F'(0) &= \sum_1^\infty \frac{(-\mu)^n}{(n-1)!}\int_{m/2}^\infty du_1 \cdots \int_{m/2}^\infty du_n \\ &\quad \times L_n(u_1, u_2, \cdots, u_n)\prod_1^n \Delta(u_i), \quad \text{(C1)} \end{aligned}$$

where $\mathfrak{D}(\mu)$ is the Fredholm determinant of (1):

$$\begin{aligned} \mathfrak{D}(\mu) &= 1 + \sum(-\mu)^n/n!\int_{m/2}^\infty du_1 \cdots \int_{m/2}^\infty du_n \\ &\quad \times \prod_i \Delta(u_i)P_n(u_1, \ldots, u_n), \end{aligned}$$

$$P_n = \begin{vmatrix} \frac{1}{2u_1} & \frac{1}{u_1+u_2} & \cdots & \frac{1}{u_1+u_n} \\ \frac{1}{u_2+u_1} & \frac{1}{2u_2} & \cdots & \frac{1}{u_2+u_n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \frac{1}{u_n+u_1} & \frac{1}{u_n+u_2} & \cdots & \frac{1}{2u_n} \end{vmatrix},$$

$$L_n = \begin{vmatrix} \frac{1}{u_1^2} & \frac{1}{u_2^2} & \cdots & \frac{1}{u_n^2} \\ \frac{1}{u_1+u_2} & \frac{1}{2u_2} & \cdots & \frac{1}{u_2+u_n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \frac{1}{u_n+u_1} & \frac{1}{u_n+u_2} & \cdots & \frac{1}{2u_n} \end{vmatrix}. \quad \text{(C2)}$$

P_n and L_n have the same elements except for the first row.

2. The Fredholm determinant of (15), (17), and (18b) is

$$\begin{aligned} \mathfrak{D}(\mu, r) &= 1 + \sum_{n=1}^\infty \frac{\mu^n}{n!}\int_{m/2}^\infty du_1 \cdots \int_{m/2}^\infty du_n \\ &\quad \times e^{-2\Sigma u_i r}\prod_i \Delta(u_i)M_n(r, u_1, \cdots, u_n) \quad \text{(C3)} \end{aligned}$$

$$M_n(r, u_1, \cdots, u_n) = \begin{vmatrix} Q_r(u_1, u_1) \cdots Q_r(u_1, u_n) \\ \cdots \\ Q_r(u_n, u_1) \cdots Q_r(u_n, u_n) \end{vmatrix},$$

[17] We investigate only $l = 1$ so that we drop the index l.

with $Q_r(u_i, u_j) = (u_i u_j r)^{-1} + (u_i + u_j)^{-1}$. M_n is singular when $r \to 0$ and it is easy to see that we get

$$M_n(r, \cdots, u_i, \cdots) = r^{-1}\bar{M}_n(u_1, \cdots, u_n) + \bar{\bar{M}}_n(u_1, \cdots, u_n).$$

When we substitute this expression of M_n in (C3) we define correspondently $\mathfrak{D}(\mu, r) = r^{-1}\bar{\mathfrak{D}}(\mu, r) + \bar{\bar{\mathfrak{D}}}(\mu, r)$, where $\bar{\mathfrak{D}}(\mu, r)$ and $\bar{\bar{\mathfrak{D}}}(\mu, r)$ go to a constant when $r \to 0$ because we recall that we have always assumed that $\Delta(y)$ is such that the Fredholm type of solution of (1) exists. It is easy to see that $\bar{\bar{M}}_n = P_n$ and

$$\bar{M}_n = \sum_{j=1} \frac{1}{u_j} \bar{M}_n^j(u_1, \cdots, u_n),$$

where $\bar{M}_n^j$ is the same determinant as P_n except that the elements of the jth column are $u_1^{-1}, u_2^{-1}, \cdots, u_n^{-1}$. We get

$$\int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n e^{-2\Sigma u_i r} \prod_i \Delta(u_i) \times \left(\frac{1}{u_j}\bar{M}_n^j - \frac{1}{u_1}\bar{M}_n^{j=1}\right) = 0$$

and finally

$$\bar{\mathfrak{D}}(\mu, r) = -\sum_1^{\infty} \frac{\mu^n}{(n-1)!} \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i) \times e^{-2\Sigma u_i r} Y_n(u_1, \cdots, u_n), \quad \text{(C4)}$$

where

$$Y_n(u_1, \cdots, u_n) = \begin{vmatrix} 0 & \frac{1}{u_1} & \frac{1}{u_2} & \cdots & \frac{1}{u_n} \\ \frac{1}{u_1} & 0 & 0 & \cdots & 0 \\ \frac{1}{u_2} & \frac{1}{u_2+u_1} & \frac{1}{2u_2} & & \frac{1}{u_2+u_n} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \frac{1}{u_n} & \frac{1}{u_n+u_1} & \frac{1}{u_n+u_2} & \cdots & \frac{1}{2u_n} \end{vmatrix}.$$

3. The Fredholm-type solution of (17) corresponding to the kernel $-\Delta(y)e^{-(x+y)r}H_r(x, y)$ with

$$H_r(x, y) = \frac{1}{y^2 r} + \frac{x}{y}\frac{1}{x+y}$$

is

$$\bar{F}(x, r) = e^{-xr}(1 + xr) + \frac{\mathcal{N}(x, \mu, r)}{\mathfrak{D}(\mu, r)},$$

where

$$\mathcal{N}(x, \mu, r) = -\sum_1^{\infty} \frac{\mu^n}{(n-1)!} \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n e^{-(x+2\Sigma u_i)r}(1 + u_1 r) \prod_i \Delta(u_i) Z_n(r, x, u_1, \cdots, u_n).$$

$$Z_n(r, x, u_1, \cdots, u_n) = \begin{vmatrix} H_r(x, u_1) & H_r(x, u_2) & \cdots & H_r(x, u_n) \\ H_r(u_2, u_1) & H_r(u_2, u_2) & \cdots & H_r(u_2, u_n) \\ H_r(u_n, u_1) & H_r(u_n, u_2) & \cdots & H_r(u_n, u_n) \end{vmatrix}. \quad \text{(C5)}$$

It is easy to see that $Z_n = (1/r)\bar{Z}_n(x, u_1, \cdots, u_n) + \bar{\bar{Z}}_n(x, u_1, u_2, \cdots, u_n)$, where

$$\bar{\bar{Z}}_n(x, u_1, \cdots, u_n) = \begin{vmatrix} \frac{x}{u_i(x+u_1)} & \frac{x}{u_2(x+u_2)} & \cdots & \frac{x}{u_n(x+u_n)} \\ \frac{u_2}{u_1(u_2+u_1)} & \frac{1}{2u_2} & \cdots & \frac{u_2}{u_n(u_2+u_n)} \\ \frac{u_n}{u_1(u_n+u_1)} & \frac{u_n}{u_2(u_n+u_2)} & \cdots & \frac{1}{2u_n} \end{vmatrix},$$

$$\bar{Z}_n(x, u_1, \cdots, u_n) = \begin{vmatrix} 0 & \frac{1}{u_1^2} & \cdots & \frac{1}{u_n^2} \\ 1 & & & \\ 1 & & \bar{\bar{Z}}_n & \\ 1 & & & \end{vmatrix}. \quad \text{(C6)}$$

We put $z_n = (1 + u_1 r)Z_n$. We define $\mathcal{N}(x, \mu, r) = r^{-1}\bar{\mathcal{N}}(x, \mu, r) + \bar{\bar{\mathcal{N}}}(x, \mu, r)$ such that $\bar{\mathcal{N}}$ and $\bar{\bar{\mathcal{N}}}$ are obtained from (C5) when z_n is replaced respectively by $\bar{Z}_n$ and $u_1\bar{Z}_n + \bar{\bar{Z}}_n + u_1 r\bar{\bar{Z}}_n$. From the assumptions made about $\Delta(y)$ we know that $\bar{\mathcal{N}}$ and $\bar{\bar{\mathcal{N}}}$ go to a constant when $r \to 0$.

Using

$$\frac{u_i}{u_j(u_i+u_j)} = \frac{1}{u_j} - \frac{1}{u_i+u_j},$$

we get with some algebra

$$(-1)^n Z_n(x, u_1, \cdots, u_n) = \begin{vmatrix} 0 & \frac{1}{u_1^2} & \frac{1}{u_2^2} & \cdots & \frac{1}{u_n^2} \\ 1 & \frac{1}{x+u_1} & \frac{1}{x+u_2} & \cdots & \frac{1}{x+u_n} \\ 1 & \frac{1}{u_2+u_1} & \frac{1}{2u_2} & & \frac{1}{u_2+u_n} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & \frac{1}{u_n+u_1} & \frac{1}{u_n+u_2} & \cdots & \frac{1}{2u_n} \end{vmatrix}. \quad \text{(C7)}$$

In the following we shall show that $e^{xr}\bar{N}(x, \mu, r)$ is x-independent. Taking into account this result in (C7) for $x = \infty$, we get $\bar{Z}_n = (-)^{n+1}L_n$. Similarly, from this result we can put $x = 0$ in (C6) and perform the following algebra: we factorize u_i^{-1} from the $(i+1)$th column $(i = 1, 2, \cdots, n)$ and u_k from the $(k+1)$th row $(k = 1, 2, 3, \cdots, n)$. We get $\bar{Z}_n = -Y_n$. Finally we have the relations

$$\bar{N}(x, \mu, 0) = \mathfrak{D}(\mu)F'(0) = -\bar{\mathfrak{D}}(\mu, 0),$$
$$e^{xr}\bar{N}(x, \mu, r) = -\bar{\mathfrak{D}}(\mu, r). \quad \text{(C8)}$$

4. We want to show that $e^{xr}\bar{N}(x, \mu, r)$ is x-independent. The x dependent part of $e^{xr}\bar{N}(x, \mu, r)$ is

$$W(x, \mu, r) = -\sum_{n=1}^{\infty} \frac{\mu^n}{(n-1)!} \times \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n e^{-2\Sigma u_i r} \prod_i \Delta(u_i)\zeta_n(x, u_1, \cdots, u_n),$$

where ζ_n is the same determinant as $\bar{Z}_n$ (C7) except that the first element of the second row is 0.

We develop ζ_n, following the elements of the first column,

$$\zeta_n = \sum_{j=2}^{n}(-1)^{j+n}\zeta_n^j(x, u_1, \cdots, u_n),$$

where

$$\zeta_n^j(x, u_1, \cdots, u_n) = \begin{vmatrix} \frac{1}{u_1^2} & \frac{1}{u_2^2} & \cdots & \frac{1}{u_j^2} & \cdots & \frac{1}{u_n^2} \\ \frac{1}{x+u_1} & \frac{1}{x+u_2} & \cdots & \frac{1}{x+u_j} & \cdots & \frac{1}{x+u_n} \\ \frac{1}{u_{j-1}+u_1} & \frac{1}{u_{j-1}+u_2} & \cdots & \frac{1}{u_{j-1}+u_j} & \cdots & \frac{1}{u_{j-1}+u_n} \\ \frac{1}{u_{j+1}+u_1} & \frac{1}{u_{j+1}+u_2} & \cdots & \frac{1}{u_{j+1}+u_j} & \cdots & \frac{1}{u_{j+1}+u_n} \\ \frac{1}{u_n+u_1} & \frac{1}{u_n+u_2} & \cdots & \frac{1}{u_n+u_j} & \cdots & \frac{1}{2u_n} \end{vmatrix}.$$

u_1 is only in the first column, u_j only in the jth, so ζ_n is antisymmetric by $1 \rightleftarrows j$ and $W(x, \mu, r) \equiv 0$.

5. From (C8) it is easy to see that

$$\lim_{r\to 0}(\bar{N}(x, \mu, r) + \mathfrak{D}(\mu, r)) = x\bar{\mathfrak{D}}(\mu, 0) + \bar{\bar{\mathfrak{D}}}(\mu, 0) + \bar{\bar{N}}(x, \mu, 0),$$

with

$$\bar{\bar{\mathfrak{D}}}(\mu, 0) = \sum_n \frac{\mu^n}{n!}\int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i) \times \bar{\bar{M}}_n(u_1, u_2, \cdots, u_n),$$

$$\bar{\bar{N}}(x, \mu, 0) = -\sum_{n=1}^{\infty} \frac{\mu^n}{(n-1)!}\int_{m/2}^{\infty} du_1 \cdots \times \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i) \times (u_i\bar{Z}_n(x, u_1, \cdots, u_n) + \bar{\bar{Z}}_n(x_1 \cdots, u_n)). \quad \text{(C9)}$$

Taking into account $H_r(u_i, u_j) = (u_i/u_j)Q_r(u_i, u_j)$, we have also

$$\bar{\bar{\mathfrak{D}}}(\mu, 0) = \sum_n \frac{\mu^n}{n!}\int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i) \times \bar{\bar{Z}}_n(u_1, u_1, \cdots, u_n) \quad \text{(C10)}$$

and

$$\bar{\mathfrak{D}}(\mu, r) = \sum_n \frac{\mu^n}{n!}\int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i) \times e^{-2\Sigma u_i r}\bar{Z}_n(u_1, u_1, \cdots, u_n), \quad \text{(C11)}$$

$\bar{Z}_n$ and $\bar{\bar{Z}}_n$ being defined in (C6).

6. We called $d(\mu)$ the nonsingular part of $\mathfrak{D}(\mu, r)$. When $r \to 0$

$$d(\mu) = \left(\frac{d}{dr}\bar{\mathfrak{D}}(\mu, r)\right)_{r=0} + \bar{\bar{\mathfrak{D}}}(\mu, 0). \quad \text{(C12)}$$

We want to show that $d(\mu)$ is also the Fredholm-denominator determinant of

$$R(x, y) = (x+y)^{-1} + x/y^2.$$

With some algebra, (C11) can be written as

$$\bar{\mathfrak{D}}(\mu, r) = -\sum_1^{\infty} \frac{\mu^n}{n!}\int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i) \times R_n(r, u_1, \cdots, u_n),$$

with

$$R_n(r, u_1, \cdots, u_n) = \begin{vmatrix} 0 & \frac{e^{-u_1 r}}{u_1} & \cdots & \frac{e^{-u_n r}}{u_n} \\ \frac{e^{-u_1 r}}{u_1} & & & \\ \vdots & & P_n^r(u_1, u_2, \cdots, u_n) & \\ \frac{e^{-u_n r}}{u_n} & & & \end{vmatrix},$$

$$(P_n^r(u_1, \cdots, u_n))_{i,j} = \frac{e^{-(u_i+u_j)r}}{u_i + u_j},$$

$$P_n^0(u_1, \cdots, u_n) = P_n(u_1, \cdots, u_n),$$

P_n being defined in (C2).

To get $(d/dr)\bar{\mathfrak{D}}(\mu, r)$ we consider $(d/dr)R_n(r_1, u_1, \cdots, u_n)$:

$$\left(\frac{d}{dr}R_n(r, u_1, \cdots, u_n)\right)_{r=0} = -S_n(u_1, \cdots, u_n) - \sum_{j=2}^{n+1} R_n^j(u_1, \cdots, u_n),$$

where $S_n(u_1, \cdots, u_n)$ has the same elements as $R_n(0, u_1, \cdots, u_n)$ except the first row which is replaced by $(0, 1, 1, \cdots, 1)$, and $R_n^j(u_1, \cdots, u_n)$ has the same elements as $R_n(0, u_1, \cdots, u_n)$ except for the jth row which is replaced by $(1, 1, \cdots, 1)$. Then

(C12) is equivalent to

$$d(\mu) = \sum_n \frac{\mu^n}{n!} \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i) \times \left(S_n(u_1, \cdots, u_n) + P_n(u_1, \cdots, u_n) + \sum_{j=2}^{n+1} R_n^j(u_1, \cdots, u_n) \right). \quad \text{(C13)}$$

We remark that the sum

$$S_n(u_1, \cdots, u_n) + \begin{vmatrix} 1 & 0 & \cdots & 0 \\ \frac{1}{u_1} & & & \\ \cdot & & P_n & \\ \cdot & & & \\ \cdot & & & \\ \frac{1}{u_n} & & & \end{vmatrix}$$

can be written $R_n^{j=1}(u_1, \cdots, u_n)$ because it has the same elements as $R_n(u_1, \cdots, u_n)$ except those of the first row which are $(1, 1, \cdots, 1)$. Then,

$$d(\mu) = \sum_n \frac{\mu^n}{n!} \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i) \times \sum_{j=1}^{n} R_n^j(u_1, u_2, \cdots, u_n). \quad \text{(C14)}$$

With some algebra one can show that

$$\sum_{j=1}^{n} R_n^j(u_1, \cdots, u_n) = (-)^n \Lambda^{(n+1)}(u_1, \cdots, u_n),$$

$$-\Lambda^{(n+1)} = \begin{vmatrix} -1 & \frac{1}{u_1^2} & \cdots & \frac{1}{u_n^2} \\ u_1 & & & \\ \cdot & & P_n(u_1, \cdots, u_n) & \\ \cdot & & & \\ \cdot & & & \\ u_n & & & \end{vmatrix}.$$

We develop $\Lambda^{n+1}(u_1, \cdots, u_n)$ following the elements of the first row; then

$$\Lambda^{(n+1)}(u_1, \cdots, u_n) = P_n(u_1, \cdots, u_n) + \sum_j P_n^j(u_1, \cdots, u_n), \quad \text{(C15)}$$

where $P_n^j(u_1, \cdots, u_n)$ has the same elements as P_n except the jth column which is $(u_1/u_j^2, u_2/u_j^2, \cdots, u_n/u_j^2)$. Now if we consider the Fredholm determinant of $R(x, y)$, we find also a term P_n coming from $(x + y)^{-1}$ alone and n other terms P_n^a.

APPENDIX D

We want to show that

$$\lim_{r \to 0} [\mathcal{N}(x, \mu, r) + \mathfrak{D}(\mu, r)] \equiv x\bar{\mathfrak{D}}(\mu, 0) + \bar{\bar{\mathfrak{D}}}(\mu, 0) + \bar{\bar{\mathcal{N}}}(x, \mu, 0) = \tilde{\mathcal{N}}(x, \mu) + \tilde{\mathfrak{D}}(\mu), \quad \text{(D1)}$$

where $\mathfrak{D}(\mu, r)$, $\mathcal{N}(x, \mu, r)$, $\tilde{\mathcal{N}}(x, \mu)$, and $\tilde{\mathfrak{D}}(\mu)$ corresponding to the Eqs. (17′) and (5′) are defined in (22) and (23); and $\bar{\mathfrak{D}}(\mu, 0)$, $\bar{\bar{\mathfrak{D}}}(\mu, 0)$, and $\bar{\bar{\mathcal{N}}}(\mu, 0)$ are defined in (C4), (C9), and (C10).

1. Using $x^2/[y^2(x + y)] = (x/y^2) - x/[y(x + y)]$, we can write

$$\tilde{\mathcal{N}}(x, \mu) = -\sum_n \frac{(-\mu)^n}{(n-1)!} \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i)(-)^n \times \left(\bar{\bar{Z}}_n(x, u_1, \cdots, u_n) - \sum_j \bar{\bar{Z}}_n^j(x, u_1, \cdots, u_n) \right),$$

where $\bar{\bar{Z}}_n(x, u_1, \cdots, u_n)$ is defined in (C6), and $\bar{\bar{Z}}_n^j(x, u_1, \cdots, u_n)$ has the same elements as

$$\bar{\bar{Z}}_n(x, u_1, \cdots, u_n),$$

except those of the jth column which are $(x/y^2, u_2/u_j^2, \cdots, u_n/u_j^2)$.

We remark that $\bar{\bar{Z}}_n(u_1, u_1, \cdots, u_n)$ appear both in $\tilde{\mathfrak{D}}(\mu)$ and $\bar{\bar{\mathfrak{D}}}(\mu, 0)$ giving the same contribution in the rhs and the lhs of (D1). Similarly, $\bar{\bar{Z}}_n(x, u_1, \cdots, u_n)$ appears both in $\tilde{\mathcal{N}}(x, \mu)$ and $\bar{\bar{\mathcal{N}}}(x, \mu, 0)$ giving also the same contribution in the rhs and in the lhs of (D1). To prove (D1) it is now sufficient to show that

$$\frac{\mu^n}{n!} \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i) \times \left(n \sum_j Z_n^j(x, u_1, \cdots, u_n) - \sum_j \bar{\bar{Z}}_n^j(u_1, u_1, \cdots, u_n) \right) = \frac{-\mu^n}{(n-1)!} \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n u_1 Z_n(x, u_1, \cdots, u_n) \times \prod_i \Delta(u_i) + x\bar{\mathfrak{D}}(\mu, 0). \quad \text{(D1′)}$$

2. We get with some algebra

$$-\sum_j \bar{\bar{Z}}_n^j(x, u_1, \cdots, u_n) = \begin{vmatrix} 0 & \frac{1}{u_1^2} & \cdots & \frac{1}{u_n^2} \\ x & & & \\ u_2 & & \bar{\bar{Z}}_n(x, u_1, \cdots, u_n) & \\ \cdot & & & \\ \cdot & & & \\ u_n & & & \end{vmatrix} = \alpha^{(n+1)}(x, u_1, \cdots, u_n).$$

This identity is also valid for $x = u_1$. We develop $\alpha^{(n+1)}(x, u_1, \cdots, u_n)$ following the elements of the first column, and remark that the minor corresponding to $(\alpha^{(n+1)})_{2,1}$ is $\bar{Z}_n(0, u_1, \cdots, u_n)$ (C6); we call $\beta^{(n+1)}(x, u_1, \cdots, u_n)$ the remaining term. $\alpha^{(n+1)}(x, u_1, \cdots, u_n)$ and $\beta^{(n+1)}(x, u_1, \cdots, u_n)$ have the same elements except that

$$(\beta^{(n+1)})_{2,1} = 0,$$

so that

$$\alpha^{(n+1)}(x, u_1, \cdots, u_n) = -x\bar{Z}_n(0, u_1, \cdots, u_n) + \beta^{(n+1)}(x, u_1, \cdots, u_n).$$

On the left-hand side of (D1'), if we consider the contribution of the first term coming from

$$\bar{Z}_n(0, u_1, \cdots, u_n) = -Y_n(u_1, \cdots, u_n),$$

we get $x\bar{\mathfrak{D}}(\mu, 0)$. Then the equality (D1') is now reduced to

$$\frac{\mu^n}{n!}\int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i) \times (-n\beta^{(n+1)}(x, u_1, \cdots, u_n) + \alpha^{(n+1)}(u_1, u_1, \cdots, u_n)) = \frac{-\mu^n}{(n-1)!}\int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_1 u_1 \times \bar{Z}_n(x, u_1, \cdots, u_n) \prod_i \Delta(u_i). \quad \text{(D1''')}$$

3. We define δ_{1k} such that

$$\delta_{1k}\beta^{(n+1)}(x, u_1, \cdots, u_k, \cdots, u_n) = \beta^{(n+1)}(x, u_k, \cdots, u_1, \cdots, u_n);$$

then we have

$$n\int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i)\beta^{(n+1)}(x, u_1, \cdots, u_n) = \sum_k \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i)\delta_{1k}\beta^{n+1}(x, u_1, \cdots, u_n),$$

where we have taken symmetry properties into account in the integration. We define

$$\sum_k \delta_{1k}\beta^{(n+1)}(x, u_1, \cdots, u_n) - \alpha^{(n+1)}(u_1, u_1, \cdots, u_n) = \gamma^{(n+2)}(x, u_1, u_2, \cdots, u_n),$$

where

$$\gamma^{(n+2)}(x, u_1, \cdots, u_n) = \begin{vmatrix} 0 & 0 & \frac{1}{u_1^2} & \cdots & \frac{1}{u_n^2} \\ 1 & 0 & \frac{x}{u_1(x+u_1)} & & \frac{x}{u_n(x+u_n)} \\ 1 & u_1 & & & \\ 1 & u_n & & \bar{\bar{Z}}_n(u_1, u_1, \cdots, u_n) & \end{vmatrix}.$$

In order to verify this identity, it is sufficient to develop $\gamma^{(n+2)}$ following the elements of the first column. We remark that the minor corresponding to $(\gamma^{(n+2)})_{k,1}$ is $(-)^{k+1}\delta_{1,k-2}\beta^{(n+1)}(x, \cdots, u_n)$ and that the minor corresponding to $(\gamma^{(n+2)})_{2,1}$ is $\alpha^{(n+1)}(u_1, u_1, \cdots, u_n)$. Now (D1'') is reduced to

$$\frac{\mu^n}{n!}\int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i)\gamma^{(n+2)}(x, u_1, \cdots, u_n) = \frac{-\mu^n}{(n-1)!}\int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n u_1 \bar{Z}_n(x, u_1, \cdots, u_n). \quad \text{(D1''')}$$

4. We develop $\gamma^{(n+2)}(x, u_1, \cdots, u_n)$ following the elements of the second column

$$\gamma^{(n+2)}(x, u_1, \cdots, u_n) = \sum_{j=1}^{n}(-)^j u_j \gamma_{j+2}^{(n+2)}(x, u_1, \cdots, u_n),$$

where $\gamma_{j+2}^{(n+2)}$ is the minor corresponding to $(\gamma^{n+2})_{j+2,2}$. Taking into account symmetry properties, each term of the sum in the left-hand side of (D1''') gives the same contribution, and because

$$\gamma_3^{(n+2)}(x, u_1, u_2, \cdots, u_n) = -\bar{Z}_n(x, u_1, \cdots, u_n),$$

the identity (D1''') is satisfied.

B. We want to show the following relation between the solutions of (1') and (5'):

$$-F'(0) = \int_{m/2}^{\infty} \frac{\Delta(x)F(x)}{x^2}\,dx = \frac{1}{\tilde{F}(0)}\int_{m/2}^{\infty} \frac{\Delta(x)\tilde{F}(x)}{x^2}\,dx. \quad \text{(D2)}$$

The Fredholm-type solution of (1') can be written $F(x) = 1 + N(x, \mu)/\mathfrak{D}(\mu)$; then by comparison with (23) we have only to show

$$\int_{m/2}^{\infty}\left(\frac{\tilde{\mathcal{N}}(x, \mu)}{x^2} - \frac{N(x, \mu)}{x^2}\right)\Delta(x)\,dx = 0 \quad \text{(D2')}$$

or

$$\sum_1^{\infty} \frac{(-\mu)^n}{(n-1)!}\Psi_n \equiv 0,$$

where

$$\Psi_n = \int_{m/2}^{\infty} dx \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_1^n \Delta(u_i) \times \phi_n(x, u_1, u_2, \cdots, u_n)\left[\frac{1}{u_1^2} - \frac{1}{x^2}\right],$$

$$\phi_n(x, u_1, u_2, \cdots, u_n) = \begin{vmatrix} \frac{1}{x+u_1} & \frac{1}{x+u_2} & \cdots & \frac{1}{x+u_n} \\ \frac{1}{u_2+u_1} & & & \\ \cdot & & & \\ \cdot & & P_{n-1}(u_2, u_3, \cdots, u_n) & \\ \cdot & & & \\ \frac{1}{u_n+u_1} & & & \end{vmatrix}.$$

But because

$$P_{n-1}(\cdots u_i, \cdots, u_j, \cdots) = P_{n-1}(\cdots u_j, \cdots, u_i, \cdots),$$

we get $\phi_n(x, u_1, \cdots) = \phi_n(u_1, x, \cdots)$ and $\Psi_n = 0$. Then (D2) follows.

APPENDIX E

We want to show that

$$d(\mu) = \mathfrak{D}(\mu)/F(0) \quad \text{when} \quad F'(0) = 0, \tag{E1}$$

$d(\mu)$, $\mathfrak{D}(\mu)$, $F(x)$ being defined by (C12), (C1), and (1'). We remark that if $F'(0) = 0$ then $F(x)$, which is a solution of (1'), is also solution of

$$F(x) = 1 + \int_{m/2}^{\infty} \mu\Delta(y)F(y)R(x, y)\, dy, \tag{E2}$$

with $R(x, y) = (x + y)^{-1} + x/y^2$, $\eta(x, \mu)$ and $d(\mu)$ being, respectively, the Fredholm numerator and denominator of the Fredholm-type solution of (E2).

Then, to prove (E1), it is sufficient to prove

$$\eta(0, \mu) + d(\mu) = \mathfrak{D}(\mu). \tag{E3}$$

From (E2) we have

$$d(\mu) + \eta(0, \mu) = \sum_1^{\infty} \frac{(-\mu)^n}{n!} \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i) \times [u_n(u_1, \cdots, u_n) - nV_n(u_1, \cdots, u_n)], \tag{E4}$$

with

$$u_n(u_1, \cdots, u_n) = \begin{vmatrix} R(u_1, u_1) & R(u_1, u_2) & \cdots & R(u_1, u_n) \\ R(u_2, u_1) & R(u_2, u_2) & \cdots & R(u_2, u_n) \\ \cdots & \cdots & \cdots & \cdots \\ R(u_n, u_1) & R(u_n, u_2) & \cdots & R(u_n, u_n) \end{vmatrix},$$

and $V_n(u_1, \cdots, u_n)$ is the same determinant as $U_n(u_1, \cdots u_n)$ except the first row which is $(u_1^{-1}, u_2^{-1}, \cdots, u_n^{-1})$. Using the same method as in Appendix D, we get

$$d(\mu) + \eta(0, \mu) = \sum_1^{\infty} \frac{(-\mu)^n}{n!} \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \times [U_n(u_1, \cdots, u_n) + \bar{V}_n(u_1, \cdots, u_n)], \tag{E5}$$

where

$$\bar{V}_n = U_n + \begin{vmatrix} 0 & \frac{1}{u_1} & \frac{1}{u_2} & \cdots & \frac{1}{u_n} \\ 1 & & U_n & & \end{vmatrix} = \begin{vmatrix} 1 & \frac{1}{u_1} & \frac{1}{u_2} & \cdots & \frac{1}{u_n} \\ 1 & & U_n & & \end{vmatrix}.$$

In $\bar{V}_n$ for the elements of the jth row ($j = 2, \cdots, n + 1$) we subtract the corresponding elements of the first one. With some algebra we get

$$\bar{V}_n + \begin{vmatrix} 1 & \frac{1}{u_1} & \cdots & \frac{1}{u_n} \\ 1 & & P_n & \end{vmatrix} = P_n.$$

Then (E1) follows.

APPENDIX F

We want to investigate the singular part (when $x_0 \to \infty$) of the Fredholm determinant of $\mathfrak{D}_{x_0}(\mu)$ corresponding to the kernel

$$K_{x_0} = \left(\frac{x^2}{y^2(x + y)} - \frac{x_0^2}{y^2(x_0 + y)}\right)\Delta(y)$$

defined on $[m/2, \infty]$. It is easy to see that

$$\mathfrak{D}_{x_0}(\mu) \underset{x_0 \to \infty}{\simeq} -x_0[A(\mu) + O(x_0^{-1})],$$

where $A(\mu)$ is the Fredholm determinant corresponding to the kernel

$$\left(\frac{x^2}{y^2(x + y)} - \frac{1}{y^2}\right)\Delta(y).$$

We get

$$A(\mu) = \sum_0^{\infty} \frac{(-\mu)^n}{n!} \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i) \times \left(\sum_{j=1}^{j=n} A_n^j(u_1, \cdots, u_n)\right),$$

where A_1^j is the determinant P_n defined in (C2) except that the elements of the jth column are $(u_1^{-2}, \cdots u_j^{-2}, \cdots, u_n^{-2})$. Taking into account symmetry properties in the integration with respect to the variables $u_1, \cdots, u_n$ we can see that each A_n^j gives the same contribution and because $A_n^1 = L_n$ [see (C12)] we get

$$A(\mu) = \sum_0^{\infty} \frac{(-\mu)^n}{(n-1)!} \int_{m/2}^{\infty} du_1 \cdots \int_{m/2}^{\infty} du_n \prod_i \Delta(u_i) L_n = \mathfrak{D}(\mu)F'(0).$$