JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Cluster Expansion in a Bose System*

M. HOWARD LEE†
Theoretical Physics Institute, Department of Physics
University of Alberta, Edmonton, Canada

(Received 7 March 1969)

When a trial function for Bose systems contains both one-particle and two-particle functions, the cluster expansion is no longer simply related to the classical expansion of Mayer and Mayer. But by observing that the permutations in one-particle functions can be represented exactly as in the high-temperature expansion of the spin-$\frac{1}{2}$ XY model, the cluster expansion is developed analogously to the classical case.

INTRODUCTION

In studying the ground-state properties of ^{3}He and ^{4}He in both liquid and solid states, the Hartree method has been shown to be quite inadequate[1] because of the hard-core interactions. The inadequacy is usually repaired by the so-called Jastrow function

$$\Psi(r_1 \cdots r_N) = \Phi(r_1 \cdots r_N) \cdot \prod_{i<j}^{N} f(|r_i - r_j|) \equiv \Phi \cdot \Psi_J, \quad (1)$$

where Φ is a one-particle (plane-wave) product function, which contains the properties of independent-particle systems such as excitations and statistics, and $f(|r_i - r_j|)$ is defined as

$$f(|r_i - r_j|) \to 0, \quad |r_i - r_j| \le r_0,$$
$$\to 1, \quad |r_i - r_j| \to \infty,$$

where r_0 is the hard-core distance.

In treating the ground state of an infinite system, Φ is usually taken to be unity for Bose systems and a Slater determinant for Fermi systems. Recent calculations[2] using these trial functions have shown some striking results. The main drawback of the Jastrow function is that an explicit expression for the expectation value of the Hamiltonian cannot be obtained. This expectation value must, therefore, be evaluated as a cluster expansion. Owing to Φ, the cluster expansions for Bose systems and Fermi systems will, in general, be quite different. In Fermi systems the presence of the Slater determinant makes the expansion extremely complicated.[3] But in Bose systems, if $\Phi = 1$, the cluster expansion can be developed exactly as the classical treatment of Mayer and Mayer. This is because the classical expansion formalism depends only on the boundary properties of $\exp -v(|r_i - r_j|)/kT$, where $v(|r_i - r_j|)$ is the two-body potential, and our $f(|r_i - r_j|)$, which plays the mathematically equivalent role to $\exp -v/kT$, meets the same boundary requirements.

From the variational point of view, the more satisfactory Jastrow function for Bose systems is $\Phi \neq 1$. Girardeau[4] has pointed out that a better trial function is one which incorporates the notion of a generalized condensation.[5] Luban[6] has further raised the possibility that for liquid ^{4}He the picture of a smeared condensation provides a more improved (i.e., lower free energy) description than the generally accepted picture of Bose–Einstein condensation in the zero-momentum state. Thus this leads us to consider a more general one-particle function for Φ:

$$\Phi = [N!\, \pi_\alpha n_\alpha!]^{-\frac{1}{2}} \sum_P P[u_{\alpha_1}(r_1) \cdots u_{\alpha_N}(r_N)], \quad (2)$$

where P denotes a permutation of the coordinates $r_1 \cdots r_N$, $u_\alpha(r)$ represents a complete orthonormal set of one-particle wavefunctions (plane waves), and n_α is the occupation number of the state α ($\sum_\alpha n_\alpha = N$). We shall assume here that the momentum distribution is smeared over a large number of states whose momenta are, however, all macroscopically negligible.

Now, the cluster expansion for Bose systems is no longer simply related to the classical one and, as in Fermi systems, the exchange effect of the

* Work supported in part by the National Research Council of Canada.

† Address after September 1, 1969: Department of Physics, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

[1] For example, see R. Jastrow, Phys. Rev. **98**, 1479 (1957); J. B. Aviles, Ann. Phys. (N.Y.) **5**, 251 (1958); D. J. Thouless, *The Quantum Mechanics of Many-Body Systems* (Academic Press Inc., New York, 1961), Chap. III.

[2] W. E. McMillan, Phys. Rev. **138**, A442 (1964); W. E. Massey, *ibid.* **151**, 153 (1966); G. N. Nosanow, *ibid.* **146**, 120 (1966); W. E. Massey and C. W. Woo, *ibid.* **164**, 256 (1967).

[3] The cluster expansions for Fermi systems were recently given by J. W. Clarke and P. Westhaus, J. Math. Phys. **9**, 131 (1968).

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

[5] The Hartree–Fock ground-state energy of a system of bosons interacting via predominantly attractive forces [i.e., Fourier transform of two-body potential $\tilde{v}(0) < 0$] is minimized when the particles undergo a smeared Bose–Einstein condensation. By a "smeared condensation," we mean that $n_\alpha = O(N^x)$ for $|\alpha| < \alpha_0$, where $0 < x < 1$, and $n_\alpha = 0$ otherwise, and $\alpha_0 = \alpha_0(N) \to 0$ as $N \to \infty$. n_α is the occupation number of state α. For other pertinent comments, see Ref. 4 and also G. Carmi, J. Math. Phys. **9**, 174 (1968).

[6] M. Luban, Phys. Rev. **128**, 965 (1962).

one-particle function must be taken into account in all of the resulting clusters. In Bose systems, however, since the exchanges bring about only the positive signs, the cluster expansion can still be developed very analogously to the classical case. It will be seen that the presence of the symmetrized product of one-particle functions leads to a simple change in the numerical coefficients of the clusters. These coefficients are very closely related to the occurrence factors of the spin clusters of the spin-$\frac{1}{2}$ XY model.[7] We present here a cluster expansion for a Bose system using a full Jastrow function.

NORMALIZATION INTEGRAL

One important quantity we can calculate with the Jastrow function is the so-called generalized normalization integral,[8]

$$D = \int d^3r_1 \cdots d^3r_N |\Psi|^2, \tag{3}$$

which is a quantum analog of the classical partition function. The presence of Φ prevents us from expanding (3) in the manner of the classical cluster expansion. But with some modifications it is still possible to take advantage of the formal relationship between (3) and the classical function.

Using (2) in (1), we can show that

$$\begin{aligned} D &= V^{-N}\int d^3r_1 \cdots d^3r_N \Psi_J^2 \cdot \Big\{1 + \tfrac{1}{2}\sum_{\alpha_1\alpha_2}{}' n_{\alpha_1}n_{\alpha_2}L^{(2)}(12) \\ &\quad + \frac{1}{3!}\sum_{\alpha_1\alpha_2\alpha_3}{}' n_{\alpha_1}n_{\alpha_2}n_{\alpha_3}L^{(3)}(123) + \cdots\Big\} \\ &\equiv D_0 + D_2 + D_3 + \cdots, \end{aligned} \tag{4}$$

where

$$L^{(2)}(12) = \exp i(\alpha_{12}\cdot r_1 + \alpha_{21}\cdot r_2), \quad \alpha_{12} \equiv \alpha_1 - \alpha_2,$$

$$\begin{aligned} L^{(3)}(123) &= \exp i(\alpha_{12}\cdot r_1 + \alpha_{23}\cdot r_2 + \alpha_{31}\cdot r_3) \\ &\quad + \exp i(\alpha_{13}\cdot r_1 + \alpha_{21}\cdot r_2 + \alpha_{32}\cdot r_3), \end{aligned}$$

and so on. D_0 is formally identical to the classical partition function. Equation (4) represents an expansion in the number of the particles permuted. The prime on the sum means that the diagonal terms are, in general, excluded.[9] When three or more particles are permuted, there is a multiplicity of the plane-wave factors (or L-factors) arising from different ways of permutation. This multiplicity makes the Bose-cluster expansion nontrivial. When the permutations are equivalent,[10] each cluster will have the same value. It will be seen that the number of equivalent permutations is directly related to the occurrence factors of the spin clusters, and the various inequivalent permutations can be represented by the spin graphs. Therein lies the connection between the Bose-cluster expansion and the high-temperature expansion of the XY model.

GRAPHIC REPRESENTATION OF L-FACTORS

$L^{(n)}(ij\cdots)$ can obviously be represented by some graphs. But, with an appropriate choice, it is possible to construct graphs for L-factors which are also graphs for the spin clusters of the high-temperature expansion, so that the multiplicities of L-factors are just the sums of the occurrence factors. Let $L^{(2)}(ij)$, which denotes the exchange of particle i in state α_i and particle j in state α_j, be represented by $G(ij)$ as shown in Fig. 1.

For two particles this is the only possible exchange. The same graph $G(1)$ also represents[11] the spin cluster $S_i^+S_j^-S_j^+S_i^-$ which has a unit occurrence factor. Three particles, say, ijk in states $\alpha_i\alpha_j\alpha_k$, respectively, can be exchanged in two ways. Hence let $L^{(3)}(ijk)$ be represented by $G(ijk)$ and $G(ikj)$. Since the two exchanges are evidently equivalent, they will contribute equally to the cluster expansion; both graphs can be represented by $G(2)$, provided that we associate 2 with it. For three particles there are no other exchanges possible. But $G(2)$ also represents the spin clusters $S_i^+S_j^-S_j^+S_k^-S_k^+S_i^-$ and $S_i^+S_j^-S_k^+S_i^-S_j^+S_k^-$, each of which has an identical trace value and the sum of their occurrence factors is 2. Four particles, say, $ijkl$ in state $\alpha_i\alpha_j\alpha_k\alpha_l$, respectively, can be exchanged in 4 inequivalent ways: $G(3)$, $G(1,1)$, $G(4)$, and $G(1\times 1)$. In $G(3)$ and $G(1,1)$ all states are distinct, but in $G(4)$ three states are distinct, and in $G(1\times 1)$ only two states are distinct. $G(3)$ represents 6 equivalent exchanges, where all particles are permuted with one another. $G(1,1)$ represents 3 equivalent exchanges, where 4 particles are permuted only pairwise. When 3 states are distinct, 4 particles can be equivalently permuted 4 ways; and when only 2 states are distinct, 4 particles can be permuted but one way. The four

[7] M. H. Lee, J. Math. Phys. (to be published); D. D. Betts and M. H. Lee, Phys. Rev. Letters **20**, 1507 (1968).

[8] F. Iwamoto and M. Yamada, Progr. Theoret. Phys. (Kyoto) **17**, 543 (1957).

[9] The permutations of particles in the same state are all included in the first term of the expansion. By the restrictions $\alpha_1 \neq \alpha_2$, $\alpha_1 \neq \alpha_2 \neq \alpha_3$, etc., we thus avoid the permutation of particles in the same state. This restriction is actually too stringent; in some cases it must be relaxed to allow certain degenerate permutations. For example, suppose there are four particles, say, $ijkl$, in state $\alpha_i\alpha_j\alpha_k\alpha_l$, respectively. If $\alpha_i = \alpha_j$, then, although particles i and j must not be permuted with each other, they can still be permuted with particles k or l, or both. When there is a degeneracy of this kind, one can show that such a term vanishes in the thermodynamic limit (N and $V \to \infty$ with ρ held fixed). Hence, it is actually sufficient to retain the permutation where there is no degeneracy.

[10] Two permutations are termed equivalent if one can be transformed into the other by exchanging the labels of particles. When permutations are inequivalent, then it is not possible to transform one permutation into another by such processes.

[11] The arrows and labels are dropped whenever convenient. Graphs are listed in Table I.

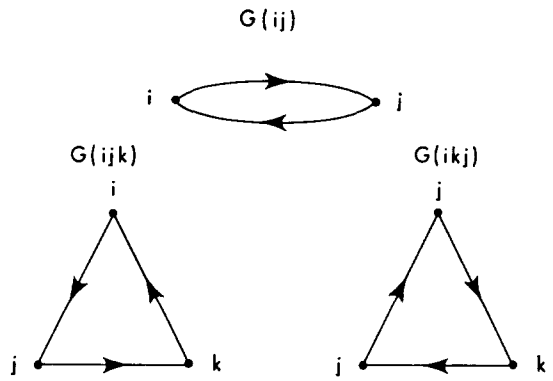
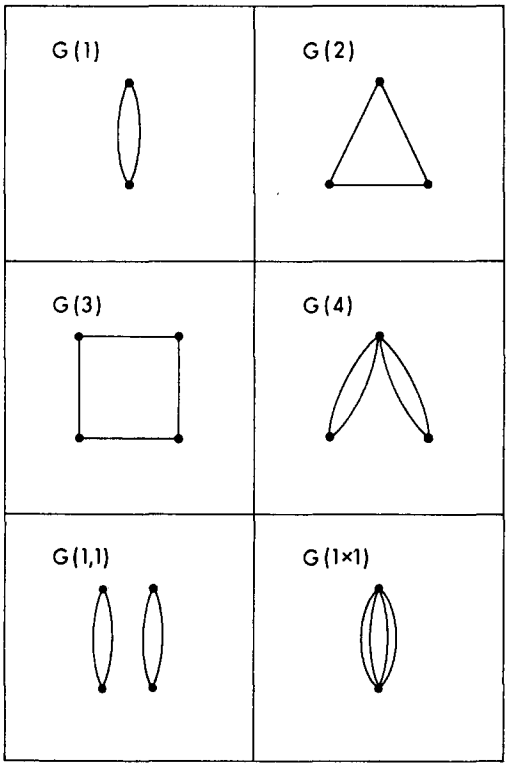


FIG. 1. Labeled graphs of the L-factors arising from the exchange of particles according to Bose statistics.

inequivalent exchanges of $L^{(4)}(ijkl)$ will contribute to the cluster expansion generally with different values. In the spin cluster expansion, we encounter exactly the same four inequivalent graphs $G(3)$, $G(1,1)$, $G(4)$, and $G(1\times 1)$ with their occurrence factors 6, 3, 4, and 1, respectively.

From these examples it is clear that the L-factors can be made to have a one-to-one correspondence to the spin clusters of the XY model. Hence we can immediately write down all the L-factors from the knowledge of the high-temperature-expansion graphs of the XY model previously given.

TABLE I. Unlabeled graphs of the L-factors which are made isomorphic to the spin graphs.[a]



[a] See Ref. 7.

CLUSTER EXPANSION

We can further write (3) in the form

$$\frac{D}{D_0} = 1 + \rho^2\binom{N}{2}^{-1}\sum{}' n_{\alpha_1}n_{\alpha_2}\int d^3r_1\, d^3r_2 L^{(2)}(12)g_J^{(2)}(12) + \rho^3\binom{N}{3}^{-1}\sum{}' n_{\alpha_1}n_{\alpha_2}n_{\alpha_3} \times \int d^3r_1\, d^3r_2\, d^3r_3 L^{(3)}(123)g_J^{(3)}(123) + \cdots, \quad (5)$$

where $\rho = N/V$ is the number density, $g_J^{(2)}$, $g_J^{(3)}$, $\cdots$ are the Jastrow 2-particle distribution function, 3-particle distribution function, $\cdots$, respectively, defined in the usual way,

$$g_J^{(2)}(r_1r_2) = \frac{N(N-1)}{\rho^2}\frac{\int d^3r_3\cdots d^3r_N\Psi_J^2}{\int d^3r_1\cdots d^3r_N\Psi_J^2} \equiv g_J(12),$$

$$g_J^{(3)}(r_1r_2r_3) = \frac{N(N-1)(N-2)}{\rho^3}\frac{\int d^3r_4\cdots d^3r_N\Psi_J^2}{\int d^3r_1\cdots d^3r_N\Psi_J^2},$$

and so on. If we decompose the many-particle distribution functions by the Kirkwood superposition approximation (KSA)[12] and introduce

$$g_J(12) = 1 + h(12),$$

where h is nearly zero everywhere except where the interactions are strong, then

$$\frac{D}{D_0} = 1 + \rho^2\binom{N}{2}^{-1}\sum{}' n_{\alpha_1}n_{\alpha_2}\int d^3r_1\, d^3r_2 L^{(2)}(12)\{h(12)\} + \rho^3\binom{N}{3}^{-1}\sum{}' n_{\alpha_1}n_{\alpha_2}n_{\alpha_3}\int d^3r_1\, d^3r_2\, d^3r_3 \times L^{(3)}(123)\{h(12)h(13) + h(12)h(23) + h(13)h(23) + h(12)h(13)h(23)\} + \cdots. \quad (6)$$

The cluster integrals in (6) will vanish unless all the coordinates over which we integrate appear explicitly in the products of h's. The formal structure of the nonvanishing products of h's is the same as that of the classical clusters.[13]

[12] S. A. Rice and P. Gray, *The Statistical Mechanics of Simple Fluids* (Interscience Publ. Co., New York, 1965), pp. 79–82, 114–134.

[13] J. de Boer, Rept. Progr. Phys. **12**, 329 (1948); J. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley & Sons, New York, 1956).

Then, dropping all the degenerate terms in the L-factors, we can write

$$\frac{D}{D_0} = 1 + \tfrac{1}{2}N\rho\beta_1 + \frac{2}{3!}N\rho^2(3\beta_1^2 + 2\beta_2) + \frac{9}{4!}N\rho^3(16\beta_1 + 12\beta_1\beta_2 + \beta_3) + \frac{3}{4!}(N^2\rho^2\beta_1^2) + \cdots = b_1 + N\rho b_2 + 2N\rho^2 b_3 + 9N\rho^3 b_4 + \tfrac{1}{2}(N\rho b_2)^2 + \cdots, \quad (7)$$

where β_i and b_i are defined as in the classical clusters.[14]

Thus, except for the numerical coefficients resulting from the multiplicities of the L-factors, (7) is formally identical to the cluster expansion of the classical partition function. Hence it is possible to expand the normalization integral to a high order immediately from the knowledge of the occurrence factors.

APPLICATION

We shall now apply our cluster-expansion formalism to obtain expressions for the ground-state energy and the pair-distribution function. In the usual way, the ground-state energy is given by

$$E = \frac{\int d^3r_1 \cdots d^3r_N \Psi^\dagger H \Psi}{\int d^3r_1 \cdots d^3r_N \Psi^\dagger \Psi}, \quad (8)$$

where

$$H = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2 + \sum_{ij}{}' V(|r_i - r_j|). \quad (9)$$

The presence of the Jastrow function in Ψ enables us to consider a very realistic two-body potential, such as the Lennard–Jones 12–6 potential in (9). We are interested in obtaining an expression for E with the single-particle momenta subject to the conditions of Girardeau's generalized condensation.

The ground-state energy can be expressed[15] as

$$E/N = \tfrac{1}{2}\rho\int d^3r[(\hbar^2/2m)\{(\nabla f)^2 - f\nabla f\}f^{-2} + V(r)]g(r), \quad (10)$$

where

$$g(r_1r_2) = \frac{N(N-1)}{\rho^2 D}\int d^3r_3 \cdots d^3r_N |\Psi|^2. \quad (11)$$

The bracketed terms in (10), which may be regarded as an effective interaction, can be explicitly given[16] since the pair function f and the pair potential V are, in principle, known. If $\Phi = 1$, then $g(r_1r_2)$ represents the usual pair-distribution function and it may be calculated by a variety of methods.[17] In our general case, where $\Phi \neq 1$, we shall show that $g(r_1r_2)$, and therefore E, can still be expressed as a cluster expansion. It may be worth pointing out here that if the ground-state energy is to have a saturation property,[16] the density dependence must be contained in $g(r_1r_2)$.

The general pair-distribution function may be further written[18] as

$$g(12) = \sum_{\alpha\beta} n_\alpha n_\beta \{A(\alpha\alpha\beta\beta) + (1 - \delta_{\alpha\beta})A(\alpha\beta\alpha\beta)\}M(\alpha\beta; 12)D^{-1}, \quad (12)$$

where

$$A(\alpha\beta\alpha\beta) = u_\alpha^*(r_1)u_\beta(r_1)u_\alpha^*(r_2)u_\beta(r_2), \quad (13)$$

$$D = \int d^3r_1 \cdots d^3r_N |\Psi|^2 = \langle n_1 \cdots n_N| \Psi_J^2 |n_N \cdots n_1\rangle_N, \quad (14)$$

and

$$M(\alpha\beta; 12) = \langle \cdots n_\alpha - 1 \cdots n_\beta - 1 \cdots | \times \Psi_J^2 |\cdots n_\beta - 1 \cdots n_\alpha - 1 \cdots\rangle_{N-2}^{12}. \quad (15)$$

By $|\cdots\rangle_n$ we mean an n-particle orthonormalized (Hartree–Fock) state. If there are unintegrated coordinates, they are explicitly shown, as in (15).

Evidently we may apply the same cluster development to $M(\alpha\beta; 12)$. First, $M(\alpha\beta; 12)$ is expanded in orders of the number of particles permuted; next, each term in the expansion is expressed in the form of the many-particle distribution function $g_J^{(n)}(r_1 \cdots r_n)$, which is then reduced by the KSA to a product of the pair-distribution functions $g_J(r_ir_j)$. Since the two coordinates (r_1 and r_2) are fixed, each term will therefore be proportional to $g_J(r_1r_2)$. The resulting cluster diagrams of $M(\alpha\beta; 12)$ consist of the same diagrams of D and diagrams with the two fixed coordinates.[19] Thus one can write

$$g(12) = g_J(12)\sum_{\alpha\beta}\frac{n_\alpha n_\beta}{N(N-1)} \times \{1 + (1 - \delta_{\alpha\beta})e^{i(\alpha-\beta)\cdot r_{12}}\}S(12; \rho), \quad (16)$$

[14] Strictly speaking, β_i and b_i are α-dependent. But since α_0, which is the maximum value of α, is vanishingly small, $\beta_i(\alpha) \simeq \beta_i(0) \equiv \beta_i$.

[15] The expression is simplified by the assumption of generalized condensation (see Ref. 4). For an analogous fermion system, see J. B. Aviles, Ref. 1.

[16] M. H. Lee and M. Luban, Phys. Rev. (to be published).

[17] For example, W. E. McMillan, Ref. 2, and M. H. Lee, Physica (to be published).

[18] This may be readily seen if we write the Hartree–Fock function as

$$\Phi(N) = [N(N-1)]^{-\frac{1}{2}}\sum_{\alpha\beta}(n_\alpha n_\beta)^{\frac{1}{2}}u_\alpha(1)u_\beta(2) \times |\cdots n_\alpha - 1 \cdots n_\beta - 1 \cdots\rangle_{N-2},$$

which is properly orthonormalized. Also see K. Huang, *Brandeis University Summer Lecture Notes* (W. A. Benjamin Inc., New York, 1959), Vol. II, pp. 2–4.

[19] Due to the momentum restriction (see Footnote 9), diagrams with only one unintegrated coordinate can be shown to vanish.

where[20]

$$S(12;\rho) = \frac{1 + N\rho b_2 + \rho^2\gamma_2(12) + \cdots}{1 + N\rho b_2 + 2N\rho^2 b_3 + \cdots} \equiv \mathcal{M}(12;\rho)/\mathcal{D}(\rho). \quad (17)$$

We have previously given the expansion for the denominator of (17), $\mathcal{D}(\rho) = D/D_0$. The $\gamma_n(12)$ are defined as in the classical statistical mechanics[21]:

$$\gamma_n(12) = (1/n!)\int \sum \pi h(ij)\, d^3r_3 \cdots d^3r_{n+2}, \quad (18)$$

where $h(ij) = g_J(ij) - 1$. The formal structure of $S(r\rho)$ is again very similar to the classical counterpart. This similarity may be exploited to put $S(r\rho)$ in a more tractable form. Perhaps the clearest way is by Kubo's method of cumulants.[22]

For this purpose it is convenient to re-express $\mathcal{D}(\rho)$ by

$$\mathcal{D} = \langle \pi(1 + l^{ij}h(ij))\rangle, \quad (19)$$

where l^{ij} are tensors for the occurrence factors (see Appendix). Then, by Kubo's method, (19) may be written as

$$\langle \pi(1 + l^{ij}h(ij))\rangle = \exp \langle \exp_L \sum l^{ij}h(ij) - 1\rangle_c, \quad (20)$$

where L is the so-called "leveling operator," which levels off a product of $h(ij)$ with a power greater than unity; c under the angular bracket denotes a cumulant average. Except for the tensors l^{ij}, our expression is formally identical to ln Z or the free energy $(-F/KT)$. It is well known in statistical mechanics that the free energy has a cluster expansion in terms of only the irreducible clusters (β_n) of the bonds $h(ij) = \exp -V(ij)/KT - 1$. Thus we can also write

$$\ln \mathcal{D} = N \sum_n \frac{\rho^n l^{(n)}\beta_n}{1+n} + O(N^{-1}), \quad (21)$$

where $l^{(n)}$ are pure numbers (see Appendix) and β_n are the nth irreducible clusters of the bonds $h(ij) = g_J(ij) - 1$.

Similarly, $\mathcal{M}(r_1r_2;\rho)$ may be expressed in the cumulant average

$$\ln \mathcal{M}(12;\rho) = \langle \exp_L \sum{}' l^{ij}h(ij) - 1\rangle_c^{1,2}, \quad (22)$$

where, by the prime, $h(12)$ is excluded from the sum and (r_1r_2) are the two unintegrated coordinates. The cluster expansion of (22), which generates two types of diagrams [β_n and $\Gamma_n(12)$], thus gives

$$\langle \exp_L \sum{}' l^{ij}h(ij) - 1\rangle_c = N \sum_n \frac{\rho^n l^{(n)}\beta_n}{1+n} + \sum_n \rho^n l^{(n)}\Gamma_n(12) + O(N^{-1}), \quad (23)$$

where $\Gamma_n(12)$ denotes the irreducible diagrams of $\gamma_n(12)$.

Finally, together with (17), (21), and (23), we obtain the density expansion

$$\ln S(12,\rho) = \sum_n \rho^n l^{(n)}\Gamma_n(12) + O(N^{-1}). \quad (24)$$

There is a theorem[23] which relates the set of irreducible diagrams to the set of connected diagrams. By aid of this theorem, (24) may be further written in a more convenient form:

$$S(12;\rho) = 1 + \sum_n \rho^n l^{(n)}\gamma_n(12), \quad (25)$$

where $\gamma_n(12)$ denotes the nth 1–2 connected clusters of the bonds $h = g_J - 1$.

Thus (10), (16), and (25) give the desired expression for the ground-state energy. With Girardeau's generalized condensation, the expression leads to an integral of the form[24]

$$E/N = \rho \int d^3r v_{\text{eff}}(r) g_J(r) S(r), \quad (26)$$

where

$$v_{\text{eff}}(r) = (\hbar/2m)\{(\nabla f)^2 - f\nabla f\}f^{-2} + V(r). \quad (27)$$

At relatively low densities, the density behavior in (26) is dominated by $g_J(r)$. Since $g_J(r)$ vanishes rigorously for small distances, the ground-state energy remains finite for any realistic two-body potential, such as the Lennard–Jones. Hence, the ground-state energy as given by (26) is expected to show a saturation at some reasonable value of the density.[16]

DISCUSSION

Since our cluster expansion is obtained by aid of the KSA, our results (7) and (26) are only approximate. In fact, the applicability of this whole formalism crucially depends on the validity of the KSA. The reliability of the KSA has been recently discussed by Alder[25] and Rahman.[26] For simple classical fluids one can numerically compare the 3-body distribution

[20] Actually, $S(r\rho)$ should depend on single-particle momenta. But in generalized condensation the dependence is negligible, as was pointed out in Footnote 14.

[21] J. E. Mayer and E. W. Montroll, J. Chem. Phys. **9**, 2, 626 (1941).

[22] R. Kubo, J. Phys. Soc. (Japan) **17**, 1100 (1962).

[23] J. M. J. Van Leeuwen, J. Groneveld, and J. de Boer, Physica **25**, 792 (1959).

[24] In model theories, Girardeau has shown that the ground-state energy with generalized condensation is lower than the ground-state energy with usual Bose condensation by a factor of 2 (see Ref. 4). The same factor is preserved in this treatment. Lowering of the ground-state energy is also shown by K. Sawada and R. Vasudevan, Phys. Rev. **124**, 300 (1961).

[25] B. J. Alder, Phys. Rev. Letters **12**, 317 (1964).

[26] A. Rahman, Phys. Rev. Letters **12**, 575 (1964).

function with the reduced triple products of pair-distribution functions at various separation distances. The results for hard spheres and a Lennard-Jones fluid show that at relatively high densities the KSA is surprisingly accurate. The agreements are almost quantitative. Furthermore, the KSA seems to give even better agreements for higher distribution functions. These results thus suggest that at fluid densities the KSA is a reasonable approximation to make.

In a sense, the use of the KSA in our cluster development is not entirely inconsistent with the essential ansatz of (1). As our N-body ground-state wavefunction contains no correlations higher than the two-body, it may be regarded as having been obtained from a true N-body correlated wavefunction by repeated applications of the KSA.

The expression (25) for $S(r\rho)$ is "exact" in the sense that all nonvanishing graphs have been accounted for. As can be seen from the virial expansion for fluids and the high-temperature expansion for spin systems, $S(r\rho)$ is not expected to be exactly summable. If, however, only a restricted class of diagrams are retained in the sum, it may be expressed in an integral equation form similar to the hypernetted chain-integral equation for fluids.[23]

Finally, in obtaining the cluster expansion, we have explicitly made use of Girardeau's generalized condensation. While this is motivated by the results of Girardeau and Luban, our formalism can be adapted to other forms of condensation. In such cases the simplicity of our cluster formalism shown here is not expected to be preserved.

ACKNOWLEDGMENTS

I am very grateful to Professor Marshall Luban for his valuable guidance and generous help on many facets of this problem. Portions of this work were done while I was at the University of Pennsylvania.

APPENDIX: OCCURRENCE FACTORS

Tensors for the occurrence factors are defined to satisfy the following relations[7]:

$$I^{ij} = I^{(2)}, \quad I^{ijk} = I^{(3)},$$
$$I^{ij}I^{kl} = (I^{(2)})^2, \quad I^{ij}I^{jk}I^{km} = I^{(4)}, \quad \text{etc.}\ ; \tag{A1}$$

$$\begin{aligned} I^{i} &= I^{(1)}, \\ I^{i}I^{j} &= I^{ij} = I^{(2)}, \\ I^{i}I^{i} &= (I^{(1)})^2, \\ I^{i}I^{ij} &= I^{ij} = I^{(2)}; \end{aligned} \tag{A2}$$

$$I^{(1)} = 0, \quad I^{(2)} = 1, \quad I^{(3)} = 2,$$
$$I^{(4)} = 9, \quad I^{(5)} = 44, \quad \text{etc.} \tag{A3}$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Formal Solutions of Inverse Scattering Problems

REESE T. PROSSER
Department of Mathematics, Dartmouth College, Hanover, New Hampshire

(Received 2 May 1968)

Formal solutions of inverse scattering problems for scattering from a potential, a variable index of refraction, and a soft boundary are developed using a method devised by Jost and Kohn.

1. INTRODUCTION

An inverse scattering problem seeks to obtain a quantitative description of an unknown scatterer from a knowledge of the scattering data. It is known in special cases (e.g., potential scattering,[1] soft-boundary scattering[2]) that under sufficiently restrictive hypotheses the totality of scattering data, as embodied by the scattering matrix, uniquely determines the form of the scatterer. The problem, then, is to implement these uniqueness results with constructive algorithms.

The most successful approach to this problem is that developed for potential scattering by Gel'fand and Levitan.[3] They obtained an algorithm which reproduces a suitably restricted spherically symmetric potential as the solution of a certain integral equation whose kernel is expressed in terms of the scattering data. But, while the Gel'fand–Levitan algorithm is an extremely elegant and highly satisfactory procedure, it does not seem to lend itself easily to more general problems. Attempts to extend it to nonspherically symmetric potentials[4] and variable indices of refraction[5] are not entirely satisfactory.

Simultaneously with the appearance of the Gel'fand–Levitan algorithm another approach to the same problem was introduced by Jost and Kohn[6] and developed by Moses.[7] They obtained an algorithm which reproduces the potential essentially by inverting the Born series for the scattering matrix. Although the Jost–Kohn algorithm is less efficient and less attractive than its more famous competitor, it is simpler in conception and easier to generalize.

This paper describes the formal extension of the Jost–Kohn algorithm to various inverse scattering problems and discusses the feasibility of practical applications. There are three sections: the first on potential scattering, the second on scattering from a variable index of refraction, and the third on scattering from a soft boundary. Only scalar wave equations are considered, although the extension of these results to vector wave equations appears to offer no difficulties in principle.

2. POTENTIAL SCATTERING

The scattering of a wavefunction $\varphi(\mathbf{x}, \mathbf{k})$ from a fixed potential $V(\mathbf{x})$ is governed by the time-independent Schrödinger equation

$$(\nabla^2 + \mathbf{k}^2)\varphi(\mathbf{x}, \mathbf{k}) = V(\mathbf{x})\varphi(\mathbf{x}, \mathbf{k}). \tag{1}$$

The solution, which is to consist of an ingoing plane wave plus an outgoing scattered wave, may be expressed as

$$\varphi(\mathbf{x}, \mathbf{k}) = e^{i\mathbf{k}\cdot\mathbf{x}} + \int \frac{e^{i|\mathbf{k}|\,|\mathbf{x}-\mathbf{y}|}}{4\pi\,|\mathbf{x}-\mathbf{y}|} V(\mathbf{y})\varphi(\mathbf{y}, \mathbf{k})\,d\mathbf{y}. \tag{2}$$

As $|\mathbf{x}| \to \infty$, the behavior of $\varphi(\mathbf{x}, \mathbf{k})$ is given by

$$\varphi(\mathbf{x}, \mathbf{k}) \to e^{i\mathbf{k}\cdot\mathbf{x}} + \frac{e^{i|\mathbf{k}|\,|\mathbf{x}|}}{4\pi\,|\mathbf{x}|} T(\mathbf{k}', \mathbf{k}) + O\left(\frac{1}{|\mathbf{x}|^2}\right). \tag{3}$$

Here, $\mathbf{k}' = (|\mathbf{k}|/|\mathbf{x}|)\mathbf{x}$, and $T(\mathbf{k}', \mathbf{k})$ is given by

$$T(\mathbf{k}', \mathbf{k}) = \int e^{-i\mathbf{k}'\cdot\mathbf{y}} V(\mathbf{y})\varphi(\mathbf{y}, \mathbf{k})\,d\mathbf{y}. \tag{4}$$

Thus, $T(\mathbf{k}', \mathbf{k})$ contains the scattering data. In fact, if $|\mathbf{k}|$ is proportional to the energy of the incoming plane wave and $\mathbf{k}/|\mathbf{k}|$ is the incoming direction, then $T(\mathbf{k}', \mathbf{k})$ gives the fraction of the outgoing wave scattered in the outgoing direction $\mathbf{k}'/|\mathbf{k}'|$.

It is known that, if $V(\mathbf{x})$ is square-integrable, then Eq. (2) has a solution $\varphi(\mathbf{x}, \mathbf{k})$ which is jointly continuous for $(\mathbf{x}, \mathbf{k}) \in E_3 \times K$, where K is any compact subset of E_3 not containing the origin.[8] It follows from (4) that if $V(\mathbf{x})$ is also integrable, then $T(\mathbf{k}', \mathbf{k})$ is jointly continuous for $(\mathbf{k}', \mathbf{k}) \in K \times K$. In particular, if $V(\mathbf{x})$ vanishes outside some sphere, then $\varphi(\mathbf{x}, \mathbf{k})$ is analytic in $\mathbf{k}$ and hence $T(\mathbf{k}', \mathbf{k})$ is analytic

[1] N. Levinson, Kgl. Danske Videnskab. Selskab, Mat.-Fys. Medd. **25**, 9 (1949).
[2] P. Lax and R. Phillips, Bull. Am. Math. Soc. **70**, 130 (1964).
[3] I. M. Gel'fand and B. M. Levitan, Izv. Akad. Nauk SSSR, Ser. Fiz. **15**, 309 (1951).
[4] I. Kay and H. Moses, Commun. Pure Appl. Math. **14**, 435 (1961).
[5] H. Moses and C. de Ridder, Lincoln Laboratory Report No. TR 322 DDC 422444, 1963.
[6] R. Jost and W. Kohn, Phys. Rev. **87**, 979 (1952).
[7] H. Moses, Phys. Rev. **102**, 559 (1956).
[8] T. Ikebe, Arch. Ratl. Mech. Anal. **5**, 1 (1960).

in $\mathbf{k}$ and $\mathbf{k}'$. Moreover, $T(\mathbf{k}', \mathbf{k})$ is jointly square-integrable in $\mathbf{k}$ and $\mathbf{k}'$.[9]

The problem now is to solve (4) for $V(\mathbf{y})$, given $T(\mathbf{k}', \mathbf{k})$. A formal solution is obtained by first solving (2) for $\varphi(\mathbf{x}, \mathbf{k})$ and substituting the result in (4). Thus,

$$\begin{aligned} T(\mathbf{k}', \mathbf{k}) = &\int e^{-i\mathbf{k}'\cdot\mathbf{y}} V(\mathbf{y}) e^{i\mathbf{k}\cdot\mathbf{y}}\, d\mathbf{y} + \iint e^{-i\mathbf{k}'\cdot\mathbf{y}_1} V(\mathbf{y}_1) \\ &\times \frac{e^{i|\mathbf{k}|\,|\mathbf{y}_1-\mathbf{y}_2|}}{4\pi\,|\mathbf{y}_1-\mathbf{y}_2|} V(\mathbf{y}_2) e^{i\mathbf{k}\cdot\mathbf{y}_2}\, d\mathbf{y}_2\, d\mathbf{y}_1 \\ &+ \cdots + \int\cdots\int e^{-i\mathbf{k}'\cdot\mathbf{y}_1} V(\mathbf{y}_1) \frac{e^{i|\mathbf{k}|\,|\mathbf{y}_1-\mathbf{y}_2|}}{4\pi\,|\mathbf{y}_1-\mathbf{y}_2|} \\ &\times V(\mathbf{y}_2)\cdots V(\mathbf{y}_n) e^{i\mathbf{k}\cdot\mathbf{y}_n} \\ &+ d\mathbf{y}_n \cdots d\mathbf{y}_1 + \cdots . \end{aligned} \tag{5}$$

If Fourier transforms are taken throughout, then

$$\begin{aligned} T(\mathbf{k}', \mathbf{k}) = V(\mathbf{k}' - \mathbf{k}) + &\int V(\mathbf{k}' - \mathbf{k}'') \frac{1}{\mathbf{k}''^2 - \mathbf{k}^2} \\ &\times V(\mathbf{k}'' - \mathbf{k})\, d\mathbf{k}'' + \cdots + \int\cdots\int V(\mathbf{k}' - \mathbf{k}'') \\ &\times \frac{1}{\mathbf{k}''^2 - \mathbf{k}^2} V(\mathbf{k}'' - \mathbf{k}''') \cdots V(\mathbf{k}^{(n)} - \mathbf{k}) \\ &\times d\mathbf{k}^{(n)} \cdots d\mathbf{k}'' + \cdots . \end{aligned} \tag{6}$$

It is known that (5), and hence (6), are convergent expansions for sufficiently weak potentials.[9]

Now if $\mathbf{k}'$ is put equal to $-\mathbf{k}$, $T(-\mathbf{k}, \mathbf{k})$ replaced by $\epsilon T(-\mathbf{k}, \mathbf{k})$, and V by $\sum_{m=1}^{\infty} \epsilon^m V_m$, then the coefficients of ϵ^m can be equated in (5) or (6). Thus

$$T(-\mathbf{k}, \mathbf{k}) = V_1(-2\mathbf{k}), \quad m = 1, \tag{7}$$

$$\begin{aligned} 0 = V_m(-2\mathbf{k}) &+ \sum_{i=2}^{m} \sum_{r_1+\cdots+r_i=m} \int\cdots\int e^{i\mathbf{y}_i\cdot\mathbf{k}} V_{r_1}(\mathbf{y}_1) \\ &\times \frac{e^{i|\mathbf{k}|\,|\mathbf{y}_1-\mathbf{y}_2|}}{4\pi\,|\mathbf{y}_1-\mathbf{y}_2|} V_{r_2}(\mathbf{y}_2)\cdots V_{r}(\mathbf{y}_i) \\ &\times e^{i\mathbf{k}\cdot\mathbf{y}_i}\, d\mathbf{y}_i \cdots d\mathbf{y}_1, \quad m > 1. \end{aligned} \tag{8}$$

Hence,

$$V_1(2\mathbf{k}) = T(\mathbf{k}, -\mathbf{k}), \tag{9}$$

$$\begin{aligned} V_m(2\mathbf{k}) = -\sum_{i=2}^{m} \sum_{r_i} &V_{r_1}(\mathbf{k} - \mathbf{k}'')\,(\mathbf{k}''^2 - \mathbf{k}^2)^{-1} \\ &\times V_{r_2}(\mathbf{k}'' - \mathbf{k}''') \cdots V_{r_i}(\mathbf{k}^{(i)} + \mathbf{k}) \\ &\times d\mathbf{k}^{(i)} \cdots d\mathbf{k}'', \end{aligned} \tag{10}$$

$$V(2\mathbf{k}) = \sum_{m=1}^{\infty} V_m(2\mathbf{k}). \tag{11}$$

[9] T. Ikebe, Pacific J. Math. **15**, 511 (1965).

Thus (9), (10), and (11) provide a formal solution to the inverse problem for potential scattering. Jost and Kohn have shown that the algorithm is convergent for sufficiently weak spherically symmetric potentials.[6] Perhaps a similar result for nonsymmetric potentials can be established along the same lines. But even if the algorithm fails to converge, it may still be summable in some sense sufficiently precise to recover the potential. Simple examples are worked out in Ref. 6.

Note especially that only a part of the scattering data is used in this algorithm; namely, only the scattering data for arbitrary incoming direction, and outgoing direction equal to the negative of the incoming direction (i.e., backscattering data at all energies and all aspects). Moreover, these data are used only in the first step of the algorithm, namely Eq. (9).

Other possibilities present themselves. If, instead of putting $\mathbf{k}' = -\mathbf{k}$ in (5), we fix the incoming direction (say along the negative x_3 axis) and put $|\mathbf{k}'| = |\mathbf{k}|$, then the same formal procedure leads to a similar algorithm with (9) replaced by

$$V_1(2\mathbf{h}) = T(\mathbf{k}', \mathbf{k}). \tag{12}$$

Here, $2\mathbf{h} = \mathbf{k}' - \mathbf{k}$. As $\mathbf{k}/|\mathbf{k}|$ is fixed, $|\mathbf{k}'|$ ranges over all energies, $\mathbf{k}'/|\mathbf{k}'|$ ranges over all outgoing directions, and $\mathbf{h}$ ranges over the half-space determined by the formula $\mathbf{h}\cdot\mathbf{k} < 0$. Hence, $V_1(2\mathbf{h})$ is not yet determined for all values of $\mathbf{h}$. But if $V(\mathbf{y})$ is real, then $V(-2\mathbf{h}) = \overline{V(2\mathbf{h})}$. Thus the algorithm is complete if (9) is replaced by

$$V_1(2\mathbf{h}) = T(\mathbf{k}', \mathbf{k}), \quad 2\mathbf{h} = \mathbf{k}' - \mathbf{k}, \tag{13}$$

$$V_1(-2\mathbf{h}) = \overline{V_1(2\mathbf{h})}. \tag{14}$$

This algorithm reproduces the potential from the scattering data for a fixed incoming direction, all energies, and all outgoing directions. The convergence properties, of course, are the same as in the preceding case.

A similar algorithm is obtained if the outgoing direction is held fixed. The roles of k and k' are now interchanged and (13) determines $V_1(2\mathbf{h})$ only in the half-space $\mathbf{h}\cdot\mathbf{k}' > 0$. Otherwise the result is the same as before.

3. SCATTERING FROM A VARIABLE INDEX OF REFRACTION

The scattering of a wavefunction $\varphi(\mathbf{x}, \mathbf{k})$ from a variable index of refraction $n(\mathbf{x})$ is governed by the wave equation

$$[\nabla^2 + \mathbf{k}^2 n(\mathbf{x})]\varphi(\mathbf{x}, \mathbf{k}) = 0. \tag{15}$$

If $w(\mathbf{x}) = 1 - n(\mathbf{x})$, Eq. (15) becomes

$$(\nabla^2 + \mathbf{k}^2)\varphi(\mathbf{x}, \mathbf{k}) = \mathbf{k}^2 w(\mathbf{x})\varphi(\mathbf{x}, \mathbf{k}). \tag{16}$$

This equation resembles Eq. (1), with $V(\mathbf{x})$ replaced by $\mathbf{k}^2 w(\mathbf{x})$. For fixed $\mathbf{k}$, the behavior of the solutions in the two cases is identical, but the dependence on $\mathbf{k}$ is different. It is just this difference that renders the Gel'fand–Levitan algorithm inapplicable.[5] However, the Jost–Kohn algorithm remains insensitive to the dependence on $\mathbf{k}$ throughout. In fact, if $V(\mathbf{x})$ is replaced by $\mathbf{k}^2 w(\mathbf{x})$ throughout, the following results are obtained for the backscattering case ($\mathbf{k}' = -\mathbf{k}$):

$$\mathbf{k}^2 w_1(2\mathbf{k}) = T(-\mathbf{k}, \mathbf{k}), \tag{17}$$

$$\begin{aligned}\mathbf{k}^2 w_m(2\mathbf{k}) = &-\sum_{i=2}^{m}\sum_{r_i} w_{r_1}(\mathbf{k} - \mathbf{k}'')\frac{\mathbf{k}^2}{\mathbf{k}'^2 - \mathbf{k}^2} \\ &\times w_{r_2}(\mathbf{k}'' - \mathbf{k}''') \cdots w_{r_i}(\mathbf{k}^{(i)} + \mathbf{k}) \\ &\times d\mathbf{k}^{(i)} \cdots d\mathbf{k}'',\end{aligned} \tag{18}$$

$$\mathbf{k}^2 w(2\mathbf{k}) = \sum \mathbf{k}^2 w_i(2\mathbf{k}). \tag{19}$$

Similar algorithms hold for the other cases.

4. SCATTERING FROM A SOFT BOUNDARY

The scattering of a wavefunction $\varphi(\mathbf{x}, \mathbf{k})$ from a soft boundary is governed by the wave equation with the boundary condition

$$\begin{aligned}(\nabla^2 + \mathbf{k}^2)\varphi(\mathbf{x}, \mathbf{k}) &= 0, \quad \mathbf{x} \in R', \\ \varphi(\mathbf{x}, \mathbf{k}) &= 0, \quad \mathbf{x} \in \partial R.\end{aligned} \tag{20}$$

Here R' is the exterior of a compact region R in E_3 with smooth boundary ∂R. Again the solution, which is to consist of an ingoing plane wave plus an outgoing scattered wave, may be expressed as the solution of an integral equation of the form

$$\varphi(\mathbf{x}, \mathbf{k}) = e^{i\mathbf{k}\cdot\mathbf{x}} + 2\int_{\partial R} \frac{e^{i|\mathbf{k}|\,|\mathbf{x}-\mathbf{y}|}}{4\pi\,|\mathbf{x} - \mathbf{y}|}\frac{\partial}{\partial n(\mathbf{y})}\varphi(\mathbf{y}, \mathbf{k})\,d\mathbf{y}. \tag{21}$$

Here $\partial/\partial n(\mathbf{y})$ denotes the exterior normal derivative and the integration is taken over the boundary ∂R of R. As $|\mathbf{x}| \to \infty$, the behavior of $\varphi(\mathbf{x}, \mathbf{k})$ is given by

$$\varphi(\mathbf{x}, \mathbf{k}) \to e^{i\mathbf{k}\cdot\mathbf{x}} + \frac{e^{i|\mathbf{k}|\,|\mathbf{x}|}}{4\pi\,|\mathbf{x}|}T(\mathbf{k}', \mathbf{k}) + O\left(\frac{1}{|\mathbf{x}|^2}\right). \tag{22}$$

Here again $\mathbf{k}' = (|\mathbf{k}|/|\mathbf{x}|)\,|\mathbf{x}|$ and $T(\mathbf{k}', \mathbf{k})$ is given by

$$T(\mathbf{k}', \mathbf{k}) = 2\int_{\partial R} e^{-i\mathbf{k}'\cdot\mathbf{y}}\frac{\partial}{\partial n(\mathbf{y})}\varphi(\mathbf{y}, \mathbf{k})\,d\mathbf{y}. \tag{23}$$

As before, $T(\mathbf{k}', \mathbf{k})$ contains the scattering data. It is known that, for compact regions with sufficiently smooth boundary, $\varphi(\mathbf{x}, \mathbf{k})$ is analytic in $\mathbf{k}$ and hence $T(\mathbf{k}', \mathbf{k})$ is analytic in $\mathbf{k}$ and $\mathbf{k}'$.[2] Moreover, $T(\mathbf{k}', \mathbf{k})$ is bounded (but no longer square-integrable) in $\mathbf{k}$ and $\mathbf{k}'$.

The problem now is to extract from (23) a quantitative description of the shape of the boundary. For this purpose it suffices to determine the characteristic function $\chi_R(\mathbf{x})$ of the region R:

$$\begin{aligned}\chi_R(\mathbf{x}) &= 1, \quad \text{if } x \in R, \\ &= 0, \quad \text{if } x \notin R,\end{aligned} \tag{24}$$

or its Fourier transform. A formal solution is obtained by first solving (21) for $\varphi(\mathbf{x}, \mathbf{k})$ and substituting the result in (23). Thus,

$$\begin{aligned}&T(\mathbf{k}', \mathbf{k}) \\ &= 2\int_{\partial R} e^{-i\mathbf{k}'\cdot\mathbf{y}}\frac{\partial}{\partial n(\mathbf{y})}e^{i\mathbf{k}\cdot\mathbf{y}}\,d\mathbf{y} \\ &\quad + 4\iint e^{-i\mathbf{k}'\cdot\mathbf{y}_1}\frac{\partial}{\partial n(\mathbf{y}_1)}\frac{e^{i|\mathbf{k}|\,|\mathbf{y}_1-\mathbf{y}_2|}}{4\pi\,|\mathbf{y}_1 - \mathbf{y}_2|} \\ &\quad \times \frac{\partial}{\partial n(\mathbf{y}_2)}e^{i\mathbf{k}\cdot\mathbf{y}_2}\,d\mathbf{y}_2\,d\mathbf{y}_1 + \cdots \\ &\quad + 2^n\int\cdots\int e^{-i\mathbf{k}'\cdot\mathbf{y}_1}\frac{\partial}{\partial n(\mathbf{y}_1)}\frac{e^{i|\mathbf{k}|\,|\mathbf{y}_1-\mathbf{y}_2|}}{4\pi\,|\mathbf{y}_1 - \mathbf{y}_2|} \\ &\quad \times \frac{\partial}{\partial n(\mathbf{y}_2)}\cdots\frac{\partial}{\partial n(\mathbf{y}_n)}e^{i\mathbf{k}\cdot\mathbf{y}_n}\,d\mathbf{y}_n\cdots d\mathbf{y}_1.\end{aligned} \tag{25}$$

This may be rewritten in terms of volume integrals involving $\chi_R(\mathbf{x})$:

$$\begin{aligned}T(\mathbf{k}', \mathbf{k}) &= 2\int_{E_3} e^{-i\mathbf{k}'\cdot\mathbf{y}}\nabla\chi_R(\mathbf{y})\cdot\nabla e^{i\mathbf{k}\cdot\mathbf{y}}\,d\mathbf{y} \\ &\quad + 4\iint e^{-i\mathbf{k}'\cdot\mathbf{y}_1}\nabla\chi_R(\mathbf{y}_1)\cdot\nabla\frac{e^{i|\mathbf{k}|\,|\mathbf{y}_1-\mathbf{y}_2|}}{4\pi\,|\mathbf{y}_1 - \mathbf{y}_2|} \\ &\quad \times \nabla\chi_R(\mathbf{y}_2)\cdot\nabla e^{i\mathbf{k}\cdot\mathbf{y}_2}\,d\mathbf{y}_2\,d\mathbf{y}_1 + \cdots \\ &\quad + 2^n\int\cdots\int e^{-i\mathbf{k}'\cdot\mathbf{y}_1}\nabla\chi_R(\mathbf{y}_1)\cdot\nabla\frac{e^{i|\mathbf{k}|\,|\mathbf{y}_1-\mathbf{y}_2|}}{4\pi\,|\mathbf{y}_1 - \mathbf{y}_2|} \\ &\quad \times \nabla\chi_R(\mathbf{y}_2)\cdots\nabla e^{i\mathbf{k}\cdot\mathbf{y}_n}\,d\mathbf{y}_n\cdots d\mathbf{y}_1.\end{aligned} \tag{26}$$

If Fourier transforms are taken throughout, then

$$\begin{aligned}&T(\mathbf{k}', \mathbf{k}) \\ &= 2\chi_R(\mathbf{k}' - \mathbf{k})(\mathbf{k}' - \mathbf{k})\cdot\mathbf{k} \\ &\quad + 4\int\chi_R(\mathbf{k}' - \mathbf{k}'')(\mathbf{k}' - \mathbf{k}'')\cdot\mathbf{k}''(\mathbf{k}''^2 - \mathbf{k}^2)^{-1} \\ &\quad \times \chi_R(\mathbf{k}'' - \mathbf{k})(\mathbf{k}'' - \mathbf{k})\cdot\mathbf{k}\,d\mathbf{k}'' + \cdots \\ &\quad + 2^n\int\cdots\int\chi_R(\mathbf{k}' - \mathbf{k}'')(\mathbf{k}' - \mathbf{k}'') \\ &\quad \cdot \mathbf{k}''(\mathbf{k}''^2 - \mathbf{k}^2)^{-1}\chi_R(\mathbf{k}'' - \mathbf{k}''')(\mathbf{k}'' - \mathbf{k}''') \\ &\quad \cdot \mathbf{k}'''(\mathbf{k}'''^2 - \mathbf{k}^2)^{-1}\cdots\chi_R(\mathbf{k}^{(n)} - \mathbf{k})(\mathbf{k}^{(n)} - \mathbf{k}) \\ &\quad \cdot \mathbf{k}\,d\mathbf{k}^{(n)}\cdots d\mathbf{k}_1.\end{aligned} \tag{27}$$

The strong resemblance between (26) and (6) suggests that the Jost–Kohn algorithm may be applicable. If $\mathbf{k}'$ is put equal to $-\mathbf{k}$, $T(-\mathbf{k},\mathbf{k})$ replaced by $\epsilon T(-\mathbf{k},\mathbf{k})$, and χ_R by $\sum \epsilon^m \chi_m$, then the coefficients of ϵ^m can be equated in (27). Thus,

$$T(-\mathbf{k},\mathbf{k}) = -4\mathbf{k}^2\chi_1(-2\mathbf{k}), \quad m = 1, \tag{28}$$

$$\begin{aligned} 0 = {} & -4\mathbf{k}^2\chi_m(-2\mathbf{k}) \\ & + \sum_{i=2}^{m} \sum_{r_i+\cdots+r_i=m} 2^i \int\cdots\int \chi_{r_1}(-\mathbf{k}-\mathbf{k}'') \\ & \times (-\mathbf{k}-\mathbf{k}'')\cdot\mathbf{k}''(\mathbf{k}''^2-\mathbf{k}^2)^{-1} \\ & \times \chi_{r_2}(\mathbf{k}''-\mathbf{k}''')(\mathbf{k}''-\mathbf{k}''') \\ & \cdot \mathbf{k}'''\cdots\chi_{r_1}(\mathbf{k}^{(i)}-\mathbf{k})(\mathbf{k}^{(i)}-\mathbf{k})\cdot(-\mathbf{k}) \\ & d\mathbf{k}^{(i)}\cdots d\mathbf{k}', \quad m > 1. \end{aligned} \tag{29}$$

Hence,

$$4\mathbf{k}^2\chi_1(2\mathbf{k}) = -T(\mathbf{k},-\mathbf{k}), \tag{30}$$

$$\begin{aligned} 4\mathbf{k}^2\chi_m(2\mathbf{k}) = {} & +\sum_{i=2}^{m} \sum_{r_i} 2^i \int\cdots\int \chi_{r_1}(\mathbf{k}-\mathbf{k}') \\ & \times (\mathbf{k}-\mathbf{k}')\cdot\mathbf{k}'\cdots\chi_{r_i}(\mathbf{k}^{(i)}+\mathbf{k}) \\ & \times (\mathbf{k}^{(i)}+\mathbf{k})\cdot\mathbf{k}\, d\mathbf{k}^{(i)}\cdots d\mathbf{k}', \end{aligned} \tag{31}$$

$$4\mathbf{k}^2\chi_R(2\mathbf{k}) = \sum_{m=1}^{\infty} 4\mathbf{k}^2\chi_m(2\mathbf{k}). \tag{32}$$

Thus (30), (31), and (32) provide a formal solution to the inverse problem for boundary scattering. The question of convergence is now under study. It appears that the algorithm may be convergent at least for regions with sufficiently smooth convex boundaries. But even if the algorithm fails to converge, it may still be summable in some sense. In any case,each step is well defined and yields quantitative information about the region which may be useful in applications.

Here again only a part of the scattering data is used, namely, backscattering data at all energies and all aspects. Similar algorithms can obviously be developed for other data, including fixed ingoing direction, fixed outgoing direction, and fixed scattering angle (e.g., bistatic data). It seems plausible that even less data will suffice, inasmuch as the boundary can be described by two independent parameters and all the cases described above involve three. This possibility is also under consideration.

ACKNOWLEDGMENT

The author gratefully acknowledges his indebtedness to Dr. Harry E. Moses of Lincoln Laboratory, whose work (Ref. 7) inspired the present paper and whose comments and suggestions contributed materially to its final form.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Two-Variable Expansion of the Scattering Amplitude: An Application of Appell's Generalized Hypergeometric Functions

JOHN M. CHARAP AND B. M. MINTON
Department of Physics, Queen Mary College, London E.1, England

(Received 20 November 1968)

Appell's polynomials in two variables orthogonal in a triangle are described and some of their properties and those of related generalized hypergeometric functions are given. An application to the expansion of the scattering amplitude is suggested, the equal-mass case being discussed in some detail. A simple crossing matrix is derived. Difficulties introduced by inequality of the particle masses are explained. A Neumann formula is presented which permits an analytic continuation in the parameters to be made of the expansion coefficients for parts of the amplitude: This is in analogy with the Froissart–Gribov continuation. A conjectured analog of the Sommerfeld–Watson transformation then suggests the existence of fixed cuts in the partial-wave scattering amplitude.

1. INTRODUCTION

The study of the partial-wave expansion of the scattering amplitude has a number of overlapping motivations. This expansion is a simple way of parameterizing the angular distribution; the parameterization has immediate physical significance because angular momentum is conserved; unitarity does not mix different angular-momentum states; resonances, etc., have well-defined angular momenta; analytic properties of the partial-wave amplitude in the angular momentum may be used to associate resonances together into families; the trajectories on which these resonances lie give information about the high-energy behavior of the crossed-channel scattering process. There is, nonetheless, a certain mismatch between the partial-wave expansion and the crossing symmetry of the scattering amplitude occasioned by the necessity to select the particular channel in which the angular-momentum decomposition is to be performed. This has motivated various searches for a crossing-symmetric analog or generalization to the partial-wave expansion. We would like to refer especially to the work of Khuri,[1] who used monomials in the Mandelstam variables, and more recently to the papers of Balachandran *et al.*[2–4] These latter papers made kind reference to the interest that one of us had in the past expressed in the use of polynomials in two variables, and it is principally from a feeling of encouragement by this reference that the present paper has been written. To a certain extent we have succeeded in making a limited advance in the direction we had set ourselves. But although we feel that some of the results we have obtained may have an intrinsic mathematical significance, we have been frustrated in our attempts to find any clear-cut physical application. There is, however, a tentative suggestion of a mechanism for the generation of *fixed cuts* in the angular-momentum plane.

There is in the Bateman Manuscript Project a section headed "Orthogonal Polynomials in the Triangle,"[5] and it was the serendipitous discovery of this section which first initiated the work embodied in this paper. In Sec. 2, we introduce the polynomials in two variables (two of the Mandelstam variables, t and u, in our application) and give an expansion of the scattering amplitude. In Sec. 3 we give further properties of these polynomials and of the functions which generalize them; in particular, we discuss the differential equations they satisfy (see also Appendix C) and their partial-wave expansions (cf. Appendix A). The crossing matrix is also defined and derived in Appendix B. Section 4 concerns the functions of the second kind, related to the polynomials in a manner analogous to that which relates the Legendre functions of the second kind to the Legendre polynomials. We give some integral representations and broach the problem of asymptotic behavior. All the foregoing discussions concern the equal-mass case, and in Sec. 5 we outline some of the serious difficulties that unequal masses introduce. Finally, in Sec. 6 we indulge in some speculations which might suggest a generalization of the Regge–Sommerfeld–Watson transformation of the amplitude, and which indicate the possible existence of the fixed cuts mentioned above. We must caution the reader that the overlap in this paper between what we have proved and derived with a degree of confidence and what we believe to have

[1] N. N. Khuri, Phys. Rev. Letters **10**, 420 (1963); Phys. Rev. **132**, 914 (1963).
[2] A. P. Balachandran and J. Nuyts, Phys. Rev. **172**, 1821 (1968).
[3] A. P. Balachandran, W. J. Meggs, and P. Ramond, ICTP, Trieste, Preprint IC/68/44,1968.
[4] A. P. Balachandran, W. J. Meggs, J. Nuyts, and P. Ramond, ICTP, Trieste, Preprint IC/68/46, 1968.
[5] A. Erdélyi, Ed., *Higher Transcendental Functions, Vol. II* (McGraw-Hill Book Co., Inc., New York, 1953).

any deep significance for the understanding of the crossing symmetric two-particle scattering amplitude is very slight.

2. EXPANSION OF THE SCATTERING AMPLITUDE IN POLYNOMIALS IN TWO VARIABLES

We shall for simplicity restrict our considerations to the elastic scattering amplitude $T(stu)$ for two spinless particles of equal mass, with only brief mention of the problems associated with generalizing this treatment. If we choose units in which the particle mass is $m = \frac{1}{2}$, the Mandelstam variables are constrained by

$$s + t + u = 1. \tag{2.1}$$

The s-channel partial-wave expansion is

$$T(stu) = \Sigma_J (2J + 1) P_J(z_s) A_J(a), \tag{2.2}$$

where

$$t = \tfrac{1}{2}a(1 - z_s),$$
$$u = \tfrac{1}{2}a(1 + z_s), \tag{2.3}$$

and

$$a = 1 - s = -p_s^2/m^2. \tag{2.4}$$

This partial-wave expansion may be thought of as an expansion of $T(stu) \equiv T(a, z_s)$ at each fixed value of a in terms of the polynomials $P_J(z_s)$, which are uniquely characterized as the orthogonal polynomials in z_s on the interval $(-1, +1)$ with unit weight function and suitable normalization. In this paper we shall be concerned with an expansion of the amplitude $T(stu)$ in terms of polynomials $E_{mn}(t, u)$ of *two* variables; these together with a related set of polynomials $F_{mn}(t, u)$ form a biorthogonal system on the triangle Δ given by $s > 0$, $t > 0$, $u > 0$, again with unit weight function.

The problem of constructing polynomials in two variables orthogonal on a triangle has been discussed by Appell,[6] who introduced generalizations of the Jacobi polynomials which are orthogonal polynomials on a line. As a special case of these generalized Jacobi polynomials (appropriate to the present problem of spinless particles, just as Legendre polynomials are the appropriate special case of one-variable Jacobi polynomials), we consider the polynomials $F_{mn}(t, u)$, defined by

$$F_{mn}(t, u) \equiv \frac{1}{m!}\frac{1}{n!}\left(\frac{\partial}{\partial t}\right)^m \left(\frac{\partial}{\partial u}\right)^n t^m u^n (1 - t - u)^{m+n}, \tag{2.5}$$

which is,of course,closely analogous with the Rodrigues formula

$$P_J(z) = \frac{1}{2^J J!}\left(\frac{d}{dz}\right)^J (z^2 - 1)^J. \tag{2.6}$$

It is clear that F_{mn} is of degree $m + n$ in each of t and u. These polynomials may be expressed in terms of one of Appell's generalizations of the hypergeometric series to two variables, namely,

$$F_2(\alpha; \beta, \beta'; \gamma, \gamma'; x, y) \equiv \sum_{pq} \frac{(\alpha)_{p+q}(\beta)_p(\beta')_q}{p!\, q!\, (\gamma)_p(\gamma')_q} x^p y^q. \tag{2.7}$$

The notation is conventional and

$$(a)_r \equiv \Gamma(a + r)/\Gamma(a) \tag{2.8}$$

for $r = 0, 1, 2, \cdots$. Explicitly, for m and n integers the series $F_2(-m - n; m + 1, n + 1; 1, 1; t, u)$ terminates and is precisely the polynomial F_{mn}. Even for arbitrary complex values of μ and ν, the series $F_2(-\mu - \nu; \mu + 1, \nu + 1; 1, 1; t, u)$ converges uniformly and absolutely for (t, u) in Δ and may be used to define a function $F_{\mu\nu}(t, u)$, which bears the same relationship to the polynomials F_{mn} as does the Legendre function $P_\nu(z)$ to the Legendre polynomials $P_J(z)$ $(J = 0, 1, 2, \cdots)$. When $\mu = -m - 1$ and $\nu = -n - 1$ with m and n nonnegative integers, the function $F_{\mu\nu}$ again reduces to a polynomial, this time of degree m in t and n in u, and we write

$$F_{-m-1,-n-1}(t, u) \equiv E_{mn}(tu). \tag{2.9}$$

The two systems of polynomials F_{mn} and E_{mn} are biorthogonal on Δ with unit weight function[7]; we have

$$\iint_\Delta dt\, du F_{mn}(t, u) E_{kl}(tu) = \delta_{mk}\delta_{nl}[2(m + n + 1)^2]^{-1}. \tag{2.10}$$

An arbitrary function of two variables defined on Δ and belonging to a suitable class of functions may be expanded in terms of either of the systems E_{mn} or F_{mn}. We shall be most interested in the first of these expansions and shall write for the expansion of the scattering amplitude $T(stu)$

$$T(stu) = \sum_{m,n} 2(m + n + 1)^2 A_{mn} E_{mn}(t, u), \tag{2.11}$$

[6] P. Appell, Arch. Math. Physik (1) **66**, 238 (1881); P. Appell and J. Kampé de Férier, *Fonctions hypergéométriques et hypersphériques, polynomes d'Hermite* (Gauthier-Villars, Paris, 1926).

[7] It is possible to construct an orthogonal system, rather than a pair of biorthogonal systems, and this has been discussed by W. Gröbner [Monatsh. Math. **52**, 38 (1948)]. With a modified notation, Gröbner's system is given in terms of our polynomials F by

$$P_{MN} = \sum_{n=0}^{N} (-)^{M+n} \frac{N!\,(2M + N + 1)!\,(M + N - n)!}{(M + N)!\, n!\, (N - n)!\, (2M + N + 1 - n)!} \times F_{M+N-n,n};$$

for these polynomials

$$\iint_\Delta dt\, du P_{MN}(t, u) P_{M'N'}(tu) = [2(M + N + 1)(2M + 1)]^{-1}\delta_{MM'}\delta_{NN'}.$$

The principal disadvantage is the fact that these polynomials do not satisfy any simple system of second-degree differential equations. They are, however, eigenfunctions of the operator discussed in Sec. 3 and Appendix C.

where the coefficients A_{mn} are obtained from Eq. (2.10) as

$$A_{mn} = \iint_{\Delta} dt\, du T(stu) F_{mn}(tu). \tag{2.12}$$

3. SOME PROPERTIES OF THE FUNCTIONS $F_{\mu\nu}$

The generalized hypergeometric function $F_{\mu\nu}(t, u)$ satisfies the pair of coupled partial-differential equations

$$\begin{aligned}\{t(1-t)\partial_t^2 - tu\partial_t\partial_u + [1 + (\nu - 2)t]\partial_t \\ - (\mu + 1)u\partial_u + (\mu + \nu)(\mu + 1)\}F_{\mu\nu}(t, u) = 0, \\ \{u(1-u)\partial_u^2 - tu\partial_t\partial_u + [1 + (\mu - 2)u]\partial_u \\ - (\nu + 1)t\partial_t + (\mu + \nu)(\nu + 1)\}F_{\mu\nu}(t, u) = 0.\end{aligned} \tag{3.1}$$

From this it follows that $F_{\mu\nu}$ is an eigenfunction of the differential operator $\mathcal{O}$,

$$\begin{aligned}\mathcal{O} \equiv -t(1-t)\partial_t^2 + 2tu\partial_t\partial_u - u(1-u)\partial_u^2 \\ + (3t - 1)\partial_t + (3u - 1)\partial_u,\end{aligned} \tag{3.2}$$

with eigenvalue $(\mu + \nu)(\mu + \nu + 2)$; that is,

$$\mathcal{O}F_{\mu\nu} = (\mu + \nu)(\mu + \nu + 2)F_{\mu\nu}. \tag{3.3}$$

Note that the polynomials F_{mn} and E_{mn} are both eigenfunctions with the eigenvalue $(m + n)(m + n + 2)$. If we change independent variables to a and $z \equiv z_s$, the operator $\mathcal{O}$ becomes

$$\mathcal{O} = -\frac{1}{a}[\partial_a a^2(1-a)\partial_a + \partial_z(1-z^2)\partial_z], \tag{3.4}$$

and this is the operator considered by Balachandran and Nuyts.[3] They showed that if $J_s^2 = -\partial_z(1-z^2)\partial_z$ is the differential operator representing the square of the s-channel angular momentum and J_t^2 and J_u^2 are similarly defined, then, with a and z as independent variables,

$$\mathcal{O} = J_s^2 + J_t^2 + J_u^2, \tag{3.5}$$

and evidently $\mathcal{O}$ commutes with J_s^2. The situation in the unequal-mass case is very different and is discussed briefly in Sec. 5. Some simple results on the operator $\mathcal{O}$ and its eigenfunctions are given in Appendix C.

Balachandran and Nuyts considered functions which were simultaneously eigenfunctions of J_s^2 and of $\mathcal{O}$. To within a numerical factor, their functions are

$$\begin{aligned}\mathcal{F}^J_{\mu+\nu}(a, z) \equiv P_J(z)a^J \\ \times {}_2F_1(\mu + \nu + J + 2, J - \mu - \nu; 2J + 2; a),\end{aligned} \tag{3.6}$$

which satisfy

$$\mathcal{O}\mathcal{F}^J_{\mu+\nu} = (\mu + \nu)(\mu + \nu + 2)\mathcal{F}^J_{\mu+\nu}, \tag{3.7}$$

$$J_s^2\mathcal{F}^J_{\mu+\nu} = J(J+1)\mathcal{F}^J_{\mu+\nu}. \tag{3.8}$$

They considered the case when μ and ν were integers m and n; and then, with $J \leq m + n$, the hypergeometric function in the expression for $\mathcal{F}^J_{m+n}$ becomes just

$$\frac{(m + n - J)!\,(2J + 1)!}{(m + n + J + 1)!} P^{(2J+1,0)}_{m+n-J}(1 - 2a), \tag{3.9}$$

the last factor being a Jacobi polynomial. The polynomials $\mathcal{F}^J_{m+n}$ were considered by Balachandran and Nuyts as the basis for an expansion of the scattering amplitude. It has also been observed[4] that these functions are particular cases of harmonic functions for the group $SU(3)$.[8]

The functions $F_{\mu\nu}(t, u)$ may be expanded "in partial waves," and we obtain (see Appendix A)

$$F_{\mu\nu}(t, u) = \Sigma_J (2J + 1) f^J_{\mu\nu}\mathcal{F}^J_{\mu+\nu}(a, z), \tag{3.10}$$

where the numerical coefficients $f^J_{\mu\nu}$ are given by

$$\begin{aligned}f^J_{\mu\nu} = \frac{(-\mu - \nu)_J(\nu + 1)_J}{(2J + 1)!} \\ \times {}_3F_2(\mu + 1, -J, -J; 1, -\nu - J; 1).\end{aligned} \tag{3.11}$$

It is amusing that the corresponding coefficients

$$e^J_{mn} = f^J_{-m-1,-n-1} \tag{3.12}$$

for the polynomials E_{mn} are related to the $3j$ symbols[9] of $O(3)$:

$$\begin{aligned}e^J_{mn} = \frac{(m + n + 2)_J}{(2J + 1)!}(-1)^n\left[\frac{(m + n + J + 1)!}{(m + n - J)!}\right]^{\frac{1}{2}} \\ \times \begin{pmatrix} J & \frac{1}{2}(m + n) & \frac{1}{2}(m + n) \\ 0 & \frac{1}{2}(m - n) & \frac{1}{2}(n - m)\end{pmatrix} \\ \text{for} \quad 0 \leq J \leq m + n, \\ = 0 \quad \text{for} \quad J > m + n.\end{aligned} \tag{3.13}$$

The presence of the "threshold factor" a^J in $\mathcal{F}^J_{\mu+\nu}$ has already been remarked upon by Balachandran and Nuyts.

The polynomials E_{mn} and F_{mn} may also be expanded in partial waves in the crossed channels. For example, if we write

$$t = 1 - b; \quad u = \tfrac{1}{2}b(1 - z_t), \tag{3.14}$$

we obtain

$$E_{mn}(t, u) = \sum_{J=0}^{n} (2J + 1)\tilde{e}^J_{mn}\mathcal{F}^J_{m+n}(b, z_t) \tag{3.15}$$

[8] M. A. B. Bég and H. Ruegg, J. Math. Phys. **6**, 677 (1965).

[9] Our notation and phase convention for the $3j$ symbols is that of M. Rotenberg, R. Bivins, N. Metropolis, and J. K. Wooten, Jr., *The 3j and 6j Symbols* (M.I.T. Press, Cambridge, Mass., 1959).

and

$$F_{mn}(t, u) = \sum_{J=0}^{n} (2J + 1)\bar{f}^{J}_{mn}\mathcal{F}^{J}_{m+n}(b, z_t), \quad (3.16)$$

where the coefficients are (see Appendix A)

$$\bar{e}^{J}_{mn} = \frac{(-1)^{m+n}n!(m + n + J + 1)!^2}{m!(n - J)!(m + n + 1)!(n + J + 1)!(2J + 1)!} \quad (3.17)$$

and what may be obtained by formally interchanging $m \leftrightarrow -m - 1$ and $n \leftrightarrow -n - 1$,

$$\bar{f}^{J}_{mn} = \frac{(-1)^{m+n}m!(n + J)!(m + n)!}{n!(n - J)!(m + n - J)!^2(2J + 1)!}. \quad (3.18)$$

One may again notice the threshold factors b^J in the terms of the expansion.

We conclude this section with a remark about the crossing matrix. Balachandran and Nuyts[2,3] derived a matrix which connects the coefficients in an expansion in terms of $\mathcal{F}^{J}_{m+n}(a, z_s)$ with those in an expansion in terms of $\mathcal{F}^{J}_{m+n}(b, z_t)$; the matrix is quite complicated, the elements involving ${}_4F_3$ functions of unit argument. If we perform a similar exercise, that is, compare the expansion

$$T(stu) = \sum_{mn} 2(m + n + 1)^2 A_{mn}E_{mn}(tu) \quad (3.19)$$

with the crossed-channel expansion

$$T(stu) = \sum_{MN} 2(M + N + 1)^2 B_{MN}E_{MN}(us), \quad (3.20)$$

the crossing matrix is

$$X^{mn}_{MN} = 2(m + n + 1)^2 \iint_{\Delta} dt\, du E_{mn}(tu)F_{MN}(us) \quad (3.21)$$

and

$$B_{MN} = \sum_{mn} X^{mn}_{MN}A_{mn}. \quad (3.22)$$

The integral in Eq. (3.21) may be evaluated (see Appendix B) and leads to the very simple expression

$$X^{mn}_{MN} = \delta_{m+n,M+N}(-1)^m\binom{M}{n}, \quad \text{for} \quad M \geq n,$$
$$= 0, \quad \text{otherwise}. \quad (3.23)$$

4. A NEUMANN FORMULA: FUNCTIONS OF THE SECOND KIND

The Neumann formula for Legendre functions is

$$Q_J(z) = \frac{1}{2}\int_{-1}^{+1} dz'(z - z')^{-1}P_J(z'), \quad (4.1)$$

which is valid for integer values of J. Here, of course, $Q_J(z)$ is the Legendre function of the second kind, which satisfies the same differential equation as does $P_J(z)$.

Let us introduce, by analogy, functions G_{mn} defined for m and n integers by

$$G_{mn}(t, u) = \iint_{\Delta} dt'\, du'(t - t')^{-1}(u - u')^{-1}F_{mn}(t', u'). \quad (4.2)$$

We may use the Rodrigues formula, Eq. (2.5), for F_{mn}, and then integrate by parts m times in t' and n times in u'. This results in

$$\begin{aligned} G_{mn}(tu) &= \frac{(-1)^{m+n}}{m!\,n!}\iint_{\Delta} dt'\, du' t'^m u'^n(1 - t' - u')^{m+n} \\ &\quad \times \left(\frac{\partial}{\partial t'}\right)^m\left(\frac{\partial}{\partial u'}\right)^n\left[\frac{1}{(t - t')(u - u')}\right] \\ &= (-1)^{m+n}t^{-m-1}u^{-n-1} \\ &\quad \times \iint_{\Delta} dt'\, du' t'^m u'^n(1 - t' - u')^{m+n} \\ &\quad \times \left(1 - \frac{t'}{t}\right)^{-m-1}\left(1 - \frac{u'}{u}\right)^{-n-1}. \end{aligned} \quad (4.3)$$

The integral may then be recognized, apart from a numerical factor, as another of Appell's generalizations of the hypergeometric function. Explicitly we have

$$\begin{aligned} \hat{G}_{mn}(t, u) &\equiv (-1)^{m+n}G_{mn}(t, u) \\ &= \frac{\Gamma(m + 1)\Gamma(n + 1)\Gamma(m + n + 1)}{\Gamma(2m + 2n + 3)}\, t^{-m-1}u^{-n-1} \\ &\quad \times F_3\Big(m + 1, n + 1; m + 1, n + 1; \\ &\qquad 2m + 2n + 3; \frac{1}{t}, \frac{1}{u}\Big). \end{aligned} \quad (4.4)$$

We shall continue to use this definition of $\hat{G}$ even when m and n become arbitrary complex numbers μ and ν.

From the differential equation satisfied by F_3, it is easy to show that $\hat{G}_{\mu\nu}$ satisfies the *same pair of differential equations* as does $F_{\mu\nu}$. The function $G_{\mu\nu}$ plays a role relative to $F_{\mu\nu}$ analogous to that which Q_ν plays relative to P_ν.

An integral representation for $\hat{G}_{\mu\nu}$ [the generalization of Eq. (4.3)] is [$\mathrm{Re}\,\mu > -1$, $\mathrm{Re}\,\nu > -1$, $\mathrm{Re}\,(\mu + \nu) > -1$]

$$\hat{G}_{\mu\nu}(t, u) = \iint_{\Delta} dt'\, du' t'^{\mu}u'^{\nu}(1 - t' - u')^{\mu+\nu} \times (t - t')^{-\mu-1}(u - u')^{-\nu-1}, \quad (4.5)$$

which is to be compared with (Re $\nu > -1$)

$$Q_\nu(z) = \frac{1}{2}\int_{-1}^{+1} dz' \left(\frac{1-z'}{2}\right)^\nu \left(\frac{1+z'}{2}\right)^\nu \left(\frac{z-z'}{2}\right)^{-\nu-1}. \tag{4.6}$$

If $0 > \mathrm{Re}\,\mu > -1, 0 > \mathrm{Re}\,\nu > -1$, we have

$$F_{\mu\nu}(t,u) = [\Gamma(\mu)\Gamma(-\mu-1)\Gamma(\nu)\Gamma(-\nu-1)]^{-1} \times \int_0^1 dx \int_0^1 dy x^\mu y^\nu (1-x)^{-\mu-1} \times (1-y)^{-\nu-1}(1-tx-uy)^{\mu+\nu}, \tag{4.7}$$

which is the analog of the representation, valid for $0 > \mathrm{Re}\,\nu > -1$,

$$P_\nu(z) = [\Gamma(\nu)\Gamma(-\nu-1)]^{-1} \times \int_0^1 dx x^\nu (1-x)^{-\nu-1}[1-\tfrac{1}{2}(1-z)x]^\nu. \tag{4.8}$$

The above integral representations for $F_{\mu\nu}$ and $\hat{G}_{\mu\nu}$ may be used to deduce that both functions have branch-cut singularities on the lines $t = 1$, $u = 1$, $t + u = 1$, $t = \infty$, $u = \infty$; and that $\hat{G}_{\mu\nu}$ has, in addition, branch-cut singularities on the lines $t = 0$, $u = 0$. Just as the branch cut from $z = -\infty$ to $z = -1$ of $Q_\nu(z)$ is a trivial logarithmic branch cut, so that $(z-1)^{\nu+1}Q_\nu(z)$ has a cut only from $z = -1$ to $z = +1$, so also the cuts from $t = 1$ to $t = \infty$ and from $u = 1$ to $u = \infty$ of $\hat{G}_{\mu\nu}$ are trivial, removable, logarithmic singularities. The function $(t-1)^{\mu+1} \times (u-1)^{\nu+1}\hat{G}_{\mu\nu}(t,u)$ has singularities only on the lines $t = 0$, $u = 0$, $t + u = 1$, i.e., on the boundaries of Δ.

The asymptotic behavior for large z of $Q_\nu(z)$ follows directly from the expression in terms of the hypergeometric function

$$Q_\nu(z) = 2^{-\nu-1}\frac{\Gamma(\nu+1)}{\Gamma(\nu+\frac{3}{2})} z^{-\nu-1} \times F(\tfrac{1}{2}\nu+1, \tfrac{1}{2}\nu+\tfrac{1}{2}; \nu+\tfrac{3}{2}; z^{-2}), \tag{4.9}$$

so that

$$Q_\nu(z) \sim 2^{-\nu-1}\frac{\Gamma(\nu+1)}{\Gamma(\nu+\frac{3}{2})} z^{-\nu-1}, \quad \text{as} \quad z \to \infty. \tag{4.10}$$

In a similar fashion, use may be made of Eq. (4.4) to deduce that

$$\hat{G}_{\mu\nu} \sim \frac{\Gamma(\mu+1)\Gamma(\nu+1)\Gamma(\mu+\nu+1)}{\Gamma(2\mu+2\nu+3)} t^{-\mu-1}u^{-\nu-1} \times {}_2F_1\left(\mu+1, \mu+1; 2\mu+2\nu+3; \frac{1}{t}\right), \quad \text{as} \quad u \to \infty,$$

$$\sim \frac{\Gamma(\mu+1)\Gamma(\nu+1)\Gamma(\mu+\nu+1)}{\Gamma(2\mu+2\nu+3)} t^{-\mu-1}u^{-\nu-1}, \quad \text{as both} \quad t \to \infty, u \to \infty. \tag{4.11}$$

For $P_\nu(z)$ the behavior for large z is, of course, also well known. It is instructive to observe that since $P_\nu(z)$, $Q_\nu(z)$, and $Q_{-\nu-1}(z)$ all satisfy the same second-degree differential equation, there must exist a linear relation with constant coefficients between the three functions. From this alone, with the behavior $z^{-\nu-1}$ for $Q_\nu(z)$, it follows that $P_\nu(z)$ has an expression as a piece with behavior $z^{-\nu-1}$ and a piece with behavior z^ν. The situation with $F_{\mu\nu}$ is more complicated, even though $F_{\mu\nu}$ and $\hat{G}_{\mu\nu}$ satisfy the same pair of partial-differential equations.

We are able though to give a derivation of the behavior for large z at fixed a of

$$F_{\mu\nu}\left[a\left(\frac{1-z}{2}\right), a\left(\frac{1+z}{2}\right)\right].$$

In this limit, the differential equations for $F_{\mu\nu}$ become

$$\left[(1-a)\left(\frac{\partial}{\partial a} - \frac{z}{a}\frac{\partial}{\partial z}\right) + (\mu+\nu-1)\right] \times \left(\frac{\partial}{\partial a} - \frac{z}{a}\frac{\partial}{\partial z}\right)F_{\mu\nu} = 0$$

and

$$\left\{\frac{1}{a}\left[\frac{\partial}{\partial a}a^2(1-a)\frac{\partial}{\partial a} - \alpha(\alpha+1)\right] + (\mu+\nu)(\mu+\nu+2)\right\}F_{\mu\nu} = 0, \tag{4.12}$$

to which the only solutions of the form

$$F \sim f(a)z^\alpha \tag{4.13}$$

are

$$F \sim (az)^\alpha, \tag{4.14}$$

with

$$\alpha = \mu+\nu \quad \text{or} \quad \alpha = -\mu-\nu-2.$$

The same conclusion may also be reached on the basis of the asymptotic expansions for the F_2 function given by Minton,[10] namely, that $F_{\mu\nu}$ behaves in this limit as the sum of two terms with behavior like $(az)^{\mu+\nu}$ and $(az)^{-\mu-\nu-2}$. More explicitly, we have

$$F_{\mu\nu} + \frac{\sin\pi\mu \sin\pi\nu}{2\pi^2\cos\pi(\mu+\nu)} G_{\mu\nu} \sim \text{const}\,(az)^{\mu+\nu}. \tag{4.15}$$

It may be worth recording that the function $H_{\mu\nu}(t,u)$, defined by

$$H_{\mu\nu}(t,u) = t^\mu u^\nu \sum_{pq} \frac{t^{-p}u^{-q}}{(p+q)!} \frac{(-\mu)_p^2(-\nu)_q^2}{(-2\mu)_p(-2\nu)_q}, \tag{4.16}$$

is a solution of the same system of equations as $F_{\mu\nu}$, which behaves asymptotically like $t^\mu u^\nu$, i.e., $(az)^{\mu+\nu}$.

[10] B. M. Minton, Proc. Cambridge Phil. Soc. **64**, 1055 (1968).

5. SOME PROBLEMS ASSOCIATED WITH UNEQUAL MASSES

The kinematics of the scattering process for the general case of unequal masses are, of course, more complicated than for the special case of equal masses we have considered. Unfortunately, these complications lead to difficulties which we have not fully overcome. If we consider the process for which the s channel has particles 1 and 2 with masses m_1 and m_2 initially and particles 3 and 4 with masses m_3 and m_4 in the final state, then the magnitude of the s-channel initial-state center-of-mass 3-momentum p_s is given in terms of s by

$$p_s^2 = [s - (m_1 + m_2)^2][s - (m_1 - m_2)^2]/4s \quad (5.1)$$

and similarly p_s', the final-state s-channel 3-momentum, is given by

$$p_s'^2 = [s - (m_3 + m_4)^2][s - (m_3 - m_4)^2]/4s. \quad (5.2)$$

The boundary of the physical region is given by

$$\begin{aligned}\Phi(stu) \equiv stu &- s\tfrac{1}{4}(m_1^2 + m_2^2 - m_3^2 - m_4^2)^2 \\ &- t\tfrac{1}{4}(m_1^2 + m_3^2 - m_2^2 - m_4^2)^2 \\ &- u\tfrac{1}{4}(m_1^2 + m_4^2 - m_2^2 - m_3^2)^2 \\ &+ \tfrac{1}{4}(m_1^2 + m_2^2 - m_3^2 - m_4^2) \\ &\times (m_1^2 + m_3^2 - m_2^2 - m_4^2) \\ &\times (m_1^2 + m_4^2 - m_2^2 - m_3^2) \\ = 0. \quad (5.3)\end{aligned}$$

In terms of s-channel variables, we may write

$$\Phi(stu) = 4sp_s^2p_s'^2(1 - z_s^2). \quad (5.4)$$

In the equal-mass case, we have simply

$$\Phi = stu, \quad (5.5)$$

and the cubic curve $\Phi = 0$ reduces to the three lines $s = 0$, $t = 0$, $u = 0$, which provide us in that case with the edges of the triangle. Along two of those edges ($t = 0$ and $u = 0$) we have $z_s^2 = 1$; this is important for the simple connection between $F_{mn}(t, u)$ and $\mathcal{F}^J_{m+n}$.

In the general-mass case, the triangle in the middle of the Mandelstam s–t–u plot may, of course, still be identified with the region Δ on which the two-variable expansion is to be made. But its boundary is now no longer part of the boundary of the physical region, and many of the attractive features of the results of Sec. 3 become obscured. In particular, the functions $F_{\mu\nu}(t, u)$ still satisfy the differential equations Eq. (3.1), and so are still eigenfunctions of the differential operator $\mathcal{O}$, as given by Eq. (3.2). But this is no longer equal to the form given in Eq. (3.4), nor to any simple generalization of that expression. Nor is there any obvious generalization of Eq. (3.5), expressing the connection between $\mathcal{O}$ and the angular-momentum operators.

Indeed, we still have (with z_s and, say, s as independent variables)

$$J_s^2 = \frac{-\partial}{\partial z_s}(1 - z_s^2)\frac{\partial}{\partial z_s}. \quad (5.6)$$

If t and u are taken to be the independent variables, with $s = \Sigma - t - u$ ($\Sigma \equiv m_1^2 + m_2^2 + m_3^2 + m_4^2$) always understood, this becomes

$$J_s^2 = -\frac{1}{s}\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial u}\right)\Phi\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial u}\right). \quad (5.7)$$

Still with t and u as independent variables, the differential operators for the total angular momentum in the t and u channels are

$$J_t^2 = -\frac{1}{t}\frac{\partial}{\partial u}\Phi\frac{\partial}{\partial u}$$

and

$$J_u^2 = -\frac{1}{u}\frac{\partial}{\partial t}\Phi\frac{\partial}{\partial t}, \quad (5.8)$$

respectively. It is no longer true that $\mathcal{O} = J_s^2 + J_t^2 + J_u^2$; furthermore,

$$[J_s^2, J_t^2 + J_u^2] \neq 0, \quad (5.9)$$

unlike the situation in the equal-mass case.

Of course, from the point of view simply of making an expansion of the scattering amplitude in terms of some complete set of functions of two variables, we are completely at liberty to choose some other variables than s, t, and u, and to define a region Δ in terms of them and so to consider polynomials in the new variables, (bi)-orthogonal on Δ. One such choice of new variables $\bar{s}$, $\bar{t}$, and $\bar{u}$ has been suggested by Balachandran *et al.*,[4] which has an elegance and symmetry that commends it for attention. They choose for $\bar{s}$, $\bar{t}$, and $\bar{u}$ the eigenvalues of a Gram matrix, which are the roots of the cubic equation

$$x^3 - \Sigma x^2 + \Psi(s, t, u)x - \Phi(s, t, u) = 0, \quad (5.10)$$

where

$$\begin{aligned}\Sigma &= m_1^2 + m_2^2 + m_3^2 + m_4^2, \\ \Psi &\equiv st - \tfrac{1}{4}(m_1^2 + m_4^2 - m_2^2 - m_3^2) \\ &\quad + tu - \tfrac{1}{4}(m_1^2 + m_2^2 - m_3^2 - m_4^2) \\ &\quad + us - \tfrac{1}{4}(m_1^2 + m_3^2 - m_2^2 - m_4^2), \quad (5.11)\end{aligned}$$

and Φ has been defined in Eq. (5.3). Then clearly

$$\bar{s} + \bar{t} + \bar{u} = \Sigma, \quad (5.12)$$

and the boundary of the physical region, $\Phi = 0$, is now given by

$$\bar{s}\bar{t}\bar{u} = 0. \tag{5.13}$$

In the equal-mass case, $\bar{s}$, $\bar{t}$, $\bar{u}$ can be identified with s, t, and u. If the region Δ is now taken to be the triangle defined by the lines $\bar{s} = 0$, $\bar{t} = 0$, $\bar{u} = 0$, the expansion suggested is in terms of $E_{mn}(\bar{t}, \bar{u})$. However, the cubic transformation between s, t, u and $\bar{s}$, $\bar{t}$, $\bar{u}$ spoils any possibility of connecting the differential operator $\mathcal{O}$ [of which $E_{mn}(\bar{t}, \bar{u})$ is an eigenfunction] with any simply identifiable differential operator in the s, t, u variables. It also prevents any application of the Neumann formula along the lines indicated in the next section. From the point of view simply of parameterizing a function defined on a region bounded by a cubic curve like $\Phi = 0$, it may still be useful to use this mapping of that region onto a triangle. For example, the Dalitz plot of a decay process with three particles in the final state has just such a boundary and so may be mapped onto a triangle, there to be parameterized by the coefficients in an expansion in terms of $E_{mn}(\bar{t}, \bar{u})$.

We think, on balance, it is preferable to keep as the region the triangle bounded by $s = 0$, $t = 0$, and $u = 0$, even at the expense of losing the relatively simple connection with the angular-momentum expansion.

6. AN ANALOG OF THE SOMMERFELD-WATSON TRANSFORMATION

The scattering amplitude $T(stu)$ may be expanded in terms of the polynomials $E_{mn}(t, u)$, as in Eq. (2.11). We would like to use such an expansion to discuss the asymptotic behavior of the scattering amplitude. A program which immediately suggests itself for this examination is indicated by the familiar procedure in the Regge analysis[11] based on the partial-wave expansion which we shall recapitulate briefly. What is done in that case is to take the expansions

$$T^{\pm}(stu) = \Sigma_J (2J + 1) P_J(z) A_J^{\pm}(s) \tag{6.1}$$

of amplitudes for which

$$T(stu) = \tfrac{1}{2}\{T^{+}(stu) + T^{+}(sut) + T^{-}(stu) - T^{-}(sut)\} \tag{6.2}$$

and for them to define analytic continuations $A^{\pm}(J, s)$ in J of the signatured partial-wave amplitudes $A_J^{\pm}(s)$. It is then possible to write the partial-wave sums as

$$T^{\pm}(stu) = \frac{1}{2i} \oint_C dJ (2J + 1) P_J(-z) A^{\pm}(J, s) \csc \pi J, \tag{6.3}$$

the integration being over a contour C in the J plane. The contour of integration is then distorted to yield

$$T^{\pm}(stu) = \frac{1}{2i} \oint_{C'} dJ (2J + 1) P_J(-z) A^{\pm}(J, s) \csc \pi J + \pi \sum_i (2\alpha_i^{\pm} + 1) P_{\alpha_i^{\pm}}(-z) \csc \pi\alpha_i^{\pm} \beta_i^{\pm}(s), \tag{6.4}$$

expressing $T^{\pm}$ as a sum over contributions from poles of $A^{\pm}(J, s)$ at positions $J = \alpha_i^{\pm}(s)$ and with residues $\beta_i^{\pm}(s)$, and a background integral. The behavior of $T^{\pm}(stu)$ for large values of z is now seen to be dominated by the contribution

$$\pi^{\frac{1}{2}}(2\alpha + 1)\Gamma(\alpha + \tfrac{1}{2})[\Gamma(\alpha + 1)]^{-1} \csc \pi\alpha\beta(s)(-2z)^{\alpha}$$

of the pole at $J = \alpha$ lying furthest to the right.

The crucially important step in this procedure is the construction of the functions $A^{\pm}(J, s)$ defined in the complex J plane, which interpolate the physical values $A_J^{\pm}(s)$ attained at $J =$ even/odd integer. And this is done by using the representation of Froissart and Gribov. The starting point is the dispersion relation in z at fixed s for $T(stu) \equiv T(sz)$, assumed to hold with, say, N subtractions, namely,

$$T(sz) = \frac{1}{\pi} \int_{-\infty}^{z_L} \frac{dz'}{z' - z} \left(\frac{z}{z'}\right)^N \sigma_L(s, z') + \frac{1}{\pi} \int_{z_R}^{\infty} \frac{dz'}{z' - z} \left(\frac{z}{z'}\right)^N \sigma_R(s, z') + \text{polynomial of degree } N - 1 \text{ in } z. \tag{6.5}$$

From the Neumann formula, Eq. (4.1), it follows that if

$$A^{\pm}(J, s) = \frac{1}{\pi} \int_{z_0}^{\infty} dz \sigma_{\pm}(s, z) Q_J(z), \tag{6.6}$$

then $A^{+}(A^{-})$ agrees with $A_J(s)$ for J an even (odd) integer greater than N, provided that

$$\sigma_{\pm}(s, z) = \sigma_R(s, z) \mp \sigma_L(s, -z) \tag{6.7}$$

and z_0 is the smaller of z_R and $-z_L$. Furthermore, from the known behavior of the Legendre functions Q_J and P_J for large J, it follows that this definition of $A^{\pm}(Js)$ is the unique interpolation consistent with the passage from Eq. (6.3) to Eq. (6.4), which is possible only if there is no contribution from the arc of the contour at infinity.

The hope that the only singularities in the J plane of the partial-wave amplitude exposed by the distortion of the contour would be simple poles (as can be proved to be the case for potential scattering with a wide class of potentials) was shown to be in vain. The

[11] See, for example, E. J. Squires, *Complex Angular Momenta and Particle Physics* (W. A. Benjamin, Inc., New York, 1963).

important distinguishing feature of the crossing-symmetric theory, which has no counterpart in potential theory, is the presence of the third double-spectral function (in the sense of the Mandelstam representation). This leads directly, as was first discussed by Gribov and Pomeranchuk,[12] to a pole in $A^+(A^-)$ at each negative odd (even) integer value of J. Such fixed poles in the J plane are inconsistent with the analytic continuation in J of the unitarity equation if the partial-wave amplitudes are meromorphic functions of J. However, as was shown by Mandelstam,[13] there is a mechanism for the generation of branch points in the J plane. Furthermore, these branch points move in such a way as to break the chain of argument which would otherwise imply, as shown by Gribov and Pomeranchuk, that the fixed poles at negative integers generated essential singularities at these points. A less agreeable feature of the Mandelstam mechanism for the generation of moving branch points is, ironically, that it also invalidates the assumption of the Mandelstam representation itself. It is not the analytic structure embodied in the double-dispersion relation which is found wanting, but the asymptotic behavior of the amplitude. In order to write the Mandelstam representation, it is necessary that behavior for large t of the amplitude $T(stu)$ shall be bounded by some power of t, say t^N, *uniformly* in s, so that we can choose N independent of s. What Mandelstam himself observed was that the moving branch points can and do move arbitrarily far to the right in the J plane, and so their contributions to the asymptotic behavior of $T(stu)$ cannot be uniformly bounded by a power t^N.

This unboundedness of the number of subtractions needed for the dispersion relations precludes an application of the analog of the Regge approach which we wish to develop to the whole scattering amplitude. Rather, it can only be applied at best to some portion of the scattering amplitude (say, a set of Feynman diagrams) which *does* satisfy the Mandelstam representation. Henceforth $T(stu)$ will denote some such portion of the amplitude. In fact, we feel obliged to limit still further the function which we can discuss, and consider what we will call $T^{(3)}(stu)$, which is that part of $T(stu)$ which arises from integrations over the "third" or t–u double-spectral function. It should be emphasized that it is with such contributions to the full amplitude that all of the special features (fixed poles, branch cuts, and the like) of the crossing-symmetric theory are associated.

[12] V. N. Gribov and I. Ya. Pomeranchuk, Phys. Letters **2**, 239 (1962).
[13] S. Mandelstam, Nuovo Cimento **30**, 1113, 1127, 1148 (1963).

Explicitly, we suppose that $T^{(3)}$ satisfies the double-dispersion relation

$$T^{(3)}(stu) = \frac{1}{\pi^2}\int \frac{dt'}{t'-t}\int \frac{du'}{u'-u}\,\rho^{(3)}(t',u'), \quad (6.8)$$

the support of $\rho^{(3)}$ lying entirely in $t' > 0$, $u' > 0$. The generalization to the less restrictive situation where a finite number of subtractions is needed presents no problems. Our real concern is to estimate the behavior of such an integral as $z \to \infty$ with s fixed.

We make the expansion

$$T^{(3)}(stu) = \sum_{mn} 2(m+n+1)^2 A^{(3)}_{mn} E_{mn}(t,u), \quad (6.9)$$

so that we have

$$A^{(3)}_{mn} = \iint_{\Delta} dt\,du F_{mn}(t,u)\frac{1}{\pi^2}\int \frac{dt'}{t'-t}\int \frac{du'}{u'-u}\,\rho^{(3)}(t',u'). \quad (6.10)$$

The convergence of Eq. (6.8) will be assumed such as to permit the interchange of the orders of integration in Eq. (6.10), and we then obtain

$$A^{(3)}_{mn} = \frac{1}{\pi^2}\int dt'\int du'\rho^{(3)}(t'u')G_{mn}(t'u'). \quad (6.11)$$

A little care is needed if the support of $\rho^{(3)}$ intersects Δ, but the $i\epsilon$ prescription for the denominators selects unambiguously the appropriate branch of G_{mn}; one must replace $G_{mn}(t'u')$ by $G_{mn}(t'-i\epsilon, u'-i\epsilon')$. We can now extend the definition to complex parameters by

$$\hat{A}(\mu,\nu) = \frac{1}{\pi^2}\int dt'\int du'\rho^{(3)}(t'u')\hat{G}_{\mu\nu}(t'-i\epsilon, u'-i\epsilon') \quad (6.12)$$

with the definition

$$\hat{A}(m,n) = (-)^{m+n}A^{(3)}_{mn} \quad (6.13)$$

for m and n integers.

The double sum of Eq. (6.9) can now be written as a contour integral:

$$T^{(3)}(stu) = -\tfrac{1}{4}\int_{\Gamma_1} d\mu\int_{\Gamma_2} d\nu 2(\mu+\nu+1)^2 \times \hat{A}(\mu,\nu)E_{\mu\nu}(tu)\csc\pi\mu\csc\pi\nu. \quad (6.14)$$

The contour $\Gamma_1(\Gamma_2)$ encircles each of the points $\mu = 0, 1, 2, \cdots$ ($\nu = 0, 1, 2, \cdots$) once in a positive sense, and no other singularities of the integrand. The integrand is meromorphic in the region for which the integral in Eq. (6.12) defining $\hat{A}(\mu\nu)$ converges, which certainly includes $\mathrm{Re}\,\mu > 0 \otimes \mathrm{Re}\,\nu > 0$ from the hypothesis of no subtractions being needed in

the double-dispersion relation. Its only poles in this latter region are those encircled by the contours at which the residues generate just the terms of the double sum.

The distortion of the contours of integration entails a discussion of a number of points, none of which we have been able fully to elucidate. Firstly, we shall be concerned with contributions from arcs of the contours at infinity, and this leads to the question of the asymptotic behavior of the integrand for $|\mu| \to \infty, |\nu| \to \infty$. A treatment of this problem, which essentially means the treatment of the behavior of $E_{\mu\nu}$ and of $\hat{G}_{\mu\nu}$ in this same limit, is inevitably extremely difficult. The analogous problem in the Regge program is solved because one knows the behavior of P_J and of Q_J for $|J| \to \infty$. More generally, thanks to the heroic efforts of Watson,[14] we know the asymptotic behavior of Gaussian hypergeometric functions of large parameter. For the generalized hypergeometric functions of the present discussion, we know of no treatment of this problem whatever.

We might trust physicists' luck and assume that the large μ, ν behavior is satisfactory. This is tantamount to the assumption that Eq. (6.12) represents the "correct" analytic continuation in μ, ν. There then arises the problem of singularities of the integrand $\hat{A}(\mu\nu) \csc \pi\mu \csc \pi\nu$ other than the ones which lead to the terms of the double sum. In particular, there are the poles of the cosecant functions when μ or ν are negative integers. Again, one might hope that these could be exploited in such a way that the $E_{\mu\nu}$ in the integrand could be replaced by that part of $E_{\mu\nu}$ which behaves for large z like $z^{\mu+\nu}$, which as we have seen [Eq. (4.15)] is

$$-\sin \pi\mu\nu \sin \pi\nu(2\pi^2)^{-1} \sec \pi(\mu + \nu)G_{-\mu-1,-\nu-1}.$$

This would be analogous to the method introduced by Mandelstam,[15] which is used in "pushing the background contour to the left" and replaces P_J by $-\pi^{-1} \tan \pi J Q_{-J-1}$.

There would then remain the singularities of the integrand which came from the analytic continuation of $\hat{A}(\mu\nu)$, and these one could presume, like the Regge poles, etc., of the analytically continued partial-wave amplitudes $A^{\pm}(J, s)$, were dynamical rather than kinematical in origin. One could also hope that $\hat{A}(\mu\nu)$ was meromorphic, unlike $A^{\pm}(J, s)$ with its branch cuts. In any case, there would be singularity surfaces in the (μ, ν) space. From our hypothesis that there were no subtractions in the Mandelstam representation, it certainly follows that the region $\mathrm{Re}\,(\mu + \nu) > 0$ is not intersected by such singularity surfaces. Indeed, there will in general be a complex number α_0 with $\mathrm{Re}\,\alpha_0 < 0$ for which the surface $\mu + \nu = \alpha_0$ osculates a singularity surface of $\hat{A}(\mu\nu)$ and for which any other osculating surface $\mu + \nu = \alpha$ will have $\mathrm{Re}\,\alpha < \mathrm{Re}\,\alpha_0$. Distortion of the contours Γ_1 and Γ_2 of integration in Eq. (6.14) would then lead to an integration over some two-dimensional surface in (μ, ν) which lies entirely inside the region $\mathrm{Re}\,(\mu + \nu) \leq \mathrm{Re}\,\alpha_0$; the integrand would be proportional to $G_{-\mu-1,-\nu-1}(t, u)$, and so behave at fixed a and large z like $(az)^{\mu+\nu}$, and the implied behavior in this limit for $T^{(3)}(stu)$ is then of the form

$$T^{(3)}(stu) \sim \int_{\Gamma} d\alpha w(\alpha)(az)^{\alpha}, \qquad (6.15)$$

where Γ is a contour which lies entirely in $\mathrm{Re}\,\alpha < \mathrm{Re}\,\alpha_0$ and which has $\alpha = \alpha_0$ as an end point. To within a logarithmic variation, this means that we have $T^{(3)} \sim (az)^{\alpha_0}$, and the important feature we wish to emphasise is that α_0 is *fixed* and is some complex number determined by the dynamics [through $\rho^{(3)}$].

The above analysis would suggest that, if all of our many assumptions are in fact justified, the piece $T^{(3)}$ of the amplitude which we have discussed behaves at large z like $(az)^{\alpha_0}$, with α_0 fixed (with some additional logarithmic factors which are also independent of a). Put in terms of the J plane, this means we have suggested a mechanism for the generation of *fixed* branch points; the contour Γ in Eq. (6.15) is, of course, just the branch cut in the J plane. It would seem to be an open question whether such fixed branch cuts are present in the full amplitude.

ACKNOWLEDGMENT

One of us (B. M. M.) would like to thank the States of Jersey for a grant.

APPENDIX A: THE PARTIAL-WAVE EXPANSION OF $F_{\mu\nu}$

The function $F_{\mu\nu}(t, u)$ is a solution of

$$\mathcal{O}F_{\mu\nu} = (\mu + \nu)(\mu + \nu + 2)F_{\mu\nu}, \qquad (A1)$$

which is regular at $t = 0$ and at $u = 0$, i.e., when $z = \pm 1$ and when $a = 0$. This differential equation is separable in z and a, and the solutions of the form $Z(z)A(a)$ with the stated regularity conditions are just $\mathcal{F}^J_{\mu+\nu}(a, z)$. It follows that, for some coefficients $f^J_{\mu\nu}$ to be determined, we must have

$$F_{\mu\nu}(t, u) = \Sigma_J(2J + 1)f^J_{\mu\nu}\mathcal{F}^J_{\mu+\nu}(a, z). \qquad (A2)$$

[14] G. N. Watson, Trans. Cambridge Phil. Soc. **22**, 2771 (1918).
[15] S. Mandelstam, Ann. Phys. (N.Y.) **19**, 254 (1962).

Then, for each integer value of J,

$$f^J_{\mu\nu} a^J {}_2F_1(\mu+\nu+J+2, J-\mu-\nu; 2J+2: a) = \tfrac{1}{2}\int_{-1}^{+1} dz P_J(z) F_{\mu\nu}(t,u), \quad \text{(A3)}$$

so that

$$f^J_{\mu\nu} = \lim_{a\to 0} a^{-J} \tfrac{1}{2}\int_{-1}^{+1} dz P_J(z) F_{\mu\nu}(t,u). \quad \text{(A4)}$$

Now from the defining series expansion, Eq. (2.7), for $F_{\mu\nu}$,

$$F_{\mu\nu}(t,u) = \sum_{p,q} \frac{(-\mu-\nu)_{p+q}(\mu+1)_p(\nu+1)_q}{p!^2\, q!^2} \times a^{p+q}\left(\frac{1-z}{2}\right)^p\left(\frac{1+z}{2}\right)^q. \quad \text{(A5)}$$

The integral

$$\tfrac{1}{2}\int_{-1}^{+1} dz P_J(z) F_{\mu\nu}(t,u) = \sum_{pq} \frac{(-\mu-\nu)_{p+q}(\mu+1)_p(\nu+1)_q}{p!^2\, q!^2} \times a^{p+q}\,\tfrac{1}{2}\int_{-1}^{+1} dz P_J(z)\left(\frac{1-z}{2}\right)^p\left(\frac{1+z}{2}\right)^q \quad \text{(A6)}$$

is thus expressed as a sum over terms which vanish unless $p+q \geq J$, and it is easy to see that

$$f^J_{\mu\nu} = \sum_{\substack{p,q \\ p+q=J}} \frac{(-\mu-\nu)_{p+q}(\mu+1)_p(\nu+1)_q}{p!^2\, q!^2} \times \tfrac{1}{2}\int_{-1}^{+1} dz P_J(z)\left(\frac{1-z}{2}\right)^p\left(\frac{1+z}{2}\right)^q. \quad \text{(A7)}$$

We need the simple integral

$$\tfrac{1}{2}\int_{-1}^{+1} dz P_J(z)\left(\frac{1-z}{2}\right)^p\left(\frac{1+z}{2}\right)^{J-p} = (-1)^p \frac{J!^2}{(2J+1)!}. \quad \text{(A8)}$$

If this result is substituted in the sum for $f^J_{\mu\nu}$, Eq. (3.11) may be derived directly.

To derive the coefficients $\bar{f}^J_{mn}$ of the crossed-channel expansion of Eq. (3.16), we start from

$$F_{mn}(t,u) = \sum_{\substack{p,q \\ p+q\leq m+n}} \frac{(-m-n)_{p+q}(m+1)_p(n+1)_q}{p!^2\, q!^2} \times (1-b)^p b^q \left(\frac{1-z_t}{2}\right)^q. \quad \text{(A9)}$$

We shall need

$$\tfrac{1}{2}\int_{-1}^{+1} dz_t P_J(z_t)\left(\frac{1-z_t}{2}\right)^q = \frac{(-1)^J q!^2}{(J+q+1)!\,(q-J)!}, \quad \text{for } q \geq J, \qquad = 0, \quad \text{for } q < J. \quad \text{(A10)}$$

This leads to

$$\tfrac{1}{2}\int_{-1}^{+1} dz_t P_J(z_t) F_{mn}(t,u) = \sum_{\substack{p,q \\ p+q\leq m+n \\ q\geq J}} \frac{(-m-n)_{p+q}(m+1)_p(n+1)_q}{p!^2\, q!^2} \times (1-b)^p b^q \frac{(-1)^J q!^2}{(J+q+1)!\,(q-J)!}, \quad \text{(A11)}$$

which vanishes for $J > m+n$. For $J \leq m+n$, after expanding the factor $(1-b)^p$ and some straightforward manipulation, there results

$$(-b)^J \sum_{r=0}^{m+n-J} \sum_{s=0}^{m+n-J-r} \frac{(-1)^s b^{r+s}(n+1)_{r+J}}{(2J+r+1)!\, r!\, s!} \times \sum_{p=s}^{m+n-J-r} \frac{(-m-n)_{p+J+r}(m+1)_p}{p!\,(p-s)!}. \quad \text{(A12)}$$

The summation on p may be performed using Vandermonde's theorem, and then, after further simplification, we obtain

$$(-b)^J \sum_{k=0}^{m+n-J} (-b)^k (-m-n)_{J+k}(-m)_{m+n-J-k} \times \frac{(n+1)_J (m+k)!}{m!\,(2J+1)!\,k!\,(m+n-J)!} \times {}_3F_2\left(\begin{matrix} n+J+1, -k, J-m-n \\ -m-, 2J+2 \end{matrix}; 1\right), \quad \text{(A13)}$$

in which the ${}_3F_2$ series is Saalschutzian and so is given by

$$(J+1-n)_k (J+m+n+2)_k / (2J+2)_k (m+1)_k.$$

A few further manipulations then lead to

$$\tfrac{1}{2}\int_{-1}^{+1} dz_t P_J(z_t) F_{mn}(t,u) = \bar{f}^J_{mn} b^J {}_2F_1(m+n+J+2, J-m-n; 2J+2; b), \quad \text{(A14)}$$

with $\bar{f}^J_{mn} = 0$ for $J > n$, and given by Eq. (3.18) otherwise. The derivation of Eq. (3.17) for $\bar{e}^J_{mn}$ proceeds similarly.

APPENDIX B: THE CROSSING MATRIX

As stated in Eq. (3.21), the crossing matrix is given by

$$X^{mn}_{MN} = 2(M+N+1)^2 \iint_\Delta dt\, du E_{mn}(t,u) F_{MN}(u,s). \quad \text{(B1)}$$

This may be rewritten as

$$X^{mn}_{MN} = 2(M+N+1)^2 \int_0^1 b\,db \frac{1}{2}\int_{-1}^{+1} dz_t \times \sum_{J=0}^{n}(2J+1)e^{-J}_{mn}\mathcal{F}^J_{m+n}(b,z_t) \times \sum_{J'}(2J'+1)f^{J'}_{MN}\mathcal{F}^{J'}_{M+N}(b,z_t). \quad \text{(B2)}$$

The presence of a factor $P_J(z_t)$ in $\mathcal{F}^J_{m+n}(b,z_t)$ means that the z_t integration is straightforward, and there results

$$X^{mn}_{MN} = 2(M+N+1)^2 \int_0^1 b\,db \sum_{J=0}^{n}(2J+1)\bar{e}^J_{mn}f^J_{MN} \times \frac{(m+n-J)!\,(2J+1)!}{(m+n+J+1)!} \times \frac{(M+N-J)!\,(2J+1)!}{(M+N+J+1)!} \times b^{2J}P^{(2J+1,0)}_{m+n-J}(1-2b)P^{(2J+1,0)}_{M+N-J}(1-2b). \quad \text{(B3)}$$

The orthogonality property

$$2(k+1)\int_0^1 b\,dbb^{2J}P^{(2J+1,0)}_{k-J}(1-2b) \times P^{(2J+1,0)}_{K-J}(1-2b) = \delta_{kK}, \quad \text{(B4)}$$

for k and K integers, shows immediately that X^{mn}_{MN} vanishes unless $m+n = M+N$.

The explicit evaluation of X^{mn}_{MN} is now most easily effected by returning to Eq. (B1) and using the power-series expansion for E_{mn} and the Rodrigues formula for F_{MN}. Then partial integration yields

$$X^{mn}_{MN} = \frac{(-1)^{M+N}2(M+N+1)^2}{M!\,N!} \times \sum_{p=0}^{m}\sum_{q=0}^{n}\frac{(m+n+2)_{p+q}(-m)_p(-n)_q}{p!^2\,q!^2} \times \iint_\Delta du\,ds u^M s^N(1-u-s)^{M+N} \times \left(\frac{\partial}{\partial u}\right)^M\left(\frac{\partial}{\partial s}\right)^N[(1-u-s)^p u^p]. \quad \text{(B5)}$$

Evidently, the integrand vanished unless $p+q \geq M+N$. But we know that $p+q \leq m+n$ and also that $M+N = m+n$ for nonzero elements of X. This means that

$$X^{mn}_{MN} = \frac{(-1)^{M+N}2(m+N+1)^2}{M!\,N!} \times \frac{(m+n+2)_{m+n}(-m)_m(-n)_n}{m!^2\,n!^2} \times \iint_\Delta du\,ds u^M s^N(1-u-s)^{M+N} \times \left(\frac{\partial}{\partial u}\right)^M\left(\frac{\partial}{\partial s}\right)^N[(1-u-s)^m u^n]\delta_{m+n,M+N}, \quad \text{(B6)}$$

and using the standard integral

$$\iint_\Delta du\,ds u^\alpha s^\beta(1-u-s)^\gamma = \frac{\alpha!\,\beta!\,\gamma!}{(\alpha+\beta+\gamma+2)!}, \quad \text{(B7)}$$

the result Eq. (3.23) follows directly.

APPENDIX C: THE OPERATOR $\mathcal{O}$ AND ITS EIGENFUNCTIONS

The differential operator $\mathcal{O}$ is given, in terms of the independent variables a and z, by

$$\mathcal{O} \equiv -\frac{1}{a}\left[\frac{\partial}{\partial a}a^2(1-a)\frac{\partial}{\partial a} + \frac{\partial}{\partial z}(1-z^2)\frac{\partial}{\partial z}\right]. \quad \text{(C1)}$$

For suitably well-behaved functions f and g, so that it is possible to integrate by parts and discard the integrated part, it is easy to see that

$$(g,\mathcal{O}f) \equiv \iint_\Delta dt\,du g^*\mathcal{O}f = \tfrac{1}{2}\int_0^1 a\,da\int_{-1}^{+1} dz g^* f = \tfrac{1}{2}\int_0^1 a\,da\int_{-1}^{+1} dz\left[a(1-a)\frac{\partial g^*}{\partial a}\frac{\partial f}{\partial a} + \frac{1-z^2}{a}\frac{\partial g^*}{\partial z}\frac{\partial f}{\partial z}\right]. \quad \text{(C2)}$$

It then follows that

$$(g,\mathcal{O}f) = (f,\mathcal{O}g)^*, \quad \text{(C3)}$$

and that

$$(f,\mathcal{O}f) \geq 0 \quad \text{(C4)}$$

with equality if, and only if, f is constant on Δ.

If f is an eigenfunction of $\mathcal{O}$, we will write

$$\mathcal{O}f = \lambda(\lambda+2)f. \quad \text{(C5)}$$

Note that $F_{\mu\nu}$, $E_{\mu\nu}$, $G_{\mu\nu}$, and $G_{-\mu-1,-\nu-1}$ are all of them eigenfunctions of $\mathcal{O}$ with

$$\lambda = \mu + \nu. \quad \text{(C6)}$$

If, in addition, f is well behaved inside and on Δ, we have

$$(f,\mathcal{O}f) = \lambda(\lambda+2)(f,f) \geq 0, \quad \text{(C7)}$$

so that $\lambda(\lambda+2)$ is real and nonnegative. If $\mathcal{O}f = 0$, f must be constant, and otherwise we have λ real and

$$\lambda > 0 \quad \text{or} \quad \lambda < -2.$$

Also, if

$$\mathcal{O}f = \lambda(\lambda+2)f \quad \text{(C8)}$$

and

$$\mathcal{O}g = \lambda'(\lambda'+2)g, \quad \text{(C9)}$$

with

$$\lambda(\lambda+2) \neq \lambda'(\lambda'+2), \quad \text{(C10)}$$

then

$$(g,f) = 0. \quad \text{(C11)}$$

The above results are, of course, standard results on self-adjoint operators.

On Continuous Eigenvalues in Neutron Thermalization

MASAMICHI TOKIZAWA
Department of Mathematics, Chuō University, Tokyo, Japan

(Received 27 December 1968)

The continuous eigenvalue spectra of the linearized Boltzmann operator describing the energy distribution of neutrons in an infinite Einstein crystal are studied. This operator consists of two terms: a multiplication operator and an integral operator with δ-function-type singular kernel. The eigenvalue problem is transformed into the solution of an inhomogeneous integral equation by applying Case's method on the one-velocity transport equation. The existence of the solution of the integral equation is examined by the Neumann-series expansion. It is found that for sufficiently low temperature the range of numerical values of the multiplication operator forms the continuous eigenvalue spectra of the Boltzmann operator and the corresponding eigenfunctions are of δ-function type.

1. INTRODUCTION

The time-dependent behavior of neutrons in matter is described by the linearized Boltzmann transport equation

$$\frac{\partial \phi}{\partial t} = -B\phi, \tag{1}$$

where $\phi(\mathbf{r}, \Omega, E, t)$ is the neutron flux per unit energy interval per unit solid angle at position $\mathbf{r}$ at time t. B is the Boltzmann operator and it usually depends on $(\mathbf{r}, \Omega, E)$. If we assume the solutions of Eq. (1) to be of the form

$$\phi(\mathbf{r}, \Omega, E, t) = \varphi(\mathbf{r}, \Omega, E)e^{-\lambda t}, \tag{2}$$

we have the eigenvalue problem

$$B\varphi(\mathbf{r}, \Omega, E) = \lambda\varphi(\mathbf{r}, \Omega, E). \tag{3}$$

One of the tractable cases for this eigenvalue problem is found in the spatially independent thermalization theory. In this theory, the energy (velocity) distribution of neutrons in an infinite homogeneous medium is studied,and the Boltzmann operator takes a rather simple form[1,2]:

$$B\varphi(E) = v[\gamma(E) + \sigma(E)]\varphi(E) - v\int_0^\infty G(E' \to E)\varphi(E')\,dE', \tag{4}$$

where v is the neutron velocity corresponding to the neutron energy E, $\gamma(E)$ is the absorption cross section at energy E, $\sigma(E)$ is the scattering cross section at energy E, and $G(E' \to E)$ is the scattering kernel representing the cross section for a neutron with initial energy E' and with final energy E after a collision. $\sigma(E)$ and $G(E' \to E)$ must satisfy the relation

$$\sigma(E) = \int_0^\infty G(E \to E')\,dE'. \tag{5}$$

The kernel $vG(E' \to E)$ of the Boltzmann operator is not a symmetric function, but it can be symmetrized by introducing the new dependent variable

$$\psi(E) = \varphi(E)/E^{\frac{1}{4}}M(E)^{\frac{1}{2}} \tag{6}$$

and taking account of the detailed balance condition

$$M(E')G(E' \to E) = M(E)G(E \to E'), \tag{7}$$

where $M(E)$ denotes the Maxwellian distribution at temperature T,

$$M(E) = Ee^{-E/T}. \tag{8}$$

The eigenvalue problem then becomes

$$H\psi(E) = \nu\psi(E), \tag{9}$$

$$H\psi(E) = Q(E)\psi(E) - \int_0^\infty K_s(E', E)\psi(E')\,dE', \tag{10}$$

$$K_s(E', E) = K_s(E, E') = \left[\frac{M(E')}{M(E)}\right]^{\frac{1}{2}} G(E' \to E)(EE')^{\frac{1}{4}}, \tag{11}$$

$$Q(E) = E^{\frac{1}{2}}[\gamma(E) + \sigma(E)]. \tag{12}$$

There are several investigations[3,4] on the eigenvalue spectra of this kind of operator. In these investigations, the integral operator with a well-behaving kernel, such as the L^2 kernel, is regarded as the perturbation to the multiplication operator $Q(E)$, but one expects that the range of numerical values of the function $Q(E)$ still forms the continuous eigenvalue spectra of the operator H. This conjecture has been proved in the framework of the theory of L^2 (square-integrable)-function space. Then the interval min $Q(E)$ to max $Q(E)$ can be considered to be the continuous eigenvalue spectra of H. This is the principal result obtained by those investigations.

[1] E. R. Cohen, *Proceedings of the International Conference on the Peaceful Uses of Atomic Energy* (United Nations, New York, 1955), Article P/611.

[2] M. M. R. Williams, *The Slowing Down and Thermalization of Neutrons* (North-Holland Publ. Co., Amsterdam, 1966).

[3] Y. Shizuta, Progr. Theoret. Phys. (Kyoto) **32**, 489 (1964).

[4] N. Corngold, Nucl. Sci. Eng. **19**, 80 (1964).

In the present paper we study the operator with the singular kernel which contains Dirac δ functions representing the Einstein crystal scatterer. The eigenvalue problem is transformed into the solution of an inhomogeneous integral equation by applying to the eigenvalue equation Case's method on the one-velocity transport equation. The existence of the solution of the integral equation is examined by the Neumann-series expansion. The formulation of this procedure is presented in Sec. 2.

The specified kernel form is given in Sec. 3. In Sec. 4 we discuss the convergence of the Neumann series; we show that under certain conditions the range of numerical values of $Q(E)$ forms the continuous eigenvalue spectra of H even in the case of the singular kernel and that the corresponding eigenfunctions are of the δ-function type.

2. AN APPROACH TO EIGENVALUE PROBLEM

If we put

$$L(E) = \int_0^\infty K_s(E', E)\psi(E')\,dE', \tag{13}$$

the eigenvalue equation (9) for the operator H defined by Eq. (10) can be rearranged as

$$[Q(E) - \nu]\psi(E) = L(E). \tag{14}$$

Although Case's method[5] was originally intended for the one-velocity transport equation, an analogous device can be applied to Eq. (14). Thus, dividing both sides of Eq. (14) by $[Q(E) - \nu]$, we have

$$\psi(E) = P\frac{L(E)}{Q(E) - \nu} + C(\nu)\delta(E - E_\nu), \tag{15}$$

where P denotes the Cauchy principal value, $C(\nu)$ is a constant, δ is the Dirac δ function, and

$$Q(E_\nu) - \nu = 0. \tag{16}$$

Equation (15) is the formal expression of the eigenfunction $\psi(E)$ of H, corresponding to the eigenvalue ν.

This procedure is based on a theorem[6,7] of distribution theory which states that, for an infinitely differentiable function $\alpha(x)$ and for a simple zero x_0 of $\alpha(x)$, a distribution T_x satisfies $\alpha(x)T_x = 0$ if and only if $T_x = C\delta(x - x_0)$, where C is a constant.

Substituting Eq. (15) into Eq. (13), we have an integral equation for $L(E)$, namely,

$$L(E) - P\int_0^\infty K_s(E', E)\frac{L(E')}{Q(E') - \nu}\,dE' = C(\nu)K_s(E_\nu, E). \tag{17}$$

If $Q(E) - \nu \neq 0$ for every E, Eq. (15) must be

$$\psi(E) = \frac{L(E)}{Q(E) - \nu} \tag{18}$$

and Eq. (17) must be

$$L(E) - P\int_0^\infty K_s(E', E)\frac{L(E')}{Q(E') - \nu}\,dE' = 0. \tag{19}$$

In the interval $0 \leq \nu < \min Q(E)$, there may be discrete eigenvalues corresponding to the eigenfunctions of the integral equation (19). For $\max Q(E) \geq \nu \geq \min Q(E)$, if there exists the unique solution of the integral equation (17), Eq. (15) represents the continuum solution of Eq. (14) and the interval forms continuous eigenvalue spectra.

The existence of the unique solution for Eq. (17) is examined by the Neumann-series expansion. Starting from

$$L_0(E) = CK_s(E_\nu, E), \tag{20}$$

we calculate successively

$$L_k(E) = P\int_0^\infty \frac{K_s(E', E)}{Q(E') - \nu}L_{k-1}(E')\,dE', \quad k = 1, 2, \cdots, \tag{21}$$

and construct the iterated solution $\sum_{k=0}^N L_k(E)$. The uniform convergence of the iterated solution ensures the existence of the unique solution for Eq. (17).

3. SINGULAR KERNEL

The scattering kernel for an Einstein crystal[2] is given by

$$G(E' \to E) = \int \frac{d^2g}{d\Omega\,dE}\,d\Omega, \tag{22}$$

$$\frac{d^2g}{d\Omega\,dE} = \frac{\sigma_b}{4\pi}\left(\frac{E}{E'}\right)^{\frac{1}{2}}S(\mathbf{k}, \epsilon), \tag{23}$$

$$S(\mathbf{k}, \epsilon) = \exp\left[-\frac{R_0}{E_r}\coth\frac{E_r}{2T}\right]\sum_{j=-\infty}^{\infty}\delta(jE_r - \epsilon) \times I_{|j|}\left[\frac{R_0/E_r}{\sinh(E_r/2T)}\right]e^{jE_r/2T}, \tag{24}$$

where σ_b is the bound-atom cross section, E_r is the value of the single-energy level, T is the temperature in energy units, M is the mass of scatterer atoms, m is the neutron mass,

$$\epsilon = E' - E, \tag{25}$$

$$R_0 = (m/M)[E' + E - 2(EE')^{\frac{1}{2}}\cos\theta], \tag{26}$$

and $I_{|j|}(z)$ denotes the modified Bessel function of order $|j|$. The modified Bessel function is defined by

$$I_\nu(z) = \sum_{k=0}^{\infty}\frac{(z/2)^{\nu+2k}}{k!\,\Gamma(\nu + k + 1)}. \tag{27}$$

[5] K. M. Case, Ann. Phys. (N.Y.) 9, 1 (1960).
[6] L. Schwartz, *Théorie des distributions* (Hermann et Cie., Paris, 1950), Vol. I.
[7] L. Schwartz, *Méthodes mathématiques pour les sciences physiques* (Hermann et Cie., Paris, 1961).

In the following analysis, we confine ourselves to a special case where this kernel takes a simpler form and yet which is based on physically reasonable assumptions and has a clear physical interest. Thus, we assume that

$$E_r/2T \gg 1, \tag{28}$$

the summation over j is taken with respect to $j = 0, \pm 1$, and the other terms are truncated. The latter assumption corresponds to the one-phonon approximation.[2] Furthermore, we can omit the $j = 0$ term because it represents the elastic scattering and does not affect the eigenvalue.

From assumption (28) we have

$$\frac{R_0/E_r}{\sinh(E_r/2T)} \approx \frac{2R_0/E_r}{\exp(E_r/2T)} \approx 0, \tag{29}$$

$$\coth(E_r/2T) \approx 1. \tag{30}$$

From the definition we have, for $z \approx 0$,

$$I_1(z) = \tfrac{1}{2}z + \mathcal{O}(z^3), \tag{31}$$

and then, combining Eq. (29) with Eq. (31), we find that

$$I_1\left[\frac{R_0/E_r}{\sinh(E_r/2T)}\right] \approx \frac{R_0}{E_r}\, e^{-E_r/2T}. \tag{32}$$

Thus,in our case, Eq. (24) is reduced to

$$\begin{aligned} &S(\mathbf{k}, \epsilon) \\ &= e^{-R_0/E_r}\left\{\delta(-E_r - \epsilon)e^{-E_r/2T} I_1\left[\frac{R_0/E_r}{\sinh(E_r/2T)}\right]\right. \\ &\quad \left. + \delta(E_r - \epsilon)e^{E_r/2T} I_1\left[\frac{R_0/E_r}{\sinh(E_r/2T)}\right]\right\} \\ &= (R_0/E_r)e^{-R_0/E_r}[\delta(-E_r - \epsilon)e^{-E_r/T} + \delta(E_r - \epsilon)]. \end{aligned} \tag{33}$$

From Eq. (22) the scattering kernel is calculated as

$$\begin{aligned} G(E' \to E) &= \frac{\sigma_b}{4\pi}\left(\frac{E}{E'}\right)^{\frac{1}{2}} \int S(\mathbf{k}, \epsilon)\, d\Omega \\ &= \frac{\sigma_b}{2}\left(\frac{E}{E'}\right)^{\frac{1}{2}} \int_0^{\pi} S(\mathbf{k}, \epsilon) \sin\theta\, d\theta \\ &= \frac{\sigma_b M}{4mE'} \int_{R_0(0)}^{R_0(\pi)} S(\mathbf{k}, \epsilon)\, dR_0, \end{aligned} \tag{34}$$

where

$$R_0(\pi) = (m/M)(E^{\frac{1}{2}} + E'^{\frac{1}{2}})^2, \tag{35}$$

$$R_0(0) = (m/M)(E^{\frac{1}{2}} - E'^{\frac{1}{2}})^2. \tag{36}$$

Substituting Eq. (33) into Eq. (34), we have

$$\begin{aligned} &G(E' \to E) \\ &= (\beta^2/E')[\delta(-E_r - E' + E)e^{-E_r/T} + \delta(E_r - E' + E)] \\ &\quad \times \int_{R_0(0)}^{R_0(\pi)} \frac{R_0}{E_r}\, e^{-R_0/E_r}\, dR_0, \end{aligned} \tag{37}$$

where

$$\beta^2 = \sigma_b M/4m. \tag{38}$$

If $\mathfrak{F}(E', E)$ denotes the integral in Eq. (37), we have

$$\begin{aligned} \mathfrak{F}(E', E) &= \int_{R_0(0)}^{R_0(\pi)} \frac{R_0}{E_r}\, e^{-R_0/E_r}\, dR_0 \\ &= E_r\left[\left(\frac{R_0(0)}{E_r} + 1\right)e^{-R_0(0)/E_r}\right. \\ &\quad \left. - \left(\frac{R_0(\pi)}{E_r} + 1\right)e^{-R_0(\pi)/E_r}\right]. \end{aligned} \tag{39}$$

For brevity, we change the scale of the energy E so that $E_r = 1$ and we assume this throughout the rest of this paper. Then the scattering kernel finally becomes

$$G(E' \to E) = \mathfrak{F}(E', E) \cdot \Delta(E', E), \tag{40}$$

where

$$\mathfrak{F}(E', E) = [R_0(0) + 1]e^{-R_0(0)} - [R_0(\pi) + 1]e^{-R_0(\pi)}, \tag{41}$$

$$\Delta(E', E) = (\beta^2/E')[\delta(E - E' - 1)e^{-1/T} + \delta(E - E' + 1)], \tag{42}$$

$$R_0(\pi) = \alpha^2(E^{\frac{1}{2}} + E'^{\frac{1}{2}})^2, \tag{43}$$

$$R_0(0) = \alpha^2(E^{\frac{1}{2}} - E'^{\frac{1}{2}})^2, \tag{44}$$

$$\alpha^2 = m/M. \tag{45}$$

Now we shall turn to the determination of $Q(E)$. From Eq. (5) we have

$$\begin{aligned} \sigma(E) &= \int_0^{\infty} \mathfrak{F}(E, E') \cdot \Delta(E, E')\, dE' \\ &= (\beta^2/E)\int_0^{\infty} \mathfrak{F}(E, E')[\delta(E' - E - 1)e^{-1/T} \\ &\qquad + \delta(E' - E + 1)]\, dE' \\ &= (\beta^2/E)[\mathfrak{F}(E, E + 1)e^{-1/T} + \mathfrak{F}(E, E - 1)]. \end{aligned} \tag{46}$$

From Eqs. (41), (43), and (44) we have

$$\begin{aligned} \mathfrak{F}(E, E + 1) &= \{\alpha^2[(E + 1)^{\frac{1}{2}} - E^{\frac{1}{2}}]^2 + 1\} \\ &\quad \times \exp\{-\alpha^2[(E + 1)^{\frac{1}{2}} - E^{\frac{1}{2}}]^2\} \\ &\quad - \{\alpha^2[(E + 1)^{\frac{1}{2}} + E^{\frac{1}{2}}]^2 + 1\} \\ &\quad \times \exp\{-\alpha^2[(E + 1)^{\frac{1}{2}} + E^{\frac{1}{2}}]^2\}, \end{aligned} \tag{47}$$

$$\begin{aligned} \mathfrak{F}(E, E - 1) &= \{\alpha^2[(E - 1)^{\frac{1}{2}} - E^{\frac{1}{2}}]^2 + 1\} \\ &\quad \times \exp\{-\alpha^2[(E - 1)^{\frac{1}{2}} - E^{\frac{1}{2}}]^2\} \\ &\quad - \{\alpha^2[(E - 1)^{\frac{1}{2}} + E^{\frac{1}{2}}]^2 + 1\} \\ &\quad \times \exp\{-\alpha^2[(E - 1)^{\frac{1}{2}} + E^{\frac{1}{2}}]^2\}. \end{aligned} \tag{48}$$

If we assume the usual $1/v$ absorption,[2] $\gamma(E)$ takes the form

$$\gamma(E) = a/E^{\frac{1}{2}}, \tag{49}$$

where a is a constant. Then $Q(E)$ becomes

$$Q(E) = a + (\beta^2/E^{\frac{1}{2}})[\mathfrak{F}(E, E+1)e^{-1/T} + \mathfrak{F}(E, E-1)]. \tag{50}$$

As $E \to \infty$, $\mathfrak{F}(E, E+1) \to 1$ and $\mathfrak{F}(E, E-1) \to 1$; consequently,

$$Q(E) \to a + (\beta^2/E^{\frac{1}{2}})[1 + e^{-1/T}]. \tag{51}$$

Because $\mathfrak{F}(1, 0) = 0$, we have

$$\begin{aligned} Q(1) &= a + \beta^2 e^{-1/T}\mathfrak{F}(1, 2) \\ &= a + \beta^2 e^{-1/T}\{[\alpha^2(\sqrt{2}-1)^2 + 1] \\ &\quad \times \exp[-\alpha^2(\sqrt{2}-1)^2] \\ &\quad - [\alpha^2(\sqrt{2}+1)^2 + 1]\exp[-\alpha^2(\sqrt{2}+1)^2]\} \end{aligned} \tag{52}$$

for the value of $Q(E)$ when $E = 1$.

It can be shown that the derivative of $Q(E)$ is

$$\begin{aligned} \frac{dQ(E)}{dE} &= -\tfrac{1}{2}\beta^2 E^{-\frac{3}{2}} \\ &\quad \times (\{[p(E)+1]e^{-p(E)} \\ &\quad - [q(E)+1]e^{-q(E)}\}e^{-1/T} \\ &\quad + [p(E-1)+1]e^{-p(E-1)} \\ &\quad - [q(E-1)+1]e^{-q(E-1)}) \\ &\quad + \frac{\beta^2}{E^{\frac{1}{2}}}\{[p^2(E)e^{-p(E)} + q^2(E)e^{-q(E)}]e^{-1/T}/ \\ &\qquad E^{\frac{1}{2}}(E+1)^{\frac{1}{2}} \\ &\quad + [p^2(E-1)e^{-p(E-1)} + q^2(E-1)e^{-q(E-1)}]/ \\ &\qquad E^{\frac{1}{2}}(E-1)^{\frac{1}{2}}\}, \end{aligned} \tag{53}$$

where

$$p(E) = \alpha^2[(E+1)^{\frac{1}{2}} - E^{\frac{1}{2}}]^2, \tag{54}$$

$$q(E) = \alpha^2[(E+1)^{\frac{1}{2}} + E^{\frac{1}{2}}]^2. \tag{55}$$

Therefore, we find that

$$\left(\frac{dQ(E)}{dE}\right)_{E=1} = \infty. \tag{56}$$

In the interval $0 \leq E \leq 1$, we find that

$$\mathfrak{F}(E, E-1) = 0, \tag{57}$$

because the energy transfer beyond $E = 0$ does not occur, and then the $\delta(E - E' + 1)$ part of $\Delta(E', E)$ in Eq. (42) must be considered to lose effect in this interval. Then,it follows that

$$Q(0) = a + \beta^2 e^{-1/T}\lim_{E\to 0}\frac{\mathfrak{F}(E, E+1)}{E^{\frac{1}{2}}}. \tag{58}$$

It is easily shown that, as $E \to 0$,

$$\begin{aligned} &\frac{\mathfrak{F}(E, E+1)}{E^{\frac{1}{2}}} \\ &\quad \to \left\{(\alpha^2+1)\frac{1 - \exp[-4\alpha^2 E^{\frac{1}{2}}(E+1)^{\frac{1}{2}}]}{E^{\frac{1}{2}}} - 4\alpha^2\right\} \\ &\qquad \times \exp(-\alpha^2) \end{aligned}$$

and,consequently,

$$\to 4\alpha^4 \exp(-\alpha^2). \tag{59}$$

Thus, we obtain

$$Q(0) = a + 4\beta^2\alpha^4 \exp[-(\alpha^2 + T^{-1})]. \tag{60}$$

The necessary knowledge of the behavior of the function $Q(E)$ is obtained from Eqs. (51), (52), (56), (57), and (60). Figure 1 shows a sketch of the curve which represents $Q(E)$.

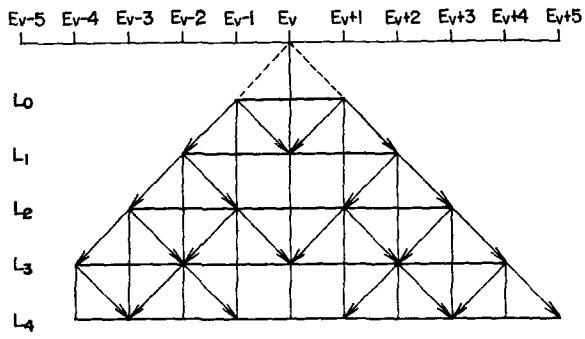


FIG. 1. The propagation of δ-functions.

4. CONVERGENCE OF THE NEUMANN SERIES

From Eqs. (11) and (40) our symmetric kernel is given by

$$K_s(E', E) = \left[\frac{M(E')}{M(E)}\right]^{\frac{1}{2}} \cdot (EE')^{\frac{1}{4}} \cdot \mathfrak{F}(E', E) \cdot \Delta(E', E). \tag{61}$$

We shall begin the construction of the iterated solution of Eq. (17) with $L_0(E)$ given by Eq. (20). Here it is assumed that $C = 1$, because the iterated solution is proportional to C. Then, from Eqs. (21), (41), (42), and (61) and repeatedly applying the relation

$$\int \delta(a - x)\delta(x - b)\,dx = \delta(a - b), \tag{62}$$

we find

$$L_1(E) = P\int_0^\infty \frac{K_s(E', E)}{Q(E') - \nu} L_0(E')\, dE'$$
$$= \frac{\mathfrak{H}(E_\nu + 1, E)}{Q(E_\nu + 1) - \nu}\left[\frac{M(E_\nu)}{M(E_\nu + 1)}\right]^{\frac{1}{2}}$$
$$\times [(E_\nu + 1)E_\nu]^{\frac{1}{4}}\mathfrak{F}(E_\nu, E_\nu + 1)$$
$$\times \frac{\beta^4}{E_\nu(E_\nu + 1)}[e^{-2/T}\delta(E - (E_\nu + 2))$$
$$+ e^{-1/T}\delta(E - E_\nu)]$$
$$+ \frac{\mathfrak{H}(E_\nu - 1, E)}{Q(E_\nu - 1) - \nu}\left[\frac{M(E_\nu)}{M(E_\nu - 1)}\right]^{\frac{1}{2}}$$
$$\times [(E_\nu - 1)E_\nu]^{\frac{1}{4}}\mathfrak{F}(E_\nu, E_\nu - 1)$$
$$\times \frac{\beta^4}{E_\nu(E_\nu - 1)}[e^{-1/T}\delta(E - E_\nu)$$
$$+ \delta(E - (E_\nu - 2))], \quad (63)$$

where

$$\mathfrak{H}(E^*, E^{**}) = \left[\frac{M(E^*)}{M(E^{**})}\right]^{\frac{1}{2}}(E^{**}E^*)^{\frac{1}{4}}\mathfrak{F}(E^*, E^{**}).$$

According to Eq. (21), a calculation is successively performed in order to obtain $L_k(E)$, $k = 2, 3, \cdots$. Equation (63) shows that two δ functions $\delta(E' - (E_\nu + 1))$ and $\delta(E' - (E_\nu - 1))$ of $L_0(E') = K_s(E_\nu, E')$ generate two pairs of δ-functions $\delta(E - (E_\nu + 2))$, $\delta(E - E_\nu)$ and $\delta(E - E_\nu)$, $\delta(E - (E_\nu - 2))$ in $L_1(E)$. Generally,at the kth step,

$$L_k(E) = P\int_0^\infty \frac{\mathfrak{H}(E', E)}{Q(E') - \nu}\cdot\frac{\beta^2}{E'}$$
$$\times [\delta(E - (E' + 1))e^{-1/T} + \delta(E - (E' - 1))]$$
$$\times L_{k-1}(E')\, dE', \quad (64)$$

it is seen that, if $L_{k-1}(E)$ contains $\delta(E - \xi)$, the term $\delta(E - (E' + 1))$ generates $\delta(E - (\xi + 1))$ and the term $\delta(E - (E' - 1))$ generates $\delta(E - (\xi - 1))$; that is, the former shifts the position of δ function on the E line by $+1$, and the latter shifts the position of δ function by -1. This suggests the introduction of the "forward-shift operator" A, and the "backward-shift operator" B, which are defined by

$$A[\delta(E - (E_\nu + n))] = A_n\delta(E - (E_\nu + n + 1)),$$
$$\text{for} \quad E_\nu + n > 0, \quad (65)$$

$$B[\delta(E - (E_\nu + n))] = B_n\delta(E - (E_\nu + n - 1)),$$
$$\text{for} \quad E_\nu + n - 1 > 0, \quad (66)$$

where

$$A_n = \frac{\mathfrak{H}(E_\nu + n, E_\nu + n + 1)}{Q(E_\nu + n) - \nu}\cdot\frac{\beta^2}{E_\nu + n}e^{-1/T}, \quad (67)$$

$$B_n = \frac{\mathfrak{H}(E_\nu + n, E_\nu + n - 1)}{Q(E_\nu + n) - \nu}\cdot\frac{\beta^2}{E_\nu + n}, \quad (68)$$

and n is an integer.

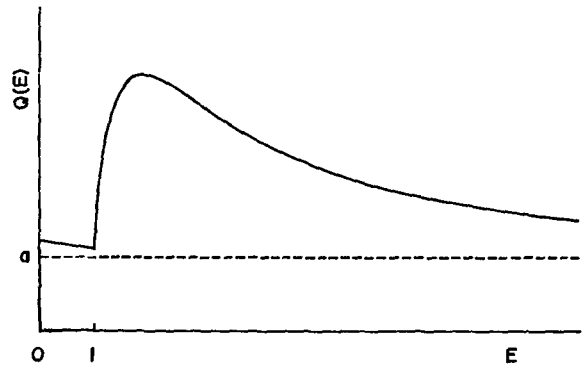


FIG. 2. The shape of $Q(E)$.

It can be shown that if $L_{k-1}(E)$ is proportional to $\delta(E - (E_\nu + n))$, then $L_k(E)$ is expressed by A and B as

$$L_k(E) \propto [A_n\delta(E - (E_\nu + n + 1))$$
$$+ B_n\delta(E - (E_\nu + n - 1))]. \quad (69)$$

Figure 2 shows the propagation of the δ function caused by the branching expressed by Eq. (69). Here two δ functions are placed at $E = E_\nu \pm 1$ for $k = 0$. Since $\delta(E - E_\nu)$ is a symmetric function with respect to $E = E_\nu$, the principal value of Eq. (64) must vanish for $\delta(E - E_\nu)$ contained in $L_{k-1}(E)$. In other words,

$$A_0 = B_0 = 0. \quad (70)$$

From the above argument we see that the Neumann series of Eq. (17) is expressed, finally, by a linear combination of δ-functions which are placed at $E = E_\nu + n\ (> 0)$ where $n = 0, \pm1, \pm2, \cdots$. Thus, what we have to show is the convergence of the series $\sum_{k=0}^\infty \mathfrak{M}_{k,n}$ for every n where $\mathfrak{M}_{k,n}$ is the coefficient of $\delta(E - (E_\nu + n))$ in $L_k(E)$.

In order to construct a majorant for the series $\sum_{k=0}^\infty |\mathfrak{M}_{k,n}|$, the following preliminary estimates are necessary:

$$\frac{M(E_\nu + n)}{M(E_\nu + n + 1)}$$
$$= \frac{(E_\nu + n)\exp[-(E_\nu + n)/T]}{(E_\nu + n + 1)\exp[-(E_\nu + n + 1)/T]}$$
$$= \left(1 - \frac{1}{E_\nu + n + 1}\right)e^{1/T} < e^{1/T}, \quad (71)$$

$$[(E_\nu + n + 1)(E_\nu + n)]^{\frac{1}{4}} < (E_\nu + n + 1)^{\frac{1}{2}}, \quad (72)$$

$$\mathfrak{F}(E_\nu + n, E_\nu + n + 1)$$
$$< \alpha^2[(E_\nu + n + 1)^{\frac{1}{2}} - (E_\nu + n)^{\frac{1}{2}}]^2 + 1$$
$$< \alpha^2\{(E_\nu + n + 1) + (E_\nu + n)$$
$$- 2[(E_\nu + n)(E_\nu + n)]^{\frac{1}{2}}\} + 1 = \alpha^2 + 1, \quad (73)$$

for $E_v + n \geq 0$. The inequality (71) follows from the definition of $M(E)$, and the inequality (73) from Eqs. (41), (43), and (44). These estimations are applied to majorize A_n. In order to majorize B_n we have the similar estimates:

$$\frac{M(E_v + n)}{M(E_v + n - 1)} = \frac{(E_v + n)\exp[-(E_v + n)/T]}{(E_v + n - 1)\exp[-(E_v + n - 1)/T]} = \left(1 + \frac{1}{E_v + n - 1}\right)e^{-1/T} \leq 2e^{-1/T}, \quad \text{for } E_v + n - 1 \geq 1, \tag{74}$$

$$[(E_v + n - 1)(E_v + n)]^{\frac{1}{4}} < (E_v + n)^{\frac{1}{2}}, \quad \text{for } E_v + n - 1 \geq 0, \tag{75}$$

$$\begin{aligned}\mathfrak{F}(E_v + n, E_v + n - 1) &< \alpha^2[(E_v + n - 1)^{\frac{1}{2}} - (E_v + n)^{\frac{1}{2}}]^2 + 1 \\ &< \alpha^2\{(E_v + n - 1) + (E_v + n) \\ &\quad - 2[(E_v + n - 1)(E_v + n - 1)]^{\frac{1}{2}}\} + 1 \\ &= \alpha^2 + 1, \quad \text{for } E_v + n - 1 \geq 0.\end{aligned} \tag{76}$$

When $E_v + n - 1 \to 0$, the inequality (74) does not hold and it is necessary to examine the behavior of B_n. In this case B_n is expressed by

$$B_n = K\frac{\mathfrak{F}(E_v + n, E_v + n - 1)}{(E_v + n - 1)^{\frac{1}{4}}}, \tag{77}$$

where K is the factor which takes a finite value as $E_v + n - 1 \to 0$. If we put $x = E_v + n - 1$, we have from Eq. (47)

$$\begin{aligned}\mathfrak{F}(x + 1, x) &= \exp[-\alpha^2(2x + 1)] \\ &\quad \times \{[\alpha^2(2x + 1) + 1 - 2\alpha^2(x(x + 1))^{\frac{1}{2}}] \\ &\quad \times \exp[2\alpha^2(x(x + 1))^{\frac{1}{2}}] \\ &\quad - [\alpha^2(2x + 1) + 1 + 2\alpha^2(x(x + 1))^{\frac{1}{2}}] \\ &\quad \times \exp[-2\alpha^2(x(x + 1))^{\frac{1}{2}}]\}.\end{aligned} \tag{78}$$

Substituting the Taylor expansion

$$\exp[\pm 2\alpha^2(x(x + 1))^{\frac{1}{2}}] = 1 \pm 2\alpha^2(x(x + 1))^{\frac{1}{2}} + 2\alpha^4 x(x + 1) + \mathcal{O}(x^{\frac{3}{2}}), \tag{79}$$

we find

$$\frac{\mathfrak{F}(x + 1, x)}{x^{\frac{1}{4}}} = \exp[-\alpha^2(2x + 1)] \times \{4\alpha^4 x^{\frac{1}{4}}(x + 1)^{\frac{1}{2}}(2x + 1) + \mathcal{O}(x^{\frac{5}{4}})\}. \tag{80}$$

Thus, we see that

$$\lim_{x\to 0} B_n = \lim_{x\to 0} K\frac{\mathfrak{F}(x + 1, x)}{x^{\frac{1}{4}}} = 0 \tag{81}$$

and then the majorant of Eq. (74) is valid for

$$E_v + n - 1 \geq 0.$$

At first we assume that $E_v \gg 1$, so that

$$Q(E) \sim a + c^*\beta^2/E^{\frac{1}{2}}, \tag{82}$$

where

$$c^* = 1 + e^{-1/T}. \tag{83}$$

From (67), (71), (72), (73), and (82) we find a majorant for $|A_n|$, $n > 0$:

$$|A_n| < \frac{e^{1/2T}(\alpha^2 + 1)}{|a + c^*\beta^2(E_v + n)^{-\frac{1}{2}} - \nu|} \cdot \frac{(E_v + n + 1)^{\frac{1}{2}}}{E_v + n}\beta^2 e^{-1/T}. \tag{84}$$

Substituting $[(E_v + n + 1)/(E_v + n)]^{\frac{1}{2}} \leq \sqrt{2}$ into (84), we have

$$|A_n| < \frac{\sqrt{2}(\alpha^2 + 1)\beta^2}{|c^*\beta^2 + (a - \nu)(E_v + n)^{\frac{1}{2}}|}e^{-1/2T} \tag{85}$$

$$\leq \frac{\sqrt{2}(\alpha^2 + 1)\beta^2}{|c^*\beta^2 + (a - \nu)(E_v + 1)^{\frac{1}{2}}|}e^{-1/2T}. \tag{86}$$

From (68), (74), (75), (76), and (82) we find a majorant for $|B_n|$, $n > 0$:

$$|B_n| < \frac{\sqrt{2}\,e^{-1/2T}(\alpha^2 + 1)}{|a + c^*\beta^2(E_v + n)^{-\frac{1}{2}} - \nu|} \cdot \frac{(E_v + n)^{\frac{1}{2}}}{E_v + n}\beta^2 \tag{87}$$

$$= \frac{\sqrt{2}(\alpha^2 + 1)\beta^2}{|c^*\beta^2 + (a - \nu)(E_v + n)^{\frac{1}{2}}|}e^{-1/2T} \tag{88}$$

$$\leq \frac{\sqrt{2}(\alpha^2 + 1)\beta^2}{|c^*\beta^2 + (a - \nu)(E_v + 1)^{\frac{1}{2}}|}e^{-1/2T}. \tag{89}$$

Putting

$$\omega = \frac{\sqrt{2}(\alpha^2 + 1)\beta^2}{|c^*\beta^2 + (a - \nu)(E_v + 1)^{\frac{1}{2}}|}, \tag{90}$$

we have, for $n > 0$, that

$$|A_n| < \omega e^{-1/2T}, \tag{91}$$

$$|B_n| < \omega e^{-1/2T}. \tag{92}$$

For $n < 0$, the behavior of A_n for $E_v + n \to 0$ must be examined. Putting $x = E_v + n$, we have an expression similar to Eq. (77):

$$A_n = K'[\mathfrak{F}(x, x + 1)/x^{\frac{1}{4}}]. \tag{93}$$

By means of (59) we find easily that, as $x \to 0$,

$$A_n \to 0 \tag{94}$$

and Eq. (67) is then valid for $E_v + n > 0$. We see that

(85) also gives a majorant for $|A_n|$, $n < 0$, and so

$$|A_n| < \frac{\sqrt{2}\,(\alpha^2+1)\beta^2}{|c^*\beta^2 + (a-\nu)(E_\nu - 1)^{\frac{1}{2}}|}\, e^{-1/2T}. \quad (95)$$

Similarly, (88) also gives a majorant for $|B_n|$, $n < 0$, so

$$|B_n| < \frac{\sqrt{2}\,(\alpha^2+1)\beta^2}{|c^*\beta^2 + (a-\nu)(E_\nu - 1)^{\frac{1}{2}}|}\, e^{-1/2T}. \quad (96)$$

Putting

$$\omega' = \frac{\sqrt{2}\,(\alpha^2+1)\beta^2}{|c^*\beta^2 + (a-\nu)(E_\nu - 1)^{\frac{1}{2}}|}, \quad (97)$$

we have, for $n < 0$, that

$$|A_n| < \omega' e^{-1/2T}, \quad (98)$$

$$|B_n| < \omega' e^{-1/2T}. \quad (99)$$

Next we treat the general case in which the assumption (82) on $Q(E)$ is not postulated. A majorant for $|A_n|$, $n > 0$, is given by replacing

$$|a + c^*\beta^2(E_\nu + n)^{-\frac{1}{2}} - \nu|$$

of Eq. (84) by

$$\gamma_\nu = \min_{n>0} |Q(E_\nu + n) - \nu| \neq 0. \quad (100)$$

Thus,

$$|A_n| < \frac{\sqrt{2}\,(\alpha^2+1)\beta^2}{\gamma_\nu (E_\nu + n)^{\frac{1}{2}}}\, e^{-1/2T} \quad (101)$$

$$\leq \frac{\sqrt{2}\,(\alpha^2+1)\beta^2}{\gamma_\nu (E_\nu + 1)^{\frac{1}{2}}}\, e^{-1/2T}. \quad (102)$$

Similarly, a majorant for $|B_n|$, $n > 0$, is given by replacing $|a + c^*\beta^2(E_\nu + n)^{-\frac{1}{2}} - \nu|$ of Eq. (87) by γ_ν defined by Eq. (100). Thus, we have the same majorant (102) for $|B_n|$ as for $|A_n|$. Putting

$$\omega'' = \sqrt{2}\,(\alpha^2+1)\beta^2/\gamma_\nu (E_\nu + 1)^{\frac{1}{2}}, \quad (103)$$

we have, for $n > 0$,

$$|A_n| < \omega'' e^{-1/2T}, \quad (104)$$

$$|B_n| < \omega'' e^{-1/2T}. \quad (105)$$

We see that majorants for $|A_n|$ and $|B_n|$, $n < 0$, are given by replacing $|a + c^*\beta^2(E_\nu + n)^{-\frac{1}{2}} - \nu|$ of Eqs. (84) and (87), respectively, by

$$\gamma_\nu' = \min_{-E\nu < n < 0} |Q(E_\nu + n) - \nu| \neq 0. \quad (106)$$

Thus, we find that

$$|A_n| < [\sqrt{2}\,(\alpha^2+1)\beta^2/\gamma_\nu'(E_\nu + n)^{\frac{1}{2}}]e^{-1/2T} \quad (107)$$

$$\leq [\sqrt{2}\,(\alpha^2+1)\beta^2/\gamma_\nu']e^{-1/2T}, \quad (108)$$

$$|B_n| < [\sqrt{2}\,(\alpha^2+1)\beta^2/\gamma_\nu']e^{-1/2T}. \quad (109)$$

Putting

$$\omega''' = \sqrt{2}\,(\alpha^2+1)\beta^2/\gamma_\nu', \quad (110)$$

we have, for $n < 0$,

$$|A_n| < \omega''' e^{-1/2T}, \quad (111)$$

$$|B_n| < \omega''' e^{-1/2T}. \quad (112)$$

Next we shall construct a majorant series for $\sum_{k=0}^{\infty} |\mathfrak{M}_{k,n}|$. It is enough to present a majorant series in the case of the asymptotic $Q(E)$ given by (82), and of $n > 0$, because all the cases have the same expression for majorants for $|A_n|$ and $|B_n|$ as we have seen before. At first, for $n = 1$, Fig. 2 shows that

$$\begin{aligned}
\mathfrak{M}_{0,1} &= 1 = 1, \\
|\mathfrak{M}_{2,1}| &< \omega^2 e^{-1/T} < 2^2\omega^2 e^{-1/T}, \\
|\mathfrak{M}_{4,1}| &< 2\omega^4 e^{-2/T} < 2^4\omega^4 e^{-2/T}, \\
|\mathfrak{M}_{6,1}| &< 5\omega^6 e^{-3/T} < 2^6\omega^6 e^{-3/T}, \\
&\cdots \\
|\mathfrak{M}_{2l,1}| &< s_l\omega^{2l} e^{-l/T} < 2^{2l}\omega^{2l} e^{-l/T},
\end{aligned} \quad (113)$$

and so on. Therefore, we find a majorant series

$$\sum_{k=0}^{\infty} |\mathfrak{M}_{k,1}| < \sum_{l=0}^{\infty} [(2\omega)^2 e^{-1/T}]^l. \quad (114)$$

This geometric series converges if

$$(2\omega)^2 e^{-1/T} < 1. \quad (115)$$

For any $n > 0$ we find that

$$\sum_{k=0}^{\infty} |\mathfrak{M}_{k,n}| < \sum_{l=0}^{\infty} (\omega e^{-1/2T})^{n-1} [(2\omega)^2 e^{-1/T}]^l. \quad (116)$$

The convergence condition (115) is also valid for this majorant series.

Similar analysis shows that majorant series for $\sum_{k=0}^{\infty} |\mathfrak{M}_{k,-1}|$ and $\sum_{k=0}^{\infty} |\mathfrak{M}_{k,n}|$, $n < 0$, with $\mathfrak{M}_{0,-1} = 1$, are obtained by replacing ω by ω' in (114) and (116), respectively, and the convergence condition for this series is

$$(2\omega')^2 e^{-1/T} < 1. \quad (117)$$

It can be shown for $n = 0$ that if we put

$$\eta = \mathfrak{M}_{0,-1}\omega' e^{-1/2T}, \tag{118}$$

$$\zeta = \mathfrak{M}_{0,1}\omega e^{-1/2T}, \tag{119}$$

then we obtain

$$\begin{aligned}
|\mathfrak{M}_{1,0}| &< \eta + \zeta, \\
|\mathfrak{M}_{3,0}| &< \eta(\omega')^2 e^{-1/T} + \zeta\omega^2 e^{-1/T}, \\
|\mathfrak{M}_{5,0}| &< 2[\eta(\omega')^4 e^{-2/T} + \zeta\omega^4 e^{-2/T}], \\
|\mathfrak{M}_{7,0}| &< 5[\eta(\omega')^6 e^{-3/T} + \zeta\omega^6 e^{-3/T}], \\
&\cdots \\
|\mathfrak{M}_{2l+1,0}| &< s_l[\eta(\omega')^{2l} e^{-l/T} + \zeta\omega^{2l} e^{-l/T}],
\end{aligned} \tag{120}$$

and so on. Therefore, a majorant series is given by

$$\sum_{k=0}^{\infty} |\mathfrak{M}_{k,0}| < \eta \sum_{l=0}^{\infty} [(2\omega')^2 e^{-1/T}]^l + \zeta \sum_{l=0}^{\infty} [(2\omega)^2 e^{-1/T}]^l. \tag{121}$$

The convergence condition for this series is Eqs. (115) and (117).

In the case of general $Q(E)$, we can easily find majorant series if we replace ω by ω'' and ω' by ω''' in related equations or inequalities. Consequently, the convergence condition for them is found to be

$$(2\omega'')^2 e^{-1/T} < 1, \tag{122}$$

$$(2\omega''')^2 e^{-1/T} < 1. \tag{123}$$

We can conclude that, for a sufficiently small T such that (122) and (123) hold, the Neumann series of Eq. (17) is convergent so long as ν belongs to the interval min $Q(E)$ to max $Q(E)$ and, hence, this interval forms a continuous eigenvalue spectra of the operator H.

Remark: A similar approach will be applicable to the kernel including more j terms in Eq. (24), or to the L^2 kernels, and similar results may be expected.

ACKNOWLEDGMENTS

The author would like to express his gratitude to Professor T. Inui, Professor H. Fujita of the University of Tokyo, and Dr. S. Iijima of NAIG research laboratory for stimulating and helpful discussion.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Solutions of the Einstein and Einstein–Maxwell Equations*

G. C. DEBNEY
Virginia Polytechnic Institute, Blacksburg, Virginia
AND
R. P. KERR AND A. SCHILD
The University of Texas at Austin, Texas

(Received 3 April 1969)

Algebraically degenerate solutions of the Einstein and Einstein–Maxwell equations are studied. Explicit solutions are obtained which contain two arbitrary functions of a complex variable, one function being associated with the gravitational field and the other mainly with the electromagnetic field.

1. INTRODUCTION

This paper is concerned with algebraically degenerate solutions of the Einstein and Einstein–Maxwell equations. It includes detailed derivations of gravitational vacuum fields, whose results were published previously by Kerr[1] and by Kerr and Schild.[2] It also includes new solutions of the Einstein–Maxwell equations.

Units are chosen such that the speed of light c and the Newtonian gravitational constant G are equal to unity:

$$c = G = 1. \tag{1.1}$$

Subject to this, the electromagnetic quantities are given in ordinary, nonrationalized, electrostatic units. The signature of space–time is chosen to be $+++-$.

Section 2 gives a short summary of the tetrad formalism which is used throughout the paper.

Section 3 gives the general theory of algebraically degenerate solutions of Einstein's vacuum equations $R_{ab} = 0$. There is only one weak restriction, that the shear, expansion, and rotation not all vanish for the null geodesic congruence which has the degenerate Debever–Penrose directions as tangents. Suitable coordinates and tetrads are obtained. The problem of solving ten partial differential equations for the ten components of the metric tensor as functions of four coordinates is reduced to solving four partial differential equations [the complex equation (3.65a) and the two real equations (3.65b), (3.65c)] for four unknown functions (the two complex functions Ω and M) of only three coordinates (the complex coordinate Y and the real coordinate ρ).

Section 4 gives a class of explicit vacuum solutions in terms of one arbitrary analytic function ϕ of one complex coordinate Y.

Sections 5 and 6 give classes of solutions of the Einstein–Maxwell equations for which the metric can be written in the form $g_{\mu\nu} = \eta_{\mu\nu} + 2he^3{}_\mu e^3{}_\nu$, where $\eta_{\mu\nu}$ is the metric of an auxiliary Minkowski space and $e^3{}_\mu$ is null. These solutions admit a one-parameter group of isometries which are translations in the auxiliary Minkowski space. The solutions contain two arbitrary analytic functions ϕ and ψ of one complex coordinate Y. In vacuum, when there is no electromagnetic field present, $\psi = 0$ and the solutions are identical with those of Sec. 4.

Section 7 examines a simple special solution which gives the gravitational and electromagnetic fields of a rotating charged body. The gyromagnetic ratio of the body turns out to be the same as that of the Dirac electron.

Most of the work described in this paper was completed in the period 1963–65. Since that time, further results have been obtained and will be published shortly. Kerr and Debney[3] have studied algebraically degenerate vacuum solutions which admit various groups of isometries. Robinson, Robinson, and Zund[4] have obtained a wider class of explicit vacuum solutions and Robinson, Schild, and Strauss[5] have found corresponding solutions with electromagnetic fields.

2. TETRAD FORMALISM

At each point of space–time we introduce a tetrad of four independent vectors $e_a{}^\mu$. The dual tetrad $e^a{}_\mu$ is

* This research was sponsored by the National Science Foundation (Grant GP-8868) and by the Aerospace Research Laboratories, Office of Aerospace Research, United States Air Force (Contract F 33615-68-C-1675).

[1] R. P. Kerr, Phys. Rev. Letters **11**, 237 (1963).

[2] R. P. Kerr and A. Schild, *Atti Del Convegno Sulla Relatività Generale: Problemi Dell' Energia E Onde Gravitazionali (Anniversary Volume, Fourth Centenary of Galileo's Birth)*, G. Barbéra, Ed. (Firenze, 1965), p. 173; *Applications of Nonlinear Partial Differential Equations in Mathematical Physics, Proceedings of Symposia in Applied Mathematics* (American Mathematical Society, Providence, R.I., 1965), Vol. XVII, p. 199.

[3] G. C. Debney and R. P. Kerr, report of work prior to publication (1968).

[4] I. Robinson, J. R. Robinson, and J. D. Zund, J. Math. Mech. **18**, 881 (1969); I. Robinson and J. R. Robinson, Intern. J. Theoret. Phys. (to be published).

[5] H. R. Strauss, Ph.D. dissertation, The University of Texas at Austin, 1969; I. Robinson, A. Schild, and H. R. Strauss, Intern. J. Theoret. Phys. (to be published).

defined by either of the equivalent relations

$$e_a{}^\mu e^b{}_\mu = \delta^b_a, \quad e_a{}^\mu e^a{}_\nu = \delta^\mu_\nu. \tag{2.1}$$

Latin suffixes are tetrad suffixes; they range and sum over 1, 2, 3, 4 and serve to label the different vectors of the tetrad. Greek suffixes are tensor suffixes; they also range and sum over 1, 2, 3, 4. A tensor $T_\mu{}^\nu{}_{\cdots}^{\cdots}$ is related to its tetrad components $T_a{}^b{}_{\cdots}^{\cdots}$ by either of the two equivalent relations

$$T_a{}^b{}_{\cdots}^{\cdots} = e_a{}^\mu e^b{}_\nu \cdots T_\mu{}^\nu{}_{\cdots}^{\cdots}, \quad T_\mu{}^\nu{}_{\cdots}^{\cdots} = e^a{}_\mu e_b{}^\nu \cdots T_a{}^b{}_{\cdots}^{\cdots}. \tag{2.2}$$

Tensor suffixes are raised and lowered by means of the metric tensors $g_{\mu\nu}$, $g^{\mu\nu}$ and tetrad suffixes by means of the tetrad components g_{ab}, g^{ab} of the metric tensors.

The directional derivative along a tetrad vector will be denoted by a comma or by the operator ∂_a:

$$T_{\cdots}^{\cdots}{}_{,a} = \partial_a T_{\cdots}^{\cdots} = e_a{}^\mu \partial T_{\cdots}^{\cdots}/\partial x^\mu. \tag{2.3}$$

The tetrad components of a covariant derivative will be denoted by a semicolon

$$T_a{}^b{}_{\cdots;c}^{\cdots} = e_a{}^\mu e^b{}_\nu \cdots e_c{}^\gamma T_\mu{}^\nu{}_{\cdots;\gamma}^{\cdots}. \tag{2.4}$$

They are given by

$$T_a{}^b{}_{\cdots;c}^{\cdots} = T_a{}^b{}_{\cdots,c}^{\cdots} - \Gamma^d{}_{ac} T_d{}^b{}_{\cdots}^{\cdots} - \cdots + \Gamma^b{}_{dc} T_a{}^d{}_{\cdots}^{\cdots} + \cdots, \tag{2.5}$$

where the $\Gamma^a{}_{bc}$ are the Ricci rotation coefficients

$$\Gamma^a{}_{bc} = -e^a{}_{\mu;\nu} e_b{}^\mu e_c{}^\nu \tag{2.6}$$

which take the place of Christoffel symbols in the tetrad formalism. The rotation coefficients also appear in the commutator of two directional derivatives along tetrad vectors

$$T_{\cdots}^{\cdots}{}_{,[ab]} = T_{\cdots}^{\cdots}{}_{,c} \Gamma^c{}_{[ab]}. \tag{2.7}$$

The tetrad vectors determine the linear differential forms

$$e^a = e^a{}_\mu \, dx^\mu, \tag{2.8}$$

in terms of which the metric form is given by

$$ds^2 = e_a e^a = g_{\mu\nu} \, dx^\mu \, dx^\nu. \tag{2.9}$$

The exterior or Grassman product of two linear differential forms $A = A_\mu \, dx^\mu$ and $B = B_\mu \, dx^\mu$ is defined by the skew-symmetric multiplication

$$A \wedge B = -B \wedge A = A_\mu B_\nu \, dx^\mu \wedge dx^\nu = A_{[\mu} B_{\nu]} \, dx^\mu \wedge dx^\nu.$$

The exterior derivative of a linear differential form is defined by

$$dA = A_{\nu,\mu} \, dx^\mu \wedge dx^\nu = A_{[\nu,\mu]} \, dx^\mu \wedge dx^\nu.$$

The rotation coefficients and the tetrad components of the curvature tensor determine the forms

$$\Gamma^a{}_b = \Gamma^a{}_{bc} e^c, \tag{2.10}$$

$$\mathcal{R}^a{}_b = R^a{}_{bcd} e^c \wedge e^d. \tag{2.11}$$

They are related to the tetrad forms by the Cartan formulas

$$de^a = e^b \wedge \Gamma^a{}_b = \Gamma^a{}_{bc} e^b \wedge e^c, \tag{2.12}$$

$$\tfrac{1}{2}\mathcal{R}^a{}_b = d\Gamma^a{}_b + \Gamma^a{}_m \wedge \Gamma^m{}_b, \tag{2.13}$$

which are easily derived.

Equations (2.12) determine the skew-symmetric part $\Gamma^a{}_{[bc]}$ of the rotation coefficients. The vanishing of the covariant derivatives of the metric tensor $g_{ab;c} = 0$ gives the symmetric part $\Gamma_{(ab)c}$ of the rotation coefficients

$$\Gamma_{(ab)c} = \tfrac{1}{2} g_{ab,c}, \quad \Gamma_{abc} = g_{am} \Gamma^m{}_{bc}. \tag{2.14}$$

The expressions for $\Gamma^a{}_{[bc]}$ and $\Gamma_{(ab)c}$ determine all rotation coefficients, and Eqs. (2.13) then determine all components of the curvature tensor.

We now limit ourselves to the case of rigid tetrads where the g_{ab} are constants. Then the rotation coefficients are skew-symmetric in the first two suffixes and are determined from Eqs. (2.12) by

$$\Gamma_{abc} = -\Gamma_{bac} = \Gamma_{a[bc]} + \Gamma_{b[ca]} - \Gamma_{c[ab]}. \tag{2.15}$$

In this paper we work throughout with null tetrads $e_a{}^\mu$ in four-dimensional space-time with signature $+++-$. e_1 and e_2 are complex conjugate null vectors,[6]

$$e_2 = \bar{e}_1, \tag{2.16}$$

e_3 and e_4 are real null vectors; their scalar products are given by

$$g_{ab} = e_a{}^\mu e_{b\mu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = g^{ab}. \tag{2.17}$$

Numerical suffixes will always refer to tetrad suffixes. These are raised or lowered by performing the permutation $1, 2, 3, 4 \to 2, 1, 4, 3$. The complex conjugate of a real geometrical object is obtained by performing the permutation $1, 2, 3, 4 \to 2, 1, 3, 4$ on the tetrad suffixes.

We now consider the congruence of null curves which have $e_4{}^\mu$ as tangents. The curves will be null geodesics if and only if

$$\Gamma_{424} = 0, \tag{2.18}$$

[6] Throughout, complex conjugation will be denoted by a bar.

which implies its complex conjugate $\Gamma_{414} = 0$. If this is satisfied, then the congruence of null geodesics will have zero shear[7] provided that

$$\Gamma_{422} = 0, \tag{2.19}$$

which implies its complex conjugate $\Gamma_{411} = 0$. The rates of expansion θ and of rotation ω of the null congruence[7] are given by

$$Z = \theta + i\omega = -\Gamma_{421}. \tag{2.20}$$

Since we shall single out the vector field e_4, those transformations from one null tetrad e_a to another e'_a, which leave the direction of e_4 unchanged, are of interest. They are given by

$$e'_1 = e^{-iB}(e_1 + Ce_4), \tag{2.21a}$$
$$e'_3 = e^{-A}(e_3 - \bar{C}e_1 - Ce_2 - C\bar{C}e_4), \tag{2.21b}$$
$$e'_4 = e^{A}e_4, \tag{2.21c}$$

where A and B are real and C is complex.

The rotation coefficients are determined by the independent forms Γ_{42}, $\Gamma_{12} + \Gamma_{34}$, Γ_{31}, which in turn determine their complex conjugates Γ_{41}, $-\Gamma_{12} + \Gamma_{34}$, Γ_{32}. The independent equations (2.13) are

$$d\Gamma_{42} + \Gamma_{42} \wedge (\Gamma_{12} + \Gamma_{34}) = \tfrac{1}{2}R_{42ab}e^a \wedge e^b, \tag{2.22a}$$
$$d(\Gamma_{12} + \Gamma_{34}) + 2\Gamma_{42} \wedge \Gamma_{31} = \tfrac{1}{2}(R_{12ab} + R_{34ab})e^a \wedge e^b, \tag{2.22b}$$
$$d\Gamma_{31} + (\Gamma_{12} + \Gamma_{34}) \wedge \Gamma_{31} = \tfrac{1}{2}R_{31ab}e^a \wedge e^b. \tag{2.22c}$$

They determine all components of the curvature tensor.

The independent components of the Ricci tensor

$$R_{ab} = R_{ba} = R^c{}_{abc} \tag{2.23}$$

are

$$R_{22} = 2R_{4223}, \tag{2.24a}$$
$$R_{24} = R_{4212} - R_{4234}, \tag{2.24b}$$
$$R_{44} = 2R_{4214}, \tag{2.24c}$$

$$R_{12} = R_{1212} + R_{3412} - 2R_{4231}, \tag{2.25a}$$
$$R_{34} = R_{1234} + R_{3434} - 2R_{4231}, \tag{2.25b}$$
$$R_{33} = -2R_{3132}, \tag{2.25c}$$
$$R_{23} = R_{1232} + R_{3432}, \tag{2.25d}$$

where R_{44}, R_{12}, R_{34}, R_{33} are real and R_{11}, R_{14}, R_{13} are the complex conjugates of R_{22}, R_{24}, R_{23}, respectively. In Eqs. (2.25a) and (2.25b), the cyclic identity $R_{4123} + R_{4231} + R_{4312} = 0$ was used.

Weyl's conformal curvature tensor C_{abcd} can be characterized[7] by the five complex quantities

$$C^{(5)} = 2R_{4242}, \tag{2.26a}$$
$$C^{(4)} = R_{1242} + R_{3442}, \tag{2.26b}$$
$$C^{(3)} = R_{1234} + R_{3434} - R_{34} + \tfrac{1}{6}R, \tag{2.26c}$$
$$C^{(2)} = R_{1231} + R_{3431}, \tag{2.26d}$$
$$C^{(1)} = 2R_{3131}. \tag{2.26e}$$

These five quantities determine the degeneracy of the conformal tensor. The condition for $e_4{}^\mu$ to be a Debever–Penrose null vector is $C^{(5)} = 0$; the conditions that it be a double, triple, or quadruple Debever–Penrose null vector are, respectively, $C^{(5)} = C^{(4)} = 0$, $C^{(5)} = C^{(4)} = C^{(3)} = 0$, $C^{(5)} = C^{(4)} = C^{(3)} = C^{(2)} = 0$. Thus the conditions for degeneracy are $C^{(5)} = C^{(4)} = 0$ or

$$R_{4242} = R_{1242} + R_{3442} = 0. \tag{2.27}$$

Einstein's equations for the vacuum gravitational field are

$$R_{ab} = 0. \tag{2.28}$$

When these equations are satisfied, the theorem of Goldberg and Sachs[7] states that space-time is algebraically degenerate if and only if it contains a shear-free congruence of null geodesics, the tangents to the congruence being multiple Debever–Penrose directions. Thus, in the vacuum, Eqs. (2.18) and (2.19) imply Eqs. (2.27), and vice versa.

3. THE DEGENERATE VACUUM SOLUTIONS

In this section we study algebraically degenerate space-times which satisfy Einstein's vacuum equations $R_{ab} = 0$. The multiple Debever–Penrose vector is taken to be the vector $e_4{}^\mu$ of a null tetrad. The Goldberg–Sachs theorem ensures that e_4 is geodesic and shear free. The only additional restriction we make is that the expansion and rotation rates of e_4 do not both vanish,[8] i.e., that

$$Z \neq 0. \tag{3.1}$$

The geodesic and shear-free condition for e_4 is $\Gamma_{424} = \Gamma_{422} = 0$. By a suitable choice of the field of null tetrads we also arrange to have

$$\Gamma_{423} = 0, \tag{3.2}$$

so that

$$\Gamma_{42} = -Ze^1. \tag{3.3}$$

That this is always possible can be seen as follows: Under the tetrad transformation (2.21),

$$\Gamma'_{423} = e^{iB}(\Gamma_{423} + \bar{C}Z). \tag{3.4}$$

[7] R. K. Sachs, Proc. Roy. Soc. (London) **A264**, 309 (1961); J. N. Goldberg and R. K. Sachs, Acta Phys. Polon. **22**, 13 (1962); I. Robinson and A. Schild, J. Math. Phys. **4**, 484 (1963).

[8] Degenerate vacuum solutions with $Z = 0$ have been studied by W. Kundt, Z. Physik **163**, 77 (1961).

Thus, choosing $\bar{C} = -Z^{-1}\Gamma_{423}$, we have $\Gamma'_{423} = 0$ and, dropping primes, we have Eq. (3.2). With Eq. (3.2) satisfied, one can still perform those tetrad transformations (2.21) for which $C = 0$. Under such a restricted change of the tetrad,

$$\Gamma'_{42} = e^{A+iB}\Gamma_{42}, \quad Z' = e^{A}Z. \tag{3.5}$$

The first half of the field equations

$$R_{22} = R_{24} = R_{44} = 0 \tag{3.6}$$

and the degeneracy conditions (2.27) state that R_{4231} is the only nonzero component of R_{42ab}, so that

$$\mathfrak{R}_{42} = R_{4231}e^3 \wedge e^1. \tag{3.7}$$

Equation (2.22a) now gives

$$\Gamma_{42} \wedge d\Gamma_{42} = 0. \tag{3.8}$$

This is the integrability condition for the existence of a complex scalar field Y such that $\Gamma_{42} = -e^{\phi}\,dY$. By using the freedom of restricted tetrad transformations (3.5), we arrange to have $\phi = 0$, so that

$$\Gamma_{42} = -dY \tag{3.9}$$

and, from Eq. (3.3),

$$e^1 = e_2 = Z^{-1}\,dY. \tag{3.10}$$

The function Y and its complex conjugate $\bar{Y}$ will serve as two space–time coordinates.

We can now compute

$$de_2 = \Gamma_{2ab}e^a \wedge e^b = Z(Z^{-1})_{,a}e^a \wedge e^1 \tag{3.11}$$

and

$$\begin{aligned}\Gamma_{42} \wedge (\Gamma_{12} + \Gamma_{34}) &= -Ze^1 \wedge (\Gamma_{12} + \Gamma_{34}) \\ &= -R_{4231}e^1 \wedge e^3,\end{aligned} \tag{3.12}$$

which follows from Eqs. (2.22a), (3.7), and from $d\Gamma_{42} = -ddY = 0$. These equations and their complex conjugates give

$$\Gamma_{124} = \Gamma_{344} = \Gamma_{311} = \Gamma_{314} = 0, \tag{3.13}$$

$$\Gamma_{122} = -\Gamma_{342} = Z(Z^{-1})_{,2}, \tag{3.14}$$

$$\Gamma_{123} + \Gamma_{312} = -\bar{Z}(\bar{Z}^{-1})_{,3}, \tag{3.15}$$

$$(Z^{-1})_{,4} = 1, \tag{3.16}$$

$$R_{4231} = Z(\Gamma_{123} + \Gamma_{343}). \tag{3.17}$$

Equations (2.18) and (3.13) imply that

$$\Gamma_{ab4} = 0. \tag{3.18}$$

Geometrically, this means that the null tetrad $e_a{}^{\mu}$ is parallelly propagated along the congruence of null geodesics which have $e_4{}^{\mu}$ as tangents.

Our results, thus far, are a consequence of the algebraic degeneracy of space–time and of the vacuum field equations (3.6), which state that five of the ten real components of the Ricci tensor R_{ab} vanish. We have also introduced two of the required four real coordinates. What remains to be done is to choose two more coordinates and to satisfy the other half of the field equations, which require that the components (2.25) of the Ricci tensor vanish.

We now introduce a third and real coordinate ρ by the conditions

$$\rho_{,3} = 1, \quad \rho_{,4} = 0. \tag{3.19}$$

The integrability conditions for this system of differential equations,

$$\rho_{,43} - \rho_{,34} = 2\Gamma^{a}_{[43]}\rho_{,a} = 0, \tag{3.20}$$

are satisfied by virtue of Eqs. (3.2) and (3.18). It follows that

$$d\rho = \rho_{,1}e^1 + \rho_{,2}e^2 + e^3,$$

so that

$$e^3 = e_4 = d\rho + \Omega\,dY + \bar{\Omega}\,d\bar{Y}, \tag{3.21}$$

where Ω is a complex scalar field.

Using the previous results, Eq. (2.12) for de_4 gives

$$\Omega_{,4} = 0, \tag{3.22}$$

$$Z(Z^{-1})_{,2} = \bar{Z}\bar{\Omega}_{,3}, \tag{3.23}$$

$$Z - \bar{Z} = \bar{Z}\bar{\Omega}_{,1} - Z\Omega_{,2}. \tag{3.24}$$

We now have the following expressions for the independent combinations of the rotation coefficients:

$$\Gamma_{42} = -Ze^1, \tag{3.25}$$

$$\Gamma_{12} + \Gamma_{34} = -2Z\Omega_{,3}e^1 + (\Gamma_{123} + \Gamma_{343})e^3, \tag{3.26}$$

$$\Gamma_{31} = \Gamma_{312}e^2 + \Gamma_{313}e^3. \tag{3.27}$$

The $e^3 \wedge e^2$ term in Eq. (2.22b) gives

$$R_{23} = -Z^{-1}[Z(\Gamma_{123} + \Gamma_{343})]_{,2}. \tag{3.28}$$

The $e^3 \wedge e^4$ term in Eq. (2.22b) gives $R_{1234} + R_{3434} = -(\Gamma_{123} + \Gamma_{343})_{,4}$, and Eq. (3.17) gives R_{4231}. Thus, by Eq. (2.25b),

$$R_{34} = -Z^2[Z^{-2}(\Gamma_{123} + \Gamma_{343})]_{,4}. \tag{3.29}$$

Similarly, the $e^1 \wedge e^2$ term in Eq. (2.22b), with Eqs. (3.17) and (3.15), gives

$$R_{12} = 2\,\mathrm{Re}\,[Z\Omega_{,32} + Z\bar{Z}(\bar{Z}^{-1})_{,3} - \bar{Z}(\Gamma_{123} + \Gamma_{343})], \tag{3.30}$$

where Re denotes the real part. We may note that R_{12} and R_{34} are not independent. The commutation

relations (2.7) for directional derivatives give

$$\Omega_{,34} = 0, \qquad \Omega_{,324} = -Z\Omega_{,32},$$
$$(Z^{-1}\Omega_{,32})_{,4} = 0, \quad (Z^{-1})_{,34} = -\Gamma_{343}. \tag{3.31}$$

Using these, it is easy to show that

$$(Z^{-1}\bar{Z}^{-1}R_{12})_{,4} = (Z^{-1} + \bar{Z}^{-1})R_{34}, \tag{3.32}$$

so that the field equation $R_{12} = 0$ implies $R_{34} = 0$.

We now introduce a fourth coordinate r which is real. At the moment we only require that $r_{,4} = 1$, so that

$$e^4 = e_3 = dr + \beta\, dY + \bar{\beta}\, d\bar{Y} + He^3, \tag{3.33}$$

where β is complex and H real. From the expressions for $e^a = e^a{}_\mu\, dx^\mu$, we can read off the covariant components $e^a{}_\mu$ of our tetrad vectors in the coordinates $x^\mu = (Y, \bar{Y}, \rho, r)$. The reciprocal of the matrix $(e^a{}_\mu)$ yields, by Eq. (2.1), the contravariant components $e_a{}^\mu$ and thus the directional derivatives $_{,a} = \partial_a = e_a{}^\mu \partial_\mu$. Since numerical suffixes always refer to tetrad components, we shall use the explicit notation $\partial_\mu = (\partial_Y, \partial_{\bar{Y}}, \partial_\rho, \partial_r)$ for differentiation with respect to the coordinates. We shall also use a dot for ∂_ρ, e.g., $\partial_\rho \Omega = \dot{\Omega}$. The directional derivatives are

$$\partial_1 = Z(D - \beta\partial_r), \tag{3.34a}$$
$$\partial_2 = \bar{Z}(\bar{D} - \bar{\beta}\partial_r), \tag{3.34b}$$
$$\partial_3 = \partial_\rho - H\partial_r, \tag{3.34c}$$
$$\partial_4 = \partial_r, \tag{3.34d}$$

where

$$D = \partial_Y - \Omega\partial_\rho, \quad \bar{D} = \partial_{\bar{Y}} - \bar{\Omega}\partial_\rho. \tag{3.35}$$

The commutation relations for the operators ∂_ρ, D, $\bar{D}$ are

$$\partial_\rho D - D\partial_\rho = -\dot{\Omega}\partial_\rho, \quad \partial_\rho \bar{D} - \bar{D}\partial_\rho = -\dot{\bar{\Omega}}\partial_\rho,$$
$$D\bar{D} - \bar{D}D = (\bar{D}\Omega - D\bar{\Omega})\partial_\rho. \tag{3.36}$$

From $(Z^{-1})_{,4} = 1$ we deduce that

$$Z^{-1} = r + \delta, \tag{3.37}$$

where δ is independent of r, i.e.,

$$\delta_{,4} = 0. \tag{3.38}$$

Equation (3.24) can be written

$$\bar{Z}^{-1} - Z^{-1} = Z^{-1}\bar{\Omega}_{,1} - \bar{Z}^{-1}\Omega_{,2}$$

and, since $\Omega_{,4} = \partial_r \Omega = 0$, this reduces to

$$\bar{\delta} - \delta = D\bar{\Omega} - \bar{D}\Omega. \tag{3.39}$$

Subject to this restriction, we can choose δ arbitrarily as a function of Y, $\bar{Y}$, and ρ, different choices of δ corresponding to different choices of the coordinate r. We make the choice[9]

$$\delta = -D\bar{\Omega}, \quad \bar{\delta} = -\bar{D}\Omega. \tag{3.40}$$

We can solve Eq. (3.23) for β and, using Eq. (3.36), we obtain

$$\beta = -Z^{-1}\dot{\Omega} - \bar{D}D\Omega. \tag{3.41}$$

Using the previous results, Eq. (2.12) for de_3 gives the remaining rotation coefficients:

$$\Gamma_{313} = -Z\beta_{,3} - ZH\dot{\Omega} + H_{,1}, \tag{3.42}$$
$$\Gamma_{343} = H_{,4}, \tag{3.43}$$
$$\Gamma_{123} = -(Z - \bar{Z})H + \tfrac{1}{2}(Z\beta_{,2} - \bar{Z}\bar{\beta}_{,1} - Z\partial_\rho D\bar{\Omega} + \bar{Z}\partial_\rho \bar{D}\Omega), \tag{3.44}$$
$$\Gamma_{312} = -\Gamma_{123} + \bar{Z}H + \bar{Z}\partial_\rho \bar{D}\Omega. \tag{3.45}$$

The last expression comes from Eq. (3.15) and

$$(\bar{Z}^{-1})_{,3} = -H + \dot{\bar{\delta}} = -H - \partial_\rho \bar{D}\Omega. \tag{3.46}$$

The field equation $R_{34} = 0$ has, by Eq. (3.29), the first integral

$$\Gamma_{123} + \Gamma_{343} = -MZ^2, \tag{3.47}$$

where M is a complex function of Y, $\bar{Y}$, ρ, i.e.,

$$M_{,4} = 0. \tag{3.48}$$

Using Eq. (3.30), the field equation $R_{12} = 0$ determines H:

$$H = \dot{\Omega}\dot{\bar{\Omega}} + \tfrac{1}{2}MZ + \tfrac{1}{2}\bar{M}\bar{Z}. \tag{3.49}$$

Substituting from Eqs. (3.43) and (3.44) into Eq. (3.47), we obtain the imaginary part of M in terms of Ω:

$$\mathrm{Im}\, M = \mathrm{Im}\, \bar{D}\bar{D}D\Omega. \tag{3.50}$$

The field equation $R_{23} = 0$ yields, by Eq. (3.28),

$$\bar{D}M = 3M\dot{\bar{\Omega}}, \tag{3.51}$$

where we have used Eq. (3.23).

There remains one field equation to be satisfied, namely, $R_{33} = 0$. The $e^3 \wedge e^2$ term in Eq. (2.22c) gives

$$-\tfrac{1}{2}R_{33} = \Gamma_{312,3} - \Gamma_{313,2} + \bar{Z}(\bar{Z}^{-1})_{,3}\Gamma_{312} + \bar{Z}\dot{\bar{\Omega}}\Gamma_{313} + \Gamma_{313}(\Gamma_{123} + \Gamma_{343}). \tag{3.52}$$

The calculation of R_{33} is straightforward but long. It can be simplified by using the identities

$$(R^{mn} - \tfrac{1}{2}g^{mn}R)_{;n} = 0.$$

We assume all previous results, so that all components of R_{mn} other than $R_{33} = R^{44}$ vanish. Then the identities reduce to $(Z^{-1}\bar{Z}^{-1}R_{33})_{,4} = 0$, so that

$$R_{33} = \mu Z\bar{Z}, \quad \mu_{,4} = 0. \tag{3.53}$$

[9] In Kerr's paper (Ref. 1), a different choice was made for r, namely $r = \mathrm{Re}\,(Z^{-1})$. The present choice of r results in somewhat simpler expressions.

Among the variables Ω, M, Z, only Z depends on the coordinate r. Thus in Eq. (3.52) we can ignore all terms of third or higher order in Z and $\bar{Z}$; these terms must vanish as a consequence of the previous field equations. We must keep terms in Z, $\bar{Z}$, Z^2, $\bar{Z}^2$, $Z\bar{Z}$, since

$$Z - \bar{Z} = Z\bar{Z}(D\bar{\Omega} - \bar{D}\Omega), \quad Z^2 = Z\bar{Z} + O_3, \quad \bar{Z}^2 = Z\bar{Z} + O_3, \tag{3.54}$$

where O_n refers to terms of order n or higher in Z and $\bar{Z}$. From Eqs. (3.49) and (3.47) we find that, to the required order,

$$\Gamma_{313} = \ddot{\Omega} + Z\partial_\rho \bar{D}D\Omega + Z\dot{\Omega}D\dot{\Omega} - Z\dot{\Omega}^2\dot{\bar{\Omega}} + O_2, \tag{3.55}$$

$$\Gamma_{343} = -\tfrac{1}{2}MZ^2 - \tfrac{1}{2}\bar{M}\bar{Z}^2, \tag{3.56}$$

$$\Gamma_{123} = -\tfrac{1}{2}MZ^2 + \tfrac{1}{2}\bar{M}\bar{Z}^2, \tag{3.57}$$

$$\Gamma_{312} = Z\bar{D}\dot{\Omega} + \tfrac{1}{2}MZ^2 + \tfrac{1}{2}MZ\bar{Z}, \tag{3.58}$$

$$Z_{,1} = -Z^2\dot{\Omega} + O_3, \quad Z_{,2} = -Z\bar{Z}\dot{\bar{\Omega}}, \quad Z_{,3} = Z^2D\dot{\Omega} + O_3. \tag{3.59}$$

Substituting into Eq. (3.52), we find that the last field equation $R_{33} = 0$ reduces to

$$\partial_\rho(M - \bar{D}\bar{D}D\Omega) = (\partial_\rho D\Omega)(\partial_\rho \bar{D}\bar{\Omega}). \tag{3.60}$$

We shall now summarize our results. The most general vacuum metric, which is algebraically degenerate and which has a nonzero complex expansion rate Z, is given by

$$ds^2 = g_{ab}e^a e^b = 2e^1e^2 + 2e^3e^4, \tag{3.61}$$

$$e^1 = Z^{-1}\,dY, \quad e^2 = \bar{Z}^{-1}\,d\bar{Y}, \tag{3.62a}$$

$$e^3 = d\rho + \Omega\,dY + \bar{\Omega}\,d\bar{Y}, \tag{3.62b}$$

$$e^4 = dr + \beta\,dY + \bar{\beta}\,d\bar{Y} + He^3, \tag{3.62c}$$

where Y and $\bar{Y}$ are complex coordinates, ρ and r are real coordinates, and where

$$Z^{-1} = r - D\bar{\Omega}, \tag{3.63a}$$

$$\beta = -Z^{-1}\dot{\Omega} - \bar{D}D\Omega, \tag{3.63b}$$

$$H = \dot{\Omega}\dot{\bar{\Omega}} + \tfrac{1}{2}MZ + \tfrac{1}{2}\bar{M}\bar{Z}, \tag{3.63c}$$

$$\dot{\Omega} = \partial_\rho\Omega, \quad D = \partial_Y - \Omega\partial_\rho. \tag{3.64}$$

The expressions Ω and M are functions of only three coordinates Y, $\bar{Y}$, ρ. They satisfy the field equations

$$\bar{D}M = 3M\dot{\bar{\Omega}}, \tag{3.65a}$$

$$\mathrm{Im}\,(M - \bar{D}\bar{D}D\Omega) = 0, \tag{3.65b}$$

$$\partial_\rho(M - \bar{D}\bar{D}D\Omega) = (\partial_\rho D\Omega)(\partial_\rho \bar{D}\bar{\Omega}). \tag{3.65c}$$

We can now compute the remaining components $C^{(3)}$, $C^{(2)}$, and $C^{(1)}$ of the curvature tensor. They are

$$C^{(3)} = -2MZ^3, \tag{3.66a}$$

$$C^{(2)} = 2Z^2\bar{D}\partial_\rho D\Omega + \text{terms which vanish if } C^{(3)} = 0, \tag{3.66b}$$

$$C^{(1)} = -2Z\partial_\rho\partial_\rho D\Omega + \text{terms which vanish if } C^{(3)} = C^{(2)} = 0. \tag{3.66c}$$

Thus our degenerate vacuum metric is flat if and only if

$$M = \bar{D}\partial_\rho D\Omega = \partial_\rho\partial_\rho D\Omega = 0. \tag{3.67}$$

4. SOME EXPLICIT VACUUM SOLUTIONS

Our field equations can be solved completely in the special case where

$$D\Omega = \partial_Y\Omega - \Omega\dot{\Omega} = 0. \tag{4.1}$$

Then Eqs. (3.65b) and (3.65c) give

$$M = \bar{M}, \quad \dot{M} = 0, \tag{4.2}$$

so that M is a real function of the two coordinates Y and $\bar{Y}$. The complex conjugate of Eq. (3.65a) gives

$$\partial_Y M^{-\frac{1}{3}} = -M^{-\frac{1}{3}}\dot{\Omega}, \tag{4.3}$$

so that

$$\ddot{\Omega} = 0. \tag{4.4}$$

Differentiating Eq. (4.3) and using Eqs. (4.1) and (4.4), we obtain

$$\partial_Y\partial_Y M^{-\frac{1}{3}} = 0. \tag{4.5}$$

This equation and its complex conjugate

$$\partial_{\bar{Y}}\partial_{\bar{Y}} M^{-\frac{1}{3}} = 0$$

imply that M has the form

$$M = mP^{-3}, \quad P = pY\bar{Y} + qY + \bar{q}\bar{Y} + c, \tag{4.6}$$

where m, p, c are real constants and q is a complex constant. Equations (4.1) and (4.3) give

$$\dot{\Omega} = -P^{-1}P_Y, \quad \partial_Y(P\Omega) = 0. \tag{4.7}$$

Since $P_Y = p\bar{Y} + q$ is a function of $\bar{Y}$ only, we can integrate these equations and obtain

$$\Omega = P^{-1}(-P_Y\rho + \bar{\phi}), \quad \bar{\Omega} = P^{-1}(-P_{\bar{Y}}\rho + \phi), \tag{4.8}$$

where ϕ is an arbitrary function of Y alone

$$\phi = \phi(Y), \tag{4.9}$$

i.e., ϕ is an arbitrary analytic function of the complex variable Y. We can now calculate Z^{-1}, β, and H.

Summarizing, we find that the metric of Eqs. (3.61) and (3.62) is given explicitly by

$$P = pY\bar{Y} + qY + \bar{q}\bar{Y} + c, \tag{4.10a}$$

$$Z^{-1} = r + P^{-1}(p\rho - \phi') + P^{-2}(P_Y\phi - P_{\bar{Y}}\bar{\phi}),$$

$$\phi' = \partial_Y\phi, \tag{4.10b}$$

$$\Omega = P^{-1}(-P_Y\rho + \bar{\phi}), \tag{4.10c}$$

$$\beta = P^{-1}P_YZ^{-1}, \tag{4.10d}$$

$$H = P^{-2}P_YP_{\bar{Y}} + \tfrac{1}{2}P^{-3}m(Z + \bar{Z}). \tag{4.10e}$$

It can be shown that each of these metrics has a field of Killing vectors and therefore admits a one-parameter group of motions. It can also be shown that, without loss of generality, P can be normalized so that it is proportional to $1 + Y\bar{Y}$, $1 - Y\bar{Y}$, or 1. We shall not establish these facts now, because they will follow from the work of the next sections.

A slightly simpler form of these explicit solutions is obtained by performing the transformations

$$e'^1 = e^1, \quad e'^2 = e^2, \quad e'^3 = P^{-1}e^3, \quad e'^4 = Pe^4, \tag{4.11a}$$

$$Z' = P^{-1}Z, \tag{4.11b}$$

$$Y' = Y, \quad \rho' = P^{-1}\rho, \quad r' = Pr + p\rho. \tag{4.11c}$$

Dropping primes, the metric is given by

$$ds^2 = 2e^1e^2 + 2e^3e^4, \tag{4.12a}$$

$$e^1 = P^{-1}Z^{-1}\,dY, \quad e^2 = P^{-1}\bar{Z}^{-1}\,d\bar{Y}, \tag{4.12b}$$

$$e^3 = d\rho + \Omega\,dY + \bar{\Omega}\,d\bar{Y}, \tag{4.12c}$$

$$e^4 = dr + \beta\,dY + \bar{\beta}\,d\bar{Y} + He^3, \tag{4.12d}$$

where Ω, β, and H are new expressions given by

$$P = pY\bar{Y} + qY + \bar{q}\bar{Y} + c, \tag{4.13a}$$

$$Z^{-1} = r - \phi' + P^{-1}(P_Y\phi - P_{\bar{Y}}\bar{\phi}), \tag{4.13b}$$

$$\Omega = P^{-2}\bar{\phi}, \tag{4.13c}$$

$$\beta = P^{-1}[p\bar{\phi} - P_Y\phi' + P^{-1}P_Y(P_Y\phi - P_{\bar{Y}}\bar{\phi})], \tag{4.13d}$$

$$H = q\bar{q} - pc + \tfrac{1}{2}m(Z + \bar{Z}). \tag{4.13e}$$

If we make the special choice[10]

$$P = 2^{-\frac{1}{2}}(1 + Y\bar{Y}), \quad \phi = -iaY, \tag{4.14}$$

we obtain the exterior gravitational field of a rotating body of mass m and angular momentum am. Replacing Y by real coordinates θ and φ, given by

$$Y = e^{i\varphi}\tan\tfrac{1}{2}\theta, \tag{4.15}$$

the metric becomes

$$ds^2 = (r^2 + a^2\cos^2\theta)(d\theta^2 + \sin^2\theta\,d\varphi^2) + 2(d\rho - a\sin^2\theta\,d\varphi)(dr - a\sin^2\theta\,d\varphi) - \left(1 - \frac{2mr}{r^2 + a^2\cos^2\theta}\right)(d\rho - a\sin^2\theta\,d\varphi)^2. \tag{4.16}$$

5. EINSTEIN-MAXWELL FIELDS FOR $\eta + 2h(e^3)^2$ METRICS

We shall here consider space-times where coordinates exist in which the metric has the form

$$g_{\mu\nu} = \eta_{\mu\nu} + 2he^3{}_\mu e^3{}_\nu, \tag{5.1}$$

where $\eta_{\mu\nu}$ is the metric of Minkowski space in Cartesian coordinates and $e^3{}_\mu$ is null:

$$g^{\mu\nu}e^3{}_\mu e^3{}_\nu = 0. \tag{5.2}$$

It is easy to see that $-g = 1$,

$$g^{\mu\nu} = \eta^{\mu\nu} - 2he^{3\mu}e^{3\nu}, \tag{5.3}$$

and that $e^3{}_\mu$ is also null with respect to the auxiliary Minkowski metric:

$$\eta^{\mu\nu}e^3{}_\mu e^3{}_\nu = 0. \tag{5.4}$$

We choose null coordinates in Minkowski space which are related to the usual rectangular Cartesians x, y, z, t by

$$2^{\frac{1}{2}}\zeta = x + iy, \quad 2^{\frac{1}{2}}\bar{\zeta} = x - iy,$$
$$2^{\frac{1}{2}}u = z + t, \quad 2^{\frac{1}{2}}v = z - t. \tag{5.5}$$

Then the metric (5.1) gives the line element

$$ds^2 = 2\,d\zeta\,d\bar{\zeta} + 2\,du\,dv + 2h(e^3)^2. \tag{5.6}$$

A general field of real null directions in Minkowski space is given by

$$e^3 = du + \bar{Y}\,d\zeta + Y\,d\bar{\zeta} - Y\bar{Y}\,dv, \tag{5.7}$$

where Y is an arbitrary complex function of position.[11] A null tetrad, with scalar products given by Eq. (2.17), is completed as follows:

$$e^1 = d\zeta - Y\,dv, \quad e^2 = d\bar{\zeta} - \bar{Y}\,dv, \tag{5.8a}$$

$$e^4 = dv + he^3. \tag{5.8b}$$

The dual tetrad $e_a{}^\mu$ and the directional derivatives $_{,a} = \partial_a = e_a{}^\mu\partial_\mu$ can now be calculated to be

$$\partial_1 = \partial_\zeta - \bar{Y}\partial_u, \quad \partial_2 = \partial_{\bar{\zeta}} - Y\partial_u, \tag{5.9a}$$

$$\partial_3 = \partial_u - h\partial_4, \tag{5.9b}$$

$$\partial_4 = \partial_v + Y\partial_\zeta + \bar{Y}\partial_{\bar{\zeta}} - Y\bar{Y}\partial_u. \tag{5.9c}$$

[10] The sign of the parameter a is opposite to that used in earlier papers (Refs. 1 and 2). The sign of a has been chosen here so that, in the asymptotically flat coordinates of Sec. 7, the z component of the angular momentum will be $+am$.

[11] The complex function Y is the ratio of the two components of the spinor which corresponds to the null vector e^3. The special case $e^3 = dv$ can be included by a limiting process, e.g., by starting with $(1 + Y\bar{Y})^{-1}e^3$ and then letting $Y \to \infty$.

We shall find solutions of the Einstein–Maxwell equations for spaces with a metric of the form (5.1) or (5.6) where the complex expansion Z is nonzero:

$$Z = -\Gamma_{421} \neq 0. \tag{5.10}$$

In the case of vacuum, when there is no electromagnetic field, we shall find all such solutions and show that they are identical with the explicit vacuum solutions of Sec. 4. When an electromagnetic field is present, we shall make some additional assumptions, so that our solutions will be more restricted.

The gravitational equations for interacting gravitational and electromagnetic fields are

$$R_{ab} - \tfrac{1}{2}g_{ab}R = -8\pi T_{ab}, \tag{5.11a}$$

$$T_{ab} = (1/4\pi)(F_{ac}F_b{}^c - \tfrac{1}{4}g_{ab}F_{cd}F^{cd}). \tag{5.11b}$$

These equations have to be supplemented by Maxwell's equations. Since $T = 0$, it follows that

$$R = 2(R_{12} + R_{34}) = 0, \tag{5.12}$$

and the gravitational equations become

$$R_{ab} = -P_{ab}, \quad P_{ab} = 2(F_{ac}F_b{}^c - \tfrac{1}{4}g_{ab}F_{cd}F^{cd}). \tag{5.13}$$

From the expressions (5.7) and (5.8), we calculate the rotation coefficients

$$de^1 = -Y_{,a}e^a \wedge (e^4 - he^3),$$

$$de^2 = -\bar{Y}_{,a}e^a \wedge (e^4 - he^3), \tag{5.14a}$$

$$de^3 = Y_{,a}e^a \wedge e^2 + \bar{Y}_{,a}e^a \wedge e^1, \tag{5.14b}$$

$$de^4 = hY_{,a}e^a \wedge e^2 + h\bar{Y}_{,a}e^a \wedge e^1 + h_{,a}e^a \wedge e^3, \tag{5.14c}$$

$$\Gamma_{42} = -Y_{,a}e^a - hY_{,4}e^3 = -dY - hY_{,4}e^3, \tag{5.15a}$$

$$\Gamma_{12} + \Gamma_{34} = h\bar{Y}_{,4}e^1 + hY_{,4}e^2 + [h_{,4} + h(\bar{Y}_{,2} - Y_{,1})]e^3, \tag{5.15b}$$

$$\Gamma_{31} = h\bar{Y}_{,1}e^1 + hY_{,1}e^2 + (h_{,1} - h\bar{Y}_{,3})e^3. \tag{5.15c}$$

The $e^1 \wedge e^4$ term in Eq. (2.22a) gives

$$\tfrac{1}{2}R_{44} = R_{4214} = 2hY_{,4}\bar{Y}_{,4}. \tag{5.16}$$

In the vacuum $R_{44} = 0$, and it follows that

$$Y_{,4} = -\Gamma_{424} = 0, \tag{5.17}$$

so that the curves which have $e_4 = e^3$ as tangents are null geodesics. We have ignored the possibility $h = 0$, since our metric then reduces to the trivial case of flat Minkowski space.

When there is an electromagnetic field present, we introduce Eq. (5.17) as an explicit assumption, and we limit ourselves to those solutions of the Einstein–Maxwell equations which satisfy this additional condition. The field equation $R_{44} = -P_{44}$ then gives $F_{41}F_{42} = 0$ and, since $F_{41} = \bar{F}_{42}$,

$$F_{42} = F_{41} = 0. \tag{5.18}$$

The expressions (5.15) simplify to

$$\Gamma_{42} = -Y_{,a}e^a = -dY, \tag{5.19a}$$

$$\Gamma_{12} + \Gamma_{34} = [h_{,4} + (\bar{Z} - Z)h]e^3, \tag{5.19b}$$

$$\Gamma_{31} = h\bar{Y}_{,1}e^1 + Zhe^2 + (h_{,1} - h\bar{Y}_{,3})e^3, \tag{5.19c}$$

where Z is the complex expansion

$$Z = -\Gamma_{421} = Y_{,1} \tag{5.20}$$

and the shear of e_4 is determined by

$$\Gamma_{422} = Y_{,2}. \tag{5.21}$$

Since $d\Gamma_{42} = -ddY = 0$, Eqs. (2.22a) become

$$R_{4241} = R_{4242} = R_{4234} = R_{4212} = 0, \tag{5.22}$$

$$R_{4231} = Y_{,1}[h_{,4} + (\bar{Z} - Z)h], \tag{5.23}$$

$$R_{4232} = Y_{,2}[h_{,4} + (\bar{Z} - Z)h]. \tag{5.24}$$

The first of these equations imply that space–time is algebraically degenerate. For vacuum, the Goldberg–Sachs theorem then states that the shear of e_4 vanishes and $Y_{,2} = 0$. Later we shall show that this is also true when there is an electromagnetic field present.

Equations (5.22) also imply $R_{24} = 0$. Since

$$P_{24} = 2(F_{21}F_{42} + F_{24}F_{43}) = 0,$$

by Eq. (5.18), the field equation $R_{24} = -P_{24}$ is now satisfied. Similarly, $P_{22} = 0$ and the field equation $R_{22} = 0$ becomes

$$Y_{,2}[h_{,4} + (\bar{Z} - Z)h] = 0. \tag{5.25}$$

Before continuing with the calculation of the gravitational equations, we turn to Maxwell's equations. The electromagnetic field will be represented by a complex tensor

$$\mathcal{F}_{\mu\nu} = -\mathcal{F}_{\nu\mu} = F_{\mu\nu} + \tfrac{1}{2}i\eta_{\mu\nu\rho\sigma}F^{\rho\sigma}, \tag{5.26}$$

where $\eta_{\mu\nu\rho\sigma}$ is completely skew-symmetric and equal to $(-g)^{\frac{1}{2}}$ when $\mu\nu\rho\sigma = 1234$. The corresponding null-tetrad components are

$$\mathcal{F}_{ab} = -\mathcal{F}_{ba} = F_{ab} + \tfrac{1}{2}i\eta_{abcd}F^{cd}, \tag{5.27}$$

where η_{abcd} is completely skew-symmetric and $\eta_{1234} = i$. Explicitly, we have

$$\mathcal{F}_{12} = \mathcal{F}_{34} = F_{12} + F_{34}, \quad \mathcal{F}_{31} = 2F_{31},$$
$$\mathcal{F}_{23} = \mathcal{F}_{14} = \mathcal{F}_{24} = 0, \tag{5.28}$$

where we have used $F_{42} = 0$. Thus the electromagnetic field is described by only two complex components, $\mathcal{F}_{12}$ and $\mathcal{F}_{31}$.

Maxwell's equations are

$$\mathcal{F}^{ab}{}_{;b} = \mathcal{F}^{ab}{}_{,b} + \Gamma^{a}{}_{cb}\mathcal{F}^{cb} + \Gamma^{b}{}_{cb}\mathcal{F}^{ac} = 0. \quad (5.29)$$

Explicitly, these equations are

$$\mathcal{F}_{12,2} - 2Y_{,3}\mathcal{F}_{12} - Y_{,2}\mathcal{F}_{31} = 0, \quad (5.30a)$$

$$\mathcal{F}_{12,1} - \mathcal{F}_{31,4} - Z\mathcal{F}_{31} = 0, \quad (5.30b)$$

$$\mathcal{F}_{12,4} + 2Z\mathcal{F}_{12} = 0, \quad (5.30c)$$

$$\mathcal{F}_{12,3} + \mathcal{F}_{31,2} - 2Zh\mathcal{F}_{12} - Y_{,3}\mathcal{F}_{31} = 0. \quad (5.30d)$$

We now calculate $R_{1212} + R_{3412}$ and $R_{1234} + R_{3434}$ by Eq. (2.22b) and obtain the gravitational field equations $R_{12} = -P_{12}$ and $R_{34} = -P_{34}$. They are

$$-(Z + \bar{Z})[h_{,4} + (\bar{Z} - Z)h] - 2Z^2h + 2Y_{,2}\bar{Y}_{,1}h = -\mathcal{F}_{12}\bar{\mathcal{F}}_{12}, \quad (5.31)$$

$$-[h_{,4} + (\bar{Z} - Z)h]_{,4} - 2Z[h_{,4} + (\bar{Z} - Z)h] = \mathcal{F}_{12}\bar{\mathcal{F}}_{12}. \quad (5.32)$$

Let us now assume that $Y_{,2} \neq 0$. Then Eq. (5.25) gives $[h_{,4} + (\bar{Z} - Z)h] = 0$, Eq. (5.32) gives $\mathcal{F}_{12} = 0$, and Eq. (5.30a) gives $\mathcal{F}_{31} = 0$. Thus we have vacuum, and we know that there the algebraic degeneracy of space-time implies $Y_{,2} = 0$. This contradiction shows that

$$Y_{,2} = 0. \quad (5.33)$$

The field equation (5.25) is now satisfied. Also, the commutation relation (2.7) for $Y_{,[14]}$ gives

$$Z_{,4} = -Z^2. \quad (5.34)$$

Adding the field equations (5.31) and (5.32), we obtain the following equation for h:

$$h_{,44} + 2(Z + \bar{Z})h_{,4} + 2Z\bar{Z}h = 0. \quad (5.35)$$

It is easy to check that $Z\bar{Z}$ and $Z + \bar{Z}$ are particular solutions and, therefore, the general solution is

$$h = \tfrac{1}{2}M(Z + \bar{Z}) + BZ\bar{Z}, \quad M_{,4} = B_{,4} = 0, \quad (5.36)$$

where M and B are real. Maxwell's equation (5.30c) can also be solved for $\mathcal{F}_{12}$ and gives

$$\mathcal{F}_{12} = AZ^2, \quad A_{,4} = 0. \quad (5.37)$$

Substituting these results into Eq. (5.31), we find that

$$B = -\tfrac{1}{2}A\bar{A}, \quad (5.38)$$

so that

$$h = \tfrac{1}{2}M(Z + \bar{Z}) - \tfrac{1}{2}A\bar{A}Z\bar{Z}. \quad (5.39)$$

Using the commutation relations

$$(AZ)_{,14} - (AZ)_{,41} = -Z(AZ)_{,1},$$

$$(AZ)_{,12} - (AZ)_{,21} = (Z - \bar{Z})(AZ)_{,3} + h(Z - \bar{Z})(AZ)_{,4},$$

$$Z_{,2} = (Z - \bar{Z})Y_{,3}, \quad Z_{,3} = Y_{,31} + hZ^2 - Y_{,3}\bar{Y}_{,3}, \quad (5.40)$$

Maxwell's equations (5.30b), (5.30a), (5.30d) give, in that order,

$$\mathcal{F}_{31} = \gamma Z - (AZ)_{,1}, \quad \gamma_{,4} = 0, \quad (5.41)$$

$$A_{,2} - 2Z^{-1}\bar{Z}Y_{,3}A = 0, \quad (5.42)$$

$$A_{,3} - Z^{-1}Y_{,3}A_{,1} - \bar{Z}^{-1}\bar{Y}_{,3}A_{,2} + \bar{Z}^{-1}\gamma_{,2} - Z^{-1}Y_{,3}\gamma = 0. \quad (5.43)$$

We can now calculate $R_{23} = R_{1232} + R_{3432}$ and $P_{23} = -\mathcal{F}_{12}\bar{\mathcal{F}}_{31}$. The gravitational field equation $R_{23} = -P_{23}$ gives

$$M_{,2} - 3Z^{-1}\bar{Z}Y_{,3}M - A\bar{\gamma}\bar{Z} = 0. \quad (5.44)$$

The last gravitational field equation is

$$R_{33} = -P_{33}, \quad (5.45a)$$

$$R_{33} = -2(Zh)_{,3} + 2\bar{Z}Zh^2 + 2(h_{,1} - h\bar{Y}_{,3})_{,2} - 2(h_{,1} - h\bar{Y}_{,3})Y_{,3} - 2Zh[h_{,4} + (\bar{Z} - Z)h], \quad (5.45b)$$

$$P_{33} = \mathcal{F}_{31}\bar{\mathcal{F}}_{31} = [\gamma Z - (AZ)_{,1}][\bar{\gamma}\bar{Z} - (\bar{A}\bar{Z})_{,2}]. \quad (5.45c)$$

As in Sec. 3, the calculation is simplified by using the equations $(R_a{}^b + P_a{}^b)_{;b} = 0$. We assume all previous results, and then these equations reduce to

$$[Z^{-1}\bar{Z}^{-1}(R_{33} + P_{33})]_{,4} = 0.$$

Thus

$$R_{33} + P_{33} = \mu Z\bar{Z}, \quad \mu_{,4} = 0, \quad (5.46)$$

and we can work with terms of order 2 in Z and $\bar{Z}$, the higher-order terms cancelling automatically. Since $(Z^{-1} - \bar{Z}^{-1})_{,4} = 0$, we have

$$Z = \bar{Z} + O_2. \quad (5.47)$$

From commutation relations we find

$$(Z^{-1}Y_{,3})_{,4} = 0, \quad (Z^{-3}Z_{,1})_{,4} = 0, \quad (Z^{-1}T_{,1})_{,4} = 0, \quad (T_{,3} - Z^{-1}Y_{,3}T_{,1} - \bar{Z}^{-1}\bar{Y}_{,3}T_{,2})_{,4} = 0, \quad (5.48)$$

for any expression T which has a zero 4-derivative, i.e., $T_{,4} = 0$. To the required order in Z and $\bar{Z}$ we now have

$$R_{33} = -2Z\bar{Z} \times (\text{part of } M_{,3} \text{ which has zero 4-derivative}) + O_3 = -2Z\bar{Z}(M_{,3} - Z^{-1}Y_{,3}M_{,1} - \bar{Z}^{-1}\bar{Y}_{,3}M_{,2}) + O_3, \quad (5.49a)$$

$$P_{33} = \gamma\bar{\gamma}Z\bar{Z} + O_3. \quad (5.49b)$$

Thus the field equation (5.45a) is

$$M_{,3} - Z^{-1}Y_{,3}M_{,1} - \bar{Z}^{-1}\bar{Y}_{,3}M_{,2} - \tfrac{1}{2}\gamma\bar{\gamma} = 0. \quad (5.50)$$

This completes the calculation of the field equations.

We shall now restrict the electromagnetic field further by the condition

$$\gamma = 0. \tag{5.51}$$

The field variables are the real function M and the complex functions Y and A. In terms of them, the metric is given by Eqs. (5.6) and (5.39) with $Z = Y_{,1}$ and the electromagnetic field by Eqs. (5.37) and (5.41) with $\gamma = 0$. The field equations are

$$Y_{,2} = Y_{,4} = 0, \quad M_{,4} = A_{,4} = 0, \tag{5.52}$$

$$M_{,2} - 3Z^{-1}\bar{Z}Y_{,3}M = 0, \tag{5.53a}$$

$$A_{,2} - 2Z^{-1}\bar{Z}Y_{,3}A = 0, \tag{5.53b}$$

$$M_{,3} - Z^{-1}Y_{,3}M_{,1} - \bar{Z}^{-1}\bar{Y}_{,3}M_{,2} = 0, \tag{5.54a}$$

$$A_{,3} - Z^{-1}Y_{,3}A_{,1} - \bar{Z}^{-1}\bar{Y}_{,3}A_{,2} = 0. \tag{5.54b}$$

Equations (5.52) and (5.54) imply that M, A, Y, and $\bar{Y}$ each satisfy the pair of partial differential equations

$$X_{,4} = 0, \quad X_{,3} - Z^{-1}Y_{,3}X_{,1} - \bar{Z}^{-1}\bar{Y}_{,3}X_{,2} = 0. \tag{5.55}$$

Also, Y and $\bar{Y}$ are functionally independent, since otherwise $Y_{,a}$ and $\bar{Y}_{,a}$ would be proportional and $\bar{Y}_{,1} = 0$ would imply $Z = Y_{,1} = 0$, contrary to our hypothesis (5.10). In four-dimensional space–time it follows that M and A are functions of Y and $\bar{Y}$:

$$M = M(Y, \bar{Y}), \quad A = A(Y, \bar{Y}). \tag{5.56}$$

Equations (5.53) can now be written

$$-Z^{-1}Y_{,3} = M^{\frac{1}{3}}(M^{-\frac{1}{3}})_{\bar{Y}} = A^{\frac{1}{2}}(A^{-\frac{1}{2}})_{\bar{Y}}, \tag{5.57}$$

since $Y_{,2} = 0$, $\bar{Y}_{,2} = \bar{Z}$. We therefore put

$$M = mP^{-3}, \tag{5.58}$$

where m is a real constant and P a real function of Y and $\bar{Y}$. Then, by Eq. (5.57),

$$A = \psi P^{-2}, \quad \psi = \psi(Y), \tag{5.59}$$

where ψ is an arbitrary analytic function of the complex variable Y, and

$$P_{\bar{Y}} = -Z^{-1}Y_{,3}P. \tag{5.60}$$

Thus, $Z^{-1}Y_{,3}$ is also a function of Y and $\bar{Y}$ and, therefore, $(Z^{-1}Y_{,3})_{,2} = \bar{Z}(Z^{-1}Y_{,3})_{\bar{Y}}$. Using the commutation relations (5.40) and $Y_{,32} = (Y_{,3})^2$, we have

$$(Z^{-1}Y_{,3})_{\bar{Y}} = Z^{-2}(Y_{,3})^2. \tag{5.61}$$

Differentiating Eq. (5.60) with respect to $\bar{Y}$, we obtain

$$P_{\bar{Y}\bar{Y}} = 0. \tag{5.62}$$

Since P is real, it follows that $P_{YY} = 0$ and that P is bilinear in Y and $\bar{Y}$:

$$P = pY\bar{Y} + qY + \bar{q}\bar{Y} + c, \tag{5.63}$$

where p, c are real constants and q is a complex constant.

From Eqs. (5.52) and (5.60), we have

$$dY = Ze^1 + Y_{,3}e^3 = P^{-1}Z(Pe^1 - P_{\bar{Y}}e^3).$$

Using the definitions (5.7) and (5.8) for the tetrad vectors, this becomes

$$dY = P^{-1}Z[(qY + c)(d\zeta - Y\,dv) - (pY + \bar{q})(du + Y\,d\bar{\zeta})]. \tag{5.64}$$

The general solution is

$$F = 0, \tag{5.65}$$

$$F \equiv \phi + (qY + c)(\zeta - Yv) - (pY + \bar{q})(u + Y\bar{\zeta}), \quad \phi = \phi(Y), \tag{5.66}$$

where ϕ is an arbitrary analytic function of the complex variable Y. Differentiating (5.66), we find that

$$Z = -PF_Y^{-1}. \tag{5.67}$$

Equation (5.64) implies that the operator

$$K \equiv K^\mu\partial_\mu = c\partial_u + \bar{q}\partial_\zeta + q\partial_{\bar{\zeta}} - p\partial_v \tag{5.68}$$

satisfies

$$KY = 0. \tag{5.69}$$

Since K is real, we also have that $K\bar{Y} = 0$ and, from Eq. (5.9a), that K commutes with the directional derivatives ∂_1 and ∂_2, so that $KZ = KZ_{,1} = K\bar{Z} = K\bar{Z}_{,2} = 0$. Thus K^μ is a Killing vector of our curved space–time with metric $g_{\mu\nu}$. Also, since K^μ has constant components, it is at the same time a translational Killing vector of the auxiliary Minkowski space with metric $\eta_{\mu\nu}$.

Let us summarize the solutions of the Einstein–Maxwell equations which we have obtained.

The metric is

$$ds^2 = 2\,d\zeta\,d\bar{\zeta} + 2\,du\,dv + P^{-3}[m(Z + \bar{Z}) - \psi\bar{\psi}P^{-1}Z\bar{Z}] \times (du + \bar{Y}\,d\zeta + Y\,d\bar{\zeta} - Y\bar{Y}\,dv)^2, \tag{5.70}$$

where

$$P = pY\bar{Y} + qY + \bar{q}\bar{Y} + c, \tag{5.71}$$

Y is given in terms of the coordinates ζ, $\bar{\zeta}$, u, v by

$$F \equiv \phi + (qY + c)(\zeta - Yv) - (pY + \bar{q})(u + Y\bar{\zeta}) = 0, \tag{5.72}$$

Z is given by

$$Z = -PF_Y^{-1}, \tag{5.73}$$

ϕ and ψ are arbitrary analytic functions of the complex variable Y, and where m, p, c are real constants and q is a complex constant.

The electromagnetic field has the following components with respect to the tetrad $e_a{}^\mu$, given by Eqs. (5.7) and (5.8):

$$F_{42} = F_{41} = 0, \tag{5.74a}$$

$$\mathcal{F}_{12} = F_{12} + F_{34} = \psi P^{-2} Z^2, \tag{5.74b}$$

$$\mathcal{F}_{31} = 2F_{31} = -(\psi P^{-2} Z)_{,1} = -\psi' F_Y{}^{-2} + \psi F_Y{}^{-3} F_{YY}, \tag{5.74c}$$

with $\bar{\mathcal{F}}_{12} = -F_{12} + F_{34}$ and $\bar{\mathcal{F}}_{31} = 2F_{32}$.

A simpler representation is given in terms of electromagnetic potentials α_μ or the corresponding form

$$\alpha = \alpha_\mu \, dx^\mu = \alpha_a e^a. \tag{5.75}$$

The two-form

$$f = \tfrac{1}{2} F_{\mu\nu} \, dx^\mu \wedge dx^\nu = \tfrac{1}{2} F_{ab} e^a \wedge e^b, \tag{5.76}$$

which describes the electromagnetic field, is obtained from α by exterior differentiation

$$f = -d\alpha. \tag{5.77}$$

The electromagnetic field (5.74) of our solutions can be obtained from the potentials

$$\alpha = -\tfrac{1}{2} P^{-2}(\psi Z + \bar{\psi} \bar{Z}) e^3 - \tfrac{1}{2}(\chi \, d\bar{Y} + \bar{\chi} \, dY), \tag{5.78}$$

where

$$\chi = \int P^{-2} \psi \, dY, \tag{5.79}$$

$\bar{Y}$ being kept constant in this integration.

The gravitational and electromagnetic fields which we have obtained admit a one-parameter group of motions. The Killing vector field K^μ is constant in the auxiliary Minkowski space. We can simplify our solutions by performing a Lorentz transformation. In the three cases where $\eta_{\mu\nu} K^\mu K^\nu$ is (a) negative, (b) positive, (c) zero, K^μ can be made to point along (a) the t or $(u - v)$ direction, (b) the z or $(u + v)$ direction, (c) the u direction. Thus, without loss of generality, we can assume P to have one of the following three forms:

$$P = 2^{-\frac{1}{2}}(1 + Y\bar{Y}), \quad p = c = 2^{-\frac{1}{2}}, \quad q = 0, \tag{5.80a}$$

$$P = 2^{-\frac{1}{2}}(1 - Y\bar{Y}), \quad -p = c = 2^{-\frac{1}{2}}, \quad q = 0, \tag{5.80b}$$

$$P = 1, \quad p = q = 0, \quad c = 1. \tag{5.80c}$$

When $\psi = 0$, the electromagnetic field vanishes and we obtain the most general vacuum solution with a metric of the form (5.6) and $Z \neq 0$.

6. COORDINATE AND TETRAD TRANSFORMATIONS

We shall now show the relationship between the two systems of coordinates used, $(\zeta, \bar{\zeta}, u, v)$ and $(Y, \bar{Y}, \rho, r)$. In particular, we shall show that the vacuum solutions of Sec. 5 are identical with those of Sec. 4.

Starting with the solutions of Sec. 5 and the null tetrad e^a, given by Eqs. (5.7) and (5.8), we have

$$dY = Ze^1 + Y_{,3} e^3 = Z(e^1 - P^{-1} P_{\bar{Y}} e^3).$$

In order to obtain a tetrad vector e'^1 of the form (4.12b), we perform the following transformation to a new null tetrad e'^a:

$$e'^1 = e^1 - P^{-1} P_{\bar{Y}} e^3, \quad e'^2 = e^2 - P^{-1} P_Y e^3, \tag{6.1a}$$

$$e'^3 = P^{-1} e^3, \tag{6.1b}$$

$$e'^4 = P(e^4 + P^{-1} P_Y e^1 + P^{-1} P_{\bar{Y}} e^2 - P^{-2} P_Y P_{\bar{Y}} e^3). \tag{6.1c}$$

Then,

$$Z' = P^{-1} Z \tag{6.2}$$

and

$$\mathcal{F}'_{12} = \mathcal{F}'_{34} = \mathcal{F}_{12}, \quad \mathcal{F}'_{31} = P\mathcal{F}_{31} - 2P_Y \mathcal{F}_{12}, \quad \mathcal{F}'_{23} = \mathcal{F}'_{14} = \mathcal{F}'_{24} = 0. \tag{6.3}$$

We introduce as new coordinates Y, $\bar{Y}$ and ρ, r, given by

$$\rho = P^{-1}(u + Y\bar{\zeta} + \bar{Y}\zeta - Y\bar{Y}v), \tag{6.4a}$$

$$r = P(v + p\rho). \tag{6.4b}$$

A straightforward calculation gives the solutions of the Einstein–Maxwell equations (5.70)–(5.74) in terms of the new tetrad and new coordinates. Dropping primes, the solutions take the form

$$ds^2 = 2e^1 e^2 + 2e^3 e^4, \tag{6.5a}$$

$$e^1 = P^{-1} Z^{-1} \, dY, \quad e^2 = P^{-1} \bar{Z}^{-1} \, d\bar{Y}, \tag{6.5b}$$

$$e^3 = d\rho + \Omega \, dY + \bar{\Omega} \, d\bar{Y}, \tag{6.5c}$$

$$e^4 = dr + \beta \, dY + \bar{\beta} \, d\bar{Y} + He^3; \tag{6.5d}$$

$$P = pY\bar{Y} + qY + \bar{q}\bar{Y} + c, \tag{6.6a}$$

$$Z^{-1} = r - \phi' + P^{-1}(P_Y \phi - P_{\bar{Y}} \bar{\phi}), \tag{6.6b}$$

$$\Omega = P^{-2} \bar{\phi}, \tag{6.6c}$$

$$\beta = P^{-1}[p\bar{\phi} - P_Y \phi' + P^{-1} P_Y (P_Y \phi - P_{\bar{Y}} \bar{\phi})], \tag{6.6d}$$

$$H = q\bar{q} - pc + \tfrac{1}{2} m(Z + \bar{Z}) - \tfrac{1}{2} \psi \bar{\psi} Z \bar{Z}; \tag{6.6e}$$

$$\mathcal{F}_{12} = \psi Z^2, \tag{6.7a}$$

$$\mathcal{F}_{31} = -\psi' P Z^2 - 2\psi P_Y Z^2 - \psi P Z^3 \phi'' + 2\psi Z^3 (P_Y r - p\bar{\phi}). \tag{6.7b}$$

Here $\phi = \phi(Y)$ and $\psi = \psi(Y)$ are arbitrary analytic functions of the complex coordinate Y; m, p, c are arbitrary real constants; q is an arbitrary complex constant.

The electromagnetic field can also be obtained from the potentials

$$\alpha = -\tfrac{1}{2}(\psi Z + \bar{\psi}\bar{Z})e^3 - \tfrac{1}{2}(\chi\, d\bar{Y} + \bar{\chi}\, dY),$$
$$\chi = \int_{(\bar{Y}=\text{const})} P^{-2}\psi\, dY, \tag{6.8}$$

as in Eqs. (5.76) and (5.77).

From the results of the previous section it follows that each of these solutions admits a one-parameter group of motions and that,without loss of generality, P can be assumed to have one of the three forms

$$P = 2^{-\frac{1}{2}}(1 + Y\bar{Y}), \tag{6.9a}$$
$$P = 2^{-\frac{1}{2}}(1 - Y\bar{Y}), \tag{6.9b}$$
$$P = 1. \tag{6.9c}$$

The vacuum solutions, for which $\psi = 0$, are identical with those of Sec. 4, Eqs. (4.12) and (4.13).

The coordinates $(Y, \bar{Y}, \rho, r)$ have the advantage that the solutions (6.5)–(6.7) are given explicitly in terms of the two arbitrary functions ϕ and ψ, whereas in the coordinates $(\zeta, \bar{\zeta}, u, v)$, Y is determined by the implicit relation $F = 0$ [Eq. (5.72)].

The coordinates $(\zeta, \bar{\zeta}, u, v)$ have the advantage of being asymptotically inertial for the special solution corresponding to a rotating body. This is shown in the next section.

7. THE FIELD OF A ROTATING CHARGED BODY

In this section we shall study the particular solution[10]

$$P = 2^{-\frac{1}{2}}(1 + Y\bar{Y}), \quad \phi = -iaY, \quad \psi = -e, \tag{7.1}$$

where a and e are real constants. As we shall see, this simple solution is asymptotically flat and represents the exterior gravitational and electromagnetic field of a rotating charged body of mass m, angular momentum $L = am$, charge e, and with a magnetic moment $\mu = ae$ parallel to the angular momentum.

The mass m, the angular momentum L (or, equivalently, $a = L/m$), and the charge e are arbitrary parameters in the solution. The magnetic moment μ, however, is then determined in this solution by the Einstein–Maxwell field equations. It is a remarkable fact that the gyromagnetic ratio

$$L/\mu = m/e \tag{7.2}$$

is the same as that of the Dirac electron.

This result suggests that the gyromagnetic ratio of the electron may be a consequence of just special relativity and Maxwell's electrodynamics, the only ingredients which are common to the present results and to the Dirac theory. It suggests that the gyromagnetic ratio may, on the one hand, be independent of Einstein's transition from special to general relativity which embodies his theory of gravitation and, on the other hand, be independent of the transition from classical to quantum theory.

The field of a rotating charged body was first obtained by Newman and his collaborators.[12] They used unconventional units and coordinates which were not asymptotically flat, so that it was not possible to identify all the physical quantities m, L, e, and μ. Carter[13] was the first to notice the gyromagnetic ratio (7.2).

We first write down the solution (7.1) in the coordinates $(Y, \bar{Y}, \rho, r)$ of Eqs. (6.5)–(6.8), but replace the complex coordinate Y by real coordinates θ and φ, given by

$$Y = e^{i\varphi}\tan\tfrac{1}{2}\theta. \tag{7.3}$$

The metric is

$$ds^2 = (r^2 + a^2\cos^2\theta)(d\theta^2 + \sin^2\theta\, d\varphi^2) + 2(d\rho - a\sin^2\theta\, d\varphi)(dr - a\sin^2\theta\, d\varphi) - \left(1 - \frac{2mr - e^2}{r^2 + a^2\cos^2\theta}\right)(d\rho - a\sin^2\theta\, d\varphi)^2. \tag{7.4}$$

It reduces to Eq. (4.16) for $e = 0$. The vector field

$$e^3 = d\rho - a\sin^2\theta\, d\varphi \tag{7.5}$$

points in a double Debever–Penrose direction of Weyl's conformal curvature tensor and is tangent to a congruence of shear-free null geodesics. The rates of expansion and of rotation, Eq. (2.20), of this null congruence are given by Sachs's complex parameter

$$Z = (r + ia\cos\theta)^{-1}. \tag{7.6}$$

The electromagnetic field is most easily obtained from the potentials α. These are further simplified by dropping the second term in Eq. (6.8),which, for real ψ, is a perfect differential, since $d(\chi\, d\bar{Y} + \bar{\chi}\, dY) = P^{-2}(\psi - \bar{\psi})\, dY \wedge d\bar{Y}$. We then obtain

$$\alpha = [er/(r^2 + a^2\cos^2\theta)](d\rho - a\sin^2\theta\, d\varphi). \tag{7.7}$$

By differentiation we obtain the components of the electromagnetic field:

$$\begin{aligned} f &= -d\alpha \\ &= F_{r\rho}\, dr \wedge d\rho + F_{r\varphi}\, dr \wedge d\varphi \\ &\quad + F_{\rho\theta}\, d\rho \wedge d\theta + F_{\varphi\theta}\, d\varphi \wedge d\theta \\ &= [e/(r^2 + a^2\cos^2\theta)^2] \\ &\quad \times [(r^2 - a^2\cos^2\theta)\, dr \wedge (d\rho - a\sin^2\theta\, d\varphi) \\ &\quad + 2ra\cos\theta\sin\theta(a\, d\rho - (r^2 + a^2)\, d\varphi) \wedge d\theta], \end{aligned} \tag{7.8}$$

with $F_{r\theta} = F_{\rho\varphi} = 0$.

[12] E. T. Newman, E. Couch, K. Chinnapared, A. Exton, A. Prakash, and R. Torrence, J. Math. Phys. 6, 918 (1965).

[13] B. Carter, report of work prior to publication, 1968 .

The metric (7.4) is not asymptotically flat in the coordinates used. In order to show that our space-time is in fact asymptotically flat, we now write down the solution (7.1) in real coordinates (x, y, z, t),which are related to the coordinates $(\zeta, \bar{\zeta}, u, v)$ of Sec. 5 by Eqs. (5.5).

Eqs. (5.72) and (6.4) give the coordinate transformation from $(\theta, \varphi, \rho, r)$ to (x, y, z, t):

$$x + iy = (r + ia)e^{i\varphi}\sin\theta, \tag{7.9a}$$

$$z = r\cos\theta, \tag{7.9b}$$

$$t = r - \rho. \tag{7.9c}$$

Elimination of θ and φ from Eqs. (7.9a) and (7.9b) gives

$$\frac{x^2 + y^2}{r^2 + a^2} + \frac{z^2}{r^2} = 1. \tag{7.10}$$

It is useful to retain r as a function of x, y, z defined by this implicit relation.

In the new coordinates the metric is

$$\begin{aligned} ds^2 = {} & dx^2 + dy^2 + dz^2 - dt^2 \\ & + \frac{2mr^3 - e^2r^2}{r^4 + a^2z^2}\left[dt + \frac{z}{r}\,dz\right. \\ & + \frac{r}{r^2 + a^2}(x\,dx + y\,dy) \\ & \left. - \frac{a}{r^2 + a^2}(x\,dy - y\,dx)\right]^2. \end{aligned} \tag{7.11}$$

The double Debever–Penrose null vector e^3 and Sachs's complex parameter Z of Sec. 5 are, by Eqs. (6.1b) and (6.2), P times the expressions given by Eqs. (7.5) and (7.6). They are

$$\begin{aligned} e^3 = 2^{\frac{1}{2}}\frac{r}{r + z}\left[dt + \frac{z}{r}\,dz + \frac{r}{r^2 + a^2}(x\,dx + y\,dy)\right. \\ \left. - \frac{a}{r^2 + a^2}(x\,dy - y\,dx)\right], \end{aligned} \tag{7.12}$$

$$Z = 2^{\frac{1}{2}}r^2/[(r + z)(r^2 + iaz)]. \tag{7.13}$$

The electromagnetic potentials, obtained from Eq. (7.7) after dropping a term $er(r^2 + a^2)^{-1}\,dr$ which is a perfect differential, are given by

$$\alpha = e\frac{r^3}{r^4 + a^2z^2}\left[dt - \frac{a}{r^2 + a^2}(x\,dy - y\,dx)\right]. \tag{7.14}$$

The electromagnetic field is

$$\begin{aligned} & (F_{xt} - iF_{yz}, F_{yt} - iF_{zx}, F_{zt} - iF_{xy}) \\ & \quad = [er^3/(r^2 + iaz)^3](x, y, z + ia). \end{aligned} \tag{7.15}$$

In the asymptotic region,r, defined by Eq. (7.10), is

$$r = R + O(R^{-1}), \quad R^2 = x^2 + y^2 + z^2 \tag{7.16}$$

and the metric (7.11) is asymptotically flat. By examining the metric in the asymptotic region, it can be shown that it corresponds to the gravitational field of a body of mass m and angular momentum

$$\mathbf{L} = (L_x, L_y, L_z) = (0, 0, L), \quad L = ma. \tag{7.17}$$

In the asymptotic region, the electromagnetic field reduces to an electric field

$$\mathbf{E} = (F_{xt}, F_{yt}, F_{zt}) = (e/R^3)(x, y, z) \tag{7.18}$$

and a magnetic field

$$\mathbf{H} = (F_{yz}, F_{zx}, F_{xy}) = (ea/R^5)(3xz, 3yz, 3z^2 - R^2). \tag{7.19}$$

This is the electromagnetic field of a body with charge e and magnetic moment

$$\boldsymbol{\mu} = (\mu_x, \mu_y, \mu_z) = (0, 0, \mu), \quad \mu = ea. \tag{7.20}$$

The gyromagnetic ratio (7.2) now follows.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Particle-Hole Matrix: Its Connection with the Symmetries and Collective Features of the Ground State

CLAUDE GARROD*
University of California, Davis, California
AND
MITJA ROSINA
University of Ljubljana
and
Nuclear Institute, "J. Stefan," Ljubljana, Yugoslavia

(Received 21 November 1968)

A detailed study is made of the properties of the "particle-hole matrix"

$$G_{abcd} \equiv \langle g|\,(a^\dagger b - \rho_{ab})^\dagger (c^\dagger d - \rho_{cd})\,|g\rangle,$$

where $|g\rangle$ is a many-fermion ground state and ρ_{ab} is the 1-body density matrix. It is shown that the zero eigenvalues of the particle-hole matrix are intimately and simply related to the one-body symmetries of the ground state. It is also shown that the large eigenvalues of G_{abcd} are closely related to the collective features of the ground state.

1. INTRODUCTION

The ground state of a system of many fermions is generally a quite complicated mathematical object. It has often been noted that many important features of the ground state are contained in the eigenvalue spectrum of the 1- and 2-body density matrices. These are defined as follows:

$$\rho_{ab} \equiv \langle g|\,a^\dagger b\,|g\rangle \tag{1}$$

and

$$\rho_{abcd} \equiv \langle g|\,b^\dagger a^\dagger cd\,|g\rangle, \tag{2}$$

where $|g\rangle$ is the many-particle ground state and $a^\dagger$, $b^\dagger$, a, b, etc., are creation and annihilation operators for a sequence of single-particle states. (In the more common notation, $a^\dagger$, $b^\dagger$, a, and b are written as $a_a^\dagger$, $a_b^\dagger$, a_a, and a_b, respectively.)

It is the purpose of this paper to point out the fact that many physically important features of the ground state are more simply and intimately associated with the "particle–hole matrix"

$$G_{abcd} \equiv \langle g|\,(a^\dagger b - \rho_{ab})^\dagger (c^\dagger d - \rho_{cd})\,|g\rangle. \tag{3}$$

The particle–hole matrix (which we shall henceforth call the G matrix) is the metric or Gramm matrix associated with the linear subspace of quantum states spanned by the following particle–hole states:

$$|ab\rangle \equiv (a^\dagger b - \rho_{ab})\,|g\rangle. \tag{4}$$

The purpose of including the constant ρ_{ab} in the definitions of G and $|ab\rangle$ is to make the state $|ab\rangle$ orthogonal to the ground state $|g\rangle$. The fact that, associated with each ground state, there is a corresponding space of particle–hole states offers an opportunity of getting additional insight into the ground state through the study of these excited states.

The two-body density matrix can likewise be interpreted as the metric matrix of a linear subspace of quantum states. The corresponding states $ab\,|g\rangle$ belong, however, to a system of $N-2$ particles. Thus, the properties of the 2-body density matrix are related to those aspects of the ground state which are relevant for processes changing particle number (double ionization, 2-body transfer reactions). When one is interested, however, in processes and properties which conserve particle number, the corresponding operators do not connect the N-particle ground state with the $(N-2)$-particle states but rather with the particle–hole states. It is for this reason that in the latter case the properties of the G matrix appear to have a more direct connection with physically important aspects of the ground state than do those of the 2-body density matrix.

The G matrix is, of course, not independent of the density matrices. The relationship between them is

$$G_{abcd} = \rho_{bcad} + \delta_{ac}\rho_{bd} - \rho^*_{ab}\rho_{cd}. \tag{5}$$

If one wishes to invert this relationship and calculate the 2-body density matrix from the G matrix, one must first express the 1-body density matrix in terms of G. The needed relation is easily obtained by taking a partial trace in the following way:

$$\sum_c G_{bcac} = N\delta_{ab} - (N-1)\rho_{ab} - (\rho^2)_{ab}. \tag{6}$$

To solve this, one diagonalizes the quantity on the left and gets independent quadratic equations for the eigenvalues of the 1-body density matrix. There is no

* Supported in part by the U.S. Atomic Energy Commission.

ambiguity in the solution since only one solution lies between 0 and 1. If the dimension L of the single-particle basis is finite, one can utilize a simpler relation:

$$\rho_{ab} = \left[N\delta_{ab} + \sum_c (G_{cacb} - G_{bcac})\right] \Big/ L. \quad (7)$$

It follows from the definition of the G matrix that it is Hermitian and that it satisfies the following symmetry relations:

$$G_{abcd} + G_{acbd} = (\delta_{ab} - \rho^*_{ab})\rho_{cd} + (\delta_{ac} - \rho_{ac})\rho_{bd}. \quad (8)$$

The matrix $\delta_{ab} - \rho_{ab}$ is just the 1-body density matrix of holes.

As the expectation value of the G matrix for any vector is the squared norm of a certain particle–hole state, it must be nonnegative. Thus G is a non-negative matrix.

2. ZERO EIGENVALUES OF THE G-MATRIX AND ONE-BODY SYMMETRIES

If the ground state $|g\rangle$ is an eigenstate of some 1-body operator

$$\hat{T}\,|g\rangle = t\,|g\rangle, \quad (9)$$

where

$$\hat{T} = \sum T_{ab} a^\dagger b, \quad (10)$$

then we shall say that it has a "1-body symmetry." Examples of "1-body symmetries" are translational, rotational, and isospin symmetry. The ground state is an eigenstate of the symmetry operators generating infinitesimal translations (momentum operator), rotations (angular-momentum operator), and isospin rotation (isospin operator).

In such a case, T_{ab} is an eigenvector of the G matrix with eigenvalue zero. Let us demonstrate this. Using the fact that

$$t = \sum T_{ab}\rho_{ab}, \quad (11)$$

we can see that

$$\sum_{cd} T_{cd}(c^\dagger d - \rho_{cd})\,|g\rangle = (\hat{T} - t)\,|g\rangle = 0. \quad (12)$$

Therefore, for any choice of a and b,

$$\langle g|\,(a^\dagger b - \rho_{ab})^\dagger \sum_{cd} T_{cd}(c^\dagger d - \rho_{cd})\,|g\rangle = 0 \quad (13)$$

or

$$\sum_{cd} G_{abcd} T_{cd} = 0. \quad (14)$$

The converse is also true. If T_{ab} is an eigenvector of G with zero eigenvalue, then

$$\sum_{abcd} T^*_{ab} G_{abcd} T_{cd} = 0. \quad (15)$$

Using the definition of G, this can be written as

$$\langle g|\,[\sum T_{ab}(a^\dagger b - \rho_{ab})]^\dagger [\sum T_{cd}(c^\dagger d - \rho_{cd})]\,|g\rangle. \quad (16)$$

This expression is the squared norm of the state $(\hat{T} - t)\,|g\rangle$. Since its norm is zero, this state must be identically zero, which proves that $|g\rangle$ is an eigenstate of $\hat{T}$.

Thus one may obtain, from a study of G, all 1-body symmetries of the ground state. If some eigenvalues of the G matrix are nonzero but very small, one obtains approximate symmetries of the system. The effects of 1-body symmetries on the two-particle density matrix has been studied extensively in the past.[1–8] The major result of these investigations is a theorem which states that whenever the ground state $|g\rangle$ is an eigenstate of some 1-body operator $\hat{T}$, then the eigenvectors of the 1- and 2-particle density matrices can be chosen as simultaneous eigenfunctions of $\hat{T}$. Below we shall give a compact derivation of the G-matrix form of this theorem. We shall then see what additional information is obtained from the theorem presented in this paper, which establishes a relationship between the 1-body symmetry operators and the null vectors of G.

Suppose $|g\rangle$ is an eigenstate of a Hermitian operator $\hat{T}$:

$$\hat{T}\,|g\rangle = \sum T_{nm} n^\dagger m\,|g\rangle = t\,|g\rangle. \quad (17)$$

It is then obvious that the expectation value of the commutator of $\hat{T}$ and any other operator $\hat{O}$ vanishes, i.e.,

$$\langle g|\,(\hat{T}\hat{O} - \hat{O}\hat{T})\,|g\rangle = 0. \quad (18)$$

If we make the following choice for $\hat{O}$,

$$\hat{O} = b^\dagger a\, c^\dagger d, \quad (19)$$

we obtain the equation

$$\sum_{m,n} T_{nm} \langle g|\,[n^\dagger m,\, b^\dagger a\, c^\dagger d]\,|g\rangle = 0. \quad (20)$$

In order to evaluate the expectation value of the commutator, we need the following two identities:

$$[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}], \quad (21)$$

$$[\hat{A}, \hat{B}\hat{C}] = \{\hat{A}, \hat{B}\}\hat{C} - \hat{B}\{\hat{A}, \hat{C}\}, \quad (22)$$

[1] W. A. Bingel, J. Chem. Phys. **32**, 1522 (1960); **34**, 1066 (1961); **36**, 2842 (1962); *Report of Density Matrix Conference* (Queen's University, Kingston, Ontario, 1967).

[2] A. J. Coleman, Preprint No. 23, Quantum Chemistry Project, University of Florida, 1963.

[3] W. Kutzelnigg, Z. Naturforsch. **18a**, 1058 (1963); **20a**, 168 (1965).

[4] P.-O. Löwdin, Phys. Rev. **97**, 1474 (1955); Rev. Mod. Phys. **35**, 629 (1963).

[5] D. Micha, J. Chem. Phys. **41**, 3648 (1964).

[6] R. McWeeny and Y. Mizuno, Proc. Roy. Soc. (London) **A259**, 554 (1961); J. Chem. Phys. **42**, 1717 (1965).

[7] D. Smith, J. Chem. Phys. **43**, S258 (1965).

[8] Claude Garrod, Phys. Rev. **172**, 173 (1968).

where the curly brackets signify the anticommutator. Using these identities and the definition of G, we obtain the result that

$$\sum_n (G_{nacd}T_{an} - G_{ancd}T_{nb} - G_{abnd}T_{nc} + G_{abcn}T_{dn}) = 0. \tag{23}$$

This equation is most transparent in the basis in which T_{ab} is diagonal. In that basis,

$$T_{ab} = t_a\delta_{ab} \tag{24}$$

and the equation takes the form

$$G_{abcd}(t_{ab} - t_{cd}) = 0, \tag{25}$$

where $t_{ab} \equiv t_a - t_b$. Thus, G_{abcd} is zero unless $t_{ab} = t_{cd}$. If the eigenvalues of T are not degenerate, we can write G as a function of three variables, namely,

$$G_{abcd} = \delta(t_{ab} - t_{cd})g(t_{ab} + t_{cd}, \tfrac{1}{2}(t_a + t_b), \tfrac{1}{2}(t_c + t_d)). \tag{26}$$

The fact that T_{ab} is a null vector of G then implies that

$$\sum_c g(0, t_a, t_c)t_c = 0. \tag{27}$$

The relationship between the zero eigenvectors of G and 1-body eigenoperators of the ground state implies an interesting theorem regarding the algebraic structure of the null space of G. (The null space of G, denoted Ω_0, is the linear subspace spanned by the eigenvectors of G which correspond to the eigenvalue zero.)

If $x_{ab} \in \Omega_0$ and $y_{ab} \in \Omega_0$, then $z_{ab} \in \Omega_0$, where

$$z_{ab} = \sum_c (x_{ac}y_{cb} - y_{ac}c_{cb}). \tag{28}$$

Thus the null space is a closed algebra with respect to the commutator product. The proof is straightforward. If $\hat{X}$ and $\hat{Y}$ are the 1-body operators associated with the matrices x_{ab} and y_{ab}, then $\hat{Z} = [\hat{X}, \hat{Y}]$ is also a 1-body operator and is given by

$$\hat{Z} = \sum z_{ab}a^\dagger b; \tag{29}$$

since

$$\hat{X}\,|g\rangle = x\,|g\rangle \tag{30}$$

and

$$\hat{Y}\,|g\rangle = y\,|g\rangle, \tag{31}$$

it is obvious that

$$[\hat{X}, \hat{Y}]\,|g\rangle = 0. \tag{32}$$

Thus $\hat{Z}$ is also an eigenoperator of $|g\rangle$ and hence $z_{ab} \in \Omega_0$.

3. LARGE EIGENVALUES OF THE G MATRIX AND THE COLLECTIVE STATES

It is well known that a large eigenvalue of the 2-body density matrix implies strong particle–particle correlations such as are found in superconductivity or nuclear pairing. Equally interesting is the interpretation of large eigenvalues of the G matrix. It will be shown that large eigenvalues of the G matrix mean strong particle–hole correlations leading to enhanced transition amplitudes between the ground state and certain collective states.

The term particle–particle correlation is used with the meaning that whenever a single-particle state a is occupied, then the probability of finding another particle in an associated state $\bar{a}$ is much larger than it is on the average. The term particle–hole correlations means, on the other hand, that whenever a single-particle state a is occupied, an associated state $\bar{a}$ is much more likely to be empty than on the average.

We shall denote by the term "transition operator" any single-particle operator $\hat{T} = \sum_{ab} T_{ab}a^\dagger b$ with coefficients normalized to unity:

$$\sum_{ab} |T_{ab}|^2 = 1. \tag{33}$$

Then the *enhanced collective states* can be defined as those excited states for which there exists some transition operator having a large transition amplitude between such state and the ground state:

$$\langle \text{coll}|\,\hat{T}\,|g\rangle \gg 1. \tag{34}$$

If the G matrix has a large eigenvalue λ, it is easy to construct an enhanced collective state and the corresponding transition operator using the eigenvector x_{ab} belonging to λ:

$$\hat{T} = \sum x_{ab}a^\dagger b, \tag{35}$$

$$\begin{aligned} |\text{coll}\rangle &= \lambda^{-\frac{1}{2}} \sum_{ab} x_{ab}(a^\dagger b - \rho_{ab})\,|g\rangle \\ &= \lambda^{-\frac{1}{2}}(\hat{T} - \langle g|\,\hat{T}\,|g\rangle)\,|g\rangle. \end{aligned} \tag{36}$$

For this state, the transition amplitude has the value

$$\begin{aligned} \langle \text{coll}|\,\hat{T}\,|g\rangle &= \langle \text{coll}|\,\hat{T} - \langle g|\,\hat{T}\,|g\rangle\,|g\rangle \\ &= \lambda^{\frac{1}{2}}\langle \text{coll}\,|\,\text{coll}\rangle = \lambda^{\frac{1}{2}}, \end{aligned} \tag{37}$$

which is much larger than one if λ is large.

The normalization (33) is chosen so that the transition amplitude $\langle \text{exc}|\,\hat{T}\,|g\rangle$ cannot exceed unity if $|g\rangle$ is uncorrelated, because, for a Slater determinant, $\lambda = 1$ or 0. So the enhancement of transition amplitudes (the size of large eigenvalues λ) can be considered as a measure of the amount of particle–hole correlation.

So far, the collective state and the corresponding transition operator were constructed mathematically. The described analysis is physically significant only if

the constructed collective state has a large overlap with a physical excited state or, at least, if the constructed collective state is distributed between a small number of physical states so that it can still be recognized as a representative of a bunch of experimental states lying close by.

The second condition for physical significance is that the constructed transition operator must have a large overlap with a physically available one (e.g., the operator for a multipole photon transition); otherwise one cannot test experimentally the enhancement of the transition amplitude to the corresponding collective states. It is a lucky coincidence that for some well-known collective transitions nature provides us just with the suitable transition operator.

A. Upper Bounds on Eigenvalues

In deriving an upper bound on eigenvalues of the G matrix, the procedure used will be analogous to the derivation of an upper bound of the 2-body density matrix in Ref. 9. Let us first derive an upper bound in the subspace of Hermitian vectors spanning the G matrix. A vector x_{ab} will be called Hermitian if $x_{ab} = x^*_{ba}$. Such a vector can be "diagonalized" in the form

$$x_{ab} = \sum_i f_i v^{i*}_a v^i_b. \tag{38}$$

If the vector x_{ab} is normalized so that

$$\sum_{ab} |x_{ab}|^2 = 1, \tag{39}$$

then the "eigenvalues" are also normalized: $\sum_i f_i^2 = 1$. The expectation value of the G matrix with respect to the vector x_{ab} is then given by

$$\begin{aligned}\Lambda &= \sum_{abcd} x^*_{ab} G_{abcd} x_{cd} \\ &= \sum_{ij} f_i \Big[\sum_{abcd} v^i_a v^{i*}_b G_{abcd} v^{j*}_c v^j_d \Big] f_j \\ &\equiv \sum_{ij} f_i \tilde{G}_{ij} f_j. \end{aligned} \tag{40}$$

The new basis of the Hermitian vector x_{ab} enabled the expectation value Λ to be expressed in terms of a much smaller matrix $\tilde{G}_{ij}$ rather than in terms of the full matrix G_{abcd}. As the matrix $\tilde{G}_{ij}$ is essentially a submatrix of the G matrix in a specific basis, it is also a nonnegative matrix. Therefore, the expectation of $\tilde{G}_{ij}$ for any normalized vector f_i is equal to or smaller than its trace. Thus,

$$\Lambda \leq \sum_i \tilde{G}_{ii}. \tag{41}$$

In order to estimate this trace, we shall use the antisymmetry relation of the G matrix:

$$G_{abcd} = -G_{acbd} + \delta_{ac}\rho_{bd} + \delta_{ab}\rho_{cd} - \rho^*_{ab}\rho_{cd} - \rho^*_{ac}\rho_{bd}. \tag{42}$$

Thus,

$$\begin{aligned}\tilde{G}_{ii} &= \sum_{abcd} v^i_a v^{i*}_b G_{abcd} v^{i*}_c v^i_d \\ &= -\sum_{abcd} v^i_a v^{i*}_c G_{acbd} v^{i*}_b v^i_d + 2\rho_{ii} - 2(\rho^2)_{ii}. \end{aligned} \tag{43}$$

The first sum on the right-hand side equals the sum on the left-hand side (they differ only in the names of summation indices), so that

$$\tilde{G}_{ii} = \rho_{ii} - (\rho^2)_{ii} \tag{44}$$

and, therefore,

$$\Lambda \leq \sum_i \rho_{ii} - \sum_i (\rho^2)_{ii} = N - \sum_i (\rho^2)_{ii} \leq N - N^2/L, \tag{45}$$

where L is the dimension (rank) of the single-particle basis. Hence, we obtain an upper bound of

$$\Lambda \leq N(L - N)/L \tag{46}$$

for the Hermitian subspace.

It is a simple matter to obtain the same upper bound for the subspace of anti-Hermitian vectors ($x^*_{ba} = -x_{ab}$) by using the fact that ix_{ab} is Hermitian if x_{ab} is anti-Hermitian. We shall use these two results in our proof that the same upper bound is in fact valid for the whole space. We shall make one assumption in this proof: namely, that G_{abcd} is a real matrix. This implies no loss in generality if, as is usually the case, the Hamiltonian of the system is real. If G_{abcd} is real, then it can be diagonalized in terms of real eigenvectors. Thus we may restrict our attention to real vectors x_{ab}. We write x_{ab} as a sum of a symmetric and an antisymmetric part:

$$x_{ab} = S_{ab} + A_{ab}, \tag{47}$$

where $S_{ba} = S_{ab}$ and $A_{ba} = -A_{ab}$. Then $\hat{S} = \sum S_{ab} a^\dagger b$ is Hermitian and $\hat{A}$ is anti-Hermitian. In terms of $\hat{S}$ and $\hat{A}$, the eigenvalue can be written as

$$\begin{aligned}\Lambda &= \sum x_{ab} G_{abcd} x_{cd} \\ &= \langle g| (\hat{S}^\dagger + \hat{A}^\dagger)(\hat{S} + \hat{A}) |g\rangle \\ &\quad - \langle g| (\hat{S}^\dagger + \hat{A}^\dagger) |g\rangle \langle g| (\hat{S} + \hat{A}) |g\rangle \\ &= \langle g| \hat{S}\hat{S} |g\rangle - \langle g| \hat{A}\hat{A} |g\rangle \\ &\quad + \langle g| [\hat{S}, \hat{A}] |g\rangle - (\langle g| \hat{S} |g\rangle)^2, \end{aligned} \tag{48}$$

where we have used the fact that $\langle g| \hat{A} |g\rangle = 0$. The commutator $[\hat{S}, \hat{A}]$ is a 1-body operator

$$[\hat{S}, \hat{A}] = \sum_{abj} (S_{aj}A_{jb} - A_{aj}S_{jb}) a^\dagger b. \tag{49}$$

[9] Claude Garrod and J. K. Percus, J. Math. Phys. **5**, 1756 (1964).

Thus the expression for Λ takes the form

$$\Lambda = \sum S_{ab}G_{abcd}S_{cd} + \sum A_{ab}G_{abcd}A_{cd} + \sum (S_{aj}A_{jb} - A_{aj}S_{jb})\rho_{ab}. \quad (50)$$

Since the normalization sum for x_{ab} breaks into two parts

$$\sum x_{ab}^2 = \sum S_{ab}^2 + \sum A_{ab}^2, \quad (51)$$

$\Lambda - \alpha \sum A_{ab}^2$ must be stationary with respect to variation of A_{ab} for some value of the Lagrange parameter α. This gives the following equation:

$$\sum_{cd} G_{abcd}A_{cd} + \tfrac{1}{2}\sum_{n}(S_{na}\rho_{nb} - S_{bn}\rho_{an}) - \alpha A_{ab} = 0. \quad (52)$$

Both the second and third terms in the equation are antisymmetric under interchange of a and b. Thus, if we multiply by S_{ab} and sum, they give no contribution. We then obtain the result

$$\sum S_{ab}G_{abcd}A_{cd} = 0. \quad (53)$$

Thus,

$$\Lambda = \sum x_{ab}G_{abcd}x_{cd} = \sum (S_{ab}G_{abcd}S_{cd} + A_{ab}G_{abcd}A_{cd}). \quad (54)$$

Since

$$\sum S_{ab}G_{abcd}S_{cd} \leq L^{-1}N(L-N)\sum S_{ab}^2 \quad (55)$$

and

$$\sum A_{ab}G_{abcd}A_{cd} \leq L^{-1}N(L-N)\sum A_{ab}^2, \quad (56)$$

we obtain the desired upper bound on Λ:

$$\Lambda \leq L^{-1}N(L-N). \quad (57)$$

It is interesting to note that the upper bound was derived by utilizing only the nonnegativity and the antisymmetry relation of the G matrix. Thus, the G matrix is "self-taming" in the sense that if it is kept nonnegative and properly antisymmetric, its eigenvalues cannot exceed a certain upper bound.

B. Vibrations

In order to illustrate the idea of the connection between large eigenvalues of the G matrix and collective motions, let us describe a crude model of shape vibrations and of rotations. Only a monopole vibration ("breathing mode") will be considered. In nuclear physics, one is interested mostly in quadrupole and octupole shape vibrations; the monopole vibration is at too high an energy and is difficult to observe. Nevertheless, the formal properties are simplest for the monopole case.

We consider the N particles to be located in two N-fold degenerate levels separated by an energy D.

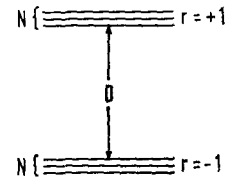


FIG. 1. Single-particle levels in the simple vibrational model.

The single-particle states are characterized by an internal quantum number i and by a radial quantum number r ($r = -1$ for the lower level and $r = +1$ for the upper level). The radial quantum number implies the "size" of the single-particle wavefunction, so that a coherent oscillation of all particles between the two levels (see Fig. 1) produces an over-all dilatation vibration ("breathing mode"). Let the particles interact by a monopole interaction which conserves the internal quantum number of each particle; it scatters two particles from one shell to another without changing i:

$$H = D\sum_i i^\dagger i + V\left[\sum_{i>j}(i^\dagger j^\dagger \bar{j}\bar{i} + i^\dagger \bar{j}^\dagger j i) + \sum_{i,j} i^\dagger \bar{j}^\dagger j\bar{i}\right]. \quad (58)$$

Here, an abridged notation was used: $i = (i, +1)$, $\bar{i} = (i, -1)$. As the interaction simultaneously scatters two particles upwards or downwards, the evenness (or oddness) of the number of particles in each level is also conserved. One can consider r as a multiplicative quantum number and the total $R = \prod_\alpha r_\alpha$ is a good quantum number.

In the limit of strong interaction ($V \gg D$), the considered schematic Hamiltonian gives the following ground state:

$$|g\rangle = \mathcal{N}\left(1 + \sum_{i>j} i^\dagger \bar{i} j^\dagger \bar{j} + \sum_{i>j>\cdots} i^\dagger \bar{i} j^\dagger \bar{j} k^\dagger \bar{k} l^\dagger \bar{l} + \cdots\right)|\Phi\rangle = \mathcal{N}P_R\Pi_i(i^\dagger + \bar{i}^\dagger)|0\rangle. \quad (59)$$

Here $|0\rangle$ means vacuum, $|\phi\rangle$ is the lowest configuration (lower level fully occupied), $\mathcal{N}$ is a normalization constant, and P_R is a projector on a state with pure $R = +1$.

The considered schematic Hamiltonian also gives rise to one strongly collective state corresponding to a monopole vibration:

$$|\text{coll}\rangle = \hat{T}|g\rangle, \quad \text{where} \quad \hat{T} = \sum_i i^\dagger \bar{i} + \bar{i}^\dagger i. \quad (60)$$

This collective state has an enhanced transition amplitude just for the monopole operator $\hat{T}$.

The G matrix of the ground state has, in fact, one large eigenvalue ($\lambda = N/2$) and the corresponding

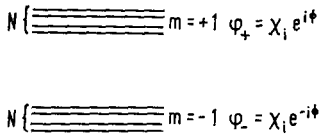


FIG. 2. Single-particle states in the simple rotational model.

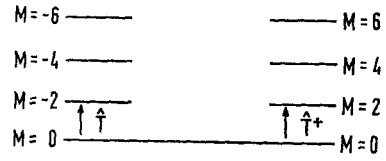


FIG. 4. The "rotational band."

eigenvector is just T, so the large eigenvalue of the G matrix really indicates a collective state. It is also connected with a strong particle–hole correlation (as we assumed it should be): Whenever a state i is occupied, the corresponding state $\bar{i}$ is certainly empty, and vice versa.

C. Rotations

For simplicity, only a simple model of a 2-dimensional rotating system will be considered (see Fig. 2). Again we have N particles to be located in two N-fold degenerate levels, which, however, differ now by the value of the magnetic quantum number m, rather than by a radial quantum number r. Besides the magnetic quantum number, the single-particle wavefunctions differ also in an internal quantum number i. In two dimensions, the total angular momentum M is just $M = \sum m$.

We shall imitate the construction of rotational bands, which proved quite successful in nuclear physics. In nonclosed-shell nuclei, the Hartree–Fock calculations describe the ground state by a Slater determinant corresponding to a deformed mass distribution. Such a Slater determinant defines the direction of the body axis of the nucleus and,therefore,cannot possess a definite angular momentum (according to Heisenberg's indeterminacy principle). The physical states, however, do possess a definite angular momentum. The deformed Slater determinant can then be considered to describe the "internal motion" of the nucleus, while one obtains physical states by projecting components with definite angular momentum out of such a Slater determinant. In the calculations of light nuclei, such projections really produced rotational bands in quite good agreement with experiment.

It is well known that,the more deformed the system is, the lower lie the rotational collective states (see Fig. 3) and the more enhanced transition amplitude they have. So let us construct in our simple model the most deformed Slater determinant. It consists of single-particle states

$$(\varphi_{i+} + \varphi_{i-}) = \chi_i(e^{i\Phi} + e^{-i\Phi}) = 2\chi_i \cos\Phi. \quad (61)$$

The rotational band is obtained by applying the projectors P_M, and the state $M = 0$ will be considered as ground state. The members of the rotational band are then the following:

$$|\Psi_M\rangle = P_M \Pi_i (i^\dagger + \bar{i}^\dagger)\,|0\rangle, \quad (62)$$

as given in Fig. 4. In this case, the ground-state G matrix has two large eigenvalues $\lambda_1 = \lambda_2 = \frac{1}{4}(N + 2)$. Their eigenvectors correspond to the following two operators:

$$\hat{T}^\dagger = \sum_i i^\dagger \bar{i}, \quad (63)$$

$$\hat{T} = \sum_i \bar{i}^\dagger i. \quad (64)$$

These two operators generate from the ground state the neighboring rotational states:

$$|\Psi_2\rangle = \hat{T}^\dagger|\Psi_0\rangle, \quad |\Psi_{-2}\rangle = \hat{T}|\Psi_0\rangle. \quad (65)$$

At the same time, these are the transition operators for which the transition amplitude is strongly enhanced.

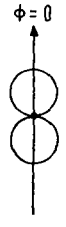


FIG. 3. The polar diagram of the mass distribution in the simple rotational model.

4. USE OF THE G MATRIX IN VARIATIONAL CALCULATIONS

The G matrix was first introduced into the literature on density matrices by Garrod and Percus[9] in an analysis of the "N-representability problem." This is the problem of finding what restrictions must be placed upon the 2-body density matrix in order to guarantee that there exists at least one fermion N-particle wavefunction from which the given density matrix can be derived.

The objective of the research done on the N-representability problem was the derivation of a variational technique which would permit the variational calculation of the ground-state 2-body density matrix without any explicit calculation of the very complicated N-particle wavefunction.

It was shown in Ref. 9 that the demand (as a

restriction on ρ_{abcd}) that the associated G matrix be nonnegative very much improved the results obtained by variational calculations utilizing the 2-body density matrix. However, repeated attempts to utilize the G-matrix nonnegativity restriction in numerical calculations of atomic density matrices by Garrod ran into serious calculational problems. The source of the difficulty is the fact that the G-matrix nonnegativity condition includes the Pauli principle. Why this fact leads to difficulties can be seen from the following line of reasoning. Suppose we use as the starting point of our variational calculation the Hartree–Fock density matrix associated with the atom. The most obvious way to change the density matrix so as to lower the energy is simply to increase the occupation of one of the lower-lying H–F states. But any such increase in occupation of those states would violate the Pauli principle and thus lead to at least one negative eigenvalue of G. Thus all these possible energy-lowering moves lead into the forbidden region of whatever space of variational parameters one is using. Thus, in order to lower the energy without violating the nonnegativity restriction on G, one must move along the edge of a multidimensional "wedge" in a direction almost perpendicular to the energy gradient in the space. When the number of parameters is large, this can become an extremely difficult calculational feat.

The obvious (by hindsight) way of avoiding the above problem was discovered by Mihailović and Rosina.[10] They expressed the energy in terms of the G matrix which they parameterized in such a way as to make it automatically nonnegative. The more complicated symmetry conditions which must then be placed on G [see Eq. (8)] rather than on ρ lead to some difficulties, but these do not seem insurmountable. The results they obtained show a great improvement over the traditional Hartree–Fock and RPA methods.

ACKNOWLEDGMENT

Most of this work was formulated while the authors were guests of Queen's University, Kingston, Ontario. They would like to thank Professor A. J. Coleman for the fine hospitality they received. Thanks are also due to M. V. Mihailović for critical remarks on collective states.

[10] M. V. Mihailović and M. Rosina, "Nuclear Many-Body Problems," in *Proceedings of the 12th International Summer Meeting in Physics at Herceg Novi, Yugoslavia*, M V. Mihailović, M. Rosina, and J. Strnad, Eds. (Federal Nuclear Energy Commission of Yugoslavia, Beograd, 1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Diffusion in Nonlinear Multiplicative Media

JOSÉ CANOSA
IBM Scientific Center, Palo Alto, California 94304

(Received 8 July 1968)

The time-dependent behavior of the nonlinear distributions defined by the diffusion equation with several nonlinear source terms is studied. The nonlinear diffusion equation is solved by an eigenfunction-expansion method, which is in principle independent of geometry or number of dimensions. The qualitative time behavior of the distributions and their steady states can be ascertained from a simple analysis of the fundamental mode approximation only. Explicit solutions are presented in one- and two-dimensional geometries.

I. INTRODUCTION

The diffusion equation with nonlinear source terms describes approximately the behavior of the neutron population in a nuclear reactor, where the neutron multiplication is a nonlinear function of the population itself. Similar problems also arise in the study of heat conduction in material media with nonlinear heat generation. In this paper, we discuss the time-dependent solutions and steady-state distributions of several problems of this type, which have in common certain time-independent distributions.

The solution is obtained by a general eigenfunction-expansion method, which in principle is independent of geometry and number of dimensions. In practice, its application is straightforward only in multidimensional domains where the coordinates are separable. Explicit solutions are presented for plane and finite cylindrical geometries. An interesting result of this study is that the stability and time behavior of the distributions are qualitatively the same as the stability and time behavior of the fundamental mode amplitude.

II. NONLINEAR DIFFUSION EQUATIONS

The nonlinear diffusion equations studied arose in the description of the time behavior of a nuclear reactor. We consider three different problems, characterized by different types of nonlinearities. In dimensionless form, they read as follows:

$$\nabla^2 y(\mathbf{x}, t) + p^2(\mathbf{x})y - \delta E(\mathbf{x}, t)y = \frac{\partial y}{\partial t}, \qquad \frac{\partial E(x, t)}{\partial t} = y - \mu E; \quad (1)$$

$$\nabla^2 y(\mathbf{x}, t) + p^2(\mathbf{x})y - \delta E(\mathbf{x}, t)y = \frac{\partial y}{\partial t}, \qquad \frac{\partial E(x, t)}{\partial t} = y; \quad (2)$$

$$\nabla^2 y(\mathbf{x}, t) + p^2(\mathbf{x})y - \delta y^2 = \frac{\partial y}{\partial t}. \quad (3)$$

The initial and boundary conditions are

$$y(\mathbf{x}, 0) = f(\mathbf{x}) > 0, \quad y(\Gamma, t) = 0, \quad (4)$$

$$E(\mathbf{x}, 0) = 0, \quad E(\Gamma, t) = 0, \quad (5)$$

and (5) only applies to (1) and (2). The multiplication factor $p^2(\mathbf{x})$ which appears in the linear source term of (1), (2), and (3) can be considered as an external input quantity such that $p^2(\mathbf{x})y$ gives the initial multiplication of the neutron distribution y; $p^2(\mathbf{x})$ is positive, at least piecewise continuous, and in all cases it is sufficiently large so that at time zero the neutron distribution is increasing with time and not in a steady state. In (1), (2), and (3), δ is a positive constant; and in (4) and (5), Γ is the domain boundary. The variable $E(\mathbf{x}, t)$ in (1) and (2) has the meaning of an energy density. According to (1), its rate of change at any point is the difference of a production term and a destruction term; the production term is proportional to the neutron distribution (to which the reactor power is proportional), and the destruction term, proportional to the energy density itself, represents some energy-removal mechanism ($\mu > 0$). In (2), the energy density increases monotonically with time as long as the neutron distribution is not zero, because there is no energy-removal mechanism. Formally, Eq. (2) is a particular case of (1), when in this latter we set $\mu = 0$.

The reason for considering the three problems (1), (2), and (3) simultaneously becomes apparent below. It suffices now to anticipate that although the time behavior of the distributions is quite different in each case, all three problems have in common certain steady-state distributions defined essentially by the following nonlinear boundary-value problem:

$$\nabla^2 Z(\mathbf{x}) + p^2(\mathbf{x})Z - CZ^2 = 0, \qquad Z(\Gamma) = 0, \quad (6)$$

where C is a positive constant characteristic of each case. In the following sections, we first discuss briefly

the steady-state distributions and then present the formal method of solution to obtain the time behavior. Finally, we conclude with a discussion of the results obtained.

III. STEADY-STATE DISTRIBUTIONS

In nonlinear problems such as those considered here, the steady-state or stationary solutions are obtained by setting the time derivatives equal to zero. This is obviously a necessary condition for the existence of such solutions. In this way, by inspection of (1), we obtain

$$\nabla^2 Z(\mathbf{x}) + p^2(\mathbf{x})Z - \delta\mu^{-1}Z^2 = 0, \qquad Z(\Gamma) = 0, \tag{7}$$

and

$$\lim_{t\to\infty} E(\mathbf{x}, t) = \mu^{-1} \lim_{t\to\infty} y(\mathbf{x}, t) = \mu^{-1}Z(\mathbf{x}). \tag{8}$$

From (2), we get by inspection

$$\lim_{t\to\infty} \frac{\partial E(\mathbf{x}, t)}{\partial t} = \lim_{t\to\infty} y(\mathbf{x}, t) = 0. \tag{9}$$

The asymptotic distribution of the energy density was obtained by Ergen[1] by reducing (2) to an equation in the energy density alone, i.e.,

$$\nabla^2 \dot{E} + p^2(\mathbf{x})\dot{E} - \delta E\dot{E} = \ddot{E}, \tag{10}$$

where the dots designate partial derivatives with respect to time. Equation (10) is now integrated with respect to t from 0 to ∞, and we get

$$\nabla^2 Z(\mathbf{x}) + p^2(\mathbf{x})Z - \tfrac{1}{2}\delta Z^2 = -f(\mathbf{x}), \qquad Z(\Gamma) = 0, \tag{11}$$

having used the initial conditions (4) and (5). We have

$$\lim_{t\to\infty} E(\mathbf{x}, t) = Z(\mathbf{x}). \tag{12}$$

Equation (11) is an inhomogeneous nonlinear boundary-value problem whose homogeneous part is identical to (7). Actually, for large $p^2(\mathbf{x})$, the inhomogeneous term hardly affects the solution of (11), and it can be neglected.[2]

By inspection of (3), we get for its steady-state solution

$$\nabla^2 Z(\mathbf{x}) + p^2(\mathbf{x})Z - \delta Z^2 = 0, \qquad Z(\Gamma) = 0, \tag{13}$$

where

$$\lim_{t\to\infty} y(x, t) = Z(\mathbf{x}). \tag{14}$$

Therefore, the nonzero steady-state solutions of (1), (2), and (3), defined by (7), (11), and (13), are given essentially by the same nonlinear boundary-value problem. However, it will be shown that the time behavior of the distributions is quite different, even qualitatively.

It must be noticed that rigorous mathematical proofs of (8), (9), (12), and (14), are quite difficult to obtain. The results (9), (12), and (14) have actually been proved by Kastenberg and Chambré.[3] They are important because they show conclusively that the steady-state solutions are asymptotically stable against any perturbation in the distributions. This is quite apparent because the steady-state solutions, with the single exception of (11), are rigorously *independent* of the initial distributions. Physically, it means that the steady-state solutions are solely determined by the material properties of the medium $p^2(\mathbf{x})$, δ, and μ, and that the initial distributions are entirely forgotten at the end of the transient [with the single exception of the steady-state solution defined by (11); however, see discussion following (12)]. It must be noticed that it is essential to know that the steady-state distributions of (1), (2), and (3) are unique. This requires that the positive solutions defined by (7), (11), and (13) also be unique. For one-dimensional slab problems (6), the uniqueness of the positive solution has been proved rigorously[4] and is now assumed on physical grounds for multidimensional problems. In some applications, we might only be interested in the steady-state distributions; then it is only necessary to solve nonlinear boundary-value problems of the type (6), a subject which has been treated in earlier work.[5] The emphasis of this paper is on the time-dependent solutions of (1), (2), and (3).

IV. EIGENFUNCTION EXPANSION METHOD

The formalism is now developed briefly for problem (1). In essence, it is an extension of the method developed for nonlinear boundary-value problems[5] to the solution of time-dependent nonlinear partial-differential equations. Its application to (2) and (3) proceeds along the same lines and we only mention the results. The solution of (1) is expressed in series form:

$$y(\mathbf{x}, t) = \sum_{n=1}^{\infty} P_n(t)\varphi_n(\mathbf{x}),$$

$$E(\mathbf{x}, t) = \sum_{n=1}^{\infty} T_n(t)\varphi_n(\mathbf{x}), \tag{15}$$

[1] W. K. Ergen, Trans. Am. Nucl. Soc. **8**, 221 (1965).

[2] Assume that in (11), $f(\mathbf{x}) = O(1)$ and $p^2(\mathbf{x}) = \nu\rho(\mathbf{x})$, $\rho(\mathbf{x}) = O(1)$ and $\nu \gg 1$. Define a new variable $Z(\mathbf{x}) = \nu z(\mathbf{x})$; then (11) becomes $(1/\nu)\nabla^2 z(\mathbf{x}) + \rho(\mathbf{x})z - (\delta/2)z^2 = -f(\mathbf{x})/\nu^2$. For $\nu \to \infty$, the inhomogeneous term can be neglected and the problem is of the singular perturbation type; for slab geometry, this problem has been solved by means of matched asymptotic expansions by J. Canosa and J. Cole, J. Math. Phys. **9**, 1915 (1968).

[3] W. E. Kastenberg and P. L. Chambré, Nucl. Sci. Eng. **31**, 67 (1968).

[4] Canosa and Cole, Ref. 2; the same proof applies to inhomogeneous equations such as (11).

[5] J. Canosa, J. Math. Phys. **8**, 2180 (1967); **9**, 2032 (1968).

where φ_n satisfy the eigenvalue problem associated with (1), i.e.,

$$\nabla^2\varphi(\mathbf{x}) + p^2(\mathbf{x})\varphi - \lambda\varphi = 0, \qquad \varphi(\Gamma) = 0. \tag{16}$$

It should be noticed that the eigenvalues and eigenfunctions defined by (16) give the solution of the linear problem associated with (1), i.e.,

$$y_L(\mathbf{x}, t) = \sum_{n=1}^{\infty} A_n\varphi_n(\mathbf{x})e^{\lambda_n t}, \tag{17}$$

where A_n would be determined by the initial conditions. The expansions (15) are now substituted into (1) and truncated after j terms; with the use of (16), we obtain

$$\sum_{n=1}^{j}(\lambda_n P_n(t) - \dot{P}_n)\varphi_n(\mathbf{x}) = \delta\sum_{m=1}^{j} P_m\varphi_m\sum_{n=1}^{j} T_n\varphi_n,$$
$$\sum_{n=1}^{j}\dot{T}_n\varphi_n = \sum_{n=1}^{j}(P_n - \mu T_n)\varphi_n, \tag{18}$$

where the dots represent time derivatives. We assume that (16) defines a complete set of orthogonal functions in the domain, i.e.,

$$\int_D \varphi_m\varphi_n\, dV = \delta_{mn}b^{nn}, \tag{19}$$

where dV is the elementary volume element, δ_{mn} is the Kronecker symbol, and b^{nn} is the nth-norm. The justification of (19) is that, in one-dimensional geometries, (16) is a Sturm–Liouville problem when $p^2(\mathbf{x})$ is a continuous function[6] and (19) obtains. As we limit ourselves to separable geometries, the resulting eigenvalue problems in each space variable are also of the Sturm–Liouville type and, therefore, (19) holds also for the multi-dimensional eigenfunctions. The orthogonality property (19) is now used to eliminate the spatial dependence from (18), by multiplying it by φ_j and integrating over the domain. In this way, we obtain a system of $2j$ nonlinear ordinary differential equations in the expansion coefficients $P_n(t)$ and $T_n(t)$ $(n = 1, 2, \cdots, j)$ of (15). Explicitly,

$$\begin{aligned}
dP_1/dt = {} & \lambda_1 P_1 - (\delta/b^{11})[b_1^{11}P_1T_1 + b_1^{22}P_2T_2 \\
& + \cdots + b_1^{jj}P_jT_j + b_2^{11}(P_1T_2 + P_2T_1) \\
& + \cdots + b_j^{11}(P_1T_j + P_jT_1) \\
& + b^{1,2,3}(P_2T_3 + P_3T_2) \\
& + \cdots + b^{1,2,j}(P_2T_j + P_jT_2) \\
& + \cdots + b^{1,j-1,j}(P_{j-1}T_j + P_jT_{j-1})],
\end{aligned}$$

$$\begin{aligned}
dP_2/dt = {} & \lambda_2 P_2 - (\delta/b^{22})[b_2^{11}P_1T_1 + b_2^{22}P_2T_2 \\
& + \cdots + b_2^{jj}P_jT_j + b_1^{22}(P_1T_2 + P_2T_1) \\
& + \cdots + b^{1,2,j}(P_1T_j + P_jT_1) \\
& + b_3^{22}(P_2T_3 + P_3T_2) \\
& + \cdots + b_j^{22}(P_2T_j + P_jT_2) \\
& + \cdots + b^{2,j-1,j}(P_{j-1}T_j + P_jT_{j-1})],
\end{aligned}$$

$$\vdots$$

$$\begin{aligned}
dP_j/dt = {} & \lambda_j P_j - (\delta/b^{jj})[b_j^{11}P_1T_1 + b_j^{22}P_2T_2 \\
& + \cdots + b_j^{jj}P_jT_j + b^{1,2,j}(P_1T_2 + P_2T_1) \\
& + \cdots + b_1^{jj}(P_1T_j + P_jT_1) \\
& + b^{2,3,j}(P_2T_3 + P_3T_2) \\
& + \cdots + b_2^{jj}(P_2T_j + P_jT_2) \\
& + \cdots + b_{j-1}^{jj}(P_{j-1}T_j + P_jT_{j-1})],
\end{aligned} \tag{20}$$

$$dT_1/dt = P_1 - \mu T_1,$$
$$dT_2/dt = P_2 - \mu T_2,$$
$$\vdots$$
$$dT_j/dt = P_j - \mu T_j.$$

For problem (2), the resulting system is (20) with $\mu = 0$. For problem (3), we get a system similar to (20), whose explicit form can be obtained from (20) by substituting the T's by P's (e.g., P_jT_l becomes P_jP_l) and deleting that last j equations. The notation in (20), valid for any geometry and number of dimensions, is as follows:

$$b_m^{ll} = \int_D \varphi_l^2\varphi_m\, dV, \quad b^{l,m,n} = \int_D \varphi_l\varphi_m\varphi_n\, dV,$$
$$b^{jj} = \int_D \varphi_j^2\, dV, \tag{21}$$

where the eigenfunctions are normalized so that the maximum value of the fundamental mode is unity. In separable two- and three-dimensional problems the eigenfunctions and eigenvalues require two and three indexes; however, for convenience of notation they have been ordered in (20) and (21) with only one index according to the magnitude of the eigenvalues.

V. ONE-DIMENSIONAL PROBLEMS

1. Homogeneous Slab Geometry

The focal point of this work is the study of the time-dependent behavior of the nonlinear distributions defined by (1), (2), and (3). This behavior is in the main independent of geometry. For this reason, we first present the results and discussion for a homogeneous slab of thickness π. In this case, the

[6] For $p^2(\mathbf{x})$ piecewise continuous, the orthogonality can be shown in the same way as in a standard Sturm–Liouville problem.

multiplication factor $p^2(\mathbf{x})$ is a constant, and the eigenfunctions and eigenvalues (16) are obtained in closed form together with all the integrals (21).[5]

The numerical solution of (1), (2), and (3) was obtained for the following values of the parameters:

$$\delta = 2, \quad \mu = 1, \quad p^2 = 10, \tag{22}$$

which lead to a violent transient with strong nonlinear effects. To fix ideas, we say that the nonlinear effects are strong when the steady-state distributions defined by (6) depart strongly from the fundamental mode of (16), which in these examples is $\varphi_1 = \sin x$. It has been shown[4] that the nonlinear effects are strong if p^2 is appreciably higher than unity as in (22). As the steady-state distributions are independent of the initial condition, for simplicity we took

$$y(x, 0) = f(x) = 0.1 \sin x. \tag{23}$$

The system of nonlinear ordinary differential equations (20) was solved without difficulty by a standard fourth-order Runge–Kutta method.

Once the coefficients $P_n(t)$ and $T_n(t)$ are computed as a function of time, the distributions $y(x, t)$ and $E(x, t)$ can be obtained at any instant t by computing the linear combinations (15). The calculations were performed by keeping the first three terms in the expansions (15),[7] and the results were displayed on an IBM 2250 Display Unit and photographed. The results obtained are shown in the figures. Their accuracy was considered satisfactory because the steady-state distributions obtained from these time-dependent calculations were in excellent agreement with a direct and independent numerical solution of the corresponding nonlinear boundary-value problem (6). Figure 1(a) shows as a function of time the amplitude of the fundamental mode defined by the coupled system (20), $P_1(t)$. The calculation was terminated after 10 dimensionless time units, when the steady state was essentially reached. The maximum value reached was almost 300 times the initial value. The plot is typical of a nonlinear oscillation. The other plots in Fig. 1 show that the distribution $y(x, t)$ also reaches its steady-state value in an oscillatory manner. Figures 1(b) and 1(c) show the first cycle of oscillation of the distribution; in Fig. 1(d), the distribution is shown again in a rising phase, and in Figure 1(e) we see the approach to the asymptotic steady state. The approximately corresponding time behavior of the fundamental mode amplitude and of the distribution must be noted. With a one-mode approximation,

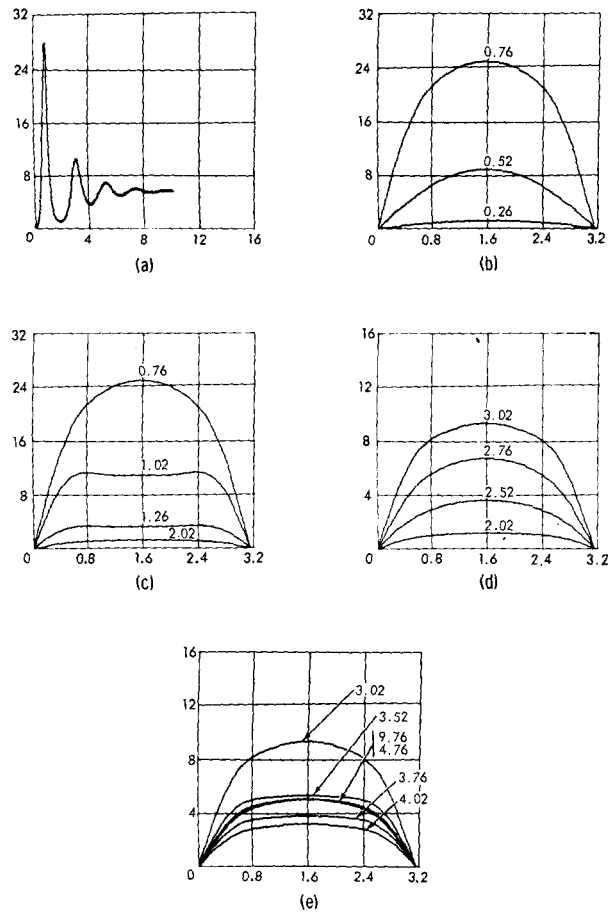


FIG. 1. Time history of the distribution showing the oscillatory approach to the steady state. Problem (1). Dimensionless units. The numbers on the curves are values of time. (a) Fundamental mode amplitude vs time. (b) Rising phase of distribution. (c) Decreasing phase of distribution. (d) Rising phase of distribution, second cycle. (e) Approach to the steady state.

i.e., keeping only the first term in (15), system (20) reduces to

$$dP_1/dt = \lambda_1 P_1 - (\delta b_1^{11}/b^{11}) P_1 T_1,$$
$$dT_1/dt = P_1 - \mu_1 T_1. \tag{24}$$

It should be stressed that the fundamental mode amplitude P_1 defined by (24) has the same qualitative behavior as shown in Fig. 1. Therefore, an analysis of the fundamental mode alone gives the main features of the time behavior and asymptotic stability of the whole distribution.

In Fig. 2, the time history of the distribution defined by (2) is shown. We used the same initial condition and parameter values as in the previous calculation for problem (1), i.e.,

$$\delta = 2, \quad p^2 = 10, \quad y(x, 0) = 0.1 \sin x. \tag{25}$$

Also, only three terms were kept in (15). The accuracy of the results was checked as in the previous calculation. Figure 2(a) shows as a function of time the

[7] For problems (1) and (2), the resulting systems (20) have six equations and six variables; for problem (3) the corresponding system has three equations and three variables.

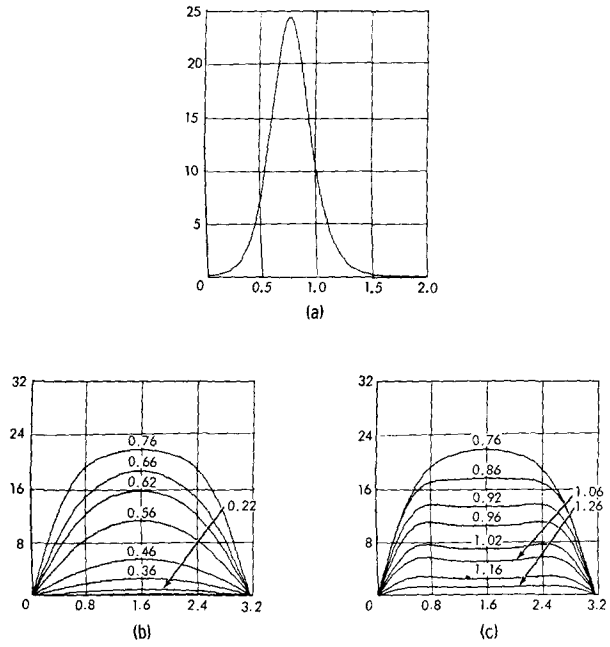


FIG. 2. Time history of the distribution. Problem (2). (a) Fundamental mode amplitude vs time. (b) Rising phase of distribution. (c) Decreasing phase of distribution.

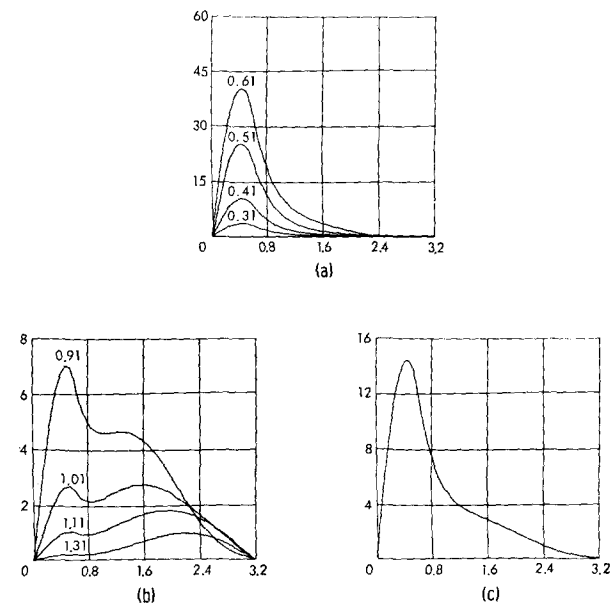


FIG. 4. Time history of the distribution. Problem (2) for a composite asymmetric slab. (a) Rising phase of distribution. (b) Decreasing phase of distribution. (c) Asymptotic energy density distribution.

fundamental mode amplitude defined by the coupled system (20), with $\mu = 0$. Figures 2(b) and 2(c) show the time behavior of the whole distribution. The corresponding behavior of the fundamental mode and of the distribution is again to be noticed. Figure 2(b) shows that at the beginning of the transient the behavior is linear, i.e., the distribution is given essentially by the fundamental mode ($\sin x$) whose amplitude grows with time. As the distribution approaches its maximum, the nonlinear effects due to the deposition of energy begin to be felt, and the shape starts flattening. After the maximum value is reached, the flattening increases and then a dip appears at the center with two maxima at the sides. Thereafter, the nonlinear effects progressively smooth the distribution, although the last discernible shapes are still dipped at the center.

In Fig. 3, we show the results for problem (3) with

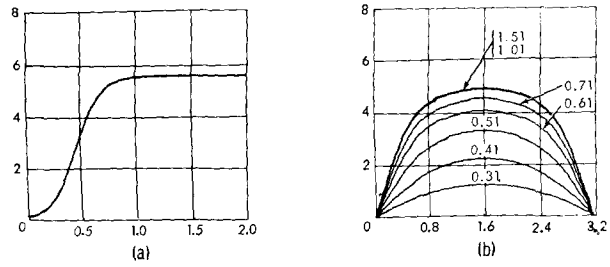


FIG. 3. Time history of the distribution. Problem (3). (a) Fundamental mode amplitude vs time. (b) Monotonic approach of the distribution to the steady state.

the same initial conditions and parameters as before [Eq. (25)]. In this case, the fundamental mode amplitude and the distribution approach monotonically their asymptotic steady states.

2. Composite Slab Geometry

In order to study the shape changes of the distribution in a more complex situation, we treated problem (2) with the following asymmetric multiplication factor:

$$p^2(x) = 25 \quad (0 \leq x \leq \pi/5),$$
$$= 1 \quad (\pi/5 \leq x \leq \pi), \tag{26}$$

and with the following initial condition and parameter value:

$$y(x, 0) = 0.1\varphi_1(x), \quad \delta = 2, \tag{27}$$

where $\varphi_1(x)$ is the fundamental eigenfunction of (16). The solution of eigenvalue problems (16) with piecewise constant coefficients as in (26) is obtained analytically by elementary means. The actual calculation of the eigenvalues and eigenfunctions and of the integrals (21) is done numerically in a straightforward way.[8]

In Fig. 4 we show the time history of the distribution. Only certain representative distributions were chosen so as not to confuse the figure. Figure 4(a) shows the rising phase of the distribution during which the

[8] J. Canosa, IBM Scientific Center (Palo Alto) Report No. 320-3240, 1968.

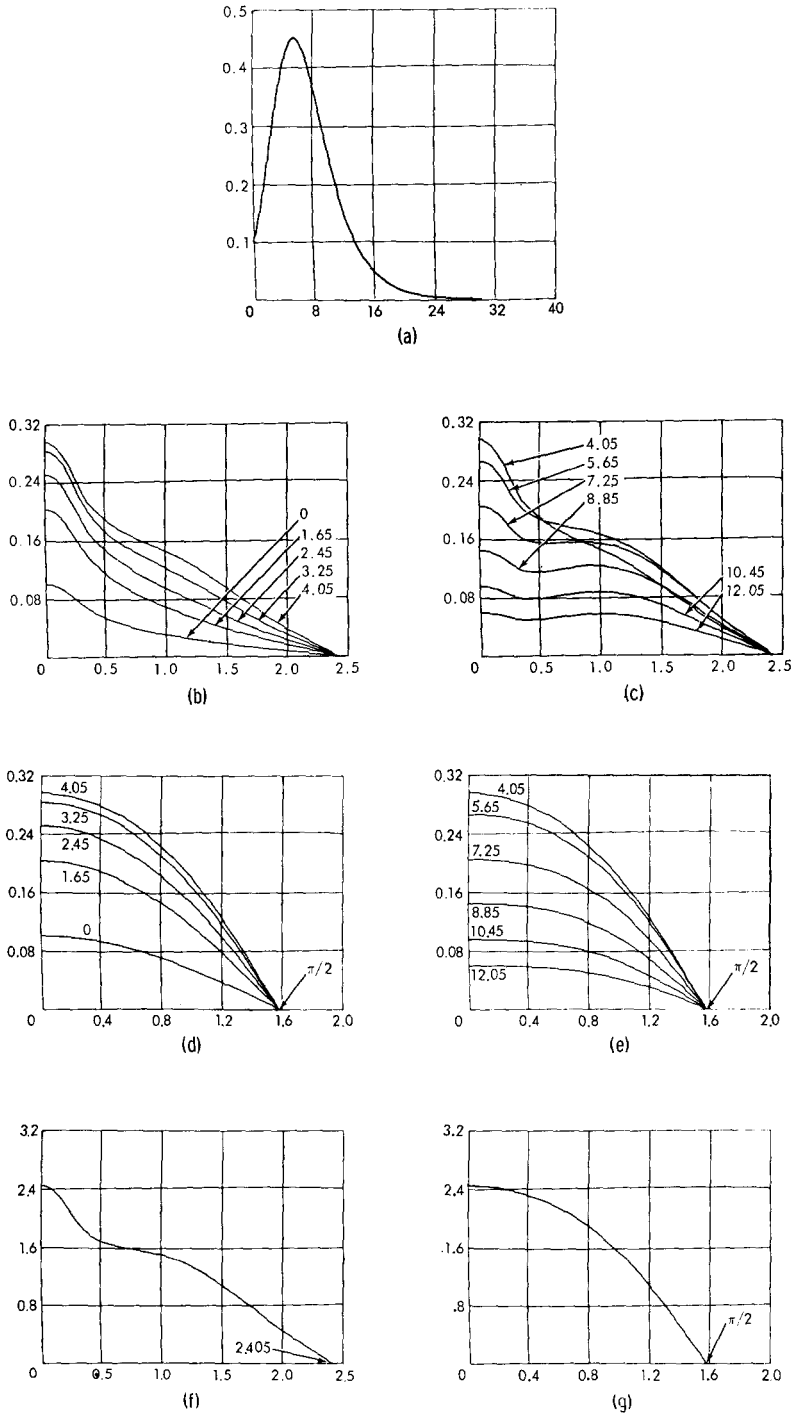


FIG. 5. Time history of the two-dimensional distribution defined by (29) and (30). Radially-composite cylinder. (a) Fundamental mode amplitude vs time. (b) Rising phase of radial distribution on plane $z = 0$. (c) Decreasing phase of radial distribution. (d) Rising phase of axial distribution on axis $r = 0$. (e) Decreasing phase of axial distribution. (f) Steady-state radial energy distribution. (g) Steady-state axial energy distribution.

behavior is linear, i.e., the shape is that of the fundamental mode of (16) and only its amplitude changes with time. The maximum reached by the peak of the distribution is approximately 420 times the initial value. After the maximum is reached, the nonlinear effects begin to be felt in the form of the shape changes shown in Fig. 4(b). The maximum of the distribution at the left disappears gradually, and the last vanishing shapes have a broad maximum at the right-hand side. In Fig. 4(c), we show the asymptotic energy distribution defined by (2), i.e.,

$$E(x, \infty) = \int_0^\infty y(x, t)\, dt. \tag{28}$$

This distribution is the record of $y(x, t)$ during the transient; indeed, it is the sum of all the distributions $y(x, t)$ weighted by the time of their persistence.

In doing the calculations shown in Fig. 4, we kept only the first three terms in the expansions (15). The accuracy was checked again by observing the excellent agreement between the steady-state energy distribution (28) obtained from the time-dependent calculation (20) and from the corresponding nonlinear boundary-value problem (11).

VI. TWO-DIMENSIONAL PROBLEM

In two- and three-dimensional geometries where the eigenvalue problem (16) can be solved by separation of variables, the solution of (1), (2), and (3) does not require a much greater effort than for one-dimensional problems. In fact, the nonlinear ordinary differential equation system (20) has the same form for any geometry or number of dimensions. All that is required is to solve (16) and obtain the integrals (21) involving the eigenfunctions. As an illustration, we have studied (2) for a two-dimensional $(r - z)$ composite cylinder, i.e.,

$$\nabla^2 y(r, z, t) + 2p^2(r)y - 2E(r, z, t)y = 2\frac{\partial y}{\partial t},$$

$$\frac{\partial E}{\partial t} = y, \tag{29}$$

$$p^2(r) = 8, \quad \text{for} \quad 0 \le r \le \tfrac{1}{10}k_{01},$$

$$= 1, \quad \text{for} \quad \tfrac{1}{10}k_{01} \le r \le k_{01},$$

$$k_{01} = \text{first zero of } J_0(r).$$

$J_0(r)$ is the Bessel function of the first kind, of order zero. The cylinder dimensions are normalized such that its height is π and its radius is k_{01}. The coefficient 2 in (29) is used for convenience of notation. The boundary and initial conditions are as before [(4) and (5)]; specifically,

$$f(r, z, 0) = 0.1\varphi_1(r, z), \tag{30}$$

where $\varphi_1(r, z)$ is the fundamental eigenfunction of (16).

The eigenvalue problem (16) was solved by separation of variables. Numerically, the eigenvalues are given by a somewhat complicated transcendental equation involving Bessel functions, and the eigenfunctions are given as products of a radial part involving Bessel functions and an axial part involving circular functions. The numerical calculations are quite straightforward with the use of a computer and are described elsewhere.[8] The first five two-dimensional eigenfunctions were kept in the expansions (15). It was observed that when going from the two-mode to the five-mode approximation, the quantitative differences obtained in the solutions were quite small; in the lack of an independent numerical check, this was considered satisfactory proof of the accuracy of the five-mode approximation used. The results obtained are shown in Fig. 5.

The time behavior of the fundamental mode amplitude [Fig. 5(a)] is of course the same as for the one-dimensional problems (2). Figures 5(b) and 5(c) show the time-history of the distribution in the radial plane $z = 0$. In the rising phase of the distribution [Fig. 5(b)], the behavior is, as before, mostly linear; only the amplitude of the distribution (the fundamental radial mode essentially) varies with time. After it reaches the maximum, nonlinear effects manifest themselves as shape changes in the decreasing distributions shown in Fig. 5(c). Figures 5(d) and 5(e) show, respectively, the increasing and decreasing phases of the axial distribution on axis $r = 0$. Hardly any shape changes are detected. The axial shape is essentially the fundamental axial eigenfunction (cos z) during all the transient. The reason for the predominance of the nonlinear effects in the radial direction is clear, because the cylinder is homogeneous axially and composite radially. Finally, in Figs. 5(f) and 5(g), the steady-state energy distributions are shown in the radial ($z = 0$) and axial ($r = 0$) planes, respectively. It is recalled that, in the present two-dimensional problem, the asymptotic energy distribution is given by

$$E(r, z, \infty) = \int_0^\infty y(r, z, t)\, dt. \tag{31}$$

ACKNOWLEDGMENT

The author is deeply grateful to Dr. Ralph M. Warten for the computer graphics program used for the plotting of the solutions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Observables for Massive Relativistic Particles of Arbitrary Spin

J. A. DE AZCÁRRAGA AND L. OLIVER*
Instituto de Física Teórica, Facultad de Ciencias, Barcelona, Spain

(Received 7 February 1969)

The expressions for the dynamical operators corresponding to a Bargmann–Wigner particle of mass κ and arbitrary spin s are given, making use of the Foldy–Wouthuysen and Chakrabarti transformations for the Bargmann–Wigner equations given by Pursey, and Azcárraga and Boya, respectively. The operators selecting the positive and negative energy states are also given in the two representations. The operators for arbitrary spin are studied and a condition for an operator be constant of motion is given. Finally, the Newton–Wigner position operator and those obtained in the Foldy–Wouthuysen and Chakrabarti generalized representations are compared.

1. INTRODUCTION

In the last 15 years some effort has been made to obtain transformations which separate the components of positive and negative energy of the solutions of the manifestly covariant equations; these transformations allow us to define new representations which have some interesting features. Foldy and Wouthuysen[1] and Tani[2] were the first to give a unitary transformation [the Foldy–Wouthuysen (FW) transformation] which diagonalizes the Hamiltonian obtained from the manifestly covariant form of the Dirac equation. This transformation provides a new representation, the FW representation, which gives, in a very natural manner, the transition to the description in the nonrelativistic limit of a spin-$\frac{1}{2}$ particle, and, in the usual representation, associates convenient operators to some dynamical variables, such as the mean position and the mean velocity.

After this work,[1] much else has been accomplished. Case[3] extended the FW transformation to the spin-0 and spin-1 particles by considering a Hamiltonian form which does not come from a manifestly covariant equation; Foldy[4] applied the FW transformation to the Klein–Gordon, Dirac, and Proca equations to give a "canonical" form of them, and Feshbach and Villars[5] emphasized the physical content of it for the two-component meson and the Dirac particle. Also, Garrido and Pascual,[6] starting from the general relativistic equation given by Umezawa,[7] gave a transformation for the Dirac and Duffin–Kemmer[8] cases, obtaining a natural generalization of the FW transformation. Recently, Bryden[9] has obtained the transformed form of the Rarita–Schwinger[10] equations for spin $\frac{3}{2}$, and Pursey[11] has generalized the FW transformation for the Bargmann–Wigner (BW) equations[12] of arbitrary spin.

Giambiagi[13] was the first to remark that there exists an analogy between the FW transformation and a Lorentz transformation. This fact was also studied by Bollini and Giambiagi[14] and Good and Rose,[15] who showed that, in the Dirac theory of the free electron, the FW transformation corresponds to the Lorentz one which brings the lab system to the rest system plus a factor which is related to the norm of the wavefunction. Recently, Sankaranarayanan[16] has shown the connection between the Garrido–Pascual[6] transformation for a Kemmer particle and that of Lorentz which goes from the lab to the rest system; and Saavedra[17] has studied for spin $\frac{1}{2}$ the FW transformation, making explicit the results of Refs. 14 and 15.

Chakrabarti[18,19] goes beyond the mentioned results; in his paper he proposes a new transformation which is precisely that which brings the lab system to the rest system. This transformation allows the separation of the two-sign energy states and also allows the definition of an operator in the usual representation which corresponds more directly to the classical definition of the position operator given by Pryce[20] than the one defined by means of the FW transformation, in spite of the fact that Pryce's method leads to the latter. On the other hand, this transformation gives rise to a new representation of the Poincaré

* Present address: Laboratoire de Physique Théorique et Hautes Énergies, Faculté des Sciences, Orsay.

[1] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950).
[2] S. Tani, Progr. Theoret. Phys. (Kyoto) **6**, 267 (1951).
[3] K. M. Case, Phys. Rev. **95**, 1323 (1954).
[4] L. L. Foldy, Phys. Rev. **102**, 568 (1956).
[5] H. Feshbach and F. Villars, Rev. Mod. Phys. **30**, 24 (1958).
[6] L. M. Garrido and P. Pascual, Nuovo Cimento **12**, 181 (1959).
[7] H. Umezawa, *Quantum Field Theory* (North-Holland Publ. Co., Amsterdam, 1956), Chap. 5.
[8] K. M. Case, Phys. Rev. **100**, 1513 (1955).
[9] A. D. Bryden, Nucl. Phys. **53**, 165 (1964).
[10] W. Rarita and J. Schwinger, Phys. Rev. **60**, 61 (1941).
[11] D. L. Pursey, Nucl. Phys. **53**, 174 (1964).
[12] V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. U.S. **34**, 211 (1948).
[13] J. J. Giambiagi, Nuovo Cimento **16**, 202 (1960).
[14] C. G. Bollini and J. J. Giambiagi, Nuovo Cimento **21**, 107 (1961).
[15] R. H. Good and E. M. Rose, Nuovo Cimento **24**, 864 (1962).
[16] A. Sankaranarayanan, Nuovo Cimento **34**, 443 (1964).
[17] I. Saavedra, Nucl. Phys. **74**, 677 (1965).
[18] A. Chakrabarti, J. Math. Phys. **4**, 1215 (1963).
[19] A. Chakrabarti, J. Math. Phys. **4**, 1223 (1963).
[20] M. H. L. Pryce, Proc. Roy. Soc. (London) **A195**, 62 (1948).

group which has to be considered as the *canonical* one obtained by Wigner[21] making use of the little group of the vector $(\kappa, 0, 0, 0)$, which is the typical one for the class $p^\mu p_\mu = \kappa^2 > 0$. This has been made explicitly for the general case[22] using the generalized Chakrabarti transformation for the Proca and BW equations which leads to the canonical form of them. On the other hand, Chakrabarti[19] has remarked that the Newton–Wigner (NW) position operator[23] does not coincide exactly with the mean-position operator deduced by means of the FW transformation, and has established, for spin $\frac{1}{2}$, the relation between them and between the NW position operator and its own operator. Lately, the Chakrabarti transformation has been applied also for the spin-0 and spin-1 cases by Sesma.[24]

In this paper we obtain in an easy way the explicit form of the operators for an arbitrary-spin BW particle by making use of the generalized FW and Chakrabarti transformations for the BW equations. The BW equations have no appropriate Hamiltonian formulation, as the Dirac and Duffin–Kemmer equations do, but it will be shown that, even in absence of this Hamiltonian form, it is possible to define operators of mean position, mean velocity, mean spin, etc., with suitable physical properties. Recently, Sankaranarayanan and Good[25] have also given formulas for position operators of any spin; however, his approach is different from ours, being based on a previous description of an arbitrary-spin particle, closely related to Weinberg's,[26] made by Weaver, Hammer, and Good.[27] The explicit form of the operators is complicated for high spin, but of course for $s = \frac{1}{2}$ they obtain the FW mean-position operator.

In order to get the mentioned results, Sec. 2 is devoted to the introduction of the named transformations for the BW equations, performing explicitly the separation of the eigenstates of the energy. The operators corresponding to the dynamical variables in the FW and BW representations are obtained in Sec. 2, where we find in particular an operator for the mean position which shows the "Zitterbewegung" and operators for mean-spin and mean-orbital angular momentum which are separately constant of motion in a sense which is adequately defined. Section 4 is dedicated to BW and Chakrabarti representations in an analogous way; the obtained operators have the convenient physical properties. Finally, Sec. 5 is dedicated to the comparison of the mean-position operators obtained in the previous sections with the corresponding NW position operator for arbitrary spin.

2. THE FOLDY-WOUTHUYSEN-PURSEY AND THE GENERALIZED CHAKRABARTI TRANSFORMATIONS

The BW equations for particles of mass κ and arbitrary spin take the well-known form[12]

$$(\gamma^{\mu(r)}p_\mu - \kappa)\psi(p, \xi_1 \cdots \xi_r \cdots \xi_{2s}) = 0, \quad (2.1)$$

where ψ is a symmetric function of the $2s$ four-valued variables ξ having in all $2(2s+1)$ independent components because it includes the two signs of the energy. If ψ is considered as an element of the 4^{2s}-dimensional space obtained by taking the tensor product of the $2s$ four-dimensional spaces, the $\gamma^{\mu(r)}$ can be written in the form

$$\gamma^{\mu(r)} = 1 \otimes \cdots \otimes 1 \otimes \gamma^\mu \otimes 1 \otimes \cdots \otimes 1, \quad (2.2)$$

the γ^μ being placed in the rth position. From (2.2) one easily verifies that

$$\gamma^{\mu(r)}\gamma^{\nu(r)} + \gamma^{\nu(r)}\gamma^{\mu(r)} = 2g^{\mu\nu(r)} \quad (2.3a)$$

and

$$\gamma^{\mu(r)}\gamma^{\nu(s)} = \gamma^{\nu(s)}\gamma^{\mu(r)}, \quad (2.3b)$$

where $g^{\mu\nu} = 0$ for $\mu \neq \nu$ and $g^{00} = -g^{ii} = 1$, $i = 1, 2, 3$.

Equations (2.1) may be written

$$p_0\psi = (\boldsymbol{\alpha}^{(r)}\mathbf{p} + \beta^{(r)}\kappa)\psi, \quad (2.4)$$

where

$$\alpha^{i(r)} = \gamma^{0(r)}\gamma^{i(r)} = (\gamma^0\gamma^i)^{(r)}, \qquad \beta^{(r)} = \gamma^{0(r)}. \quad (2.5)$$

In these conditions, the generalized FW transformation or Foldy–Wouthuysen–Pursey (FWP) transformation takes the form[11]

$$F = \prod_{r=1}^{2s} F^{(r)}, \quad (2.6a)$$

where

$$F^{(r)} = [2\omega(\omega + \kappa)]^{-\frac{1}{2}}[\omega + \kappa + \beta^{(r)}\boldsymbol{\alpha}^{(r)}\mathbf{p}] \quad (2.6b)$$

or

$$F^{(r)} = \exp\left(\tfrac{1}{2}\beta^{(r)}\frac{\boldsymbol{\alpha}^{(r)}\mathbf{p}}{p}\tan^{-1}\frac{p}{\kappa}\right) \quad (2.6c)$$

with $\omega = +(\kappa^2 + p^2)^{\frac{1}{2}}$ and $p = (\mathbf{pp})^{\frac{1}{2}}$.

---

[21] E. P. Wigner, Ann. Math. **40**, 149 (1939).

[22] J. A. de Azcárraga and L. J. Boya, J. Math. Phys. **9**, 1689 (1968). The notation used in this paper is different from that in the present one.

[23] T. D. Newton and E. P. Wigner, Rev. Mod. Phys. **21**, 400 (1949).

[24] J. Sesma, J. Math. Phys. **7**, 1300 (1966).

[25] A. Sankaranarayanan and R. H. Good, Jr., Phys. Rev. **140**, B509 (1965).

[26] S. Weinberg, Phys. Rev. **133**, B1318 (1964).

[27] D. L. Weaver, C. L. Hammer, and R. H. Good, Jr., Phys. Rev. **135**, B241 (1964). In this paper a Hamiltonian form of wave equations for arbitrary spin is derived, starting from a postulated canonical form of it by means of a procedure which is, in some way, inverse to that of Foldy and Wouthuysen. The complexity of the Hamiltonians obtained increases with the spin.

Under this transformation, Eqs. (2.4) become

$$p_0\psi_F = \beta^{(r)}\omega\psi_F, \tag{2.7}$$

with

$$\psi_F = F\psi. \tag{2.8}$$

From (2.7) it follows that

$$\beta^{(r)}\psi_F = \beta^{(s)}\psi_F \tag{2.9}$$

and that

$$\psi_F = [\tfrac{1}{2}(1+\beta)]^{(r)}\psi_F + [\tfrac{1}{2}(1-\beta)]^{(s)}\psi_F. \tag{2.10}$$

Then, taking into account that

$$[\tfrac{1}{2}(1\pm\beta)]^{(r)}[\tfrac{1}{2}(1\mp\beta)]^{(r)} = 0, \tag{2.11}$$

from (2.10) we obtain

$$\psi_F = \psi_F^{(+)} + \psi_F^{(-)}, \tag{2.12}$$

where

$$\psi_F^{(\pm)} = \Lambda^{(\pm)}\psi_F. \tag{2.13}$$

$\Lambda^{(\pm)}$ are the projectors

$$\Lambda^{(+)} = \prod_{r=1}^{2s}[\tfrac{1}{2}(1+\beta^{(r)})], \quad \Lambda^{(-)} = \prod_{r=1}^{2s}[\tfrac{1}{2}(1-\beta^{(r)})], \tag{2.14}$$

which satisfy

$$\Lambda^{(+)} + \Lambda^{(-)} \approx I. \tag{2.15}$$

[$A \approx B$ means $(A-B)\psi = 0$.]

The relations

$$\begin{aligned}\Lambda^{(+)}\beta^{(r)} &= +\Lambda^{(+)},\\ \Lambda^{(-)}\beta^{(r)} &= -\Lambda^{(+)}\end{aligned} \tag{2.16}$$

allow us to write from Eqs. (2.7) the dynamical equations

$$\begin{aligned}p_0\psi_F^{(+)} &= \omega\psi_F^{(+)},\\ p_0\psi_F^{(-)} &= -\omega\psi_F^{(-)}.\end{aligned} \tag{2.17}$$

Then, $\psi^{(+)}$ and $\psi^{(-)}$ are, according to (2.17), the solutions of positive and negative energy of (2.7). As is well known, $\psi^{(+)}$ and $\psi^{(-)}$ have $2s+1$ independent components.

We consider now the Chakrabarti generalized transformation for Eqs. (2.1).[22] The explicit form of this transformation is

$$Q = \prod_{r=1}^{2s} Q^{(r)}, \tag{2.18a}$$

where

$$Q^{(r)} = [2m(m+p_0)]^{-\frac{1}{2}}[p_0 + m - \boldsymbol{\alpha}^{(r)}\mathbf{p}] \tag{2.18b}$$

or

$$Q^{(r)} = \exp\left(\frac{1}{2}\frac{\boldsymbol{\alpha}^{(r)}\mathbf{p}}{p}\tanh^{-1}\frac{p}{p_0}\right), \tag{2.18c}$$

with $p_0 = \pm\omega$ and $m = \pm(p_\mu p^\mu)^{\frac{1}{2}}$.

This transformation applied to Eqs. (2.1) gives

$$(m\gamma^{0\,(r)} - \kappa)\psi_Q = 0, \tag{2.19}$$

where

$$\psi_Q = Q\psi. \tag{2.20}$$

Proceeding in an analogous way as we did in the FWP case, we define

$$\psi_Q^{(\pm)} = \Lambda^{(\pm)}\psi_Q \tag{2.21}$$

and obtain the relations

$$\begin{aligned}m\psi_Q^{(+)} &= \kappa\psi_Q^{(+)},\\ m\psi_Q^{(-)} &= -\kappa\psi_Q^{(-)},\end{aligned} \tag{2.22}$$

which are similar to Eqs. (2.17).

Finally, we consider the relation between ψ_Q and ψ_F. From the definition it is evident that

$$\psi_F = FQ'^{-1}\psi_Q, \tag{2.23}$$

with

$$Q' = \prod_{r=1}^{2s}[2\omega(\omega+\kappa)][\omega + \kappa - \epsilon\boldsymbol{\alpha}^{(r)}\mathbf{p}], \tag{2.24}$$

which from (2.18b) involves the substitution for p_0 of $\epsilon\omega$ and for m of $\epsilon\kappa$ (where $\epsilon = p_0/|p_0|$). Then we obtain the relation

$$\psi_F = (\omega/\kappa)^s\psi_Q, \tag{2.25}$$

which for $s = \frac{1}{2}$ coincides with the well-known one.[15,18] (In Ref. 15, $\kappa = 1$.)

3. OBSERVABLES IN THE BW AND FWP REPRESENTATIONS

We consider now the explicit form of the observables in the BW and FWP representations. The Bargmann–Wigner equations for any spin have no suitable Hamiltonian formulation; however, the form of Eq. (2.4) shows that

$$p_0\psi = i\frac{\partial}{\partial t}\psi = h^{(r)}\psi, \tag{3.1}$$

with

$$h^{(r)} = \boldsymbol{\alpha}^{(r)}\mathbf{p} + \beta^{(r)}\kappa. \tag{3.2}$$

Then it is natural, in analogy with the $s = \frac{1}{2}$ case, to define

$$\dot{O} = i[h^{(r)}, O] = i[\boldsymbol{\alpha}^{(r)}\mathbf{p} + \beta^{(r)}\kappa, O], \quad \forall r. \tag{3.3}$$

We show now that this definition is consistent with the dynamical properties of the observables; in particular, let us prove the unicity, from the dynamical point of view, of the operator associated to the velocity of the particle. According to (3.3) we associate to the dynamical variable velocity $\dot{\mathbf{x}}$ any of the operators

$$\dot{\mathbf{x}} = i[h^{(r)}, \mathbf{x}] = \boldsymbol{\alpha}^{(r)}. \tag{3.4}$$

TABLE I. Operators in the BW and the FWP representations.

	Operators in the BW representation	Operators in the FWP representation
Position	$\mathbf{x}$	$\mathbf{x}_F = \mathbf{x} - \frac{i\beta\boldsymbol{\alpha}}{2\omega} - \frac{(\boldsymbol{\sigma}\wedge\mathbf{p})\omega - i\beta(\boldsymbol{\alpha}\mathbf{p})\mathbf{p}}{2\omega^2(\omega+\kappa)}$
Velocity	$\dot{\mathbf{x}} = \boldsymbol{\alpha}^{(r)} \quad (r = 1, \cdots, 2s)$	$\dot{\mathbf{x}}_F = \boldsymbol{\alpha}^{(r)} + \beta^{(r)}\frac{\mathbf{p}}{\omega} - \frac{(\boldsymbol{\alpha}^{(r)}\mathbf{p})\mathbf{p}}{\omega(\omega+\kappa)}; \quad (r = 1, \cdots, 2s)$
Momentum	$\mathbf{p}$	$\mathbf{p}_F = \mathbf{p}$
Energy	$p^0 \approx h^{(r)} = (\boldsymbol{\alpha}^{(r)}\mathbf{p} + \beta^{(r)}\kappa) \quad (r = 1, \cdots, 2s)$	$h_F^{(r)} = \beta^{(r)}\omega \quad (r = 1, \cdots, 2s)$
Spin	$\mathbf{s} = \frac{1}{2}\boldsymbol{\sigma}, \quad \sigma^k = i\sum_{r=1}^{2s}\gamma^{i(r)}\gamma^{j(r)}i, j, k, \text{ cycl.}$	$S_F = \frac{\kappa}{\omega}\frac{\boldsymbol{\sigma}}{2} + \frac{(\boldsymbol{\sigma}\mathbf{p})\mathbf{p}}{2\omega(\omega+\kappa)} + \frac{i\beta(\boldsymbol{\alpha}\wedge\mathbf{p})}{2\omega}$
Orbital angular momentum	$\mathbf{x}\wedge\mathbf{p}$	$\mathbf{x}_F\wedge\mathbf{p}$
Mean position	$\mathbf{X} = \mathbf{x} + \frac{i\beta\boldsymbol{\alpha}}{2\omega} - \frac{(\boldsymbol{\sigma}\wedge\mathbf{p})\omega + i\beta(\boldsymbol{\alpha}\mathbf{p})\mathbf{p}}{2\omega^2(\omega+\kappa)}$	$\mathbf{X}_F = \mathbf{x}$
Mean velocity	$\dot{\mathbf{X}} = \frac{\mathbf{p}}{\omega}\frac{\boldsymbol{\alpha}^{(r)}\mathbf{p} + \beta^{(r)}\kappa}{\omega} \approx \frac{\mathbf{p}}{p_0} \quad (r = 1, \cdots, 2s)$	$\dot{\mathbf{X}}_F = \beta^{(r)}\frac{\mathbf{p}}{\omega} \approx \frac{\mathbf{p}}{p_0} \quad (r = 1, \cdots, 2s)$
Mean spin	$\mathbf{S} = \frac{\kappa}{\omega}\frac{\boldsymbol{\sigma}}{2} + \frac{(\boldsymbol{\sigma}\mathbf{p})\mathbf{p}}{2\omega(\omega+\kappa)} - \frac{i\beta(\boldsymbol{\alpha}\wedge\mathbf{p})}{2\omega}$	$\mathbf{S}_F = \mathbf{s}$
Mean orbital angular momentum	$\mathbf{X}\wedge\mathbf{p}$	$\mathbf{x}\wedge\mathbf{p}$
Projection operator	$\Lambda^{(\pm)} = \prod_{r=1}^{2s}\left[\frac{1}{2}\left(1 \pm \frac{\boldsymbol{\alpha}^{(r)}\mathbf{p} + \beta^{(r)}\kappa}{\omega}\right)\right]$	$\Lambda_F^{(\pm)} = \prod_{r=1}^{2s}[\frac{1}{2}(1 \pm \beta^{(r)})]$

All of these operators satisfy the same differential equation

$$\ddot{\boldsymbol{\alpha}}^{(r)} + 4\omega^2\boldsymbol{\alpha}^{(r)} = 4h^{(r)}\mathbf{p} \approx 4p_0\mathbf{p} \tag{3.5}$$

or, equivalently,

$$\dddot{\mathbf{x}} + 4\omega^2\dot{\mathbf{x}} \approx 4p_0\mathbf{p}, \tag{3.6}$$

from which is found that

$$\dot{\mathbf{x}} \approx \mathbf{p}/p_0 + \mathbf{A}\exp(2i\omega t) + \mathbf{B}\exp(-2i\omega t), \tag{3.7}$$

an expression which shows the dynamical equivalence between the operators $\boldsymbol{\alpha}^{(r)}$ associated to the velocity $\dot{\mathbf{x}}$. (3.7) also shows the "Zitterbewegung" for particles of any spin. The different operators in the BW and FWP representations are given in Table I; in the tables[28]

$$\boldsymbol{\sigma} \equiv \sum_{r=1}^{2s}\boldsymbol{\sigma}^{(r)}, \quad \boldsymbol{\alpha} \equiv \sum_{r=1}^{2s}\boldsymbol{\alpha}^{(r)}, \quad \beta\boldsymbol{\alpha} \equiv \sum_{r=1}^{2s}\beta^{(r)}\boldsymbol{\alpha}^{(r)}.$$

The mean position is defined in the usual way[1]; its time derivative, the mean velocity $\dot{\mathbf{X}}$, is equal to the first term appearing on the right-hand side of (3.7). The position $\mathbf{x}$ can be split in two parts: $\mathbf{X}$ describes in average the evolution of the particle, and $\mathbf{x} - \mathbf{X}$ oscillates around this mean position with an amplitude of order of the Compton wavelength of the particle $1/m$. On the other hand, the mean spin and the mean orbital angular momentum are separately constants of motion, because

$$[h^{(r)}, S] = 0, \quad [h^{(r)}, \mathbf{X}\wedge\mathbf{p}] = 0, \quad \forall r. \tag{3.8}$$

It is interesting to note that our condition for an operator O to be constant of motion

$$[h^{(r)}, O] = 0, \quad \forall r, \tag{3.9}$$

is the same as in the Dirac case and compatible with the spin-1 case, taking into account the Hamiltonian formulation of the Duffin–Kemmer (DK) wave equation. As it is well known, a free Kemmer particle of mass κ can be described by the equation

$$(\beta_\mu\partial_\mu + \kappa)\psi = 0, \tag{3.10}$$

where the β's satisfy the structural commutation rules of the DK ring,

$$\beta_\lambda\beta_\mu\beta_\nu + \beta_\nu\beta_\mu\beta_\lambda = \beta_\lambda\delta_{\mu\nu} + \beta_\nu\delta_{\mu\lambda}. \tag{3.11}$$

Using only the algebraic properties satisfied by the β's, Case has shown[8] that (3.10) is equivalent to the Hamiltonian equation

$$p_0\psi = (\boldsymbol{\alpha}'\mathbf{p} + \beta_4)\psi = H\psi, \tag{3.12}$$

where

$$\boldsymbol{\alpha}' = -i(\boldsymbol{\beta}\beta_4 - \beta_4\boldsymbol{\beta}), \tag{3.13}$$

[28] For instance, in the case of the position operator,

$$\begin{aligned} F\partial F^{-1} &= (F\otimes F\cdots\otimes F)\partial(F^{-1}\otimes F^{-1}\cdots\otimes F^{-1}) \\ &= (F\otimes F\cdots\otimes F)(\partial F^{-1}\otimes F^{-1}\cdots\otimes F^{-1}) \\ &\quad + \cdots + (F\otimes F\cdots\otimes F)(F^{-1}\otimes F^{-1}\cdots\otimes\partial F^{-1}) \\ &= (F\partial F^{-1}\otimes 1\cdots\otimes 1) + (1\otimes 1\otimes\cdots\otimes F\partial F^{-1}) \\ &= \sum_{r=1}^{2s} F^{(r)}\partial F^{(r)-1}. \end{aligned}$$

with the initial condition that

$$(H\beta_4 - \kappa)\psi = 0 \tag{3.14}$$

independently of the particular representation chosen for the β's.

It is not difficult to see that the BW equations for a spin-1 particle,

$$p_0\psi = h^{(1)}\psi,$$
$$p_0\psi = h^{(2)}\psi, \tag{3.15}$$

where ψ is *symmetric* in the two indexes ξ_r, imply (3.12) with the initial condition (3.14). The representation of the β's is in this case

$$\beta'_\mu = \tfrac{1}{2}[(\gamma'_\mu \otimes 1) + (1 \otimes \gamma'_\mu)], \quad \mu = 1, 2, 3, 4, \tag{3.16}$$

where $\gamma'_4 = \gamma^0$ and $\boldsymbol{\gamma}' = i\boldsymbol{\gamma}$, and with this representation H takes the form

$$H = \tfrac{1}{2}(h^{(1)} + h^{(2)}), \tag{3.17}$$

which shows that if an observable for a spin-1 BW particle is constant of motion in the sense of (3.9) it is also constant of motion for the DK case.

4. OBSERVABLES IN THE BARGMANN-WIGNER AND IN THE CHAKRABARTI GENERALIZED REPRESENTATIONS

Following a procedure similar to that used in Sec. 3, we obtain the observables for both BW and CG representations; we take $p_0 = \epsilon(\kappa^2 + \mathbf{p}^2)^{\frac{1}{2}}$ and $m = \epsilon\kappa$. The results are shown in Table II.

We note that for the mean velocity we obtain

$$\dot{\mathbf{X}} \approx \mathbf{p}/p_0, \tag{4.1}$$

which justifies the denomination "mean velocity" given to this operator [see (3.7)]. Also in that case, the mean spin and the mean orbital angular momentum are separately constant of motion in the defined sense.

Again in Table II, $\boldsymbol{\alpha}$ and $\boldsymbol{\sigma}$ have the same meaning as in Table I.

5. THE NEWTON-WIGNER POSITION OPERATOR AND THE FWP AND CG REPRESENTATIONS

The Newton–Wigner (NW) position operator[23] for the scalar product defined in Ref. 12, Eq. (18a), has the form

$$q_N^k = E \prod_{r=1}^{2s} (1 + \gamma^{0(r)}) \times \frac{p_0^{2s+\frac{1}{2}}}{(p_0 + \kappa)^s}\left(-i\frac{\partial}{\partial p_k}\right)\frac{p_0^{-\frac{1}{2}}}{(p_0 + \kappa)^s} E, \tag{5.1}$$

where E is the operator

$$E = \prod_{r=1}^{2s}\left[\frac{1}{2}\left(1 + \frac{\boldsymbol{\alpha}^{(r)}\mathbf{p} + \beta^{(r)}\kappa}{\omega}\right)\right], \tag{5.2}$$

TABLE II. Operators in the BW and the CG representations.

Position	$\mathbf{x}$	$\mathbf{x}_Q = \mathbf{x} + \frac{i\boldsymbol{\alpha}}{2m} + \frac{1}{2m(m+p^0)}\left[(\boldsymbol{\sigma}\wedge\mathbf{p}) - \frac{i(\boldsymbol{\alpha}\mathbf{p})\mathbf{p}}{p^0}\right]$
Velocity	$\dot{\mathbf{x}} = \boldsymbol{\alpha}^{(r)} \quad (r = 1, \cdots, 2s)$	$\dot{\mathbf{x}}_Q = \frac{p_0}{m}\boldsymbol{\alpha}^{(r)} - \frac{i(\boldsymbol{\sigma}^{(r)}\wedge\mathbf{p})}{m} - \frac{(\boldsymbol{\alpha}^{(r)}\mathbf{p})\mathbf{p}}{m(m+p_0)}$ $(r = 1, \cdots, 2s)$
Momentum	$\mathbf{p}$	$\mathbf{p}$
Mass	$\beta^{(r)}p^0 - \beta^{(r)}\boldsymbol{\alpha}^{(r)}\mathbf{p} \quad (r = 1, \cdots, 2s)$	$\beta^{(r)}m \quad (r = 1, \cdots, 2s)$
Spin	$\mathbf{s} = \frac{1}{2}\boldsymbol{\sigma}$	$\mathbf{s}_Q = \frac{1}{2m}\left[p_0\boldsymbol{\sigma} - i(\boldsymbol{\alpha}\wedge\mathbf{p}) - \frac{(\boldsymbol{\sigma}\mathbf{p})\mathbf{p}}{(p_0+m)}\right]$
Orbital angular momentum	$\mathbf{x}\wedge\mathbf{p}$	$\mathbf{x}_Q\wedge\mathbf{p}$
Mean position	$\mathbf{X} = \mathbf{x} - \frac{i\boldsymbol{\alpha}}{2m} + \frac{1}{2m(m+p_0)} \times \left[(\boldsymbol{\sigma}\wedge\mathbf{p}) + \frac{i(\boldsymbol{\alpha}\mathbf{p})\mathbf{p}}{p_0}\right]$	$\mathbf{X}_Q = \mathbf{x}$
Mean velocity	$\dot{\mathbf{X}} = \left(\frac{(\boldsymbol{\alpha}^{(r)}\mathbf{p})\mathbf{p}}{p_0 m(m+p_0)} - \frac{\boldsymbol{\alpha}^{(r)}}{m}\right)(\beta^{(r)}\kappa - p^0) + \frac{i(\boldsymbol{\sigma}^{(r)}\wedge\mathbf{p})}{m} \approx \frac{\mathbf{p}}{p_0}$	$\dot{\mathbf{X}}_Q = \boldsymbol{\alpha}^{(r)}\left(1 - \frac{\beta^{(r)}\kappa}{m}\right) + \frac{\mathbf{p}}{p_0 m}\beta^{(r)}\kappa \approx \frac{\mathbf{p}}{p_0} \quad (r = 1, \cdots, 2s)$
Mean spin	$\mathbf{S} = \frac{1}{2m}\left[p^0\boldsymbol{\sigma} + i(\boldsymbol{\alpha}\wedge\mathbf{p}) - \frac{(\boldsymbol{\sigma}\mathbf{p})\mathbf{p}}{(p^0+m)}\right]$	$\mathbf{S}_Q = \mathbf{s}$
Mean orbital angular momentum	$\mathbf{X}\wedge\mathbf{p}$	$\mathbf{x}\wedge\mathbf{p}$
Projection operator	$\Lambda^{(\pm)} = \prod_{r=1}^{2s}\left(1 \pm \frac{\beta^{(r)}\omega - \beta^{(r)}\boldsymbol{\alpha}^{(r)}\mathbf{p}}{\kappa}\right)$	$\Lambda_Q^{(\pm)} = \prod_{r=1}^{2s}\frac{1}{2}(1 \pm \beta^{(r)})$

which projects an arbitrary function $\phi(p, \xi_1 \cdots \xi_{2s})$ with $\omega^2 - p^2 = \kappa^2$ and symmetric in the ξ indexes over a positive-energy solution of (2.4).

It is interesting to see its form in the FWP representation. After calculations, one obtains

$$\mathbf{q}_N = F^{-1}[\omega^{(s+\frac{1}{2})}\mathbf{x}\omega^{-(s+\frac{1}{2})}] \prod_{r=1}^{2s} [\tfrac{1}{2}(1 + \beta^{(r)})]F, \quad (5.3)$$

where

$$\prod_{r=1}^{2s} [\tfrac{1}{2}(1 + \beta^{(r)})]$$

is the positive-energy-projection operator in the FWP representation, or

$$F\mathbf{q}_N F^{-1} = \left(\mathbf{x} - i(s + \tfrac{1}{2})\frac{\mathbf{p}}{\omega^2}\right) \prod_{r=1}^{2s} [\tfrac{1}{2}(1 + \beta^{(r)})]. \quad (5.4)$$

The terms in this expression are not difficult to interpret:

$$\prod_{r=1}^{2s} [\tfrac{1}{2}(1 + \beta^{(r)})]$$

is the energy-projection operator, and the additional term $-i(s + \frac{1}{2})(\mathbf{p}/\omega^2)$ makes the resulting operator Hermitian.

In order to consider now the CG representation and taking into account the other scalar product which can be defined for the solutions of the BW equations [Eq. (16) of Ref. 12], we write

$$\mathbf{q}'_N = E' \prod_{r=1}^{2s} (1 + \gamma^{0(r)}) \frac{p_0^{\frac{1}{2}}}{(p_0 + \kappa)^s} \mathbf{x} \frac{p_0^{-\frac{1}{2}}}{(p_0 + \kappa)^s} E', \quad (5.5)$$

where E' is the covariant projector

$$E' = \prod_{r=1}^{2s} \left(\frac{\gamma^\mu p_\mu + \kappa}{2\kappa}\right)^{(r)}. \quad (5.6)$$

We now obtain

$$\mathbf{q}'_N = Q^{-1}(\omega^{\frac{1}{2}}\mathbf{x}\omega^{-\frac{1}{2}}) \prod_{r=1}^{2s} [\tfrac{1}{2}(1 + \gamma^{0(r)})]Q \quad (5.7)$$

and

$$Q\mathbf{q}'_N Q^{-1} = \left(\mathbf{x} - i\frac{\mathbf{p}}{\omega^2}\right) \prod_{r=1}^{2s} [\tfrac{1}{2}(1 + \gamma^{0(r)})]. \quad (5.8)$$

Again, the term $-i(\mathbf{p}/\omega^2)$ makes the resulting operator Hermitian owing to the presence of the volume element $d\mathbf{p}/\omega$ in the scalar product.

ACKNOWLEDGMENTS

The authors acknowledge supporting grants from the Ministerio de Educación y Ciencia (J. A.) and from the Fundación Juan March (L. O.).

APPENDIX

Besides the more familiar properties of the α and γ Dirac matrices, in calculating operators we have made use of the relations

$$(\boldsymbol{\alpha}\mathbf{p})\boldsymbol{\alpha} = \mathbf{p} + i(\boldsymbol{\sigma} \wedge \mathbf{p}), \quad (A1)$$

$$[\boldsymbol{\sigma}, \boldsymbol{\alpha}\mathbf{p}] = -2i(\boldsymbol{\alpha} \wedge \mathbf{p}), \quad (A2)$$

$$(\boldsymbol{\alpha}\mathbf{p})(\boldsymbol{\alpha} \wedge \mathbf{p}) = i((\boldsymbol{\sigma}\mathbf{p})\mathbf{p} - \mathbf{p}^2\boldsymbol{\sigma}), \quad (A3)$$

$$(\boldsymbol{\sigma} \wedge \mathbf{p})(\boldsymbol{\alpha}\mathbf{p}) = i\mathbf{p}^2\boldsymbol{\alpha} - i\mathbf{p}(\boldsymbol{\alpha}\mathbf{p}), \quad (A4)$$

$$[\boldsymbol{\alpha}, (\boldsymbol{\alpha}\mathbf{p})] = -2i(\boldsymbol{\sigma} \wedge \mathbf{p}), \quad (A5)$$

whose generalization to the BW case is immediate.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Properties of Generalizations to Padé Approximants

A. K. COMMON
Department of Applied Mathematics, University of Kent at Canterbury

(Received 24 January 1968)

The Padé-approximant method has recently been generalized in such a way that convergence of the new approximants can be proved for a larger class of functions than the series of Stieltjes and conditions can be obtained for these approximants to form sequences of converging upper and lower bounds. In this paper properties of these approximants to a class of functions, which correspond to a series of Stieltjes with nonzero radius of convergence, are considered. It is shown that under certain conditions they form only sequences of lower bounds to the exact function, but one can then define new approximants, which give upper bounds. Conditions are also obtained under which an upper bound can be put on the error of an approximant, using only the information necessary to calculate this approximant.

1. INTRODUCTION

The Padé-approximant method is very useful for the approximate analytic continuation of a series outside its circle of convergence. Although sequences of such approximants seem to converge for a large class of functions, it can only be proved for a series of Stieltjes. In that case, the Padé approximants have moreover the very important property that they form monotonically converging sequences of upper and lower bounds to the exact function for points on the positive real axis,[1] so that one can then put strict limits on the error of the approximants.

Recently Baker[2] has discussed a generalization of the Padé method to obtain converging upper and lower bounds for a wider class of functions. This new class of functions is of the type

$$g(z) = \int_0^\infty b(z, s)\, d\phi(s), \tag{1.1}$$

where $d\phi(s) \geq 0$ and the corresponding approximants are

$$B_{n,j}(z) = \sum_{m=1}^{n} \alpha_m b(z, \sigma_m) + \sum_{k=0}^{j} \beta_k \left[\frac{\partial^k b(z, s)}{\partial s^k}\right]_{s=0} \tag{1.2}$$

The constants α_m, σ_m, and β_k are obtained from the $[n, n+j]$ Padé approximant to the corresponding series of Stieltjes

$$f(z) = \int_0^\infty \frac{d\phi(s)}{1 + zs}. \tag{1.3}$$

Approximants of the above type have been suggested previously by Gammel[3] for the case when $b(z, s) \equiv b(zs)$. Baker has therefore generalized these latter approximants and has proved a number of theorems on them which we will now describe.[4] The latter author proves that a sufficient condition for the $B_{n,j}(z)$ to converge as $n \to \infty$ for a particular value of z is that $b(z, s)$ is regular in a uniform neighborhood of the positive real s axis and $(\log s)^{1+\eta} b(z, s)$ is bounded as $s \to \infty$, for some $\eta > 0$. He then shows that the $B_{n,j}(z)$ form monotonically converging sequences of upper and lower bounds on the positive real axis if and only if

$$(-1)^j \frac{\partial^j b(z, s)}{\partial s^j} \geq 0, \tag{1.4}$$

for all $z, s \geq 0$ and $j = 0, 1, 2 \cdots$. These results and their derivation are discussed in greater detail in Sec. 2.

The Padé approximants to a series of Stieltjes, with radius of convergence R, converge in the complex plane, cut from $-R$ to $-\infty$. In a recent paper[5] we have discussed the properties of these approximants on the negative real axis between 0 and $-R$. At these points they form converging sequences of lower bounds to the exact function, but not upper bounds, so one cannot put strict limits on the error of the approximants from only a knowledge of the first few coefficients of the series of Stieltjes needed to calculate the approximants.

To remedy this, we defined modified approximants, sequences of which converge to the series of Stieltjes and form upper bounds on the negative real axis between 0 and $-R$. Using these new approximants, which can be obtained from the coefficients of the power series, with the Padé approximants, one can put strict limits on the error of each approximant for the above points, and this result can be extended

[1] A review of the properties of Padé approximants has been given by G. Baker, Jr., Advan. Theoret. Phys. **1**, 1 (1965).

[2] G. Baker, Jr., "Convergent, Bounding Approximation Procedures with Applications to the Ferromagnetic Ising Model," Brookhaven preprint.

[3] For an application, see D. P. Taylor, J. L. Gammel, and C. Rousseau, Bull. Am. Phys. Soc., Ser. (11), **12**, 83 (1967).

[4] Because of the work of these two authors, we would like to suggest the name "Gammel–Baker" approximants for the functions $B_{n,j}(z)$ and will use it for the remainder of this paper.

[5] A. K. Common "Padé Approximants and Bounds to Series of Stieltjes" (to be published in J. Math. Phys.).

to all points in the circle of convergence. The above work is discussed in Sec. 3, where we show that the modified approximant is closely related to a special case of the Gammel–Baker approximants.

The main object of this paper is to show how these results for a series of Stieltjes may be generalized to the functions $g(z)$, given by

$$g(z) = \int_0^{1/R} b(z, s)\, d\phi(s). \tag{1.5}$$

For definiteness we consider the case when $b(z, s)$ is regular for z in the complex plane, cut from $-R$ to $-\infty$, when $0 \leq s \leq 1/R$. The results obtained are easily extended to cases where the region of analyticity of $b(z, s)$ is different from this. The $B_{n,j}(z)$ then converge to $g(z)$ as $n \to \infty$ in the cut plane and in particular for $-R < z \leq 0$. We will suppose that for these values of z

$$\frac{\partial^j b(z, s)}{\partial s^j} \geq 0, \tag{1.6}$$

when $0 \leq s \leq 1/R$ and $j = 0, 1, 2, \cdots$. It is a simple matter to prove from this condition that the $B_{n,j}(z)$ form converging lower bounds to $g(z)$, as is shown in Sec. 4.

To obtain upper bounds, we define a new type of generalized approximant to $g(z)$ which we denote by $C_{n,j}(z)$. It is of the form

$$C_{n,j}(z) = f(0)b(z, 1/R) - \sum_{m=1}^{n} \alpha_m' \left[\frac{\partial b(z, s)}{\partial s}\right]_{s=\sigma_m'} - \sum_{k=0}^{j} \frac{\beta_k'}{k!}\left[\frac{\partial^{k+1} b(z, s)}{\partial s^{k+1}}\right]_{s=0}, \tag{1.7}$$

where α_m', σ_m', and β_k' are obtained from the modified approximant to the corresponding series of Stieltjes $f(z)$. In Sec. 4 we show that $C_{n,j}(z)$ converge to $g(z)$ as $n \to \infty$ in the cut plane and that they give converging upper bounds to $g(z)$ for $-R < z \leq 0$ if condition (1.6) holds.

Therefore one can put strict limits on the error of $B_{n,j}(z)$ and $C_{n,j}(z)$ for these values of z. We will also prove that (1.6) is both a necessary and sufficient condition for the approximants to give the above bounds to $g(z)$, and that if

$$(-1)^N \left[\frac{\partial^{N+j} b(z, s)}{\partial z^N \partial s^j}\right]_{z=0} \geq 0 \tag{1.8}$$

for $0 \leq s \leq 1/R$ and $N, j = 0, 1, 2 \cdots$, one can put strict bounds on the error of the approximants for all z such that $|z| < R$.

In Sec. 5 we summarize the results of this paper and give examples of functions to which they may be applied.

2. PROPERTIES OF GAMMEL–BAKER APPROXIMANTS

The $[n, n + j]$ Padé approximants to the series of Stieltjes

$$f(z) = \sum_{i=0}^{\infty} f_i(-z)^i, \tag{2.1}$$

with

$$f_i = \int_0^{\infty} s^i\, d\phi(s), \tag{2.2}$$

may be written in the form

$$[n, n + j] = \sum_{m=1}^{n} \frac{\alpha_m}{1 + \sigma_m z} + \sum_{k=0}^{j} \frac{\beta_k(-z)^k}{k!}, \tag{2.3}$$

where σ_m, α_m, and β_k are obtained from the f_i in the usual way. The approximant $B_{n,j}(z)$ given by (1.2) with the same constants σ_m, etc., as in (2.3), may be written as

$$B_{n,j}(z) = \int_c b(x, s)[n, n + j](-1/s)\, ds/s, \tag{2.4}$$

where C is a contour encircling the poles of $[n, n + j](-1/s)$ Baker proves that $\lim_{n\to\infty} B_{n,j}(z)$ exists if $b(z, s)$ satisfies the conditions given in the Introduction and that the limit is independent of j and equal to $g(z)$ if $\sum_{i=1}^{\infty} (f_i)^{-1/(2i+1)}$ diverges. The result is true in particular when the upper limit of the integral in the definition (1.1) of $g(z)$ is finite and equal to $1/R$, when one can take C to be a contour encircling the line segment $0 \leq s \leq 1/R$.

To show that the $B_{n,j}(z)$ form monotonically converging sequences of upper and lower bounds to $g(z)$, Baker first proves the relations

$$B_{n+1,j}(z) - B_{n,j}(z) = (+ve \times \text{const}) \times \left[\frac{\partial^{2n+j+1} b(z, s)}{\partial s^{2n+j+1}}\right]_{s=\sigma}, \tag{2.5}$$

$$B_{n,j}(z) - B_{n-1,j+2}(z) = (+ve \times \text{const}) \times \left[\frac{\partial^{2n+j+1} b(z, s)}{\partial s^{2n+j+1}}\right]_{s=\sigma'}, \tag{2.6}$$

where σ' lies between 0 and the largest σ_m appearing in $B_{n,j}(z)$ and $B_{n-1,j+2}(z)$ on the left-hand side of (2.6). σ is similarly defined, except when $j = -1$, and then it must be greater than the smallest σ_m appearing in $B_{n+1,j}(z)$ and $B_{n,j}(z)$ as well. Since all the $\sigma_m \geq 0$, it follows from (1.4) that for $z \geq 0$

$$(-1)^{j+1}\{B_{n+1,j}(z) - B_{n,j}(z)\} \geq 0, \tag{2.7}$$

$$(-1)^{j+1}\{B_{n,j}(z) - B_{n-1,j+2}(z)\} \geq 0, \tag{2.8}$$

and hence that

$$B_{n,0}(z) \geq g(z) \geq B_{n,-1}(z). \tag{2.9}$$

It follows that the $B_{n,0}(z)$ and $B_{n,-1}(z)$ sequences form the best upper and lower bounds obtainable with a given number of coefficients f_i and that the use of additional coefficients improves the bound.

To prove that (1.4) is also a necessary condition for the approximants to give converging upper and lower bounds, Baker first uses (2.7) with $j = -1$, to prove that it is true for even-order derivatives. The odd-order derivatives are then proved to be negative in the following manner. If for instance $\partial b/\partial s$ is positive for some value of s, it remains positive for all greater values of s, since $\partial^2 b/\partial s^2$ is positive. Therefore $b(z, s) \to \infty$ as $s \to \infty$, which contradicts the assumption on its asymptotic behavior, and hence $\partial b/\partial s \leq 0$ for all $s \geq 0$. Similarly, (1.4) may be proved for all odd-order derivatives.

This method of proof for odd derivatives cannot be used when the upper limit of the integral in the definition of $g(z)$ is finite, since we no longer have a condition on the asymptotic behavior of $b(z, s)$. We have not been able to find a proof that condition (1.4) is necessary for odd-order derivatives in this case.

3. APPROXIMANTS TO CONVERGENT SERIES OF STIELTJES

In a recent paper[5] we studied series of Stieltjes $f(z)$ with radius of convergence R for which the Padé approximants converge in the complex plane cut from $-R$ to $-\infty$. They form monotonically increasing sequences of lower bounds to $f(z)$ when $-R < z \leq 0$; and we also proved that, for these values of z, derivatives of the Padé approximants of any order have similar properties, which may be written as

$$(-1)^m \frac{d^m}{dz^m}\{f(z) - [N, N+j]\} \geq 0 \qquad (3.1)$$

for $m = 0, 1, 2 \cdots$.

To put strict limits on the error of an approximant when $-R < z \leq 0$, we had to define a new sequence of approximants $(N, N+j)$, which can be obtained from the first $2N + j + 2$ coefficients of the power series. They converge to $f(z)$ as $N \to \infty$ in the cut plane and, for the above values of z, form monotonically converging upper bounds and satisfy the inequalities

$$(-1)^m \frac{d^m}{dz^m}\{f(z) - (N, N+j)\} \leq 0 \qquad (3.2)$$

for $m = 0, 1, 2 \cdots$.

From (3.1) and (3.2) we obtained the result that if w lies in the circle of convergence of $f(z)$, then

$$|\{f(z) - [N, N+j]\}_{z=w}| \leq |\{[N, N+j] - (N, N+j)\}_{z=-|w|}|. \qquad (3.3)$$

Thus for these values of w, one can put strict limits on the error of the approximant $[N, N+j]$ from the first few coefficients of the power series and similarly for $(N, N+j)$.[6]

We now explain how the approximants $(N, N+j)$ were defined and show their connection with the Gammel–Baker approximants. One can write $f(z)$ in the form

$$f(z) = \int_0^{1/R} \frac{d\phi(u)}{1+uz} = \frac{f_0 R}{R+z} + z[zK'(z) + K(z)], \qquad (3.4)$$

where

$$K(z) = \int_0^{1/R} \frac{\phi(u)\, du}{1+uz}. \qquad (3.5)$$

Without loss of generality we can take $\phi(0) = 0$ and hence $\phi(u) \geq 0$ for $u \geq 0$, so that $K(z)$ is a series of Stieltjes. The coefficients k_i in its power-series expansion

$$K(z) = \sum_{i=0}^{\infty} k_i (-z)^i \qquad (3.6)$$

are given in terms of the f_i by

$$k_i = (i+1)^{-1}\{f_0/(Ri+1) - f_{(i+1)}\}. \qquad (3.7)$$

The $(N, N+j)$ approximant to $f(z)$ is given by replacing $K(z)$ by its $[N, N+j]$ Padé approximant in (3.4). If this Padé approximant is written in the form

$$[N, N+j] = \sum_{m=1}^{N} \frac{\alpha'_m}{1+z\sigma'_m} + \sum_{k=0}^{j} \beta'_k (-1)^k z^k, \qquad (3.8)$$

then the $(N, N+j)$ approximant to $f(z)$ is given by

$$(N, N+j) = \frac{f_0 R}{R+z} + z\left\{\sum_{m=1}^{N} \frac{\alpha'_m}{(1+z\sigma'_m)^2} + \sum_{k=0}^{j} (1+k)\beta'_k (-z)^k\right\} = \frac{f_0 R}{R+z} + z B_{N,j}(z), \qquad (3.9)$$

[6] This result can be extended to any point in the cut plane. So that, given the first few coefficients of a series of Stieltjes, one can define an approximant and also put an upper bound on the error of this approximant at any such point. However, an unsatisfactory feature still remains in that at some points no bound at all can be given until a fixed number of coefficients of the series of Stieltjes is known. It is hoped to discuss this work in a future paper.

where $B_{N,j}(z)$ is the Gammel–Baker approximant to the function $h(z)$, defined by

$$h(z) = \int_0^{1/R} \frac{\phi(u)\,du}{(1+zu)^2}. \tag{3.10}$$

Since

$$f(z) = f_0 R/(R+z) + zh(z), \tag{3.11}$$

we see from (3.9) that the $(N, N+j)$ approximant to $f(z)$ is obtained by replacing $h(z)$ in (3.11) by its Gammel–Baker approximant.

4. PROPERTIES OF THE APPROXIMANTS $B_{n,j}(z)$ AND THE DEFINITION OF NEW APPROXIMANTS $C_{n,j}(z)$

In this section we consider approximants to the function

$$g(z) = \int_0^{1/R} b(z, s)\,d\phi(s), \tag{4.1}$$

where $b(z, s)$, as a function of z, is regular in at least the complex z plane cut from $-R$ to $-\infty$, when $0 \le s \le 1/R$ and, as a function of s, is regular in a neighborhood of the above segment in the complex s plane for values of z in the above cut plane. An example with this analytic behavior would be

$$b(z, s) = (1+zs)^{-m}, \quad m > 0.$$

The $B_{n,j}(z)$ then converge to $g(z)$ in the cut z plane.

We consider the case when

$$\frac{\partial^j b(z, s)}{\partial s^j} \ge 0 \tag{4.2}$$

for $-R < z \le 0$, $0 \le s \le 1/R$, and $j = 0, 1, 2, \cdots$. It follows immediately from (2.5) and (2.6), which are again true for the above values of z, that

$$B_{n+1,j}(z) - B_{n,j}(z) \ge 0, \tag{4.3}$$

$$B_{n,j}(z) - B_{n-1,j+2}(z) \ge 0, \tag{4.4}$$

and hence that

$$B_{n,j}(z) \le g(z). \tag{4.5}$$

The Gammel–Baker approximants thus give converging lower bounds to $g(z)$, but not upper bounds.

We now show how a new sequence of generalized approximants can be defined which gives converging upper bounds to $g(z)$ when condition (4.2) is satisfied. Since $b(z, s)$ is regular in a neighborhood of $0 \le s \le 1/R$, we may write

$$g(z) = \frac{1}{2\pi i}\int_0^{1/R} d\phi(s) \int_c \frac{b(z, y)}{(y-s)}\,dy, \tag{4.6}$$

where C is a contour encircling the above segment of the real s axis such that $b(z, y)$ as a function of y is regular inside and on it. Therefore

$$g(z) = \frac{1}{2\pi i}\int_c b(z, y)\left[\int_0^{1/R} \frac{d\phi(s)}{1+s(-1/y)}\right]\frac{dy}{y}$$
$$= \frac{1}{2\pi i}\int_c b(z, y)\left[\lim_{n\to\infty}(n, n+j)(-1/y)\right]\frac{dy}{y}, \tag{4.7}$$

where $(n, n+j)$ is the modified approximant to $f(z)$. Since the limit is obtained uniformly on C, we may write, on substituting for $(n, n+j)$ from (3.9),

$$g(z) = \frac{1}{2\pi i}\lim_{n\to\infty}\int_c b(z, y)\left\{\frac{f_0 R}{R - 1/y} - \frac{1}{y}\left[\sum_{m=1}^{n}\frac{\alpha'_m}{(1-\sigma'_m/y)^2} + \sum_{k=0}^{j}(1+k)\beta'_k\left(\frac{1}{y}\right)^k\right]\right\}\frac{dy}{y}$$
$$= \lim_{n\to\infty} C_{n,j}(z),$$

where

$$C_{n,j}(z) = f_0 b(z, 1/R) - \sum_{m=1}^{n}\alpha'_m\left[\frac{\partial b(z, s)}{\partial s}\right]_{s=\sigma_m} - \sum_{k=0}^{j}\frac{\beta'_k}{k!}\left[\frac{\partial^{k+1} b(z, s)}{\partial s^{k+1}}\right]_{s=0}. \tag{4.8}$$

The functions $C_{n,j}(z)$ form sequences of approximants which converge to $g(z)$ in the complex plane cut from $-R$ to $-\infty$. They depend on the parameters α'_m, σ'_m, and β'_k which are obtained from the modified approximant to $f(z)$, and so can be calculated for any particular values of n and j from the first $2n + j + 2$ coefficients of this series of Stieltjes.

Using exactly the same methods employed by Baker to obtain (2.5) and (2.6), one can prove that

$$C_{n+1,j}(z) - C_{n,j}(z) = -(+ve\ \text{const}) \times \left[\frac{\partial^{2n+j+2} b(z, s)}{\partial s^{2n+j+2}}\right]_{s=\sigma}, \tag{4.9}$$

$$C_{n,j}(z) - C_{n-1,j+2}(z) = -(+ve\ \text{const}) \times \left[\frac{\partial^{2n+j+2} b(z, s)}{\partial s^{2n+j+2}}\right]_{s=\sigma'}, \tag{4.10}$$

where σ' lies between 0 and the largest σ'_m appearing in $C_{n,j}(z)$ and $C_{n-1,j+2}(z)$, and σ is similarly defined except for $j = -1$, when it must also be greater than the smallest σ'_m appearing in the corresponding approximants. It follows immediately that if $b(z, s)$ satisfies (4.2), then

$$C_{n+1,j}(z) - C_{n,j}(z) \le 0, \tag{4.11}$$

$$C_{n,j}(z) - C_{n-1,j+2}(z) \le 0, \tag{4.12}$$

and hence that

$$C_{n,j}(z) \geq g(z), \tag{4.13}$$

for $-R < z \leq 0$.

We will prove that if $b(z, s)$ satisfies (1.8), then

$$\left[(-1)^N \frac{d^N}{dz^N}\{g(z) - B_{n,j}(z)\}\right]_{z=0} \geq 0, \tag{4.14}$$

and

$$\left[(-1)^N \frac{d^N}{dz^N}\{g(z) - C_{n,j}(z)\}\right]_{z=0} \leq 0, \tag{4.15}$$

for $N = 0, 1, 2 \cdots$, so that if w is any number such that $|w| < R$,

$$|\{g(z) - B_{n,j}(z)\}_{z=w}| \leq |\{B_{n,j}(z) - C_{n,j}(z)\}_{z=-|w|}|. \tag{4.16}$$

Thus we can in this case put strict limits on the error of the approximants $B_{n,j}(z)$ for $|z| < R$, given only the coefficients f_i necessary to calculate this approximant and $C_{n,j}(z)$.

Inequalities (4.14) and (4.15) are obtained using (1.8), from the following relations:

$$\frac{d^N}{dz^N}\{B_{n+1,j}(z) - B_{n,j}(z)\} = \{+ve \text{ const}\} \times \left\{\frac{\partial^{2n+j+1+N} b(z, s)}{\partial z^N \partial s^{2n+j+1}}\right\}_{s=\sigma}, \tag{4.17}$$

$$\frac{d^N}{dz^N}\{B_{n,j}(z) - B_{n-1,j+2}(z)\} = \{+ve \text{ const}\} \times \left\{\frac{\partial^{2n+j+1+N} b(z, s)}{\partial z^N \partial s^{2n+j+1}}\right\}_{s=\sigma'}, \tag{4.18}$$

where σ and σ' lie between the same limits as the σ and σ' appearing in (2.5) and (2.6), but are not in general the same values. The reason for this is that both these quantities are functions of z and so (4.17) and (4.18) cannot be obtained by just differentiating (2.5) and (2.6) N times with respect to z. One can, however, prove (4.17) and (4.18) by replacing $b(z, s)$ with $\partial^N b(z, s)/\partial z^N$ at each step in Baker's proof of (2.5) and (2.6), and there is no reason why σ or σ' should have the same value in the two cases.

Finally, we prove that

$$\frac{\partial^j b(z, s)}{\partial s^j} \geq 0 \tag{4.2}$$

is both a necessary and sufficient condition for the inequalities (4.3) to (4.5) to hold. Sufficiency has been proved, and to prove necessity we use the method employed by Baker to show that (1.4) is a necessary condition for (2.7)–(2.9) to hold. We thus use (2.5) with $j = -1$, i.e.,

$$B_{n+1,-1}(z) - B_{n,-1}(z) = (+ve \times \text{const}) \times \left[\frac{\partial^{2n} b(z, s)}{\partial s^{2n}}\right]_{s=\sigma}, \tag{4.19}$$

where σ lies between the maximum and minimum values of σ_m in the definition (1.2) of $B_{n+1,-1}(z)$ and $B_{n,-1}(z)$. The σ_m may be chosen arbitrarily in the range 0 to $1/R$ by selecting $d\phi$ to be a sum of delta functions. By making them lie in a sufficiently narrow interval and sweeping it past any desired point, we can make σ any point in the range 0 to $1/R$.[7] Therefore, if inequality (4.3) holds, (4.2) is true for all even j and $1/R \geq s \geq 0$.

Similarly for any $0 \leq s_1 \leq 1/R$, we may choose $d\phi$ as a sum of delta functions such that all the α_m are less than or equal to s_1. Therefore, if (4.3) holds for $j = 0$,

$$\left[\frac{\partial^{2n+1} b(z, s)}{\partial s^{2n+1}}\right]_{s=\sigma} \geq 0 \tag{4.20}$$

from (2.5), where σ is less than or equal to the maximum σ_m, and hence less than or equal to s_1. But

$$\frac{\partial^{2n+2} b(z, s)}{\partial s^{2n+2}} \geq 0$$

for all $s_1 \geq s \geq \sigma$, so that

$$\left[\frac{\partial^{2n+1} b(z, s)}{\partial s^{2n+1}}\right]_{s=s_1} \geq 0.$$

Since this is true for all $1/R \geq s_1 \geq 0$, we have proved that (4.2) is a necessary condition for (4.3)–(4.5) to hold.

5. CONCLUSIONS

We have in this paper investigated properties of the Gammel–Baker approximants to the function $g(z)$ given by (1.5), and in particular under the assumption that for $-R < z \leq 0$

$$\frac{\partial^j b(z, s)}{\partial s^j} \geq 0, \tag{1.6}$$

where $j = 0, 1, 2 \cdots$. The $B_{n,j}(z)$ then give for these values of z converging lower bounds to $g(z)$. To obtain upper bounds, we defined a new sequence of approximants $C_{n,j}(z)$, which, like the $B_{n,j}(z)$, can be obtained from the coefficients of the corresponding series of Stieltjes. Obviously, $B_{n,j}(z)$ and $C_{n,j}(z)$ give converging lower and upper bounds for all values of z for which (1.6) is satisfied, and so for these values

[7] The poles of the $[N, N + j]$ approximant to a series of Stieltjes with radius of convergence R must be on the negative real axis between $-R$ and $-\infty$. Therefore in this case, the σ_m appearing in (2.3) satisfy $0 \leq \sigma_m \leq 1/R$.

one may put strict limits on the error of the approximants.

The case when $b(z, s)$ is analytic in the complex z plane cut from $-R$ to $-\infty$ when $0 \leq s \leq 1/R$, and hence $g(z)$ regular for the same values of z, has been considered. We have shown that if

$$(-1)^N\left[\frac{\partial^{N+j}b(z, s)}{\partial z^N \partial s^j}\right]_{z=0} \geq 0 \tag{1.8}$$

for $0 \leq s \leq 1/R$ and $N, j = 0, 1, 2 \cdots$, then one can put strict limits on the error of the approximants $B_{n,j}(z)$ for $|z| \leq R$. Once again this result may be extended in an obvious way to the case when the region of analyticity of $g(z)$ is different from that assumed above.

Several examples of the function $b(z, s)$ were discussed by Baker, and we now demonstrate how the results of this paper may be applied to them. The first example considered was when

$$b(z, s) = L_\zeta(z, s), \tag{5.1}$$

where $L_\zeta(z)$ is the LeRoy function[8] defined by

$$L_\zeta(z) = \int_0^\infty \exp\left[-(t + zt^\zeta)\right] dt, \quad 0 \leq \zeta < 1. \tag{5.2}$$

It is not difficult to show in this case that $b(z, s)$ satisfies (1.6) and (1.8) for $z \leq 0$, so that $B_{n,j}(z)$ and $C_{n,j}(z)$ have the properties proved in Sec. 3 and Sec. 4. In particular, since $L_\zeta(z)$ is an entire function, we see that the approximants converge to $g(z)$ in the whole complex plane and that strict limits can be put on the error of the approximants in this region. A special case is when $\zeta = 0$, for then $b(z, s) = e^{-zs}$.

Another family of functions considered by Baker was of the form

$$b(z, s) = [1 + zs/(n + 1)]^{-n}, \quad 0 < n < \infty. \tag{5.3}$$

In this case $b(z, s)$ is regular in the complex z plane cut from $-(n + 1)R$ to $-\infty$ when $0 \leq s \leq 1/R$. Again $b(z, s)$ satisfies (1.6) and (1.8), this time when $-(n + 1)R < z \leq 0$. Therefore one can put strict limits on the approximants $B_{n,j}(z)$ and $C_{n,j}(z)$ for all values of z such that $|z| < (n + 1)R$.

We may sum up briefly by saying that in this paper we have shown that the generalized approximants to the class of functions $g(z)$ defined by (1.5) have properties completely analogous to those of the Padé approximants to the corresponding series of Stieltjes with nonzero radius of convergence. An important property, which should prove very useful in practice, is that if $b(z, s)$ satisfies certain defined inequalities, then one can put strict limits on the error of the approximants in prescribed regions of the complex plane.

ACKNOWLEDGMENTS

The author would like to thank Dr. G. Baker for a helpful and encouraging correspondence and Professor J. S. R. Chisholm for his comments.

[8] See G. H. Hardy, *Divergent Series* (Oxford University Press, London, 1956), p. 197.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

One More Technique for the Dimer Problem

JEROME K. PERCUS*
Courant Institute of Mathematical Sciences and Department of Physics
New York University, New York

(Received 13 August 1968)

The problem of counting the dimer coverings of a square lattice is recast as a counting of coverings by oriented closed loops. Thus the answer is expressed as the value of a suitable permanent. This permanent is transformed to a determinant, which on evaluation recovers the familiar result.

1. INTRODUCTION

The many-body problem on a lattice has come under increasing scrutiny in recent years. This has not been hindered by the close relationship between the lattice problem and a number of real physical systems. On the other hand, much of the theoretical motivation has been supplied by the realization in practice of the tempting relative ease of solution, compared to continuum problems. Two somewhat different approaches have been used for this purpose: the nominally straightforward transfer-matrix method[1] and the Pfaffian method[2] with its Green's function[3] cousins. A much older combinatorial technique, that of counting permutations by evaluating a permanent,[4] seems not to have been used to advantage, presumably because the older methods of evaluating permanents are woefully inadequate. It is the purpose of this paper to repair this deficiency, in particular, in the case of one of the most basic counting problems, that of close-packing dimers on a square grid.

In Sec. 2, we show how the dimer-covering problem on a square lattice can be converted to that of a covering by oriented closed loops. In Sec. 3, a now standard theorem of Kasteleyn[5] is used to convert the resulting permanent to a determinant, which is evaluated in Sec. 4.

2. CONVERSION TO A PERMANENT

We consider a square grid of $2M \times 2N$ vertices. A dimer is a horizontal or vertical link which joins two vertices. We seek the number

$$\mathcal{N}(2M, 2N) \tag{2.1}$$

of nonoverlapping dimer configurations which fully occupy the vertices of the lattice. Generalization to odd row or column number and to unequal weights for horizontal and vertical dimers is immediate but serves no useful purpose in the context of this presentation. To make contact with classical combinatorial methods,[4] we must represent an allowed configuration as a permutation of vertices. This is most easily accomplished by isolating the A and B sublattices of the lattice, where an A vertex has an even sum of row and column indices and a B vertex an odd row and column sum, and then putting the A and B sublattices in 1-to-1 correspondence, each in lexicographic order. Now a dimer necessarily joins an A site to a B site. If we then draw an arrow from each A vertex to its dimer-joined B vertex and finally identify (collapse) numerically corresponding A and B sites, a set of oriented closed loops results (see Fig. 1). This set of course constitutes a permutation on the $2MN$ collapsed sites.

In order to express the number of permutations—and hence of dimer configurations—as a permanent, an allowed permutation must be specified as any which is constructed from a fixed set of allowed

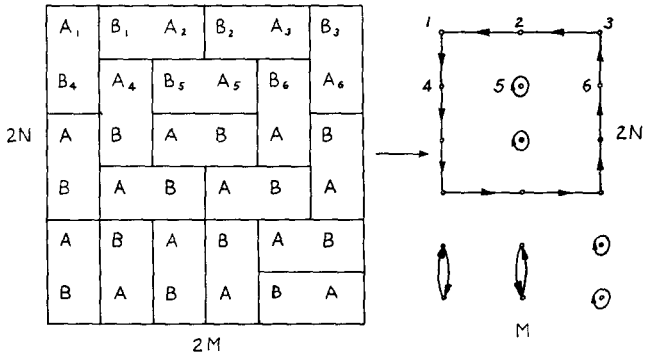


FIG. 1. The transformation from dimers on a $2M \times 2N$ lattice to directed bonds on the $M \times 2N$ lattice created by identifying A_i and B_i.

* Supported in part by the U.S. Atomic Energy Commission, Contract AT(30-1)-1480.

[1] Ranging from the original application to the Ising model to a return to the dimer problem. See, respectively, L. Onsager, Phys. Rev. **65**, 117 (1944); E. H. Lieb, J. Math. Phys. **8**, 2339 (1967).

[2] For Ising and dimer problems, respectively, see C. A. Hurst and H. S. Green, J. Chem. Phys. **33**, 1059 (1960); M. E. Fisher, Phys. Rev. **124**, 1664 (1961).

[3] See, e.g., C. A. Hurst, J. Math. Phys. **7**, 305 (1966).

[4] See, e.g., P. A. MacMahon, *Combinatory Analysis* (Chelsea Publ. Co., New York, 1960), Vol. I, p. 93.

[5] P. W. Kasteleyn, J. Math. Phys. **4**, 287 (1963).

vertex–vertex transitions. In other words,

$$\text{perm}\,(K_{ij}) = \sum_P K_{1P(1)}K_{2P(2)}\cdots K_{nP(n)}, \quad (2.2)$$

summed over all permutations P of the $n\,(=2MN)$ vertices, performs such a counting if each K_{ij} is either 0 or 1. For the situation under discussion, we see that after the vertex condensation, an arrow can go left only on an odd row, go right only on an even row (hence, a 2-vertex loop must be vertical), and cover a single vertex alone at any location. Any vertical arrow can occur. We conclude that K can be written as the $2N \times 2N$ compartmentalized matrix of transitions from one row to another:

$$K = \begin{pmatrix} C & I & & & & & \\ I & C^{\mathrm{T}} & I & & & & \\ & I & C & I & & & \\ & & I & C^{\mathrm{T}} & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ 0 & & & & \ddots & \ddots & I \\ & & & & & I & C^{\mathrm{T}} \end{pmatrix}_{2N\times 2N}, \quad (2.3)$$

where

$$C = \begin{pmatrix} 1 & & & & \\ 1 & 1 & & 0 & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ 0 & & & \ddots & 1 \\ & & & & 1 \end{pmatrix}_{M\times M},$$

$$C^{\mathrm{T}} = \begin{pmatrix} 1 & 1 & & & \\ & 1 & 1 & 0 & \\ & & 1 & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & 1 \end{pmatrix}_{M\times M},$$

$$I = \begin{pmatrix} 1 & & & & \\ & 1 & & 0 & \\ & & 1 & & \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix}_{M\times M}. \quad (2.4)$$

Here C is the $M \times M$ submatrix connecting an odd row to itself (via a left arrow or single vertex arrow), C^{T} for an even row to itself, and I connects adjacent rows (via a vertical arrow).

3. EQUIVALENT DETERMINANT

There are few effective methods for evaluating permanents of involved structure. One which works superbly when it works at all[6] is that of converting the permanent to a determinant, since the latter can usually be evaluated by more conventional means. Since the determinant "simply" inserts a $(-1)^P$ in (2.2), we indeed will have

$$\text{perm}\,(K_{ij}) = \det\,(\bar{K}_{ij})$$

if

$$\bar{K}_{ij} = \tau_{ij}K_{ij}, \quad (3.1)$$

where

$$\tau_{1P(1)}\tau_{2P(2)}\cdots\tau_{nP(n)} = (-1)^P.$$

Now a permutation can be written as a product of cycles, each one of which represents a single oriented

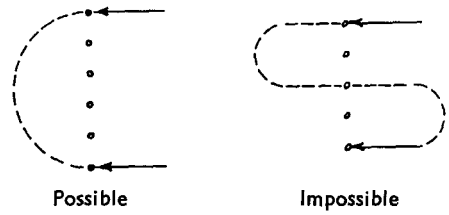


FIG. 2. An interior vertical segment must be even, since bond directions at its boundary must oppose.

closed curve or loop in the vertices. In the present case of a square lattice, each loop has an even number of vertices (up and down arrows are paired, as are right and left) except for the one-vertex loops which we term trivial. Since the parity of an even-vertex cycle is odd (it requires an odd number of transpositions), we conclude that the parity of any P is given by the number of nontrivial loops it contains. The question then is whether we can assign values τ_{ij} whose product for a nontrivial loop is -1.

This question is answered at once by a theorem of Kasteleyn[5]:

Theorem: Bond signs can be chosen on a planar graph such that every oriented loop of even interior opposes an odd number of bonds.

To use the theorem, we require that our loops have even interiors. But in fact we have the stronger result that any connected sequence of interior points along a vertical line must be bounded by one left and one

[6] See, e.g., J. K. Percus, Lecture notes on combinatorial methods, Courant Institute, New York, 1968.

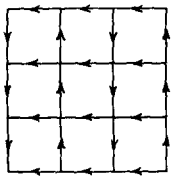


FIG. 3. Canonical bond directions for Kasteleyn's theorem on a square lattice.

right arrow, since otherwise (see Fig. 2) they could not all be interior. Thus each such connected sequence consists of an even number of points, as must a complete interior. There remains only the realization of the bond signs, and this is given by Fig. 3, which is well known.[7] If we now choose τ_{ij} as $+1$ if $i \to j$ has the bond direction, -1 if not, a loop will therefore accumulate a weight of -1; for the trivial loop, we must retain $\tau_{ii} = 1$. We conclude then from (2.3), (2.4), (3.1) that

$$\mathcal{N}(2M, 2N) = \det(\bar{K}_{ij}),$$

where

$$\bar{K} = \begin{pmatrix} C & -I & & & & & 0 \\ I & C^{\mathrm{T}} & -I & & & & \\ & I & C & -I & & & \\ & & I & C^{\mathrm{T}} & \cdot & & \\ & & & \cdot & \cdot & & \\ & & & & \cdot & \cdot & \\ 0 & & & & & I & \\ & & & & & I & C^{\mathrm{T}} \end{pmatrix}_{2N \times 2N} . \tag{3.2}$$

4. EVALUATION OF DETERMINANT

The evaluation of (3.2) is essentially mechanical. A useful preliminary reduction consists of multiplying every even column by C on the right, adding to it its left neighboring column, and subtracting from it its right neighboring column. Then every odd row has only one entry, namely C; expanding by these entries and using the fact that $\det C = 1$, we find

$$\mathcal{N}(2M, 2N) = \det(\bar{\bar{K}}_{ij}),$$

where

$$\bar{\bar{K}} = \begin{pmatrix} C^{\mathrm{T}}C + 2I & -I & & & & 0 \\ -I & C^{\mathrm{T}}C + 2I & -I & & & \\ & -I & C^{\mathrm{T}}C + 2I & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ 0 & & & & -I & \\ & & & & -I & C^{\mathrm{T}}C + I \end{pmatrix}_{N \times N} \tag{4.1}$$

and

$$C^{\mathrm{T}}C = \begin{pmatrix} 2 & 1 & & & & \\ 1 & 2 & 1 & & & \\ & 1 & 2 & 1 & & 0 \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot \\ 0 & & & & 2 & 1 \\ & & & & 1 & 1 \end{pmatrix}_{M \times M}$$

[7] See, e.g., E. W. Montroll, in *Applied Combinatorial Mathematics*, E. F. Beckenbach, Ed. (John Wiley & Sons, Inc., New York, 1964), p. 114.

Now define the matrix

$$J_p = \begin{pmatrix} 2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & 0 & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & \cdot & \cdot & \cdot \\ & & 0 & & & 2 & -1 \\ & & & & & -1 & 1 \end{pmatrix}_{p\times p}, \tag{4.2}$$

with eigenvalues $\lambda_1^{(p)}, \cdots, \lambda_p^{(p)}$. If $\bar{\bar{K}}$ is transformed by the transformation which diagonalizes J_N (this, of course, is a transformation among rows, not elements), we have

$$\det \bar{\bar{K}} = \det \begin{pmatrix} C^TC + \lambda_1^{(N)} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & C^TC + \lambda_N^{(N)} \end{pmatrix}$$

$$= \prod_{i=1}^{N} \det (C^TC + \lambda_i^{(N)}).$$

But C^TC is similar to J_M—under the orthogonal transformation

$$\begin{pmatrix} 1 & & & & \\ & -1 & & & \\ & & 1 & & \\ & & & -1 & \\ & & & & \cdot \\ & & & & & \cdot \\ & & & & & & \cdot \end{pmatrix},$$

whence

$$\mathcal{N}(2M, 2N) = \prod_{i=1}^{N} \prod_{j=1}^{M} (\lambda_i^{(N)} + \lambda_j^{(M)}). \tag{4.3}$$

The eigenvalues are easily found to be

$$\lambda_j^{(p)} = 4 \sin^2 \theta_j^{(p)},$$

where $\theta_j^{(p)} = [(2j + 1)/(2p + 1)]\pi$, and so we obtain the familiar form[8]

$$\mathcal{N}(2M, 2N) = 2^{2MN} \prod_{i=1}^{N} \prod_{j=1}^{M} \left(\sin^2 \frac{2i+1}{2N+1} \pi + \sin^2 \frac{2j+1}{2M+1} \pi \right). \tag{4.4}$$

In the special case $M = N$,

$$\prod_{i,j} (\lambda_i + \lambda_j) = 2^N \left(\prod_{i>j} (\lambda_i + \lambda_j) \right)^2$$

and $\prod_{i>j} (\lambda_i + \lambda_j)$ (as a symmetric function of the roots of the characteristic polynomial of J_N) is readily evaluated for small N. Since

$$\prod_{i \neq j} (\lambda_i + \lambda_j) = \prod_{i \neq j} (\lambda_i^2 - \lambda_j^2)/(\lambda_i - \lambda_j),$$

we also have that

$$\mathcal{N}(2N, 2N) = 2^N \frac{\operatorname{disc} \det (x - J_N^2)}{\operatorname{disc} \det (x - J_N)}, \tag{4.5}$$

where $\operatorname{disc} f(x)$ refers to the discriminant of the polynomial $f(x)$.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the sympathetic attention of my students in "Combinatorial Methods" at the Courant Institute, 1967–68, during the initial presentation of this material, and the critical but friendly reading of the manuscript by Professor O. E. Percus.

[8] M. E. Fisher, Phys. Rev. **124**, 1664 (1961).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Plane-Symmetric Gauge Field

C. A. UZES*
Physics Department, University of Georgia, Athens, Georgia

(Received 9 July 1968)

Plane-symmetric solutions to the gauge field equations are considered on the classical level. The conditions of plane symmetry are gauge-invariantly defined in terms of the internal holonomy group. It is shown that, if analytic gauge fields satisfy the plane-symmetry conditions and if there exists at least one source-free region in event space, then the internal holonomy group is Abelian and a gauge exists in which the gauge field satisfies Maxwell's equations.

I. INTRODUCTION

In local gauge theory one may consider that at each event[1] there exists a distinct n-dimensional internal space. Under coordinate-dependent internal base (gauge) transformations $S(x)$, the $n \times n$ matrix fields of the gauge potential and gauge field transform according to

$$\Gamma'_\mu = S^{-1}\Gamma_\mu S - S^{-1}\partial_\mu S,$$
$$\phi'_{\mu\nu} = S^{-1}\phi_{\mu\nu}S, \tag{1}$$

where

$$\phi_{\mu\nu} = \partial_\mu\Gamma_\nu - \partial_\nu\Gamma_\mu - [\Gamma_\mu, \Gamma_\nu].$$

The action principle applied to the Lagrangian[2]

$$L = -\tfrac{1}{4}\,\mathrm{Tr}\,\phi_{\mu\nu}\phi^{\mu\nu}$$

gives rise to the free gauge field equation

$$\partial_\lambda\phi^{\kappa\lambda} - [\Gamma_\lambda, \phi^{\kappa\lambda}] = 0. \tag{2}$$

Except for those having the pseudo-self-dual property, for which the free gauge field's energy–momentum tensor density vanishes everywhere in event space,[3] little is known about classical solutions to these equations.[4] Only point-charge solutions, for which the external source current is nonvanishing on a world line, have been studied in some detail.[5–8] This has been due primarily to the complexity and nonlinearity of the equations. One manner of simplifying them is to impose event-space symmetry conditions on the solutions. The class exhibiting plane symmetry is particularly simple and it is of interest because it includes the free gauge field plane wave.

A possible ambiguity in the definition of the gauge-invariant plane-symmetry conditions is resolved by defining them in terms of the internal holonomy group $\mathcal{H}$. The procedure will be first to show that any analytic gauge field satisfying the plane-symmetry conditions belongs to the Lie algebra of an Abelian $\mathcal{H}$, and then to show that a gauge exists in which such a field satisfies Maxwell's equations. It will be evident that the results apply equally well to plane-symmetric analytic solutions to the inhomogeneous equation, so long as there exists at least one region in event space (which may be unbounded) free of external sources.

II. THE INTERNAL GEOMETRY

Under the gauge transformation $S(x^\kappa)$, covariant internal vectors $\mathbf{v}(x^\kappa)$ and linear internal operators $\mathbf{P}(x^\kappa)$ transform according to the rules

$$\mathbf{v}' = S^{-1}\mathbf{v},$$
$$\mathbf{P}' = S^{-1}\mathbf{P}S.$$

One can easily verify from Eq. (1) that the gauge-covariant derivatives of covariant internal vectors and linear internal operators are, respectively,

$$\nabla_\mu\mathbf{v} = \partial_\mu\mathbf{v} - \Gamma_\mu\mathbf{v},$$
$$\nabla_\mu\mathbf{P} = \partial_\mu\mathbf{P} - [\Gamma_\mu, \mathbf{P}],$$

and, thus, that the gauge potential can play the role of an internal linear connection.

The vector $\mathbf{v}(x^\kappa + dx^\kappa)$ is equivalent to $\mathbf{v}(x^\kappa)$ provided the gauge-invariant condition

$$dx^\kappa\nabla_\kappa\mathbf{v} = 0$$

is satisfied at x^κ. It follows that

$$\mathbf{v}(x^\kappa + dx^\kappa) = (I + dx^\beta\Gamma_\beta)\mathbf{v}(x^\kappa), \tag{3}$$

where I is the identity transformation in internal space, may be viewed as the result of the equivalence

* This work was supported in part by the University of California, Riverside, California.

[1] The event space taken here is Minkowskian. In the following, the x^κ stand for the Cartesian inertial coordinates of events, while ∂_κ denotes the partial derivatives $\partial/\partial x^\kappa$.

[2] The gauge group here is unrestricted. This Lagrangian is a generalization of that employed by C. N. Yang and R. L. Mills [Phys. Rev. **96**, 191 (1954)] for the case of the isotopic spin gauge group. "Tr" denotes the trace.

[3] H. G. Loos, Nuovo Cimento **52A**, 1085 (1967).

[4] Only classical gauge fields are considered here. The quantized version of local gauge theory remains incomplete. See J. Schwinger, Phys. Rev. **125**, 1043 (1962); R. L. Arnowitt and S. I. Fickler, Phys. Rev. **127**, 1821 (1962).

[5] R. P. Treat, Nuovo Cimento **50A**, 871 (1967).

[6] M. Ikeda and Y. Miyachi, Progr. Theoret. Phys. (Kyoto) **27**, 474 (1962).

[7] H. G. Loos, Nucl. Phys. **72**, 677 (1965).

[8] H. G. Loos, J. Math. Phys. **8**, 1870 (1967).

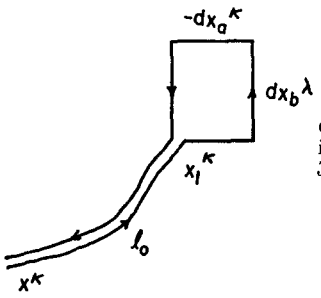


FIG. 1. A closed loop in event space corresponding to an infinitesimal transformation of $\mathcal{H}(x^\kappa)$.

transport of $\mathbf{v}(x^\kappa)$ along dx^κ to $x^\kappa + dx^\kappa$. By carrying out a succession of such equivalence displacements along a curve ℓ in event space which connects the events x^κ and x'^κ, one can construct at x'^κ a vector equivalent to $\mathbf{v}(x^\kappa)$.

Starting at x^κ, the element $\mathbf{H}_\ell(x^\kappa)$ of the internal holonomy group $\mathcal{H}(x^\kappa)$ defines, through the transformation

$$\mathbf{v}' = \mathbf{H}_\ell(x^\kappa)\mathbf{v},$$

the equivalence transport of arbitrary vectors about the *closed* loop ℓ.[9] The gauge field and its covariant derivatives (evaluated at x^κ) in regions where they are analytic span the Lie algebra of $\mathcal{H}(x^\kappa)$. Infinitesimal transformations of $\mathcal{H}(x^\kappa)$ are defined by loops ℓ, as in Fig. 1, consisting of the following parts: an arbitrary path ℓ_0 connecting the events x_1^κ and x^κ, the infinitesimal closed loop ℓ_1 at x_1^κ, described by the bivector $df^{\kappa\lambda} = dx_a^{[\kappa}dx_b^{\lambda]}$, and the reverse of ℓ_0 from x_1^κ back to x^κ. Such a transformation may be expressed in the form

$$\mathbf{H}_\ell(x^\kappa) = I + \tfrac{1}{2}\,df^{\delta\lambda}(I - q^\mu\nabla_\mu + q^{\mu\nu}\nabla_\mu\nabla_\nu - \cdots - q^{\nu\mu\cdots\beta}\nabla_\nu\nabla_\mu\cdots\nabla_\beta)\phi_{\delta\lambda}(x^\kappa), \quad (4)$$

where

$$q^\mu = x^\mu(s_1) - x^\mu, \qquad q^{\nu\mu} = \int_0^{s_1} q'^\nu q^\mu(s)\,ds,$$

$$q^{\sigma\mu\cdots\nu} = \int_0^{s_1} q'^\sigma q^{\mu\cdots\nu}\,ds, \quad q'^\nu = \frac{dq^\nu}{ds}, \quad (5)$$

and where $x^\kappa(s)$ is the parametric description of ℓ_0, with $x^\kappa(0) = x^\kappa$.

III. DEFINITION OF PLANE SYMMETRY

For the electromagnetic field, plane symmetry is imposed by requiring that the field tensor be constant along the plane of symmetry at a fixed instant of time. For instance, for the plane normal to the x direction this means

$$\partial_y F_{\kappa\lambda} = 0, \quad \partial_z F_{\kappa\lambda} = 0. \quad (6)$$

This condition is gauge-invariant since the electromagnetic field tensor is invariant under electromagnetic gauge transformations.

For more general gauge fields, plane symmetry cannot be defined in this manner, since

$$\partial_y\phi_{\kappa\lambda} = 0, \quad \partial_z\phi_{\kappa\lambda} = 0$$

are not invariant conditions under the coordinate dependent transformations (1). One may consider replacing the partial derivatives by covariant derivatives and taking the resulting expression as the *gauge-invariant* statement of plane symmetry. Although these conditions are necessary, their implications are uncertain, since the identity

$$(\nabla_\kappa\nabla_\lambda - \nabla_\lambda\nabla_\kappa)\mathbf{P} = -[\phi_{\kappa\lambda}, \mathbf{P}] \quad (7)$$

implies that the higher-order covariant derivatives of $\phi_{\kappa\lambda}$ do not necessarily satisfy

$$\nabla_y\nabla_\beta\cdots\nabla_\gamma\phi_{\kappa\lambda} = 0, \quad \nabla_z\nabla_\beta\cdots\nabla_\gamma\phi_{\kappa\lambda} = 0,$$

while (6) implies that

$$\partial_y\partial_\beta\cdots\partial_\gamma F_{\kappa\lambda} = 0, \quad \partial_z\partial_\beta\cdots\partial_\gamma F_{\kappa\lambda} = 0.$$

A proper way to define plane symmetry is to require that for symmetry in the (y, z) plane the internal geometry be homogeneous along the surface $x =$ const, $t =$ const, i.e., *that moving along any such plane one is unable to distinguish geometrically between points on the plane.* The only geometric material available is the set of the $\phi_{\kappa\lambda}$ and their covariant derivatives or, equivalently, the internal holonomy group. Thus, plane symmetry will be defined in terms of the homogeneity conditions satisfied by the internal holonomy group.

Consider the element $\mathbf{H}_\ell(x^\kappa)$ at the event $x^\kappa = (t, x, y, z)$. Let $\mathbf{H}_{\ell'}(x'^\kappa)$ be the element of $\mathcal{H}(x'^\kappa)$ defined by the loop ℓ' resulting from the parallel displacement of ℓ from the event x^κ to $x'^\kappa = (t, x, y', z')$ and let $\ell\!\!\!/$ be any path lying in the plane ($t =$ const, $x =$ const) which connects the events x^κ and x'^κ, as in Fig. 2. The internal geometry will be said to have plane symmetry if and only if for every value of t and x and every such curve $\ell\!\!\!/$ the transformation $\mathbf{H}_{\ell'}(x'^\kappa)$ is identical to that obtained by the equivalence transport along $\ell\!\!\!/$ of $\mathbf{H}_\ell(x^\kappa)$ to the event x'^κ.

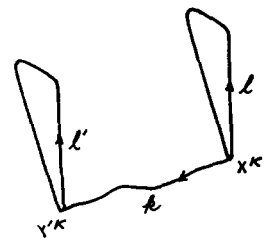


FIG. 2. Possible curves in event space through which the plane-symmetry conditions are defined. $\ell\!\!\!/$ lies in the (y, z) plane.

[9] For a complete discussion of the following and other interesting properties of the internal holonomy group, the reader is referred to H. G. Loos, J. Math. Phys. 8, 2114 (1967).

Plane symmetry states that

$$\mathbf{H}_{\ell'}(x'^{\kappa}) = \mathbf{T}_{\ell}\mathbf{H}_{\ell}(x^{\kappa})\mathbf{T}_{\ell}^{-1}, \tag{8}$$

where $\mathbf{T}_{\ell}(x'^{\kappa}, x^{\beta})$ defines, through

$$\mathbf{v}' = \mathbf{T}_{\ell}(x'^{\kappa}, x^{\beta})\mathbf{v},$$

the equivalence transport of arbitrary vectors from x^{κ} along ℓ to x'^{κ}. Let $x'^{\kappa} = x^{\kappa} + dx^{\kappa}$ and let ℓ be given by the increment dx^{κ}. Using (3), Eq. (8) reduces to the requirement

$$\mathbf{H}_{\ell'}(x'^{\kappa}) = \mathbf{H}_{\ell}(x^{\kappa}) + dx^{\kappa}[\Gamma_{\kappa}, \mathbf{H}_{\ell}(x^{\mu})]. \tag{9}$$

Let ℓ be given by the infinitesimal transformation (4). Since ℓ' is the result of the parallel displacement of ℓ from x^{κ} to x'^{κ}, the events on ℓ' must be related to those on ℓ by a point transformation of the form

$$x'^{\kappa}(s) = x^{\kappa}(s) + a^{\kappa},$$

where a^{κ} is independent of s. But $df^{\kappa\lambda}$ and, as can be seen from (5), the $q^{\mu\cdots\nu}$ are invariant under translations. Thus, the matrix operator $\mathbf{H}_{\ell'}(x'^{\kappa})$ may be obtained from $\mathbf{H}_{\ell}(x^{\kappa})$ through replacement of the argument x^{κ} in (4) by x'^{κ}, and can be written as $\mathbf{H}_{\ell}(x'^{\kappa})$. Then (9) states that, since $dx^{\kappa} = (0, 0, dy, dz)$ is arbitrary,

$$\nabla_y\mathbf{H}_{\ell}(x^{\kappa}) = 0, \quad \nabla_z\mathbf{H}_{\ell}(x^{\kappa}) = 0, \tag{10}$$

for curves ℓ defining infinitesimal loops.

The bivector $df^{\kappa\lambda}$ and the curve ℓ_0 may be chosen arbitrarily, so that the expansion coefficient $df^{\kappa\lambda}$ and the $q^{\mu\cdots\nu}$ may be considered as independent variables. This implies that

$$\nabla_y\nabla_{\mu}\cdots\nabla_{\nu}\phi_{\kappa\lambda} = 0, \quad \nabla_z\nabla_{\mu}\cdots\nabla_{\nu}\phi_{\kappa\lambda} = 0 \tag{11}$$

are necessary conditions for plane symmetry of the gauge field. They are also sufficient. Since a general transformation of $\mathfrak{K}(x^{\kappa})$ may be written in the form

$$\exp(q^{\mu\nu}\phi_{\mu\nu} + q^{\mu\beta\nu}\nabla_{\nu}\phi_{\mu\beta} + \cdots + q^{\mu\beta\lambda\cdots\nu}\nabla_{\nu}\cdots\nabla_{\lambda}\phi_{\mu\beta})$$

and the parameters $q^{\mu\cdots\nu}$ are invariant under parallel transport of ℓ in event space, (10) must hold for arbitrary loops. However, only the statement of necessity is needed below.

IV. PLANE SYMMETRY AND THE LIE ALGEBRA OF $\mathfrak{K}$

From (11), it follows that

$$(\nabla_y\nabla_y - \nabla_y\nabla_y)\nabla_{\lambda}\cdots\nabla_{\beta}\phi_{\mu\nu} = 0,$$
$$(\nabla_z\nabla_y - \nabla_y\nabla_z)\nabla_{\lambda}\cdots\nabla_{\beta}\phi_{\mu\nu} = 0.$$

Applying the identity (7), for the commutator between ϕ_{yy} and ϕ_{zy} and any of the $\phi_{\kappa\lambda}$ or its covariant derivatives one obtains

$$[\phi_{yy}, \nabla_{\lambda}\cdots\nabla_{\beta}\phi_{\mu\nu}] = 0, \quad [\phi_{zy}, \nabla_{\lambda}\cdots\nabla_{\beta}\phi_{\mu\nu}] = 0.$$

Repeated covariant differentiation of these expressions yields

$$[\nabla_{\alpha}\cdots\nabla_{\kappa}\phi_{yy}, \nabla_{\lambda}\cdots\nabla_{\beta}\phi_{\mu\nu}] = 0,$$
$$[\nabla_{\alpha}\cdots\nabla_{\kappa}\phi_{zy}, \nabla_{\lambda}\cdots\nabla_{\beta}\phi_{\mu\nu}] = 0. \tag{12}$$

Now all the commutation relations have been derived except those of ϕ_{tx} with its covariant derivatives and those between the covariant derivatives of ϕ_{tx}. It is clear from (11) that only the commutation relations between the various covariant derivatives of ϕ_{tx} with respect to the coordinates t and x need be determined.

In the presence of external sources the free-field equation (2) must be replaced by

$$J^{\kappa} = \nabla_{\lambda}\phi^{\kappa\lambda}, \tag{13}$$

where J^{κ} is the external source current. Under the conditions of plane symmetry (11), this reduces to

$$J^{\kappa} = \nabla_t\phi^{\kappa t} + \nabla_x\phi^{\kappa x}.$$

Substituting t and x for κ and employing the antisymmetry of $\phi_{\kappa\lambda}$ in the event-space indices, it is evident that in any source-free region of event space all the covariant derivatives of ϕ_{tx} vanish. The commutation relations of the gauge field and its covariant derivatives are then completely specified by (12) and thus the Lie algebra of $\mathfrak{K}(x^{\kappa})$ in this region is Abelian.

The internal holonomy groups at different events are isomorphic.[9] Hence, if analytic gauge fields satisfy the plane-symmetry conditions (11), and if there exists at least one source-free region in event space, $\mathfrak{K}(x^{\kappa})$ is Abelian at all events.

V. GAUGE FIELDS WITH ABELIAN $\mathfrak{K}$

It has recently been shown that a gauge exists in which the expansion[10]

$$\Gamma_{\kappa} = b_{\kappa}^{i}(x^{\beta})L_i$$

holds, where the L_i are constant matrices belonging to the Lie algebra of $\mathfrak{K}(x^{\kappa})$ at every event. In this case, the partial derivatives of Γ_{κ} also belong to this Lie algebra. Thus for Abelian $\mathfrak{K}$

$$J^{\kappa} = \partial_{\lambda}\phi^{\kappa\lambda},$$
$$\phi_{\mu\nu} = \partial_{\mu}\Gamma_{\nu} - \partial_{\nu}\Gamma_{\mu}, \tag{14}$$

and a gauge exists in which analytic gauge fields with Abelian internal holonomy groups satisfy Maxwell's

[10] H. G. Loos, "Expansion Theorem for Gauge Potentials" (unpublished).

equations,[11] since (14) insures that

$$\partial_\kappa \phi_{\mu\nu} + \partial_\nu \phi_{\kappa\mu} + \partial_\mu \phi_{\nu\kappa} = 0$$

is identically satisfied.

CONCLUSION

The above results are similar to those obtained by Loos[7] for gauge fields with spherical symmetry. There the gauge-invariant conditions of spherical symmetry were defined in terms of the homogeneity and isotropy of the internal geometry on a constant-time spherical surface in event space. In both cases the symmetry conditions require that the corresponding holonomy groups be Abelian.

For gauge fields with Abelian $\mathcal{K}$, a gauge exists in which each component of $\phi_{\kappa\lambda}$ in the Lie algebra of $\mathcal{K}$ satisfies Maxwell's equations. Thus the apparent nonlinearity produced by the commutator terms in the gauge field equation is not essential when the field is analytic and has plane symmetry and there can exist no self-interaction. The additional condition, that there must exist at least one external source-free region in event space (which may be unbounded), of course, includes the free gauge field plane wave.

ACKNOWLEDGMENTS

The author wishes to express his appreciation to Dr. H. G. Loos for suggesting this investigation, and to Dr. Loos and Dr. D. R. Tompkins for helpfully commenting on the text.

[11] A previous proof by the author (Ph.D. thesis, University of California, Riverside, California, 1967) has been quoted in Ref. 9, before the expansion theorem of Ref. 10 was known. The present version is more straightforward.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Representation Theory of $SO(4, 1)$ and $E(3, 1)$: An Explicit Spinor Calculus*

WAYNE J. HOLMAN III†
Center for Theoretical Studies, University of Miami, Coral Gables, Florida

(Received 29 July 1968)

The explicit spinor basis states for the irreducible unitary representations of the spinor covering group of the de Sitter group $SO(4, 1)$ are obtained by analytic continuation from those of the compact group $Sp(4)$. Some features of the contraction of these basis states to those of the Poincaré group are discussed; the explicit asymptotic form of the basis state of a continuous representation of $SO(4, 1)$ in the limit of the contraction to a basis state of a finite-mass discrete-spin representation of $E(3, 1)$ is investigated.

1. INTRODUCTION

In a previous publication,[1] the basis states of the irreducible unitary representations of the simply-connected compact group $Sp(4)$, which is locally isomorphic to the group $SO(5)$ of rotations in five dimensions, were constructed in terms of boson creation operators. In the present work we obtain the basis states of all the irreducible unitary representations of $Sp(2, 2)$, the spinor covering group of the de Sitter group $SO(4, 1)$, by analytic continuation from those of $Sp(4)$, which were obtained in Paper I.

The representation theory of $SO(4, 1)$ was first treated by Thomas,[2] whose work was corrected by Newton[3]; subsequently the representations have been discussed by Dixmier[4] and Kihlberg and Ström.[5] All of these authors used infinitesimal methods. Takahashi[6] has determined the representations by means of Mackey's theory of induced representations in global form. More recently, Kuriyan, Mukunda, and Sudarshan[7] treated the representations of both $SO(5)$ and $SO(4, 1)$ from the standpoint of their theory of master analytic representations. The de Sitter groups $SO(4, 1)$ and $SO(3, 2)$ are of interest in

* This work was supported by National Aeronautics and Space Administration Contract No. NASA NGR 10-007-010, and by Air Force Office of Scientific Research Contract No. AF AFOSR 1268-67.

† Present address: Theoretical Physics Department, Middle East Technical University, Ankara, Turkey.

[1] W. J. Holman III, J. Math. Phys. **12**, 1710 (1969), hereafter designated as Paper I.

[2] L. H. Thomas, Ann. Math. **42**, 113 (1941).

[3] T. D. Newton, Ann. Math. **51**, 730 (1950).

[4] J. Dixmier, Bull. Soc. Math. France **89**, 9 (1960).

[5] A. Kihlberg and S. Ström, Arkiv Fysik **31**, 491 (1966).

[6] R. Takahashi, Bull. Soc. Math. France **91**, 289 (1963).

[7] J. G. Kuriyan, N. Mukunda, and E. C. G. Sudarshan, Commun. Math. Phys. **8**, 204 (1968).

equations,[11] since (14) insures that

$$\partial_\kappa \phi_{\mu\nu} + \partial_\nu \phi_{\kappa\mu} + \partial_\mu \phi_{\nu\kappa} = 0$$

is identically satisfied.

CONCLUSION

The above results are similar to those obtained by Loos[7] for gauge fields with spherical symmetry. There the gauge-invariant conditions of spherical symmetry were defined in terms of the homogeneity and isotropy of the internal geometry on a constant-time spherical surface in event space. In both cases the symmetry conditions require that the corresponding holonomy groups be Abelian.

For gauge fields with Abelian $\mathcal{K}$, a gauge exists in which each component of $\phi_{\kappa\lambda}$ in the Lie algebra of $\mathcal{K}$ satisfies Maxwell's equations. Thus the apparent nonlinearity produced by the commutator terms in the gauge field equation is not essential when the field is analytic and has plane symmetry and there can exist no self-interaction. The additional condition, that there must exist at least one external source-free region in event space (which may be unbounded), of course, includes the free gauge field plane wave.

ACKNOWLEDGMENTS

The author wishes to express his appreciation to Dr. H. G. Loos for suggesting this investigation, and to Dr. Loos and Dr. D. R. Tompkins for helpfully commenting on the text.

[11] A previous proof by the author (Ph.D. thesis, University of California, Riverside, California, 1967) has been quoted in Ref. 9, before the expansion theorem of Ref. 10 was known. The present version is more straightforward.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Representation Theory of $SO(4, 1)$ and $E(3, 1)$: An Explicit Spinor Calculus*

WAYNE J. HOLMAN III†
Center for Theoretical Studies, University of Miami, Coral Gables, Florida

(Received 29 July 1968)

The explicit spinor basis states for the irreducible unitary representations of the spinor covering group of the de Sitter group $SO(4, 1)$ are obtained by analytic continuation from those of the compact group $Sp(4)$. Some features of the contraction of these basis states to those of the Poincaré group are discussed; the explicit asymptotic form of the basis state of a continuous representation of $SO(4, 1)$ in the limit of the contraction to a basis state of a finite-mass discrete-spin representation of $E(3, 1)$ is investigated.

1. INTRODUCTION

In a previous publication,[1] the basis states of the irreducible unitary representations of the simply-connected compact group $Sp(4)$, which is locally isomorphic to the group $SO(5)$ of rotations in five dimensions, were constructed in terms of boson creation operators. In the present work we obtain the basis states of all the irreducible unitary representations of $Sp(2, 2)$, the spinor covering group of the de Sitter group $SO(4, 1)$, by analytic continuation from those of $Sp(4)$, which were obtained in Paper I.

The representation theory of $SO(4, 1)$ was first treated by Thomas,[2] whose work was corrected by Newton[3]; subsequently the representations have been discussed by Dixmier[4] and Kihlberg and Ström.[5] All of these authors used infinitesimal methods. Takahashi[6] has determined the representations by means of Mackey's theory of induced representations in global form. More recently, Kuriyan, Mukunda, and Sudarshan[7] treated the representations of both $SO(5)$ and $SO(4, 1)$ from the standpoint of their theory of master analytic representations. The de Sitter groups $SO(4, 1)$ and $SO(3, 2)$ are of interest in

* This work was supported by National Aeronautics and Space Administration Contract No. NASA NGR 10-007-010, and by Air Force Office of Scientific Research Contract No. AF AFOSR 1268-67.

† Present address: Theoretical Physics Department, Middle East Technical University, Ankara, Turkey.

[1] W. J. Holman III, J. Math. Phys. **12**, 1710 (1969), hereafter designated as Paper I.

[2] L. H. Thomas, Ann. Math. **42**, 113 (1941).

[3] T. D. Newton, Ann. Math. **51**, 730 (1950).

[4] J. Dixmier, Bull. Soc. Math. France **89**, 9 (1960).

[5] A. Kihlberg and S. Ström, Arkiv Fysik **31**, 491 (1966).

[6] R. Takahashi, Bull. Soc. Math. France **91**, 289 (1963).

[7] J. G. Kuriyan, N. Mukunda, and E. C. G. Sudarshan, Commun. Math. Phys. **8**, 204 (1968).

physics principally because they can be contracted in the sense of Inönü and Wigner[8] to the Poincaré group. This contraction has been discussed by Ström[9] for the representation theory of $SO(4,1)$, and by Evans[10] for $SO(3,2)$. Dirac[11] early treated the problem of wave equations in de Sitter space, and Roman and Aghassi[12–16] discussed the problems of a field theory in a de Sitter world and found that the O'Raifeartaigh theorem does not hold if the space–time symmetry is that of $SO(4,1)$. Tagirov and Chernikov,[17] Nachtmann,[18] and Gürsey and Lee[19,20] have also treated the problems of field theory in this space.

In Sec. 2 we derive the spinor basis states for all the irreducible unitary representations of $Sp(2,2)$. We follow, in general, the procedure of Barut and Fronsdal,[21] who extended the boson calculus for $SO(3)$ representations by analytic continuation to a spinor calculus for those of $SO(2,1)$. In Sec. 3 we discuss features of the contraction of the $Sp(2,2)$ states to those of the representations of the Poincaré group $E(3,1)$, and we obtain the asymptotic form of the basis state of $Sp(2,2)$ contracted to that of a finite-mass discrete-spin representation of $E(3,1)$ in the helicity-angular-momentum basis; i.e., the $SO(4)$ subgroup of $SO(4,1)$ is simultaneously contracted to the $E(3)$ subgroup of $E(3,1)$. The representation theory of the Poincaré group in this basis has been discussed by Lomont and Moses.[22,23]

2. SPINOR BASIS STATES FOR IRREDUCIBLE UNITARY REPRESENTATIONS OF $SP(2,2)$

In Paper I the states of the irreducible unitary representations of $Sp(4) \cong SO(5)$ were derived by means of a calculus of boson operators and were given in Eq. (2.21) of that paper. We may transform the generators of $Sp(4)$ to those of $Sp(2,2)$ by means of the following relations:

$$\begin{aligned}
\sqrt{\tfrac{1}{2}}(L_{52}+iL_{15}) &= \sqrt{\tfrac{1}{2}}(E_{14}+E_{32}) \to i\sqrt{\tfrac{1}{2}}(E_{14}+E_{32}) \\
&= \sqrt{\tfrac{1}{2}}(M_{52}+iM_{15}), \\
\sqrt{\tfrac{1}{2}}(L_{52}-iL_{15}) &= \sqrt{\tfrac{1}{2}}(E_{41}+E_{23}) \to i\sqrt{\tfrac{1}{2}}(E_{41}+E_{23}) \\
&= \sqrt{\tfrac{1}{2}}(M_{52}-iM_{15}), \\
\sqrt{\tfrac{1}{2}}(L_{45}+iL_{53}) &= \sqrt{\tfrac{1}{2}}(E_{13}-E_{42}) \to i\sqrt{\tfrac{1}{2}}(E_{13}-E_{42}) \\
&= \sqrt{\tfrac{1}{2}}(M_{45}+iM_{53}), \\
\sqrt{\tfrac{1}{2}}(L_{45}-iL_{53}) &= \sqrt{\tfrac{1}{2}}(E_{31}-E_{24}) \to i\sqrt{\tfrac{1}{2}}(E_{31}-E_{24}) \\
&= \sqrt{\tfrac{1}{2}}(M_{45}-iM_{53}), \qquad (2.1)
\end{aligned}$$

while the generators of the compact subgroup $SU(2) \otimes SU(2)$ remain unchanged. The matrix elements of these generators in $Sp(4)$ are given by

$$\begin{aligned}
&\sqrt{\tfrac{1}{2}}(E_{41}+E_{23})\,|\Phi,\Delta;J,\Lambda;M_J,M_\Lambda\rangle \\
&\quad= -\sqrt{\tfrac{1}{2}}\,A_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J-\frac{1}{2}}_{M_J\ -\frac{1}{2}\ M_J-\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda-\frac{1}{2}}_{M_\Lambda\ -\frac{1}{2}\ M_\Lambda-\frac{1}{2}}\,|\Phi,\Delta;J-\tfrac{1}{2},\Lambda-\tfrac{1}{2};M_J-\tfrac{1}{2},M_\Lambda-\tfrac{1}{2}\rangle \\
&\qquad-\sqrt{\tfrac{1}{2}}\,B_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J-\frac{1}{2}}_{M_J\ -\frac{1}{2}\ M_J-\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda+\frac{1}{2}}_{M_\Lambda\ -\frac{1}{2}\ M_\Lambda-\frac{1}{2}}\,|\Phi,\Delta;J-\tfrac{1}{2},\Lambda+\tfrac{1}{2};M_J-\tfrac{1}{2},M_\Lambda-\tfrac{1}{2}\rangle \\
&\qquad+\sqrt{\tfrac{1}{2}}\,C_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J+\frac{1}{2}}_{M_J\ -\frac{1}{2}\ M_J-\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda-\frac{1}{2}}_{M_\Lambda\ -\frac{1}{2}\ M_\Lambda-\frac{1}{2}}\,|\Phi,\Delta;J+\tfrac{1}{2},\Lambda-\tfrac{1}{2};M_J-\tfrac{1}{2},M_\Lambda-\tfrac{1}{2}\rangle \\
&\qquad-\sqrt{\tfrac{1}{2}}\,D_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J+\frac{1}{2}}_{M_J\ -\frac{1}{2}\ M_J-\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda+\frac{1}{2}}_{M_\Lambda\ -\frac{1}{2}\ M_\Lambda-\frac{1}{2}}\,|\Phi,\Delta;J+\tfrac{1}{2},\Lambda+\tfrac{1}{2};M_J-\tfrac{1}{2},M_\Lambda-\tfrac{1}{2}\rangle, \qquad (2.2a)
\end{aligned}$$

$$\begin{aligned}
&\sqrt{\tfrac{1}{2}}(E_{14}+E_{32})\,|\Phi,\Delta;J,\Lambda;M_J,M_\Lambda\rangle \\
&\quad= -\sqrt{\tfrac{1}{2}}\,A_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J-\frac{1}{2}}_{M_J\ \frac{1}{2}\ M_J+\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda-\frac{1}{2}}_{M_\Lambda\ \frac{1}{2}\ M_\Lambda+\frac{1}{2}}\,|\Phi,\Delta;J-\tfrac{1}{2},\Lambda-\tfrac{1}{2};M_J+\tfrac{1}{2},M_\Lambda+\tfrac{1}{2}\rangle \\
&\qquad-\sqrt{\tfrac{1}{2}}\,B_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J-\frac{1}{2}}_{M_J\ \frac{1}{2}\ M_J+\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda+\frac{1}{2}}_{M_\Lambda\ \frac{1}{2}\ M_\Lambda+\frac{1}{2}}\,|\Phi,\Delta;J-\tfrac{1}{2},\Lambda+\tfrac{1}{2};M_J+\tfrac{1}{2},M_\Lambda+\tfrac{1}{2}\rangle \\
&\qquad+\sqrt{\tfrac{1}{2}}\,C_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J+\frac{1}{2}}_{M_J\ \frac{1}{2}\ M_J+\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda-\frac{1}{2}}_{M_\Lambda\ \frac{1}{2}\ M_\Lambda+\frac{1}{2}}\,|\Phi,\Delta;J+\tfrac{1}{2},\Lambda-\tfrac{1}{2};M_J+\tfrac{1}{2},M_\Lambda+\tfrac{1}{2}\rangle \\
&\qquad-\sqrt{\tfrac{1}{2}}\,D_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J+\frac{1}{2}}_{M_J\ \frac{1}{2}\ M_J+\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda+\frac{1}{2}}_{M_\Lambda\ \frac{1}{2}\ M_\Lambda+\frac{1}{2}}\,|\Phi,\Delta;J+\tfrac{1}{2},\Lambda+\tfrac{1}{2};M_J+\tfrac{1}{2},M_\Lambda+\tfrac{1}{2}\rangle, \qquad (2.2b)
\end{aligned}$$

[8] E. Inönü and E. P. Wigner, Proc. Natl. Acad. Sci. (U.S.) **39**, 510 (1953).
[9] S. Ström, Arkiv Fysik **30**, 455 (1965).
[10] N. T. Evans, J. Math. Phys. **8**, 170 (1967).
[11] P. A. M. Dirac, Ann. Math. **36**, 657 (1935).
[12] P. Roman and J. J. Aghassi, Phys. Letters **14**, 68 (1965).
[13] P. Roman and J. J. Aghassi, Nuovo Cimento **36**, 1062 (1965).
[14] P. Roman and J. J. Aghassi, Nuovo Cimento **37**, 354 (1965).
[15] P. Roman and J. J. Aghassi, Nuovo Cimento **38**, 1092 (1965).
[16] P. Roman and J. J. Aghassi, J. Math. Phys. **7**, 1273 (1966).
[17] E. Tagirov and N. Chernikov, report of work prior to publication.
[18] O. Nachtmann, Commun. Math. Phys. **6**, 1 (1967).
[19] F. Gürsey and T. D. Lee, Proc. Natl. Acad. Sci. (U.S.) **49**, 179 (1963).
[20] F. Gürsey, in *Group Theoretical Concepts and Methods in Elementary Particle Physics* (Gordon & Breach, Science Publishers, Inc., New York, 1964), p. 365.
[21] A. O. Barut and C. Fronsdal, Proc. Roy. Soc. (London) **A287**, 532 (1965).
[22] J. S. Lomont and H. E. Moses, J. Math. Phys. **5**, 294 (1964).
[23] J. S. Lomont and H. E. Moses, J. Math. Phys. **5**, 1438 (1964).

$$\sqrt{\tfrac{1}{2}}(E_{13}-E_{42})\,|\Phi,\Delta;J,\Lambda;M_J,M_\Lambda\rangle$$
$$=-\sqrt{\tfrac{1}{2}}\,A_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J-\frac{1}{2}}_{M_J\ \frac{1}{2}\ M_J+\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda-\frac{1}{2}}_{M_\Lambda\ -\frac{1}{2}\ M_\Lambda-\frac{1}{2}}\,|\Phi,\Delta;J-\tfrac{1}{2},\Lambda-\tfrac{1}{2};M_J+\tfrac{1}{2},M_\Lambda-\tfrac{1}{2}\rangle$$
$$-\sqrt{\tfrac{1}{2}}\,B_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J-\frac{1}{2}}_{M_J\ \frac{1}{2}\ M_J+\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda+\frac{1}{2}}_{M_\Lambda\ -\frac{1}{2}\ M_\Lambda-\frac{1}{2}}\,|\Phi,\Delta;J-\tfrac{1}{2},\Lambda+\tfrac{1}{2};M_J+\tfrac{1}{2},M_\Lambda-\tfrac{1}{2}\rangle$$
$$+\sqrt{\tfrac{1}{2}}\,C_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J+\frac{1}{2}}_{M_J\ \frac{1}{2}\ M_J+\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda-\frac{1}{2}}_{M_\Lambda\ -\frac{1}{2}\ M_\Lambda-\frac{1}{2}}\,|\Phi,\Delta;J+\tfrac{1}{2},\Lambda-\tfrac{1}{2};M_J+\tfrac{1}{2},M_\Lambda-\tfrac{1}{2}\rangle$$
$$-\sqrt{\tfrac{1}{2}}\,D_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J+\frac{1}{2}}_{M_J\ \frac{1}{2}\ M_J+\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda+\frac{1}{2}}_{M_\Lambda\ -\frac{1}{2}\ M_\Lambda-\frac{1}{2}}\,|\Phi,\Delta;J+\tfrac{1}{2},\Lambda+\tfrac{1}{2};M_J+\tfrac{1}{2},M_\Lambda-\tfrac{1}{2}\rangle, \quad (2.2c)$$

$$\sqrt{\tfrac{1}{2}}(E_{31}-E_{24})\,|\Phi,\Delta;J,\Lambda;M_J,M_\Lambda\rangle$$
$$=+\sqrt{\tfrac{1}{2}}\,A_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J-\frac{1}{2}}_{M_J\ -\frac{1}{2}\ M_J-\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda-\frac{1}{2}}_{M_\Lambda\ \frac{1}{2}\ M_\Lambda+\frac{1}{2}}\,|\Phi,\Delta;J-\tfrac{1}{2},\Lambda-\tfrac{1}{2};M_J-\tfrac{1}{2},M_\Lambda+\tfrac{1}{2}\rangle$$
$$+\sqrt{\tfrac{1}{2}}\,B_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J-\frac{1}{2}}_{M_J\ -\frac{1}{2}\ M_J-\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda+\frac{1}{2}}_{M_\Lambda\ \frac{1}{2}\ M_\Lambda+\frac{1}{2}}\,|\Phi,\Delta;J-\tfrac{1}{2},\Lambda+\tfrac{1}{2};M_J-\tfrac{1}{2},M_\Lambda+\tfrac{1}{2}\rangle$$
$$-\sqrt{\tfrac{1}{2}}\,C_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J+\frac{1}{2}}_{M_J\ -\frac{1}{2}\ M_J-\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda-\frac{1}{2}}_{M_\Lambda\ \frac{1}{2}\ M_\Lambda+\frac{1}{2}}\,|\Phi,\Delta;J+\tfrac{1}{2},\Lambda-\tfrac{1}{2};M_J-\tfrac{1}{2},M_\Lambda+\tfrac{1}{2}\rangle$$
$$+\sqrt{\tfrac{1}{2}}\,D_{J,\Lambda}\,C^{J\ \frac{1}{2}\ J+\frac{1}{2}}_{M_J\ -\frac{1}{2}\ M_J-\frac{1}{2}}\,C^{\Lambda\ \frac{1}{2}\ \Lambda+\frac{1}{2}}_{M_\Lambda\ \frac{1}{2}\ M_\Lambda+\frac{1}{2}}\,|\Phi,\Delta;J+\tfrac{1}{2},\Lambda+\tfrac{1}{2};M_J-\tfrac{1}{2},M_\Lambda+\tfrac{1}{2}\rangle, \quad (2.2d)$$

where

$$A_{J,\Lambda}=\left(\frac{(\Phi-J-\Lambda)(\Delta+J+\Lambda+1)(-\Delta+J+\Lambda)(\Phi+J+\Lambda+1)}{(2J)(2\Lambda)}\right)^{\frac{1}{2}}, \quad (2.3a)$$

$$B_{J,\Lambda}=\left(\frac{(\Delta-J+\Lambda+1)(\Phi-J+\Lambda+1)(\Phi+J-\Lambda)(\Delta+J-\Lambda)}{(2J)(2\Lambda+2)}\right)^{\frac{1}{2}}, \quad (2.3b)$$

$$C_{J,\Lambda}=\left(\frac{(\Delta-J+\Lambda)(\Phi+J-\Lambda+1)(\Phi-J+\Lambda)(\Delta+J-\Lambda+1)}{(2J+2)(2\Lambda)}\right)^{\frac{1}{2}}, \quad (2.3c)$$

$$D_{J,\Lambda}=\left(\frac{(\Phi-J-\Lambda-1)(\Delta+J+\Lambda+2)(\Phi+J+\Lambda+2)(-\Delta+J+\Lambda+1)}{(2J+2)(2\Lambda+2)}\right)^{\frac{1}{2}}, \quad (2.3d)$$

and where we have written $\Phi\equiv J_m+\Lambda_m+1$ and $\Delta\equiv J_m-\Lambda_m$. We use the Condon-Shortley conventions for the phases of the $SU(2)$ Wigner coefficients and the conventions of Paper I for the phases of the basis states of $Sp(4)$. These matrix elements of the generators may be verified (tediously) in our spinor calculus for the basis states by use of our lowering operators O_{--} and O_{-+} for the semimaximal states and also the raising operators O_{++} and O_{+-} which have the properties

$$O_{++}=E_{14}+E_{32}, \quad (2.4a)$$

$$O_{+-}=(E_{33}-E_{44}+1)(E_{13}-E_{42})+E_{43}(E_{14}+E_{32}), \quad (2.4b)$$

$$O_{++}\,|\Phi,\Delta;J,\Lambda;J,\Lambda\rangle = c\,|\Phi,\Delta;J+\tfrac{1}{2},\Lambda+\tfrac{1}{2};J+\tfrac{1}{2},\Lambda+\tfrac{1}{2}\rangle, \quad (2.5a)$$

$$O_{+-}\,|\Phi,\Delta;J,\Lambda;J,\Lambda\rangle = c'\,|\Phi,\Delta;J+\tfrac{1}{2},\Lambda-\tfrac{1}{2};J+\tfrac{1}{2},\Lambda-\tfrac{1}{2}\rangle. \quad (2.5b)$$

In $Sp(4)$, the betweenness conditions

$$\Phi-1\geq J+\Lambda\geq\Delta\geq|J-\Lambda| \quad (2.6)$$

are satisfied. When we apply the relations (2.1) in order to go from the operators of $Sp(4)$ to those of $Sp(2,2)$, we find that the basis states $|\Phi,\Delta;J,\Lambda;M_J,M_\Lambda\rangle$ are states of irreducible unitary representations of $Sp(2,2)$ for different values of Φ. The states $|\Phi,\Delta\rangle$ are unitary in $Sp(2,2)$ for the following values of Φ and Δ:

(A) The continuous class:

(a) The representations $v_{\Phi,\Delta}$, with $\Delta=1,2,3,\cdots$ and $-\Phi(\Phi+1)>0$.

(b) The representations $v_{\Phi,\Delta}$, with $\Delta=\frac{1}{2},\frac{3}{2},\frac{5}{2},\cdots$ and $-\Phi(\Phi+1)>\frac{1}{4}$.

(c) The representations $v_{\Phi,0}$, with $\Delta=0$ and $-\Phi(\Phi+1)>-2$. Here, $J=\Lambda$.

(B) The discrete class:

(a) The representations $\pi^{+}_{\Phi,\Delta}$, with $\Delta=\frac{1}{2},1,\frac{3}{2},2,\cdots$ and $-\Phi=\Delta-1,\Delta-2,\cdots,1$ or $\frac{1}{2}$, where $J+\Lambda\geq\Delta\geq\Lambda-J\geq\Phi+1$.

(b) The representations $\pi^{-}_{\Phi,\Delta}$, with $\Delta=\frac{1}{2},1,\frac{3}{2},2,\cdots$ and $-\Phi=\Delta-1,\Delta-2,\cdots,1$ or $\frac{1}{2}$, where $J+\Lambda\geq\Delta\geq J-\Lambda\geq\Phi+1$.

(c) The representations $\pi_{0,\Delta}$, with $\Delta=1,2,3,\cdots$ and $\Phi=0$. Here, $J=\Lambda$.

In all of these cases the betweenness conditions $J+\Lambda\geq\Delta\geq|J-\Lambda|$ hold. In going from the basis states of $Sp(4)$ to those of $Sp(2,2)$, we must perform the continuation only in Φ. We must now examine the $Sp(4)$ basis states derived in Paper I and find their analytic continuation in Φ. The semimaximal states

$|\Phi, \Delta; J, \Lambda; J, \Lambda\rangle$, given by Eq. (2.18) of Paper I, can be written in the form

$$\left(\frac{(2\Delta+1)!\,\Gamma(\Phi+J-\Lambda+1)\Gamma(\Phi+J+\Lambda+2)(2\Lambda+1)!}{(\Delta-J+\Lambda)!\,\Gamma(\Phi-J-\Lambda)\Gamma(2\Phi+1)\Gamma(\Phi-J+\Lambda+1)(\Delta+J+\Lambda+1)!\,(\Delta+J-\Lambda)!}\right)^{\frac{1}{2}}$$
$$\times\left(\frac{1}{(J+\Lambda-\Delta)!\,(2J+1)!}\right)^{\frac{1}{2}}{}_2F_1\left(-\Phi+J+\Lambda+1,\,-\Phi+J-\Lambda\,|2J+1|-\frac{a_{12}}{a_{34}}\right)$$
$$\times\left(\frac{a_{34}}{a_{13}}\right)^{\Phi-J-\Lambda-1}\left(\frac{a_3}{a_1}\right)^{\Delta-J+\Lambda}a_{13}^{\Phi-\Delta-1}a_1^{2\Delta}\,|0\rangle. \tag{2.7}$$

We now omit the vacuum state $|0\rangle$ and consider the operators a_k^i, not as boson creation operators, but as spinors which can be assigned any complex number as exponent. The conjugates of these spinors we write as $\partial/\partial a_k^i$, so that the generators can be expressed in terms of them as

$$E_{ik}=\sum_{p=1}^{2}a_i^p\frac{\partial}{\partial a_k^p}. \tag{2.8}$$

We now perform the analytic continuation of the state (2.7) in Φ; this continuation is unique since both the gamma function and the hypergeometric function ${}_2F_1$ possess unique analytic continuations in the complex planes of all arguments and indices. We adopt the convention that the phase of each of the gamma functions in the normalization factor is zero at the commencement of the continuation.

When Φ is continued to an arbitrary complex value, the series representation of the hypergeometric function becomes nonterminating. We may, however, continue to use the series representation (or the Pochhammer or the Barnes contour-integral representation) since the ${}_2F_1$ function has an abstract spinor as its argument and is itself merely a symbolic recipe which describes the behavior of the basis state under the operation of the generators (2.8), i.e., since we have not defined a norm for the argument of the ${}_2F_1$ function.

We may now apply the lowering operators E_{21} and $-E_{43}$ to lower the magnetic quantum numbers from their maximal to general values. We note that the semimaximal basis state in the form (2.7) has the important property that all the generators of the maximal compact subgroup $Sp(2)\otimes Sp(2)$ commute with the ${}_2F_1$ function, i.e., with a_{12} and a_{34}. Hence we may write the normalized general state as

$$(-1)^{\Lambda-M_\Lambda}\left(\frac{(2\Delta+1)!\,\Gamma(\Phi+J-\Lambda+1)\Gamma(\Phi+J+\Lambda+2)(\Delta-J+\Lambda)!\,(\Delta+J-\Lambda)!}{\Gamma(2\Phi+1)\Gamma(\Phi-J-\Lambda)\Gamma(\Phi-J+\Lambda+1)(\Delta+J+\Lambda+1)!\,(2J+1)!}\right)^{\frac{1}{2}}$$
$$\times\left(\frac{(J+\Lambda-\Delta)!\,(2\Lambda+1)!\,(J-M_J)!\,(J+M_J)!\,(\Lambda-M_\Lambda)!\,(\Lambda+M_\Lambda)!}{(2J)!\,(2\Lambda)!}\right)^{\frac{1}{2}}$$
$$\times\sum_{k_1z_1z_2}\frac{1}{k_1!\,(\Delta+J-\Lambda-k_1)z_1!\,(\Delta-J+\Lambda-z_1)!\,z_2!\,(-\Delta+M_\Lambda+M_J+k_1+z_1+z_2)!}$$
$$\times\frac{1}{(\Lambda-M_\Lambda-z_1-z_2)!\,(J-M_J-k_1-z_2)!}\,{}_2F_1\left(-\Phi+J+\Lambda+1,\,-\Phi+J-\Lambda\,|2J+1|-\frac{a_{12}}{a_{34}}\right)$$
$$\times\left(\frac{a_{34}}{a_{13}}\right)^{\Phi-J-\Lambda-1}\left(\frac{a_3}{a_1}\right)^{\Delta-J+\Lambda}\left(\frac{a_{23}}{a_{13}}\right)^{J-M_J}\left(\frac{a_{14}}{a_{13}}\right)^{\Lambda-M_\Lambda}\left(\frac{a_{13}a_2}{a_{23}a_1}\right)^{k_1}\left(\frac{a_{13}a_4}{a_{14}a_3}\right)^{z_1}\left(\frac{a_{13}a_{24}}{a_{14}a_{23}}\right)^{z_2}. \tag{2.9}$$

The states (2.9) now serve as basis states for the irreducible unitary representations of $Sp(2,2)$ under the generators on the right of (2.1), where the E_{ik} are defined as in (2.8).

We now constitute our representation space (2.9) as a Hilbert space by defining abstractly the inner product and norm

$$\langle\Phi\Delta; J'\Lambda'; M'_JM'_\Lambda\,|\,\Phi\Delta; J\Lambda; M_JM_\Lambda\rangle = \delta_{JJ'}\delta_{\Lambda\Lambda'}\delta_{M_JM_J'}\delta_{M_\Lambda M_\Lambda'} \tag{2.10}$$

for all cases in which $(\Phi\Delta)$ label one of the irreducible unitary representations given above. We may further define the state

$$|\Phi\Delta; J\Lambda; LM\rangle = \sum_{M_J+M_\Lambda=M}C^{J\ \Lambda\ L}_{M_J\ M_\Lambda\ M}\,|\Phi\Delta; J\Lambda; M_JM_\Lambda\rangle, \tag{2.11}$$

diagonalizing the invariant of the subgroup $SO(3)$ of rotations in the dimensions (1, 2, 3).

Here we have followed the phase conventions of Paper I. The basis states defined by Hecht for the $Sp(4)$ case[24] differ from these by the phase $(-1)^{\Lambda-M_\Lambda}$, and those defined by Ström differ from them by the phase $(-1)^{J-\Lambda}$.[9]

[24] K. T. Hecht, Nucl. Phys. **63**, 177 (1965).

3. CONTRACTION TO REPRESENTATIONS OF THE POINCARÉ GROUP

The representations of the de Sitter group are of interest in physics primarily because they can be contracted in the sense of Inönü and Wigner to those of the Poincaré group. This contraction has been studied in detail for the $SO(4, 1)$ case by Ström, upon whose work the following section is based. The contraction is effected by the following mapping of the generators of the de Sitter group:

$$
\begin{aligned}
&\tfrac{1}{2}(E_{11} - E_{22} + E_{33} - E_{44}) = L_{12} \to L_{12}, \\
&\tfrac{1}{2}(E_{11} - E_{22} - E_{33} + E_{44}) = L_{34} \to \frac{1}{R} L_{34} \xrightarrow[R\to\infty]{} P_3, \\
&E_{12} - E_{34} = L_{23} + iL_{31} \to L_{23} + iL_{31}, \\
&E_{12} + E_{34} = L_{14} + iL_{24} \to \frac{1}{R}(L_{14} + iL_{24}) \xrightarrow[R\to\infty]{} P_1 + iP_2, \\
&E_{21} - E_{43} = L_{23} - iL_{31} \to L_{23} - iL_{31}, \\
&E_{21} + E_{43} = L_{14} - iL_{24} \to \frac{1}{R}(L_{14} - iL_{24}) \xrightarrow[R\to\infty]{} P_1 - iP_2, \\
&\tfrac{1}{2}i(E_{14} + E_{41} + E_{32} + E_{23}) = M_{52} \to M_{02}, \\
&\tfrac{1}{2}(E_{14} - E_{41} + E_{32} - E_{23}) = M_{15} \to M_{10}, \\
&\tfrac{1}{2}(E_{13} - E_{31} - E_{42} + E_{24}) = M_{53} \to M_{03}, \\
&\tfrac{1}{2}i(E_{13} + E_{31} - E_{42} - E_{24}) = M_{45} \to \frac{1}{R} M_{45} \xrightarrow[R\to\infty]{} P_0.
\end{aligned}
\tag{3.1}
$$

The contraction of the first six of these generators expresses the contraction of the generators of the $SO(4)$ subgroup of $SO(4, 1)$ to the $E(3)$ subgroup of the Poincaré group $E(3, 1)$. Our momentum operators P_i $(i = 1, 2, 3)$ differ from those of Ström by a sign and our generators of Lorentz transformations M_{02}, M_{10}, M_{03} have the correspondences

$$
\begin{aligned}
M_{10} &= -N_1, \\
M_{02} &= +N_2, \\
M_{03} &= +N_3,
\end{aligned}
\tag{3.2}
$$

with his and our P_0 is identical with his. We obtain the matrix elements of the generators of the de Sitter group as given by Ström [Eq. (2.3) of Ref. 9] when we insert the phase $(-1)^{J-\Lambda}$ into the expression for the basis state (2.9) and continue Φ into the *upper* half plane. When we continue Φ into the lower half plane, then the signs of the matrix elements of the de Sitter group generators [except those of the $Sp(2) \otimes Sp(2)$ subgroup] are reversed. This convention is important since in the contraction of the basis states (2.11), we find that it distinguishes between the positive- and the negative-energy representations of the Poincaré group. The contraction of the $Sp(2) \otimes Sp(2)$ subgroup to the $E(3)$ subgroup of $E(3, 1)$ has been discussed elsewhere.[25]

[25] W. J. Holman III, Ann. Phys. (N.Y.) **52**, 176 (1969).

The representations of the Poincaré group are labeled by the invariants μ^2, the square of mass, and by $-\mu^2 s(s + 1)$, the eigenvalue of the Pauli–Lubansky operator, where s is the spin; furthermore, the sign of energy ϵ, the helicity δ, and the sign ϵ_1 of the operator which corresponds to the rotation in the little group in the case $\mu^2 < 0$ are invariants for special representations.

We have the following representations of the Poincaré group, which are given along with the representations of the de Sitter group from which they are reached by contraction as determined by Ström. In each case the contraction parameter is R, and we perform the contraction by taking the asymptotic limit $R \to \infty$. We denote $\varphi \equiv J + \Lambda + 1$ and $\delta \equiv J - \Lambda$. In all cases, $\varphi = Rp$, where p is the absolute value of the three-momentum, and δ is the helicity, so that in the limit of the contraction $p\delta$ is the eigenvalue of the $E(3)$ invariant operator $\mathbf{P} \cdot \mathbf{L}$:

(A) The real-mass discrete-spin representations $P(\mu^2, s, \epsilon)$, $\mu^2 > 0, s = 0, \frac{1}{2}, 1, \cdots, \epsilon = \pm 1$, reached from the continuous representation $\nu_{\Phi,\Delta}$, where $\Phi = -\frac{1}{2} + i\epsilon R\mu$, $\varphi = Rp$, $\Delta = s$.

(B) The zero-mass infinite-spin representations $P(0, a^2, \epsilon)$, $\mu^2 = 0, -\mu^2 s(s + 1) = -a^2 < 0, \epsilon = \pm 1$, reached from the continuous representation $\nu_{\Phi,\Delta}$, where $\Phi = -\frac{1}{2} + i\epsilon(Ra)^{\frac{1}{2}}$, $\Delta = (Ra)^{\frac{1}{2}}$.

(C) The zero-mass finite-spin representations

$$P(0, 0, \epsilon, \delta), \quad \mu^2 = 0, \quad -\mu^2 s(s+1) = 0,$$
$$\epsilon = \pm 1, \quad \delta = 0, \pm\tfrac{1}{2}, \pm 1, \cdots,$$

obtained as special cases of $P(\mu^2, s, \epsilon)$ (for $\mu = 0$), or of $P(0, a^2, \epsilon)$ (for $a = 0$).

(D) The imaginary-mass continuous-spin representations $P(-\mu^2, \alpha)$, $\mu^2 < 0$, $\alpha = -s(s+1) > 0$ for single-valued representations of $E(3, 1)$, $\alpha > \frac{1}{4}$ for double-valued representations, reached from the continuous representation $v_{\Phi,\Delta}$, where $\Delta = R\mu$, $-\Phi(\Phi+1) = \alpha$.

(E) The imaginary-mass discrete-spin representations $P(-\mu^2, s, \epsilon_1)$, $\mu^2 < 0$, $s = -\frac{1}{2}, 0, \frac{1}{2}, 1, \cdots$, $\epsilon_1 = \pm 1$, reached from the discrete representation $\pi^{\epsilon_1}_{\Phi,\Delta}$, where $\Phi = s$, $\Delta = R\mu$.

We shall consider in detail only case (A), which is both the easiest and physically the most important. In the limit $R \to \infty$, the basis state of the representation $v_{\Phi,\Delta}$ approaches that of $P(\mu^2, s, \epsilon)$:

$$|\Phi\Delta; \varphi\delta; LM\rangle \xrightarrow[R\to\infty]{} R^{-\frac{1}{2}} |\mu, s, \epsilon; p\delta; LM\rangle. \quad (3.3)$$

By the insertion of the factor $R^{-\frac{1}{2}}$ on the right of (3.3), we pass from the set of basis states normalized in the sense of the Kronecker delta (2.10) to a set of states orthonormal in the sense of the Dirac delta function in momentum space:

$$\langle \mu, s, \epsilon; p'\delta'; L'M' \,|\, \mu, s, \epsilon; p\delta; LM\rangle = \delta_{\delta\delta'}\delta_{LL'}\delta_{MM'}\delta(p - p'). \quad (3.4)$$

We now investigate the limit of the contraction of the basis state (2.11). We note that, if we operate with the state (2.9) rather than with (2.11), then in the limit of the contraction the representations of the Poincaré group are labeled by the $E(3)$ invariants p^2 and $p\delta$, by the third component of the three-momentum $p\cos\chi$, where $\cos\chi$ is given by (3.6) below, and by the third component of the total angular momentum M. It is well known[26,27] that the $SO(3)$ Wigner coefficient approaches the form of the $SO(3)$ rotation function asymptotically:

$$C^{J\ \Lambda\ L}_{M_J\ M_\Lambda\ M} \xrightarrow[J|\gg M_J|\gg L,\delta]{} \left(\frac{2L+1}{\varphi}\right)^{\frac{1}{2}} (-1)^{J-M_J} d^L_{M\delta}(\chi), \quad (3.5)$$

where

$$\cos\chi = M_J/J. \quad (3.6)$$

Applying the Euler–Maclaurin theorem to (2.11), we find that the discrete remainder terms vanish in the limit $\varphi \to \infty$, and that in this limit the projection of the (LM) state of the $SO(3)$ subgroup passes from summation over the discrete indices of a Wigner coefficient to integration over the continuous argument of a representation function.

Next, we note that all three indices of the ${}_2F_1$ function become large in (2.9). We may easily find the asymptotic limit of the function by the standard WKB method, keeping terms of both first and zero order in R. This provides us, actually, with little gain in the simplicity of the expression and we shall often find it more convenient to use the Pochhammer or Barnes integral representation or the series expansion. The results of the WKB approximation are given for the sake of possible convenience. First we apply the identity

$$\begin{aligned}
&{}_2F_1\left(-\Phi+\varphi, -\Phi+\delta \,|\varphi+\delta+1| -\frac{a_{12}}{a_{34}}\right) \\
&= \left(\frac{a_{12}+a_{34}}{a_{34}}\right)^{\Phi-\varphi} \\
&\quad \times {}_2F_1\left(-\Phi+\varphi, \Phi+\varphi+1 \,|\varphi+\delta+1| \frac{a_{12}}{a_{12}+a_{34}}\right) \\
&= (a_{34})^{-\Phi+\varphi}\, {}_2F_1(-\Phi+\varphi, \Phi+\varphi+1 \,|\varphi+\delta+1|\, a_{12}),
\end{aligned} \quad (3.7)$$

where we may discard the factors $(a_{12} + a_{34})$ which are invariant under all the generators of $Sp(2, 2)$. The WKB approximation, to zero order in R, then gives us

$$(a_{34})^{-\Phi+\varphi}\left[A_+ \exp\left(\int da_{12}(R\Omega^{(0)}_+ + \Omega^{(1)}_+)\right) + A_- \exp\left(\int da_{12}(R\Omega^{(0)}_- + \Omega^{(1)}_-)\right)\right] \quad (3.8)$$

for (3.7), where

$$\begin{aligned}
\Omega^{(0)}_\pm &= -\frac{p - 2pa_{12}}{2a_{12}(1-a_{12})} \pm \frac{[p^2 + 4\mu^2 a_{12} - 4\mu^2(a_{12})^2]^{\frac{1}{2}}}{2a_{12}(1-a_{12})}, \\
\Omega^{(1)}_\pm &= -\frac{\delta}{2a_{12}(1-a_{12})} - \frac{\mu^2(1-2a_{12})}{p^2 + 4\mu^2 a_{12} - 4\mu^2(a_{12})^2} \\
&\quad \pm \frac{\delta(p - 2pa_{12})}{2a_{12}(1-a_{12})[p^2 + 4\mu^2 a_{12} - 4\mu^2(a_{12})^2]^{\frac{1}{2}}},
\end{aligned} \quad (3.9)$$

and the constants $A_\pm$ are easily determined since the two terms of (3.8) correspond to the two terms of the asymptotic expansion of the ${}_2F_1$ function on the right

[26] P. Brussaard and J. H. Tolhoek, Physica **23**, 955 (1957).

[27] E. P. Wigner, *Group Theory and Its Applications to the Quantum Mechanics of Atomic Spectra* (Academic Press Inc., New York, 1959), p. 351.

of (3.7):

$$A_+ = \frac{\Gamma(-i\epsilon R\mu)\Gamma(Rp+\delta+1)}{\Gamma(\frac{1}{2}+R(p-i\epsilon\mu))\Gamma(\frac{1}{2}+\delta-i\epsilon R\mu)},$$
$$A_- = \frac{\Gamma(+i\epsilon R\mu)\Gamma(Rp+\delta+1)}{\Gamma(\frac{1}{2}+R(p+i\epsilon\mu))\Gamma(\frac{1}{2}+\delta+i\epsilon R\mu)}, \quad (3.10)$$

and the gamma functions may be treated by Stirling's approximation.

The sums over k_1 and z_1 in (2.9) remain finite as $R\to\infty$, but we must look more closely at the sum over z_2, whose range becomes infinite. We must examine the sums over M_J, M_Λ, and z_2 together, i.e., the constituent

$$\sum_{M_J M_\Lambda z_2} C^{J\ \Lambda\ L}_{M_J\ M_\Lambda\ M}(-1)^{\Lambda-M_\Lambda} \times \frac{[(J-M_J)!\,(J+M_J)!\,(\Lambda-M_\Lambda)!\,(\Lambda+M_\Lambda)!]^{\frac{1}{2}}}{z_2!\,(\Lambda-M_\Lambda-z_1-z_2)!\,(J-M_J-k_1-z_2)!} \times \frac{1}{(-\Delta+M+k_1+z_1+z_2)!} \times \left(\frac{a_{23}}{a_{13}}\right)^{J-M_J}\left(\frac{a_{14}}{a_{13}}\right)^{\Lambda-M_\Lambda}\left(\frac{a_{13}a_{24}}{a_{14}a_{23}}\right)^{z_2} \quad (3.11)$$

of (2.11). The WKB approximation for the ${}_2F_1$ series over z_2 in the limit $R\to\infty$ becomes complicated, and it is more convenient to deal with the asymptotic limit in other ways. We may write the hypergeometric series over z_2 as a Pochhammer contour integral[28] and project the angular-momentum state (LM) by means of the Euler–Maclaurin integral taken over the asymptotic limit (3.5) of the $SO(3)$ Wigner coefficient. Thus (3.11) becomes [we treat here only the case $(M+\delta)\le 0$]

$$\tfrac{1}{2}[Rp(2L+1)]^{\frac{1}{2}}(Rp)^{\Delta}\int_{-1}^{+1} d(\cos\chi)\, d^L_{M\delta}(\chi)\left(\frac{a_{23}}{a_{13}}\right)^{\frac{1}{2}(Rp+\delta-1)(1-\cos\chi)} \times \left(-\frac{a_{14}}{a_{13}}\right)^{\frac{1}{2}(Rp+\delta-1)(1+\cos\chi)-M-\delta} (\sin^2\tfrac{1}{2}\chi)^{\Delta-z_1-\frac{1}{2}(M+\delta)}(\cos^2\tfrac{1}{2}\chi)^{z_1+\frac{1}{2}(M+\delta)} \times (-1)^{k_1}\frac{1}{2\pi i}\int_1^{(0+)} dt\, t^{-\frac{1}{2}(Rp+\delta-1)(1-\cos\chi)+k_1-1} \times (1-t)^{\frac{1}{2}(Rp+\delta-1)(1-\cos\chi)-\Delta+M+z_1}(1-Wt)^{\frac{1}{2}(Rp+\delta-1)(1+\cos\chi)-M-\delta-z_1}, \quad (3.12)$$
$$W = a_{13}a_{24}/a_{14}a_{23},$$

where we have applied Stirling's approximation to the gamma functions outside the hypergeometric function. We now note that

$$(\sin^2\tfrac{1}{2}\chi)^{\Delta-z_1-\frac{1}{2}(M+\delta)}(\cos^2\tfrac{1}{2}\chi)^{z_1+\frac{1}{2}(M+\delta)} = (-1)^{\Delta-\delta-z_1}\sum_j [(2j+1)]^{\frac{1}{2}}\, d^{\ j}_{-M\ -\delta}(\chi) \times \left[\frac{(\delta+z_1+M)!\,(\Delta-z_1-M)!\,(\Delta-\delta-z_1)!\,z_1!}{(j+\Delta+1)!\,(\Delta-j)!}\right]^{\frac{1}{2}} C^{\frac{1}{2}(\Delta+\delta)\ \frac{1}{2}(\Delta-\delta)\ j}_{\frac{1}{2}(\delta-\Delta)+z_1+M\ \frac{1}{2}(\Delta-\delta)-z_1\ M}, \quad (3.13)$$

and that

$$\left[\left(-\frac{a_{14}}{a_{23}}\right)\frac{t(1-Wt)}{(1-t)}\right]^{\frac{1}{2}(Rp+\delta-1)\cos\chi}$$
$$= \exp\left\{\tfrac{1}{2}(Rp+\delta-1)\ln\left[\left(-\frac{a_{14}}{a_{23}}\right)\frac{t(1-Wt)}{(1-t)}\right]\cos\chi\right\}$$
$$= \sqrt{\frac{\pi}{2}}\sum_{k=0}^{\infty}(2k+1)\, d^k_{0\ 0}(\chi)\left\{(\tfrac{1}{2}Rp)\ln\left[\left(-\frac{a_{14}}{a_{23}}\right)\frac{t(1-Wt)}{(1-t)}\right]\right\}^{-\frac{1}{2}} I_{k+\frac{1}{2}}\left\{\tfrac{1}{2}(Rp+\delta-1)\ln\left[\left(-\frac{a_{14}}{a_{23}}\right)\frac{t(1-Wt)}{(1-t)}\right]\right\}$$
$$\approx \left(\frac{1}{Rp}\right)\sum_{k=0}^{\infty}(2k+1)\, d^k_{0\ 0}(\chi)\left\{\ln\left[\left(-\frac{a_{14}}{a_{23}}\right)\frac{t(1-Wt)}{(1-t)}\right]\right\}^{-1} \times \left\{\left[\left(-\frac{a_{14}}{a_{23}}\right)\frac{t(1-Wt)}{(1-t)}\right]^{\frac{1}{2}(Rp+\delta-1)} + (-1)^{k+1}\left[\left(-\frac{a_{14}}{a_{23}}\right)\frac{t(1-Wt)}{(1-t)}\right]^{-\frac{1}{2}(Rp+\delta-1)}\right\}, \quad (3.14)$$

where we have used the asymptotic form of the hyperbolic Bessel function and kept only the leading terms.

[28] L. J. Slater, *Generalized Hypergeometric Functions* (Cambridge University Press, Cambridge, 1966), p. 22. The contour of integration in the t plane in (3.12) and thereafter starts at $+1$, circles 0 once in the counterclockwise (positive) sense, then closes at $+1$.

Substituting (3.13) and (3.14) into (3.12) and performing the integration, we obtain, for (3.12),

$$
\begin{aligned}
&(Rp)^{\Delta-\frac{1}{2}}(-1)^{\Delta-\delta+z_1+k_1}\left(\frac{a_{23}}{a_{13}}\right)^{\frac{1}{2}(Rp+\delta-1)}\left(-\frac{a_{14}}{a_{13}}\right)^{\frac{1}{2}(Rp+\delta-1)-M-\delta}\\
&\quad\times\sum_{k,j}\left[\frac{(2L+1)(2j+1)(\delta+z_1+M)!\,(\Delta-z_1-M)!\,(\Delta-\delta-z_1)!\,z_1!}{(j+\Delta+1)!\,(\Delta-j)!}\right]^{\frac{1}{2}}\\
&\quad\times C^{\frac{1}{2}(\Delta+\delta)}_{\frac{1}{2}(\delta-\Delta)+z_1+M\ \frac{1}{2}(\Delta-\delta)-z_1\ M}\,C^{L\ \ j\ \ k}_{M\ -M\ 0}\,C^{L\ \ j\ \ k}_{\delta\ -\delta\ 0}\,\frac{1}{2\pi i}\int_1^{(0+)}dt\,t^{-\frac{1}{2}(Rp+\delta-1)+k_1-1}(1-t)^{\frac{1}{2}(Rp+\delta-1)-\Delta+M+z_1}\\
&\quad\times(1-Wt)^{\frac{1}{2}(Rp+\delta-1)-M-\delta-z_1}\left\{\ln\left[\left(-\frac{a_{14}}{a_{23}}\right)\frac{t(1-Wt)}{(1-t)}\right]\right\}^{-1}\\
&\quad\times\left\{\left[\left(-\frac{a_{14}}{a_{23}}\right)\frac{t(1-Wt)}{(1-t)}\right]^{\frac{1}{2}(Rp+\delta-1)}+(-1)^{k+1}\left[\left(-\frac{a_{14}}{a_{23}}\right)\frac{t(1-Wt)}{(1-t)}\right]^{-\frac{1}{2}(Rp+\delta-1)}\right\}.
\end{aligned}
\tag{3.15}
$$

We insert (3.15) into (3.3) and perform the summation over k_1 by means of the binomial theorem. The complete result is

$$
\begin{aligned}
&|\mu,s,\epsilon;p\delta;LM\rangle\\
&\quad=\lim_{R\to\infty}\left[\frac{(2L+1)(2\Delta+1)!\,\Gamma(\frac{1}{2}+\delta+i\epsilon R\mu)\Gamma(\frac{1}{2}+R(p+i\epsilon\mu))}{\Gamma(2i\epsilon R\mu)\Gamma(\frac{1}{2}-R(p-i\epsilon\mu))\Gamma(\frac{1}{2}-\delta+i\epsilon R\mu)(\Delta+\delta)!}\,\frac{(\Delta-\delta)!}{(Rp+\delta)!\,(Rp-\delta-1)!}\right]^{\frac{1}{2}}\frac{(Rp)^{-\delta}}{p}\\
&\quad\times(-1)^{\Delta-\delta}\,{}_2F_1(\tfrac{1}{2}+R(p-i\epsilon\mu),\tfrac{1}{2}+R(p+i\epsilon\mu)\,|\,\tfrac{1}{2}+\delta+Rp|\,a_{12})\\
&\quad\times(a_{13})^{\frac{1}{2}+R(p-i\epsilon\mu)}\left(\frac{a_3}{a_1}\right)^{\Delta-\delta}\left(-\frac{a_{14}}{a_{13}}\right)^{-M-\delta}\sum_{z_1}\frac{1}{z_1!\,(\Delta-\delta-z_1)!}\left(-\frac{a_{13}a_4}{a_{14}a_3}\right)^{z_1}\\
&\quad\times\sum_{kj}\left[\frac{(2L+1)(2j+1)(\delta+M+z_1)!\,(\Delta-M-z_1)!\,(\Delta-\delta-z_1)!\,z_1!}{(j+\Delta+1)!\,(\Delta-j)!}\right]^{\frac{1}{2}}\\
&\quad\times C^{L\ \ j\ \ k}_{M\ -M\ 0}\,C^{L\ \ j\ \ k}_{\delta\ -\delta\ 0}\,C^{\frac{1}{2}(\Delta+\delta)\ \ \frac{1}{2}(\Delta-\delta)\ \ j}_{\frac{1}{2}(\delta-\Delta)+z_1+M\ \frac{1}{2}(\Delta-\delta)-z_1\ M}(2\pi i)^{-1}\\
&\quad\times\int_1^{(0+)}dt\,t^{-1}(1-t)^{-\Delta+M+z_1}(1-Wt)^{-M-\delta-z_1}\left[1-\left(\frac{a_{13}a_2}{a_{23}a_1}\right)t\right]^{\Delta+\delta}\left\{\ln\left[\left(-\frac{a_{14}}{a_{23}}\right)\frac{t(1-Wt)}{(1-t)}\right]\right\}^{-1}\\
&\quad\times\left\{\left[\left(-\frac{a_{14}}{a_{13}}\right)(1-Wt)\right]^{Rp+\delta-1}+(-1)^{k+1}\left[\left(\frac{a_{23}}{a_{13}}\right)\frac{(1-t)}{t}\right]^{Rp+\delta-1}\right\}.
\end{aligned}
\tag{3.16}
$$

A somewhat less complicated expression may be obtained if we perform the summation over the magnetic quantum numbers M_J, M_Λ before proceeding to the limit $R\to\infty$. This may be accomplished by the use of the addition theorem for binomial coefficients to reduce the number of gamma functions containing these quantum numbers in the expression (2.11). It was by this method that Racah was able to perform the summation over the magnetic quantum numbers of four Wigner coefficients to obtain an expression for the $SO(3)$ (6-j) symbol as a simple ${}_4F_3$ series.[29] By this method we may write (3.11) as

$$
\begin{aligned}
&(-1)^{2J+\Delta+M+k_1+z_1}[(2L+1)(L+M)!\,(L-M)!\,(L+\delta)!\,(L-\delta)!]^{\frac{1}{2}}(Rp)^{-L-\frac{1}{2}}\\
&\quad\times\sum_{z,\rho_1,\rho_2,\rho_3}\frac{1}{\rho_1!\,(L+M-z-\rho_1)!\,(L-\delta-z-\rho_1)!}\,\frac{1}{\rho_2!\,(z-\rho_2)!\,(\delta-M+z-\rho_2)!}\\
&\quad\times\frac{(2J-\Delta-\delta+k_1+z-\rho_2-\rho_3)!}{(2J-\Delta-\delta)!}\,\frac{(k_1+z-\rho_2)!\,(L+\Delta-z-k_1-\rho_1+\rho_3)!}{\rho_3!\,(k_1+z-\rho_2-\rho_3)!}\left(\frac{a_{23}}{a_{13}}\right)^{2J}\left(\frac{a_{13}}{a_{14}}\right)^{\delta+M}\\
&\quad\times\left(\frac{a_{14}}{a_{23}}\right)^{\Delta+\delta-k_1+\rho_3}\frac{1}{2\pi i}\int_1^{(0+)}dt\,t^{M-\Delta+k_1+z_1-\rho_3-1}(1-t)^{\rho_3}\\
&\quad\times(1-Wt)^{2J-\Delta-\delta-\rho_3}\left[1+\left(\frac{a_{14}}{a_{23}}\right)\frac{(1-t)}{t(1-Wt)}\right]^{-L-\Delta+z+k_1+\rho_1-\rho_3-1},
\end{aligned}
\tag{3.17}
$$

where z is the summation index of the ${}_3F_2$ series contained in the $SO(3)$ Wigner coefficient indicated in (3.11), and ρ_1, ρ_2, and ρ_3 are indices of summation for the auxiliary expansions which we have used, and we have

[29] G. Racah, Phys. Rev. **62**, 438 (1942).

performed the summation over M_J, M_Λ. We note now that the factor

$$\frac{(2J-\Delta-\delta+k_1+z-\rho_2-\rho_3)!}{(2J-\Delta-\delta)!} \xrightarrow[R\to\infty]{} (Rp)^{k_1+z-\rho_2-\rho_3} \tag{3.18}$$

prescribes that we keep only the terms $\rho_2=\rho_3=0, k_1=\Delta+\delta, z=L+M$ [if $(M+\delta)\leq 0$] or $z=L-\delta$ [if $(M+\delta)\geq 0$] in the summations over these indices. In consequence, also, $\rho_1=0$. Thus (3.17) becomes

(i) For $M+\delta\geq 0$:

$$(-1)^{2J-M+L+z_1}\frac{[(2L+1)(L+M)!\,(L-M)!\,(L+\delta)!\,(L-\delta)!]^{\frac{1}{2}}}{(L-\delta)!\,(M+\delta)!\,(L-M)!}(Rp)^{\Delta-\frac{1}{2}}\left(\frac{a_{23}}{a_{13}}\right)^{2J}\left(\frac{a_{13}}{a_{14}}\right)^{M+\delta}$$
$$\times\frac{1}{2\pi i}\int_1^{(0+)}dt\,t^{\delta+M+z_1-1}(1-Wt)^{2J-\Delta-\delta}\left[1+\left(\frac{a_{14}}{a_{23}}\right)\frac{(1-t)}{t(1-Wt)}\right]^{-1}, \tag{3.16$'$}$$

(ii) For $M+\delta\leq 0$:

$$(-1)^{2J+\delta+L+z_1}\left[\frac{(2L+1)(L-M)!\,(L-\delta)!}{(L+M)!\,(L+\delta)!}\right]^{\frac{1}{2}}(Rp)^{\Delta+\delta+M-\frac{1}{2}}\left(\frac{a_{23}}{a_{13}}\right)^{2J}\left(\frac{a_{13}}{a_{14}}\right)^{M+\delta}$$
$$\times\frac{1}{2\pi i}\int_1^{(0+)}dt\,t^{M+\delta+z_1-1}(1-Wt)^{2J-\Delta-\delta}\left[1+\left(\frac{a_{14}}{a_{13}}\right)\frac{(1-t)}{t(1-Wt)}\right]^{M+\delta-1}. \tag{3.16$''$}$$

We may perform the summation over z_1 by the binomial theorem and the complete basis state may be written as follows:

(i) For $M+\delta\geq 0$:

$$|\mu,s,\epsilon;p\delta;LM\rangle$$
$$=(-1)^{s+\delta-M+L}\left[\frac{(2L+1)(2s+1)!\,\Gamma(\frac{1}{2}+\delta+i\epsilon R\mu)}{p\Gamma(2i\epsilon R\mu)\Gamma(\frac{1}{2}-R(p-i\epsilon\mu))}\right.$$
$$\left.\times\frac{\Gamma(\frac{1}{2}+R(p+i\epsilon\mu))(L+M)!\,(L+\delta)!}{\Gamma(\frac{1}{2}-\delta+i\epsilon R\mu)(Rp+\delta)!\,(Rp+\delta-1)!\,(s+\delta)!\,(s-\delta)!\,(L-M)!\,(L-\delta)!}\right]^{\frac{1}{2}}\frac{1}{(M+\delta)!}$$
$$\times\frac{1}{2\pi i}\int_1^{(0+)}dt\,t^{\delta+M-1}(Wt-1)^{Rp-s-1}\left[1-\left(\frac{a_{13}a_4}{a_{14}a_3}\right)t\right]^{s-\delta}\left[1+\left(\frac{a_{14}}{a_{23}}\right)\frac{(1-t)}{t(1-Wt)}\right]^{-1}$$
$$\times{}_2F_1(\tfrac{1}{2}+R(p-i\epsilon\mu),\tfrac{1}{2}+R(p+i\epsilon\mu)\,|\,\tfrac{1}{2}+\delta+Rp|\,a_{12})(a_{13})^{\frac{1}{2}+R(p-i\epsilon\mu)}\left(\frac{a_3}{a_1}\right)^{s-\delta}\left(\frac{a_{13}a_2}{a_{23}a_1}\right)^{s+\delta}\left(\frac{a_{23}}{a_{13}}\right)^{Rp+\delta-1}\left(\frac{a_{13}}{a_{14}}\right)^{M+\delta}, \tag{3.17$'$}$$

(ii) For $M+\delta\leq 0$:

$$|\mu,s,\epsilon;p\delta;LM\rangle$$
$$=(-1)^{s-L}\left[\frac{(2L+1)(2s+1)!\,\Gamma(\frac{1}{2}+\delta+i\epsilon R\mu)\Gamma(\frac{1}{2}+R(p+i\epsilon\mu))}{p\Gamma(2i\epsilon R\mu)\Gamma(\frac{1}{2}-R(p-i\epsilon\mu))\Gamma(\frac{1}{2}-\delta+i\epsilon R\mu)}\right.$$
$$\left.\times\frac{(L-M)!\,(L-\delta)!}{(Rp+\delta)!\,(Rp+\delta-1)!\,(L+M)!\,(L+\delta)!\,(s+\delta)!\,(s-\delta)!}\right]^{\frac{1}{2}}(Rp)^M$$
$$\times\frac{1}{2\pi i}\int_1^{(0+)}dt\,t^{M+\delta-1}(Wt-1)^{Rp-s-1}\left[1-\left(\frac{a_{13}a_4}{a_{14}a_3}\right)t\right]^{s-\delta}\left[1+\left(\frac{a_{14}}{a_{13}}\right)\frac{(1-t)}{t(1-Wt)}\right]^{M+\delta-1}(a_{13})^{\frac{1}{2}+R(p-i\epsilon\mu)}$$
$$\times{}_2F_1(\tfrac{1}{2}+R(p-i\epsilon\mu),\tfrac{1}{2}+R(p+i\epsilon\mu)\,|\,\tfrac{1}{2}+\delta+Rp|\,a_{12})\left(\frac{a_3}{a_1}\right)^{s-\delta}\left(\frac{a_{13}a_2}{a_{23}a_1}\right)^{s+\delta}\left(\frac{a_{23}}{a_{13}}\right)^{Rp+\delta-1}\left(\frac{a_{13}}{a_{14}}\right)^{M+\delta},$$
$$W=a_{13}a_{24}/a_{14}a_{23}. \tag{3.17$''$}$$

For a discussion of the properties of the asymptotic Hilbert spaces formed by these basis states and of the expression of the $E(3,1)$ generators as differential operators in momentum space, we refer the reader to the work of Ström,[9] and of Lomont and Moses.[22,23]

ACKNOWLEDGMENTS

The author would like to thank Dr. Behram Kurşunoğlu for his gracious hospitality at the Center for Theoretical Studies and Dr. Erdal Inönü for helpful discussions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Remarks on Lie-Algebra-Invariant Equations. I

MAYER HUMI
Department of Applied Mathematics, The Weizmann Institute of Science, Rehovot, Israel

(Received 27 December 1968)

We construct first-order-invariant equations with respect to the inhomogeneous Euclidean algebra $IE(n)$ (isomorphic to the n-dimensional Galilei algebra) by embedding its Euclidean subalgebra $E(n)$ in $O(n+1, 1)$. This embedding allows us to utilize Gel'fand's method [I. Gel'fand *et al.*, *Representations of the Rotation and Lorentz Groups and Their Applications* (Pergamon Press, Inc., New York, 1963)] adapted to the construction of invariant equations on semisimple algebras to nonsemisimple ones also.

1. INTRODUCTION

It was recently pointed out in the literature[1] that first-order-invariant equations with respect to the homogeneous Galilei algebra [which is isomorphic to the homogeneous Euclidean algebra in three dimensions—$E(3)$] might play the same role in nonrelativistic quantum theory that the Dirac equation plays with respect to the Klein–Gordon equation in the Lorentz-invariant theory. This means that these equations describe particles with spin in the nonrelativistic domain. The method employed in the construction of these equations is that of linearization of the Schrödinger equation. However, one might seek a more covariant and transparent method in which the group-theoretic aspects of the problem will play a more pronounced role. This can be achieved by a generalization of Gel'fand method,[2] which is applicable to the construction of invariant equations with respect to semisimple algebras. The obstacle that must be overcome to achieve this is that the second-order Casimir operator for the n-dimensional Galilei algebras [isomorphic to the inhomogeneous Euclidean algebra $IE(n)$] is not a homogeneous second-order expression in the momenta, but rather has the form $E - p_i p^i/2m$. In Sec. 2 of this paper we show that this difficulty can be overcome and the basic equations for a group-theoretic construction of finite- and infinite-component equations, invariant under the transformation of $E(n)$, are given. In Secs. 3 and 4 we give examples of the general theory. In Sec. 3 we construct the infinite-component equations of $E(2)$. The main motivation for treating this subject is to compare the resulting mass spectrum with that of the corresponding infinite-component equations of the Lorentz group. We rederive in Sec. 4 the nonrelativistic Dirac equation, but the method employed enables us to infer results which were not found in Ref. 1. It should be stressed, however, that both topics treated in Secs. 3 and 4 are special examples of the general theory developed in Sec. 2 and there is no obstacle in applying the theory to other cases, such as other types of finite-dimensional representations[3] of $E(3)$ than those considered in Sec. 4.

2. GENERAL GROUP-THEORETIC ANALYSIS

Our purpose is to construct invariant equations with respect to $IE(n)$ in the form

$$(L_\mu p^\mu + k)\psi = 0. \tag{2.1}$$

To do this, we use a method devised in a recent paper of Pinski[4] in order to introduce a nonsingular metric in $E(n)$. The method consists of embedding $E(n) = O(n) \times T_n$ in $O(n+1, 1)$ and considering vectors on $E(n)$ as having $n+2$ components, instead of $n+1$. Thus, for example, the momentum vector is defined as $p^\alpha = (p^i, E, m)$. The nonsingular metric is

$$g_{\alpha\beta} = \begin{pmatrix} -I & \vdots & 0 \\ \cdots & \vdots & \cdots \\ 0 & \vdots & \begin{matrix} 0 & 1 \\ 1 & -1 \end{matrix} \end{pmatrix}, \tag{2.2}$$

where I is the n-dimensional unit matrix.

The algebra $E(n)$ has $n(n+1)/2$ generators. These are J_{ij} [generators of $O(n)$] and K_i (the boost operators), $i, j = 1, \cdots, n$. If we introduce the notations

$$J_{i,n+1} = K_i,$$

we can write the CR (commutation relations) of $E(n)$ in the form (noting that $g_{n+1,n+1} = 0$):

$$[J_{\mu\nu}, J_{\rho\pi}] = i(g_{\mu\rho}J_{\nu\pi} + g_{\nu\pi}J_{\mu\rho} - g_{\mu\pi}J_{\nu\rho} - g_{\nu\rho}J_{\mu\pi}),$$
$$\mu, \nu, \rho, \pi = 1, \cdots, n+1. \tag{2.3}$$

Moreover, it is possible to write the CR of p_α with the generators of the group as follows:

$$[p_\lambda, J_{\mu\nu}] = i(g_{\lambda\mu}p_\nu - g_{\lambda\nu}p_\mu). \tag{2.4}$$

It is crucial to notice that, in what follows, the metric g is applied in order to lower and raise indices only of

[1] J.-M. Lévy-Leblond, Commun. Math. Phys. 6, 286 (1967).
[2] I. Gel'fand *et al.*, *Representations of the Rotation and Lorentz Groups and Their Applications* (Pergamon Press, Inc., New York, 1963).
[3] C. George and M. Levy-Nahas, J. Math. Phys. 7, 980 (1966).
[4] G. Pinski, J. Math. Phys. 9, 1927 (1968).

p_α [e.g., $p_\alpha = (-p^i, m, E - m)$] and other vector operators of the algebra. It is not applied to $J_{\mu\nu}$, since $J_{n+1,n+2}$ is not defined.

It is easily shown now that a sufficient condition for (2.1) to be invariant under $E(n)$ is that L_μ ($\mu = 1, \cdots, n+2$) satisfy similar CR as p_μ, i.e.,

$$[L_\lambda, J_{\mu\nu}] = i(g_{\lambda\mu}L_\nu - g_{\lambda\nu}L_\mu). \tag{2.5}$$

In fact, a sufficient condition for the invariance of (2.1) is that $L^\mu p_\mu$ be a scalar of the algebra, i.e.,

$$\begin{aligned} 0 &= [L^\lambda p_\lambda, J_{\mu\nu}] \\ &= L^\lambda[p_\lambda, J_{\mu\nu}] + [L_\lambda, J_{\mu\nu}]p^\lambda \\ &= iL^\lambda(g_{\lambda\mu}p_\nu - g_{\lambda\nu}p_\mu) + [L_\lambda, J_{\mu\nu}]p^\lambda \\ &= i(L_\mu g_{\nu\lambda} - g_{\mu\lambda}L_\nu)p^\lambda + [L_\lambda, J_{\mu\nu}]p^\lambda, \end{aligned}$$

from which (2.5) follows immediately.

Equations (2.5) are, therefore, the basic equations we must solve in order to write down first-order-invariant equations for $IE(n)$.

3. INFINITE-COMPONENT EQUATIONS FOR $E(2)$

The infinite-dimensional representations of $E(2)$ can be most easily found from the second Casimir operator, whose coordinate form is

$$\left[-\frac{1}{r}\frac{d}{dr}\left(r\frac{d}{dr}\right) - \frac{1}{r^2}m^2 + \lambda\right]\xi^\tau_{l,m} = 0. \tag{3.1}$$

By standard transformations, Eq. (3.1) can be brought to Bessel equation and we find that

$$H_+\xi^\tau_{l,m} = (J_{23} - iJ_{13})\xi^\tau_{l,m} = l\xi^\tau_{l,m+1}, \tag{3.2a}$$

$$H_-\xi^\tau_{l,m} = (J_{23} + iJ_{13})\xi^\tau_{l,m} = l\xi^\tau_{l,m+1}. \tag{3.2b}$$

To find the desired L_μ, let us note that

$$[L_3, J_{\mu\nu}] = 0 \tag{3.3}$$

and, therefore, if we put

$$L_3\xi^\tau_{l,m} = \sum d^{\tau\tau'}_{lml'm}\xi^{\tau'}_{l'm'}, \tag{3.4}$$

we find that

$$d^{\tau\tau'}_{lm'm'} = \delta_{mm'}\delta_{ll'}d^{\tau\tau'}_l, \tag{3.5}$$

where $d^{\tau\tau'}_l$ is a constant. To calculate the other L_μ, it is sufficient to find L_4. Other L_μ can be found from the CR:

$$[L_4, J_{13}] = -iL_1, \quad [L_4, J_{23}] = -iL_2. \tag{3.6}$$

To calculate L_4, let us put

$$L_4\xi^\tau_{l,m} = \sum c^{\tau\tau}_{lml'm}\xi^{\tau'}_{l'm'}; \tag{3.7}$$

then, from $[L_4, J_{12}] = 0$, one infers that $m = m'$, i.e.,

$$c^{\tau\tau'}_{lm,l'm'} = c^{\tau\tau'}_{ll'm}\delta_{mm'},$$

but from (2.5) one can deduce that

$$[[L_4, H_-], H_+] = -2L_3, \tag{3.8a}$$

$$[[L_4, H_+], H_-] = -2L_3. \tag{3.8b}$$

It is sufficient to satisfy either of these equations and (3.6) in order to satisfy (2.5). A short calculation shows that $c^{\tau\tau'}_{ll'm}$ must satisfy

$$c^{\tau\tau'}_{ll'm} = \delta_{ll'}c^{\tau\tau'}_{lm} \tag{3.9}$$

and

$$-2d^{\tau\tau'}_l/l^2 = 2c^{\tau\tau'}_{lm} - c^{\tau\tau'}_{lm-1} - c^{\tau\tau'}_{l,m+1}. \tag{3.10}$$

This is the basic set of equations one must solve recursively in order to write down $E(2)$-invariant equations.

It is worthwhile to make the following two observations:

(1) In $E(2)$ a representation couples to itself only by means of an invariant equation of type (2.1) [in contrast to $O(2, 1)$ or $O(3)$].

(2) When $p = 0$, we have[5] $E = m$, and the mass spectrum of (2.11) is given by

$$m = k/(c^{\tau\tau'}_{lm} + dl). \tag{3.11}$$

Thus the spectrum does not depend on the label m explicitly but can be fixed by a free choice of three parameters. This should be compared to the corresponding infinite-component equations of the Lorentz group[2] whose spectrum is fixed to the multiplicative constant, e.g., the spectrum of the Majorana equation is

$$m = k/(j + \tfrac{1}{2}). \tag{3.12}$$

4. FINITE-COMPONENT EQUATIONS OF $E(3)$

We derive in the following the nonrelativistic analog of the Dirac equation. This is done by using a special type of finite representations of $E(3)$ in the form[1]

$$\Delta^j(R, \mathbf{v}) = \begin{pmatrix} D^j(R) & 0 \\ (\mathbf{J}\cdot\mathbf{v})D^j(R) & D^j(R) \end{pmatrix}, \tag{4.1}$$

where $R \in O(3)$, $v \in T_3$, and $D^j(R)$ is a representation of $O(3)$. We stress, however, that the same procedure can be easily applied to other types of finite-dimensional representation[3] of $E(3)$.

In the following we want to fix the finite-component equations resulting from the self-coupling of the representations of type (4.1) by equations of the form

$$L_\mu p^\mu \psi = 0. \tag{4.2}$$

To accomplish this, let us write down L_μ in block form

$$L_\mu = \begin{pmatrix} A_\mu & B_\mu \\ C_\mu & D_\mu \end{pmatrix}. \tag{4.3}$$

[5] We use here the nonrelativistic limit of $E^2 - p^2 = m^2$, which is $E = m + p^2/2m$.

Since $[L_4, J_{\mu\nu}] = 0$, it follows that $B_4 = 0$ and that $D_4 = -A_4$, C_4 are in the form λI. For L_5 we have

$$[L_5, J_{ij}] = 0, \tag{4.4}$$

from which we deduce that A_5, B_5, C_5, D_5 are in the form λI.

However, L_5 must also satisfy

$$[[L_5, F_+], F_-] = -2L_4, \tag{4.5}$$

where

$$F_\pm = J_{14} \pm iJ_{24}.$$

From this equation one finds that

$$A_4 = D_4 = 0 \tag{4.6}$$

and

$$B_5\{H_+, H_-\} = 2C_4. \tag{4.7}$$

Let us fix $C_4 = 2I$ and deal with $\Delta^{\frac{1}{2}}$. Then $\{H_+, H_-\} = 4I$ and, therefore, $B_5 = I$. However, we see that A_5, D_5, and C_5 are still arbitrary multiples of the identity. We choose $A_5 = D_5 = 0$ and $C_5 = -I$ in order to ensure $L_5L_5 = -1$ and a familiar form of the L_1. The complete set of L_μ is computed to be

$$L_1 = i\begin{pmatrix} \sigma_1 & \\ & \sigma_1 \end{pmatrix}, \quad L_4 = \begin{pmatrix} 0 & 0 \\ 2I & 0 \end{pmatrix}, \quad L_5 = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \tag{4.8}$$

where the σ_1 are the Pauli matrices. Let us notice that the solutions of (4.2) satisfy the nonrelativistic relation between E, m, p:

$$2Em = m^2 + p^2, \tag{4.9}$$

as one can see immediately, if we multiply (4.2) by $L_\nu p^\nu$ and note that

$$\{L_\mu, L_\nu\} = 2g_{\mu\nu}. \tag{4.10}$$

Out of this group-theoretic treatment, one infers that $\Delta^{\frac{1}{2}}$ is the only representation of type (4.1) for which it is possible to construct an invariant equation in the form (4.2). This result follows from (4.7) and (4.4). In fact, for $j \neq \frac{1}{2}$ $\{H_+H_-\} \neq \lambda I$ and therefore, from (4.7), $B \neq \lambda I$, but this contradicts (4.4).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Remarks on Lie-Algebra-Invariant Equations. II

MAYER HUMI
Department of Applied Mathematics, The Weizmann Institute of Science, Rehovot, Israel

(Received 27 December 1968)

We deal with a generalization of Gel'fand method which provides us with a simplified and constructive derivation of some known and new results on the construction and general properties of tensor and nth-order Lie-algebra-invariant equations.

1. INTRODUCTION

An important role in relativistic physics is played by Lorentz-invariant equations of the form[1]

$$(L^{\mu}p_{\mu} + k)\psi = 0, \tag{1.1}$$

where p_{μ} is a vector operator over the Lorentz algebra, i.e., it satisfies with the generators of $O(3, 1)$ the following commutation relations (CR):

$$[p_{\lambda}, J_{\mu\nu}] = i(g_{\mu\lambda}p_{\nu} - g_{\nu\lambda}p_{\mu}). \tag{1.2}$$

Our purpose in this paper is to deal with two generalizations to equations of type (1.1).

At first we consider invariant tensor equations over $O(p, q)$, i.e., equations of type (1.1) in which p_{μ} is a tensor operator over the algebra and, therefore, satisfies the CR

$$[p_{\lambda}, J_{\mu\nu}] = \sum_{\eta} \langle \eta | J_{\mu\nu} | \lambda \rangle p_{\eta}, \tag{1.3}$$

where $\langle \eta | J_{\mu\nu} | \lambda \rangle$ are the matrix elements of the generators in a representation D of the algebra and the range of η, $\lambda = \dim(D)$.

The mathematical motivation for dealing with such equations is obvious. From the physical point of view, let us note that Dirac has already considered[2] as a special type of such equations in his attempt to describe spin-$\frac{1}{2}$ particles in a de Sitter space.

The second type of equations is the nth-order Lorentz-invariant equations of the form

$$(L^{\mu_1 \cdots \mu_n}p_{\mu_1} \cdots p_{\mu_n} + k)\psi = 0, \tag{1.4}$$

where p_{μ} is a vector operator.

Let us point out that equations of type (1.4) were treated by Weinberg,[3] who found them appropriate to the description of particles with arbitrary spin ($n = 2s$). Thus we stress that the results given in Secs. 3 and 4 are not new. What is new (and forms our motivation) is their simplified derivation based on the Gel'fand method. In fact, this method forms the basis of all the results in this and the previous paper.

2. CONSTRUCTION OF TENSOR EQUATIONS

To write down invariant tensor equations of the form (1.1), we must find the appropriate L^{μ}. To this end we follow Gel'fand and find first the CR of L^{λ} with $J_{\mu\nu}$. To accomplish this, we note that a sufficient condition for the invariance of (1.1) is that $L^{\mu}p_{\mu}$ be a scalar of the algebra (k is a numerical constant), that is,

$$[L^{\mu}p_{\mu}, J_{\lambda\rho}] = 0, \tag{2.1}$$

from which we find

$$\begin{aligned}[L^{\mu}p_{\mu}, J_{\lambda\rho}] &= L^{\mu}[p_{\mu}, J_{\lambda\rho}] + [L^{\mu}, J_{\lambda\rho}]p_{\mu} \\ &= \sum_{\nu} \langle \nu | J_{\lambda\rho} | \mu \rangle L^{\mu}p_{\nu} + [L^{\mu}, J_{\lambda\rho}]p_{\mu},\end{aligned}$$

i.e.,

$$[L^{\mu}, J_{\lambda\rho}] = -\sum \langle \mu | J_{\lambda\rho} | \nu \rangle L^{\nu}. \tag{2.2}$$

Thus, to construct an invariant tensor equation, we must solve simultaneously Eqs. (2.2).

As an example, we give now the derivation of these equations for $O(3)$. In this case it is convenient to label the desired L^{μ} by the eigenvalues of J_3 and use the raising and lowering generators of the algebra.

We then have to solve the following CR:

$$[L^{m}, J_{+}] = -\langle j, m | J_{+} | j, m - 1 \rangle L^{m-1}, \tag{2.3}$$

$$[L^{m}, J_{-}] = -\langle j, m | J_{-} | j, m + 1 \rangle L^{m+1}, \tag{2.4}$$

$$[L^{m}, J_{3}] = -mL^{m}. \tag{2.5}$$

(We deal only with one-valued representation of $O(3)$, i.e., j is an integer.)

We see immediately that it is sufficient to calculate L^0, since the other components can be calculated then by virtue of (2.3) and (2.4). For L^0, we have then to satisfy the following two equations:

$$[[L^0, J_{+}], J_{-}] = j(j + 1)L^0, \tag{2.6}$$

$$[L^0, J_3] = 0, \tag{2.7}$$

[1] I. Gel'fand *et al.*, *Representations of Rotation and Lorentz Groups and Their Applications* (Pergamon Press, Inc., New York, 1963).
[2] P. A. M. Dirac, Ann. Math. **36**, 657 (1935).
[3] S. Weinberg, Phys. Rev. **133B**, 1318 (1964).

where we have used the well-known matrix elements of the $O(3)$ generators. To solve Eqs. (2.6) and (2.7), let us put for L^0, as in Ref. 1,

$$L^0\xi^{\tau}_{l,m} = \sum C^{\tau\tau'}_{lm;l'm'}\xi^{\tau'}_{l'm'}. \tag{2.8}$$

We then deduce from (2.7) that

$$C^{\tau\tau'}_{lm;l'm'} = \delta_{mm'}C^{\tau\tau'}_{ll'm}. \tag{2.9}$$

To solve (2.6), let us introduce the notations

$$J_+|l, m\rangle = \alpha_l^{m+1}|1, m+1\rangle, \quad J_-|l, m\rangle = \alpha_l^m|l, m\rangle. \tag{2.10}$$

We then have for (2.6) that

$$[(\alpha_l^m)^2 + (\alpha_{l'}^{m+1})^2]C^{\tau\tau'}_{ll'm} - \alpha_l^m\alpha_{l'}^m C^{\tau\tau'}_{ll',m-1} - \alpha_l^{m+1}\alpha_{l'}^{m+1}C^{\tau\tau'}_{ll',m+1} = j(j+1)C^{\tau\tau'}_{ll'm}. \tag{2.11}$$

To find which representations are coupled by the tensor equation of type (1.1), we have to solve (2.11) recursively, starting at the highest value of m (or the lowest value). At the end of this process, we arrive at an equation which contains the lowest two C's only. Since these values were already defined by previous equations, this last relationship fixes whether the two representations interlock or not. This procedure then shows that l, l' interlock if and only if

$$|l' - l'| \leq j. \tag{2.12}$$

We can deduce this relation also by looking on $p_\mu\psi$ as a multiplication of the representations j by l in which l' must be contained and therefore must satisfy (2.12).[4]

Equation (2.12) is a straightforward generalization of the classical relation $|l' - l| \leq 1$ for vector equations. In the same way, one can work out the selection rules for other groups of type $O(p, q)$, although the analysis will be more involved.

3. A RELATION BETWEEN TENSOR EQUATIONS

We would like to point out here a relation between equations of the form

$$(A^{\mu\nu}J_{\mu\nu} + k)\psi = 0, \tag{3.1}$$

where the $J_{\mu\nu}$ are the generators of $O(p+1, q)$ or $O(p, q+1)$ and vector equations over $O(p, q)$. To this end we realize the algebras $O(p+1, q)$ and $O(p, q+1)$ in terms of the generators of the vector inhomogeneous $O(p, q)$ $[= IO(p, q)]$.[5] Thus if we denote the generators of $O(p+1, q)$ or $O(p, q+1)$ by $J_{\mu\nu}$ $(\mu, \nu = 1, \cdots, n+1; p+q=n)$ and those of $IO(p, q)$ by M_{ij}, p_i $(i, j = 1, \cdots, n)$, then

$$J_{\mu\nu} = M_{\mu\nu}, \quad \mu, \nu = 1, \cdots, n,$$
$$J_{\mu,n+1} = \alpha(p^iM_{\mu i} + M_{\mu i}p^i), \quad \alpha = \tfrac{1}{2}(\epsilon p^2)^{-\frac{1}{2}}, \tag{3.2}$$

where ϵ is the metric of the $(n+1)$th coordinate $\epsilon = \pm 1$.

In terms of this realization, (3.1) transforms into

$$\{A^{ij}M_{ij} + \alpha A^{j,n+1}(p^iM_{ji} + M_{ji}p^i) + k\}\psi = 0. \tag{3.3}$$

Using the CR of p_i with M_{ij}, we transform (3.3) into

$$\{A^{ij}M_{ij} + \alpha A^{j,n+1}(-ip_j + 2M_{ji}p^i) + k\}\psi = 0. \tag{3.4}$$

Following Dirac,[2] we use the realization $M_{ij} = (x_ip_j - x_jp_i)$ and take (3.4) in the limit $x = (0, \cdots, 0, R)$; Eq. (3.4) then takes the form

$$(i\alpha A^{j,n+1}p_j - k)\psi = 0, \tag{3.5}$$

which is evidently a vector invariant equation of $O(p, q)$.

But $A_{\mu\nu}$ satisfies

$$[A_{\mu\nu}, J_{\rho\lambda}] = i(g_{\mu\rho}A_{\nu\lambda} + g_{\nu\lambda}A_{\mu\rho} - g_{\mu\lambda}A_{\nu\rho} - g_{\nu\rho}A_{\mu\lambda}). \tag{3.6}$$

We deduce therefore from (3.5) an explicit relation that exists between the L^μ needed for the construction of vector equations over $O(p, q)$ and the representations of $O(p+1, q)$ and $O(p, q+1)$. This is in fact an extension of Dirac results for the spin $\frac{1}{2}$ equation over the de Sitter group. In a way this relation gives also an easy way to write down vector equations of $O(p, q)$ if the representations of $O(p+1, q)$ or $O(p, q+1)$ are known, and vice versa.

4. VECTOR nth-ORDER EQUATIONS

Let us start the discussion by noting that Weinberg derived these equations by the use of the spinor calculus. Our derivation, on the contrary, uses the Gel'fand method and shows explicitly the connection between these equations and the first-order ones.

To accomplish our purpose let us consider first the equation

$$L^\mu p_\mu\psi + k\phi = 0, \tag{4.1}$$

where ψ belongs to a representation (l_0, l_1) of $O(3, 1)$ and ϕ to (l_0', l_1'). To find which representations of $O(3, 1)$ are coupled by (4.1), we proceed in the same way as Gel'fand and find that a necessary and sufficient condition is that, for any $g \in O(3, 1)$, the L^μ satisfy

$$g_{\mu\nu}T_g(l_0', l_0')L^\mu T_g^{-1}(l_0, l_1) = L^\nu. \tag{4.2}$$

[4] The details of this argument are given in Ref. 1 for vector equations. We do not repeat them, as they are similar in the general case.

[5] J. Rosen and P. Roman, J. Math. Phys. 7, 2072 (1966).

Following the same arguments given in Ref. 1, one can now show that (l_0', l_1') and (l_0, l_1) couple if and only if they satisfy

$$(l_0', l_1') = (l_0 \pm 1, l_1) \quad \text{or} \quad (l_0', l_1') = (l_0, l_1 \pm 1). \tag{4.3}$$

Moreover, the Gel'fand formulas for the matrix elements of L^μ still hold. The difference between (4.1) and (1.1) is that for (4.1) the representations are only coupled but not interlocked, i.e., there is no backward coupling of (l_0', l_1') to (l_0, l_1). This extra freedom in (4.1) now enables us to construct the following two chains of equations (represented by the coupled representations):

$$(s, s+1) \to (s-1, s+1) \to \cdots \to (-s, s+1), \tag{4.4}$$

$$(-s, s+1) \to (-s+1, s+1) \to \cdots \to (s, s+1). \tag{4.5}$$

By eliminating the intervening ϕ's, we find a $2s$-order equation which couples $(s, s+1)$ to $(-s, s+1)$ and back, i.e., we are led to the equations

$$L^{\mu_1 \cdots \mu_{2s}} p_{\mu_1} \cdots p_{\mu_{2s}} \psi + \tilde{\psi} = 0, \tag{4.6}$$

$$\tilde{L}^{\mu_1 \cdots \mu_{2s}} p_{\mu_1 \cdots \mu_{2s}} \tilde{\psi} + k\psi = 0, \tag{4.7}$$

where $\psi \in (s, s+1)$, $\tilde{\psi} \in (-s, s+1)$, and

$$L^{\mu_1 \cdots \mu_{2s}} = L^{\mu_1}(s, s-1) L^{\mu_2}(s-1, s-2) \cdots L^{\mu_{2s}}(-s+1, s), \tag{4.8}$$

$$\tilde{L}^{\mu_1 \cdots \mu_{2s}} = \tilde{L}^{\mu_1}(-s, s+1) \tilde{L}^{\mu_2}(-s+1, -s+2) \cdots \tilde{L}^{\mu_{2s}}(s-1, s), \tag{4.9}$$

and where we have used self-explanatory notations in view of (4.4) and (4.5). One can write now $\Psi = \binom{\tilde{\psi}}{\psi}$ and couple (4.6), (4.7) into a single equation in the usual way, resulting in the Weinberg equations.

Another approach to such equations would be the direct approach presented in Secs. 2 and 3. As an example, we consider second-order equations of the form

$$(T_{\mu\nu} p^\mu p^\nu + k)\psi = 0. \tag{4.10}$$

A short analysis shows that the $T_{\mu\nu}$ must satisfy the CR

$$[T_{\mu\nu}, T_{\rho\lambda}] = -i(g_{\nu\lambda} T_{\mu\rho} + g_{\mu\lambda} T_{\rho\nu} - g_{\nu\rho} T_{\mu\lambda} - g_{\mu\rho} T_{\lambda\nu}), \tag{4.11}$$

but the direct analysis of (4.11) is cumbersome even for $O(3)$. Of course, the method presented in this section for the construction of nth-order equations is applicable to any $O(p, q)$ group.

ACKNOWLEDGMENTS

I would like to thank Professor Y. Ne'eman, who suggested to me the study of Lie-invariant equations, and Professor B. Kaufman and Dr. R. Ruskies, who read parts of the manuscripts and offered some valuable remarks.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Finite Transformations and Basis States of $SU(n)$

DOUGLAS FRANCIS HOLLAND
Physics Department, Northern Arizona University, Flagstaff, Arizona

(Received 14 February 1969)

A convenient parameterization for finite transformations of $SU(n)$ is developed which explicitly exhibits the special unitary subgroups. This is also used to parameterize the defining space. Higher-dimensional representations are discussed. The question of which representations carry the trivial representation of $SU(n-1)$ is considered, as well as the parameterization of these states. Application is made to the transformations and basis states of $SU(3)$.

I. INTRODUCTION

The groups $SU(n)$, $n = 2, 3, 4, 6, 8$, have been of considerable interest as possible symmetry groups for the elementary particles. In this paper, finite transformations of $SU(n)$ are investigated and special application is made to $SU(3)$.

The defining representation of the group $SU(n)$ is the set of all special unitary $n \times n$ matrices which can be considered to perform transformations on an n-dimensional complex space. Murnaghan[1] derived a parameterization of unitary groups. Chacon and Moshinsky[2] using this parameterization found the irreducible representations for $U(3)$. In the following, a parameterization of $SU(n)$ is derived which is of special interest for the following reasons:

(1) It exhibits the subgroups $SU(n-1), SU(n-2), \cdots$,

(2) it employs a minimum number of the infinitesimal generators (IG's) of the group, and

(3) maximum use is made of the IG's which are also generators of the above subgroups.

This greatly facilitates the derivations of the representations of the group if the representations of the subgroups above are known.

For the group $SU(2)$, the IG's are the angular-momentum operators L_x, L_y, and L_z. These can be represented as differential operators acting upon the basis states Ψ_{lm}, which can be derived from the following relation:

$$\Psi_{lm} = \langle lm| D(\alpha, \beta, \gamma) |l = 0\rangle.$$

These states do not carry all representations of $SU(2)$, since they omit the half-integer values for l. Beg and Ruegg[3] as well as Nelson[4] employ differential operators representing the IG's of $SU(3)$ to derive basis states. By specializing to a particular right-hand state in the representations as shown above for $SU(2)$, basis states of the group are found which, for $SU(3)$, reduce to those of Nelson. These basis states carry all irreducible representations of $SU(3)$. The explicit form of $SU(n-1)$ in the parameterization clearly shows the possibility of labeling the states by these quantum numbers. It is also shown that, for some representation of $SU(n)$, basis states can be obtained.

II. PARAMETERIZATION

A parameterization of $n \times n$ special unitary matrices is derived and the ranges of the parameters are determined. The following notation is used:

$S(n-i)$ is a parameterization of the matrix formed by the direct sum of the defining representation of $SU(n-i)$ and the i-dimensional identity matrix;

$C(n-i)$, $T(n-i)$ are transformations in the above matrix group.

The parameterization has the factored form

$$S(n) = S(n-1)C(n)T(n-1). \tag{1}$$

The subgroup $SU(n-1)$ is explicitly shown. It will exhibit the subgroup $SU(n-2)$ when factored as above. In this way all the $SU(m)$, $m < n$, subgroups are displayed.

The only IG's employed in the parameterization are denoted as follows:

G_j: $-i$ in the $(j-1, j)$ position; i in the $(j, j-1)$ position;

D_j: $1/(j-1)$ on the diagonal to the (j, j) position; -1 in the (j, j) position;

all other positions are zero.

One should note that G_j and D_j are also IG's of the matrix group $S(j)$. By simple exponentiation of these

[1] Francis D. Murnaghan, *The Unitary and Rotation Group* (Spartan Books, Washington, D.C., 1962).
[2] E. Chacon and M. Moshinsky, Phys. Letters **23**, 567 (1966).
[3] M. Beg and H. Ruegg, J. Math. Phys. **6**, 677 (1965).
[4] T. J. Nelson, J. Math. Phys. **8**, 857 (1967).

matrices one finds that

$$e^{-i\theta G_j} = \begin{vmatrix} \ddots & & & & & & \\ & \ddots & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & & & & \cos\theta & -\sin\theta \\ & & & & & \sin\theta & \cos\theta \\ & & & & & & & 1 \\ & & & & & & & & 1 \\ & & & & & & & & & \ddots \\ & & & & & & & & & & \ddots \\ & & & & & & & & & & & \ddots \end{vmatrix},$$

$$e^{-i\phi D_j} = \begin{vmatrix} \ddots & & & & & & \\ & \ddots & & & & & \\ & & \ddots & & & & \\ & & & e^{i\phi/(1-j)} & & & \\ & & & & e^{i\phi/(1-j)} & & \\ & & & & & e^{i\phi} & \\ & & & & & & 1 \\ & & & & & & & 1 \\ & & & & & & & & \ddots \\ & & & & & & & & & \ddots \\ & & & & & & & & & & \ddots \end{vmatrix}.$$

Let U_k be the most general special unitary matrix with zeros in the $(1, n)$, $(2, n), \cdots, (k-1, n)$ positions. Furthermore, let the argument of the (k, n) and $(k+1, n)$ positions be equal. The following parameterization of these positions in U_k can be used:

$$[U_k]_{k,n} = e^{i\alpha_k}\cos\beta_k, \quad 0 \leq \alpha \leq 2\pi.$$
$$[U_k]_{k+1,n} = e^{i\alpha_k}\cos\beta_{k+1}, \quad 0 \leq \beta \leq \tfrac{1}{2}\pi.$$

In the following, $S(n)^{-1}$ is transformed into a special unitary $n \times n$ matrix with a one in the (n, n) position and all other positions in the nth row and column equal to zero. The matrix is then a representation of $SU(n-1)$. This is accomplished with the transformations

$$e^{-i\phi_1 D_2} S(n)^{-1} = U_1, \tag{2a}$$
$$e^{-i\phi_k D_k} e^{-i\theta_k G_k} U_{k-1} = U_k. \tag{2b}$$

Proof of Eq. (2a) is trivial and requires a range for ϕ_1 of $0 \leq \phi_1 \leq 2\pi$. For Eq. (2b), a solution for (ϕ_k, θ_k) must exist in the following:

$$[U_k]_{k-1,n} = e^{i\alpha_{k-1}}(\cos\theta_k \cos\beta_{k-1} - \sin\theta_k \cos\beta_k)e^{i\phi k/(1-k)} = 0 \tag{3}$$

$$[U_k]_{k,n} = e^{i\alpha_{k-1}}(\sin\theta_k \cos\beta_{k-1} + \cos\theta_k \cos\beta_k)e^{i\phi_k} = e^{i\alpha_{k+1}}\cos\beta_k''. \tag{4}$$

These equations are satisfied by allowing the ranges

$$0 \leq \phi_k \leq 2\pi, \quad 0 \leq \theta_k \leq \tfrac{1}{2}\pi.$$

In U_n, the (n, n) position can be set equal to one. From unitarity, all other positions in the nth row and column are zero. The final form can be written as

$$C(n)T(n-1)S(n)^{-1} = S(n-1)^{-1}, \tag{5}$$

which rearranged, is Eq. (1):

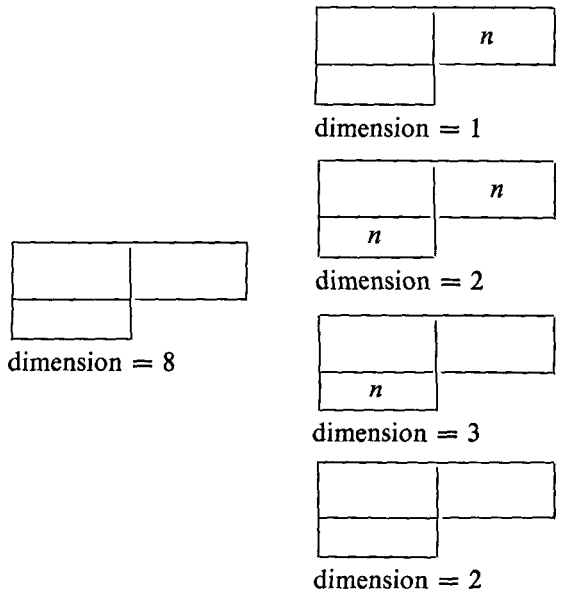
$$C(n) = e^{-i\phi_n D_n} e^{-i\theta_n G_n}, \tag{5a}$$
$$T(n-1) = e^{-i\phi_{n-1}D_{n-1}} e^{-i\theta_{n-1}G_{n-1}} \cdots e^{-i\phi_2 D_2} \times e^{-i\theta_2 G_2} e^{-i\phi_1 D_2}. \tag{5b}$$

The parameters have the ranges

$$0 \leq \phi_k \leq 2\pi, \quad 0 \leq \theta_k \leq \tfrac{1}{2}\pi.$$

III. REPRESENTATIONS

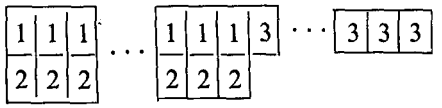
$S(n-1)$ and $T(n-1)$ are transformations in $SU(n-1)$ which do not affect the nth component of vectors in the defining space. Tensor basis states are constructed carrying the irreducible representations (IR's) according to the various Young patterns.[5] The components of this tensor are arranged into subspaces carrying the representations of $SU(n-1)$ which are invariant under the transformations $S(n-1)$ and $T(n-1)$. This is accomplished by systematically placing n in the overhang of the rows of the Young pattern and using all acceptable combinations of numbers excluding n in the other positions.[5] To illustrate, the following example is shown for $SU(3)$:



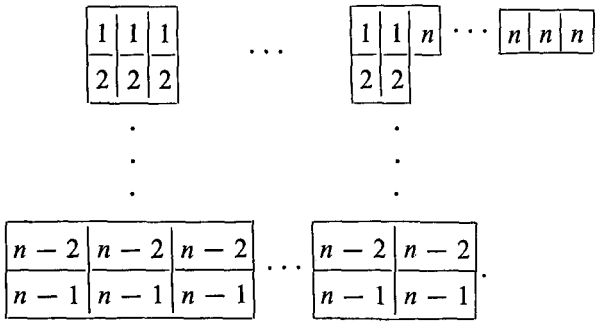
An important question for latter work is whether all irreducible tensors of $SU(n)$ carry the trivial (one-dimensional) representation of this $SU(n-1)$

[5] M. Hamermesh, *Group Theory* (Addison-Wesley Publ. Co., Reading, Mass., 1962), pp. 386–387.

transformation. The answer is yes for $SU(3)$, but not so for other SU groups. The most general $SU(3)$ tensor contains the following component which is invariant under the $SU(2)$ transformations and therefore carries the trivial representation:



For $SU(n)$ in general however, only the following patterns contain the trivial representation of $SU(n-1)$:



Using small greek letters to represent the numbers specifying the states in an IR, we have

$$\langle\sigma|\, S(n)\,|\rho\rangle = \sum_{\mu,\nu} \langle\sigma|\, S(n-1)\,|\mu\rangle\langle\mu|\, C(n)\,|\nu\rangle\langle\nu|\, T(n-1)\,|\rho\rangle. \tag{6}$$

The representations of $S(n-1)$ and $T(n-1)$ are direct sums of the IR's of $SU(n-1)$. If all IR's of $SU(n-1)$ are known, then these factors are easily evaluated. The factor in the center term must be found by direct application to the tensor states. The other factor in the center term is diagonal and poses no problem.

For $SU(3)$, using the standard notation,[6]

$$\begin{aligned}
&\left\langle \begin{matrix} IY \\ M \end{matrix} \right| S(3) \left| \begin{matrix} I'Y' \\ M' \end{matrix} \right\rangle \\
&= \sum_{M'',M'''} \left\langle \begin{matrix} IY \\ M \end{matrix} \right| S(2) \left| \begin{matrix} IY \\ M'' \end{matrix} \right\rangle \\
&\quad \times \left\langle \begin{matrix} IY \\ M'' \end{matrix} \right| C(3) \left| \begin{matrix} I'Y' \\ M''' \end{matrix} \right\rangle \left\langle \begin{matrix} I'Y' \\ M''' \end{matrix} \right| T(2) \left| \begin{matrix} I'Y' \\ M' \end{matrix} \right\rangle \\
&= \sum_{M'',M'''} D^{I}_{MM''}(\alpha,\beta,\gamma) e^{-i\frac{3}{2}\phi_3 Y} \\
&\quad \times \left\langle \begin{matrix} IY \\ M'' \end{matrix} \right| e^{-i\theta_3 G_3} \left| \begin{matrix} I'Y' \\ M''' \end{matrix} \right\rangle D^{I'}_{M'''M'}(2\phi_2, 2\theta_2, 2\phi_1).
\end{aligned} \tag{7}$$

The remaining center term is evaluated in Ref. 6.

IV. BASIS STATES

As noted in Sec. I, for some of the representations of $SU(2)$, coordinate representations of basis states can be realized which are of physical interest. Similarly, basis states for all representations of $SU(3)$ were found although their physical meaning, if any, is unknown. For the group $SU(n)$, one can parameterize the defining n-dimensional complex space as follows. Consider the constant vector $(0, 0, 0, \cdots, 1)$. $S(n)$ maps this vector into the entire space. Therefore,

$$\begin{aligned}
\Psi &\equiv S(n)(0,0,0,\cdots,1) \\
&= S(n-1)C(n)(0,0,0,\cdots,1) \\
&= S(2)C(3)C(4)\cdots C(n)(0,0,0,\cdots,1)
\end{aligned} \tag{8}$$

is the desired parameterization. The number of parameters is $3 + 2(n-2) = 2n-1$ which is the minimum necessary for a complete parameterization.

For representations that contain the trivial representation of $SU(n-1)$, let ρ_0 designate the state. Then

$$T(n-1)\,|\rho_0\rangle = |\rho_0\rangle$$

and the basis states can be represented as follows:

$$\Psi(\sigma) \equiv \sum_{\nu} \langle\sigma|\, S(n-1)\,|\nu\rangle\langle\nu|\, C(n)\,|\rho_0\rangle.$$

These states are labeled by all the numbers specifying the states and representations of $SU(n-1)$ plus the diagonal element in $C(n)$. Specifically for $SU(3)$, the basis states follow from Eq. (7):

$$\Psi^{IY}_{M} = \sum_{M''} D^{I}_{MM''}(\alpha,\beta,\gamma) e^{-i\frac{3}{2}\phi_3 Y} \left\langle \begin{matrix} IY \\ M'' \end{matrix} \right| e^{-i\theta_3 G_3} \left| \begin{matrix} OY \\ O \end{matrix} \right\rangle.$$

Using a method developed in a previous paper,[6] this becomes

$$\begin{aligned}
\Psi^{IY}_{M} &= \frac{(2I+1)^{\frac{1}{2}}}{(\lambda+1)^{\frac{1}{2}}(\mu+1)^{\frac{1}{2}}} \csc\theta_3 \\
&\quad \times e^{-i\frac{3}{2}\phi_3 Y} D^{I}_{M,-\frac{1}{2}Y-\frac{1}{3}(\lambda-\beta)}(\alpha,\beta,\gamma) \\
&\quad \times d^{\frac{1}{2}(\lambda+\mu+1)}_{\frac{1}{6}(\mu-\lambda+3Y+6I+3),\frac{1}{6}(\mu-\lambda+3Y-6I-3)}(2\theta_3)
\end{aligned}$$

to within a phase factor independent of M.

A question of considerable interest is what particles, if any, these functions represent. Nelson[7] defined a G-parity operation on these states and constructed a mass formula from which he concluded that these functions represent mesons.

If this coordinate representation of basis states has applicability, then the meaning of the parameters and the space in which they are significant is left unanswered. One eligible candidate which would eliminate the unwanted representations of $SU(3)$ is the 8-dimensional real space carrying the group $SU(3)/z3$ [a subgroup of $O(8)$].

ACKNOWLEDGMENT

I would like to show my appreciation to Dr. H. Mahmoud for many discussions of the subject.

[6] D. Holland, J. Math. Phys. **10**, 531 (1969).

[7] T. J. Nelson, Nuovo Cimento **52A**, 985 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Diffraction by a Slit or Strip*

PAUL B. BAILEY AND GEORGE E. BARR
Sandia Laboratories, Albuquerque, New Mexico

(Received 17 January 1969)

The problem of diffraction by a slit plane (or an infinite strip) is solved exactly for the case of a two-dimensional source distribution, the Green's function being given explicitly in terms of the fundamental solution of the wave operator. Also treated are the limiting case of plane waves and the asymptotic behavior with respect to large wavenumber.

I. INTRODUCTION

The purpose of this paper is to show how the Green's function can be constructed for the problem of diffraction by a slit plane (or an infinite strip) in the two-dimensional case; i.e., when the source distribution consists of line sources parallel to the edges of the slit (or strip). We shall consider only the case of zero initial and boundary conditions since to do otherwise only complicates matters and sheds no additional light on the phenomenon of diffraction.

Denote by $g_\nu(t - T, P, Q)$ the Green's function for the problem (in two space dimensions),

$$\left(\nabla^2 - \frac{\partial^2}{\partial t^2}\right)u_\nu(t, P) = f(t, P), \quad \text{for} \quad t \geq 0, \quad (1)$$

with Dirichlet boundary condition

$$u_1(t, P) = 0 \quad (2)$$

on the slit plane (or strip) if $\nu = 1$, and Neumann boundary condition

$$\frac{\partial}{\partial n} u_2(t, P) = 0 \quad (3)$$

on the slit plane (or strip) if $\nu = 2$, so that $u_\nu(t, P)$ has the form

$$u_\nu(t, P) = \int\int_0^t f(T, Q)g_\nu(t - T, P, Q)\, dT\, dQ. \quad (4)$$

Here, $\partial/\partial n$ means normal derivative on the slit plane (or strip) and $\int dQ$ means integration over all two-dimensional space.

Since the Laplace transform of the inhomogeneous wave equation (1) is, under suitable conditions, the inhomogeneous modified Helmholtz equation

$$(\nabla^2 - \gamma^2)\bar{u}_\nu(P) = \bar{f}(P), \quad (5)$$

where the bars indicate Laplace transform and γ is the transform parameter, the corresponding Green's function $G_\nu(P, Q)$ for (5) is just the Laplace transform of $g_\nu(t, P, Q)$. [We are assuming zero initial conditions for (1).[1]]

The solution of the diffraction problem for the case of a two-dimensional plane harmonic wave incident upon a perfectly conducting half-plane was given by Sommerfeld in 1896. Carslaw in 1899, and later Macdonald,[2] obtained the Green's functions G_1 and G_2, corresponding to the Dirichlet and Neumann conditions, respectively, on the boundary for both a line source and a point source. The problem has been developed much further and the list of contributors is large. Suitable references can be found in Bouwkamp's review article[3] and in Jones' recent book.[4] (The potential problem for a line source corresponds to the diffraction of cylindrical waves, whereas the point source corresponds to spherical waves. Similarly, the Dirichlet and Neumann conditions on the boundary correspond, respectively, to the perfectly conducting boundary for an electromagnetic wave and the perfectly rigid boundary for an acoustic wave.)

For diffraction of scalar waves by an infinite slit in a perfectly reflecting plane or by an infinite perfectly reflecting strip, problems of the type with an incident plane wave have been extensively studied. Solutions of these are either exact solutions expressed in the form of infinite series in elliptic cylinder functions or parabolic cylinder functions[5] or are approximate solutions for the cases that the ratio of slit (or strip) width to wavelength is rather large or rather small.[3,6] Problems of the cylindrical- and of the spherical-wave

* This work was supported by the United States Atomic Energy Commission.

[1] H. G. Garnir, Soc. Roy. Sci. (Liege) **22**, 29 (1953); L. Cagniard, E. A. Flinn, and C. H. Dix, *Reflection and Refraction of Progressive Seismic Waves* (McGraw-Hill Book Co., New York, 1962).

[2] A. Sommerfeld, Math. Ann. **47**, 317 (1896); H. S. Carslaw, Proc. London Math. Soc. **30**, 121 (1899); H. M. McDonald, Proc. London Math. Soc. (2) **14**, 410 (1915).

[3] C. J. Bouwkamp, Rept. Progr. Phys. **17**, 35 (1954).

[4] D. S. Jones, *The Theory of Electromagnetism* (The Macmillan Co., New York, 1964).

[5] M. J. O. Strutt, Z. Physik **69**, 597 (1931); P. M. Morse and P. J. Rubenstein, Phys. Rev. **54**, 895 (1938); S. Skavlem, Arch. Math. Naturvidenskab. **51**, 61 (1951).

[6] S. N. Karp and A. Russek, J. Appl. Phys. **27**, 886 (1956); E. Groschwitz and H. Hönl, Z. Physik **131**, 305 (1952); H. Hönl and E. Zimmer, Z. Physik **135**, 196 (1953).

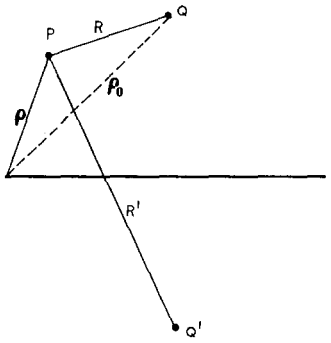


FIG. 1. A line source is located at Q, at a distance ρ_0 from the edge of a half-plane. The image of Q in the plane is Q' and the distances of an arbitrary observation point P from the edge and from Q and Q' are, respectively, ρ, R, and R'.

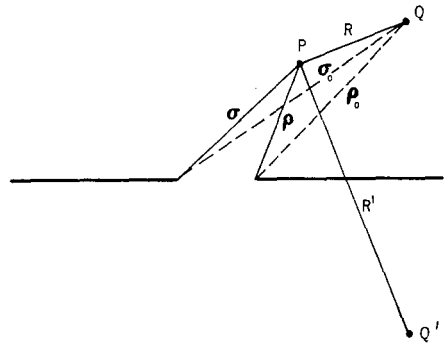


FIG. 2. A line source is located at Q, at distances ρ_0 and σ_0 from the two edges of an infinite slit in a plane. Q' is the image of Q and P is an arbitrary observation point whose distances from the edges of the slit are ρ and σ.

type have been rather less studied.[4,7] For insight into the nature of the problem and for an extensive list of literature one can consult Bouwkamp[3] and Jones.[4]

Section II is devoted to a description of the corresponding Green's function for the half-plane, the elements of which will be used in the construction for the slit and strip in Secs. III and IV.

The geometrical parameters of interest for the case of a line source Q, parallel to the edge of a half-plane (or slit plane), are those shown in Figs. 1 and 2. Q' is the image of the source Q, and P is an arbitrary point of observation. Although only two coordinates would be needed to specify the location of any point P, the Green's functions seem to be most easily described in terms of a superabundant set of coordinates: namely, the distances ρ, σ, ρ_0, σ_0 of P and Q from the edges, together with the distances R and R' of P from Q and Q', respectively. Figure 3 is intended to depict the space–time regions bounded by the various wavefronts.

The limiting case of plane waves falling upon a slit plane (or strip) are deduced in Sec. V by the usual device of multiplication by a suitable factor describing the source strength and a limit process moving the location of the source to infinity.[3,8]

Section VI briefly indicates how the asymptotic behavior with respect to large wavenumber can be obtained.

II. GREEN'S FUNCTIONS FOR THE HALF-PLANE

Since the construction of the Green's functions for the slit (strip) will use the solutions obtained by others[9] for the half-plane problems, it will be a convenience to have these listed below.

In the case of a line source at Q, the Green's function $g_\nu(t - T, P, Q)$, $\nu = 1$ or 2, for the wave equation (1) can be defined as

$$\begin{aligned} g_\nu(t-T,P,Q) &\equiv -(2\pi)^{-1}v, && \text{for } (T,Q)\in\mathfrak{D}_i,\\ &\equiv -(2\pi)^{-1}[v+(-1)^\nu v^*], && \text{for } (T,Q)\in\mathfrak{D}_r,\\ &\equiv \tfrac{1}{2}[-(2\pi)^{-1}][v+(-1)^\nu v^*], && \text{for } (T,Q)\in\mathfrak{D}_d, \end{aligned} \tag{6}$$

where

$$v \equiv [(t-T)^2 - R^2]^{-\frac{1}{2}}, \quad v^* \equiv [(t-T)^2 - R'^2]^{-\frac{1}{2}}.$$

For the Dirichlet problem, $\nu = 1$ and $g_1 = 0$ on the half-plane; for the Neumann problem, $\nu = 2$ and $\partial g_2/\partial n = 0$ on the half-plane ($\partial/\partial n$ = normal derivative). R is the distance between points P and Q, R' is the distance between P and Q' (the image of Q in the half-plane). The region $\mathfrak{D}_i$ consists of the set of points (T, Q) such that P receives at time t a direct signal which left Q at time T, $\mathfrak{D}_r$ is the set of points

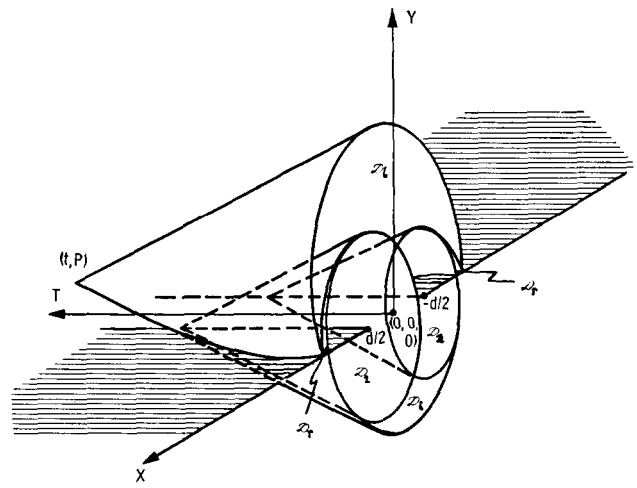


FIG. 3. The regions of integration corresponding to a point of observation (t, P) above a slit plane (slit width d): $\mathfrak{D}_i$ is the region from which only the direct radiation is received at (t, P); $\mathfrak{D}_r$ is the region from which only direct and reflected radiation are received; $\mathfrak{D}_1$ and $\mathfrak{D}_2$ are the regions from which diffracted radiation is received.

[7] E. N. Fix, Phil. Trans. Roy. Soc. (London) **A241**, 71 (1948); **A242**, 6 (1949).

[8] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Inc., New York, 1959), Chap. XI.

[9] A. P. Burger, Proc. Roy. Soc. (London) **A252**, 411 (1959); F. G. Friedlander, Quart. J. Mech. Appl. Math. **4**, 344 (1951); *Sound Pulses* (Cambridge University Press, Cambridge, England, 1958); G. E. Barr, Quart. Appl. Math. **25**, 193 (1967), and references given in these papers.

(T, Q) from which both the direct and reflected signals are received, and $\mathfrak{D}_d$ is the set of points (T, Q) from which the diffracted signal would be received. Thus, if one denotes the distances of P and Q from the edge of the half-plane by ρ and ρ_0, respectively, one has

$$\begin{aligned}
\mathfrak{D}_d &= \{(T, Q): \rho + \rho_0 \leq t - T\},\\
\mathfrak{D}_r &= \{(T, Q) \notin \mathfrak{D}_d\text{: the line } \overline{PQ'} \text{ meets the half-plane; } R' < t - T\},\\
\mathfrak{D}_i &= \{(T, Q) \notin \mathfrak{D}_d \cup \mathfrak{D}_r\text{: the line } \overline{PQ} \text{ does not meet the half-plane; } R < t - T\}. \quad (7)
\end{aligned}$$

In terms of this function g_ν, the solution u_ν of the wave equation (1) satisfying the Dirichlet or Neumann condition on the half-plane and satisfying zero initial conditions at time $t = 0$ can be written in the form

$$u_\nu(t, P) = \iiint f(T, Q) g_\nu(t - T, P, Q)\, dT\, dQ \quad (8)$$

with $\nu = 1$ or 2.

One also needs the result [Ref. 9, pp. 352–353] that to obtain the solution of the wave equation (1), which is equal to a given function $p_1(t, P_0)$ on the half-plane [or has normal derivative there equal to $p_2(t, P_0)$] rather than vanishing, one must add to (8) a boundary integral of the form

$$-\frac{1}{\pi}\frac{\partial}{\partial y}\iint p_1(T, Q_0) v(t - T, R_0)\, dT\, dQ_0, \quad (9)$$

where the subscript 0 denotes points on the half-plane, R_0 is the distance from P to Q_0, and the range of integration is over those points (T, Q_0) for which $t - \rho - \rho_0 < T < t - R_0$ (or of the form

$$-\epsilon\pi^{-1}\iint p_2(T, Q_0) v(t - T, R_0)\, dT\, dQ_0 \quad (10)$$

in the case $\nu = 2$, where $\epsilon = 1$ or -1 according as P is above or below the plane).

The Green's function $G_\nu(P, Q)$ for the Helmholtz equation (5), being the Laplace transform of g_ν, is just

$$\begin{aligned}
&G_\nu(P, Q)\\
&= -(2\pi)^{-1}\Big(\int_R^\infty e^{-\gamma\tau} v(\tau, R)\, d\tau + (-1)^\nu \int_{R'}^\infty e^{-\gamma\tau} v(\tau, R')\, d\tau\\
&\quad - \tfrac{1}{2}\int_{\rho+\rho_0}^\infty e^{-\gamma\tau}[v(\tau, R) + (-1)^\nu v(\tau, R')]\, d\tau\Big), \quad \text{in } D_1,\\
&= -(2\pi)^{-1}\Big(\int_R^\infty e^{-\gamma\tau} v(\tau, R)\, d\tau\\
&\quad - \tfrac{1}{2}\int_{\rho+\rho_0}^\infty e^{-\gamma\tau}[v(\tau, R) - (-1)^\nu v(\tau, R')]\, d\tau\Big), \quad \text{in } D_2,\\
&= -(2\pi)^{-1}(\tfrac{1}{2})\int_{\rho+\rho_0}^\infty e^{-\gamma\tau}[v(\tau, R) + (-1)^\nu v(\tau, R')]\, d\tau, \quad \text{in } D_3, \quad (11)
\end{aligned}$$

where the regions D_1, D_2, D_3 are defined as

$$\begin{aligned}
D_1 &= \{Q: \overline{PQ'} \text{ meets the half-plane}\},\\
D_2 &= \{Q: \text{neither } \overline{PQ} \text{ nor } \overline{PQ'} \text{ meets the half-plane}\},\\
D_3 &= \{Q: \overline{PQ} \text{ meets the half-plane}\}, \quad (12)
\end{aligned}$$

and $\nu = 1, 2$ correspond to the Dirichlet and Neumann boundary conditions, respectively.

III. GREEN'S FUNCTION FOR THE SLIT

A. Construction

In this construction of the solution of the wave equation in the presence of a plane with an infinite slit, one first constructs the function ${}_0u_\nu(t, P)$ which is the solution for times t small enough that the two half-planes have not yet had time to interact; that is, for $t \leq d$, d being the width of the slit. Clearly, for such t, the solution can be written down at once by considering the two half-planes separately. Thus, if one uses the same sort of notation as before, with ρ, ρ_0 denoting distances from one edge and σ, σ_0 distances from the other edge, and defines regions

$$\begin{aligned}
\mathfrak{D}_1 &= \{(T, Q): \rho + \rho_0 < t - T\},\\
\mathfrak{D}_2 &= \{(T, Q): \sigma + \sigma_0 < t - T\},\\
\mathfrak{D}_r &= \{(T, Q) \notin \mathfrak{D}_1 \cup \mathfrak{D}_2\text{: line } \overline{PQ'} \text{ meets the slit plane and } R' < t - T\},\\
\mathfrak{D}_i &= \{(T, Q) \notin \mathfrak{D}_1 \cup \mathfrak{D}_2 \cup \mathfrak{D}_r\text{: line } \overline{PQ} \text{ does not meet the slit plane and } R < t - T\}, \quad (13)
\end{aligned}$$

and also defines

$$\begin{aligned}
&{}_0g_\nu(t - T, P, Q)\\
&= -(2\pi)^{-1}v, && \text{for } (T, Q) \in \mathfrak{D}_i,\\
&= -(2\pi)^{-1}[v + (-1)^\nu v^*], && \text{for } (T, Q) \in \mathfrak{D}_r,\\
&= -(2\pi)^{-1}\{\tfrac{1}{2}[v + (-1)^\nu v^*]\}, && \text{for } (T, Q) \in \mathfrak{D}_1 \text{ or } \mathfrak{D}_2, \text{ but } \notin \mathfrak{D}_1 \cap \mathfrak{D}_2,\\
&= -(2\pi)^{-1}(-1)^\nu v^*, && \text{for } (T, Q) \in \mathfrak{D}_1 \cap \mathfrak{D}_2, \quad (14)
\end{aligned}$$

then one has

$${}_0u_\nu(t, P) = \iiint f(T, Q)\, {}_0g_\nu(t - T, P, Q)\, dT\, dQ. \quad (15)$$

Verification that (15) satisfies the differential equation (1) involves the same computations as for the case of the half-plane and are accordingly omitted here. To verify that the boundary condition is satisfied for $0 \leq t < d$, d the width of the slit, but not necessarily for $t \geq d$, one notes that, on the plane, $R = R'$, $v = v^*$, $\partial v/\partial n = -\partial v^*/\partial n$, and $\mathfrak{D}_1 \cap \mathfrak{D}_2$ is empty for all points P on the plane if and only if $0 \leq t < d$.

For $t \geq d$, however, the intersection $\mathfrak{D}_1 \cap \mathfrak{D}_2$ is not empty for all points P on the plane, implying that

for such t the function ${}_0u_\nu$ does not satisfy the desired boundary condition. It follows that what one needs to do is to construct a function ${}_1u_\nu(t, P)$ which satisfies the homogeneous wave equation

$$\left(\nabla^2 - \frac{\partial^2}{\partial t^2}\right){}_1u_\nu = 0, \tag{16}$$

and which takes on the negative of the values of ${}_0u_\nu$ on the slit plane (or has normal derivative there, which is the negative of the normal derivative of ${}_0u_2$), in which case the sum

$$u_\nu = {}_0u_\nu + {}_1u_\nu \tag{17}$$

will be the solution of the problem.

Such a function ${}_1u_\nu$ can indeed be constructed if one uses the results (9) and (10). (One needs two such boundary integrals, one for each half-plane.)

For the sake of definiteness, take rectangular coordinates (x, y) for P, (X, Y) for Q, and suppose the slit is the set of points $y = 0$, $-\frac{1}{2}d \le x \le \frac{1}{2}d$. Then, for $t \ge d$, the value of ${}_0u_\nu(t, x, y)$ on the slit plane is

$${}_0u_\nu(t, x, 0) = \tfrac{1}{2}(-1)^{\nu+1}\pi^{-1}\iiint f(T, X, Y) \times [(t - T)^2 - (x - X)^2 - Y^2]^{-\frac{1}{2}}\, dT\, dX\, dY, \tag{18}$$

with T in the range

$$0 \le T < t - x - \tfrac{1}{2}d - [(X + \tfrac{1}{2}d)^2 + Y^2]^{\frac{1}{2}}, \quad \text{for } x \ge \tfrac{1}{2}d,$$

and

$$0 \le T < t + x - \tfrac{1}{2}d - [(X - \tfrac{1}{2}d)^2 + Y^2]^{\frac{1}{2}}, \quad \text{for } x \le -\tfrac{1}{2}d.$$

Taking first the case $\nu = 1$, in view of (9), one puts

$${}_1u_1(t, x, y) = -\frac{\partial}{\partial y}\pi^{-1}\iint {}_0u_1(T, X, 0) \times [(t - T)^2 - (x - X)^2 - y^2]^{-\frac{1}{2}}\, dT\, dX \tag{19}$$

over the range

$$t - X + \tfrac{1}{2}d - [(x - \tfrac{1}{2}d)^2 + y^2]^{\frac{1}{2}} < T < t - [(x - X)^2 + y^2]^{\frac{1}{2}}, \quad X \ge \tfrac{1}{2}d,$$

$$t + X + \tfrac{1}{2}d - [(x + \tfrac{1}{2}d)^2 + y^2]^{\frac{1}{2}} < T < t - [(x - X)^2 + y^2]^{\frac{1}{2}}, \quad X \le -\tfrac{1}{2}d,$$

then

$${}_1u_1(t, x, y) = -\frac{1}{2\pi^2}\frac{\partial}{\partial y}\iint\left(\iiint f(T, X, Y) \times [(t' - T)^2 - (x' - X)^2 - Y^2]^{-\frac{1}{2}} \times [(t - t')^2 - (x - x')^2 - y^2]^{-\frac{1}{2}}\, dT\, dX\, dY\right) dt'\, dx'.$$

Here, the range of integration is

$$t - x' + \tfrac{1}{2}d - [(x - \tfrac{1}{2}d)^2 + y^2]^{\frac{1}{2}} < t' < t - [(x - x')^2 + y^2]^{\frac{1}{2}}, \quad x' \ge \tfrac{1}{2}d,$$

and

$$t + x' + \tfrac{1}{2}d - [(x + \tfrac{1}{2}d)^2 + y^2]^{\frac{1}{2}} < t' < t - [(x - x')^2 + y^2]^{\frac{1}{2}}, \quad x' \le -\tfrac{1}{2}d.$$

Upon interchanging the order of integration, this can be written in the form

$${}_1u_1(t, x, y) = \iiint f(T, X, Y)\, {}_1g_1(t - T, x, y, X, Y)\, dT\, dX\, dY \tag{20}$$

over

$$0 \le T < t - d - [(x - \tfrac{1}{2}d)^2 + y^2]^{\frac{1}{2}} - [(X + \tfrac{1}{2}d)^2 + Y^2]^{\frac{1}{2}}$$

$$0 \le T < t - d - [(x + \tfrac{1}{2}d)^2 + y^2]^{\frac{1}{2}} - [(X - \tfrac{1}{2}d)^2 + Y^2]^{\frac{1}{2}},$$

where ${}_1g_1$ is the function which must be added to ${}_0g_1$ to give the Green's function g_1 for the slit problem with the Dirichlet boundary condition. In the same way, one finds that ${}_1g_2$ can be written in the form

$$\begin{aligned} {}_1g_2(t - T, x, y, X, Y) &= -(2\pi)^{-1}w(t - T, x, y, X, Y), \\ &\quad \text{in } 0 \le T < t - d - [(x - \tfrac{1}{2}d)^2 + y^2]^{\frac{1}{2}} - [(X + \tfrac{1}{2}d)^2 + Y^2]^{\frac{1}{2}}, \\ &= -(2\pi)^{-1}w(t - T, -x, y, -X, Y), \\ &\quad \text{in } 0 \le T < t - d - [(x + \tfrac{1}{2}d)^2 + y^2]^{\frac{1}{2}} - [(X - \tfrac{1}{2}d)^2 + Y^2]^{\frac{1}{2}}, \end{aligned} \tag{21}$$

where

$$w(t - T, x, y, X, Y) = -\epsilon\pi^{-1}\iint[(t' - t)^2 - (x' - x)^2 - y^2]^{-\frac{1}{2}} \times [(t' - T)^2 - (x' - X)^2 - Y^2]^{-\frac{1}{2}}\, dt'\, dx' \tag{22}$$

over

$$x' \ge \tfrac{1}{2}d,$$

$$T + x' + \tfrac{1}{2}d + [(X + \tfrac{1}{2}d)^2 + Y^2]^{\frac{1}{2}} < t' < t - [(x' - x)^2 + y^2]^{\frac{1}{2}},$$

$$t - x' + \tfrac{1}{2}d - [(x - \tfrac{1}{2}d)^2 + y^2]^{\frac{1}{2}} < t' < t - [(x' - x)^2 + y^2]^{\frac{1}{2}}.$$

The function ${}_1g_1$ turns out to be just the derivative of ${}_1g_2$:

$${}_1g_1(t - T, x, y, X, Y) = \frac{\partial}{\partial y}\,{}_1g_2(t - T, x, y, X, Y). \tag{23}$$

B. Verification

To verify that the function $g_\nu(t-T,P,Q)$ is correct, one can substitute in (4) and show that the resulting function $u_\nu(t,P)$ does indeed satisfy the wave equation (1) and the given boundary condition. Since this verification is lengthy,[10] only the part different from the half-plane problem will be done here, namely the part involving the function w, which is the term called ${}_1u_\nu(t,P)$ and given by Eq. (20).

In view of the definition (22) of $w(t-T,x,y,X,Y)$, one sees that w vanishes identically, and so does each of its derivatives, whenever the arguments are such that

$$T \geq t - d - [(x-\tfrac{1}{2}d)^2+y^2]^{\frac{1}{2}} - [(X+\tfrac{1}{2}d)^2+Y^2]^{\frac{1}{2}},$$

since then the range of integration is an empty set. Furthermore, as $y \to 0^{\pm}$, one has

$$\frac{\partial}{\partial y} w(t-T,x,y,X,Y) \to [(t-T)^2-(x-X)^2-Y^2]^{-\frac{1}{2}}, \quad \text{if } x \geq \tfrac{1}{2}d \text{ or if } x \leq -\tfrac{1}{2}d, \qquad \to 0, \quad \text{if } -\tfrac{1}{2}d < x < \tfrac{1}{2}d, \tag{24}$$

when

$$T \leq t - d - [(x-\tfrac{1}{2}d)^2+y^2]^{\frac{1}{2}} - [(X+\tfrac{1}{2}d)^2+Y^2]^{\frac{1}{2}}.$$

To see this, note that if (for $|x| \geq \frac{1}{2}d$) the integral (20) is split into two integrals, one over the set of (t',x') for which $(t'-t)^2-(x'-x)^2-y^2 \geq \delta > 0$, δ an arbitrarily small positive number, and the other over the remaining (t',x'), then $\partial/\partial y$ can be carried under the integral sign in the first integral and the limit as $y \to 0$ is clearly zero. The second integral can be further subdivided into two, one being over the (t',x') for which $(x'-x) \geq \delta$, and the other being the remainder. After integration by parts with respect to x' in the first integral, differentiation with respect to y can be carried out without difficulty and the limit as y tends to zero is found to be zero.

As for the remaining integral, since it involves only those points (t',x') for which

$$(t'-t)^2+(x'-x)^2 \leq \delta,$$

δ an arbitrarily small positive number, the factor $[(t'-T)^2-(x'-X)^2-Y^2]^{-\frac{1}{2}}$ can be replaced by its value at $t'=t$ and $x'=x$ without changing the value of the limit as y tends to zero. Thus, proceeding along these lines, one sees that (24) holds if it can be shown that

$$\lim_{y\to 0+} \frac{2}{\pi y}\left(\int_0^{(\delta^2-y^2)^{\frac{1}{2}}}\int_{(x^2+y^2)^{\frac{1}{2}}}^{\delta}(t^2-x^2-y^2)^{-\frac{1}{2}}\,dt\,dx - \int_0^{\delta}\int_x^{\delta}(t^2-x^2)^{-\frac{1}{2}}\,dt\,dx\right) = -1.$$

The two inner integrals in the left-hand side of this expression can be carried out easily, while the outer integrals of the result can, by integration by parts, be put in a known form; that is, the entire expression becomes

$$\lim_{y\to 0+}\frac{2\delta}{\pi y}\left(\int_0^{(\delta^2-y^2)^{\frac{1}{2}}}\frac{dx}{(\delta^2-x^2-y^2)^{\frac{1}{2}}} - \int_0^{\delta}\frac{dx}{(\delta^2-x^2)^{\frac{1}{2}}} - y^2\int_0^{(\delta^2-y^2)^{\frac{1}{2}}}\frac{dx}{(x^2+y^2)(\delta^2-x^2-y^2)^{\frac{1}{2}}}\right)$$
$$= \lim_{y\to 0+}\frac{2\delta}{\pi y}\left[\left(\arcsin\frac{x}{(\delta^2-y^2)^{\frac{1}{2}}}\right)_0^{(\delta^2-y^2)^{\frac{1}{2}}} - \left(\frac{y}{\delta}\arctan\frac{x\delta}{y}\right)_0^{\infty} - \left(\arcsin\frac{x}{\delta}\right)_0^{\delta}\right] = -1.$$

The verification that w is indeed a solution of the homogeneous wave equation is exactly the same as the verification for the boundary term (9) or (10) in the case of the half-plane problem and so will not be repeated here. Finally, since w and its derivatives vanish on the surface

$$t = T - d - [(x-\tfrac{1}{2}d)^2+y^2]^{\frac{1}{2}} - [(X+\tfrac{1}{2}d)^2+Y^2]^{\frac{1}{2}},$$

it follows that the function ${}_1u_\nu(t,P)$ also satisfies the homogeneous wave equation and thus the necessary verification is complete.

C. Green's Function for the Modified Helmholtz Equation for the Slit Plane

One obtains the Green's function $G_\nu(P,Q)$ for the modified Helmholtz equation for the case of a slit plane by simply taking the Laplace transform of $g_\nu(\tau,P,Q)$. With regions D_1, D_2, D_3 defined again by (12), but with "slit plane" replacing "half-plane," one finds that

$$G_\nu(P,Q) = -(2\pi)^{-1}\left(\int_R^{\infty} e^{-\gamma\tau}v(\tau,R)\,d\tau + (-1)^\nu\int_{R'}^{\infty}e^{-\gamma\tau}v(\tau,R')\,d\tau - \tfrac{1}{2}\int_{\rho+\rho_0}^{\infty}e^{-\gamma\tau}[v(\tau,R)+(-1)^\nu v(\tau,R')]\,d\tau - \tfrac{1}{2}\int_{\sigma+\sigma_0}^{\infty}e^{-\gamma\tau}[v(\tau,R)-(-1)^\nu v(\tau,R')]\,d\tau\right) + Z, \quad \text{in } D_1,$$

[10] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Co., New York, 1953), Vol. 1, Chap. 7.

$$= -(2\pi)^{-1}\Big(\int_R^\infty e^{-\gamma\tau}v(\tau, R)\,d\tau$$
$$- \tfrac{1}{2}\int_{\rho+\rho_0}^\infty e^{-\gamma\tau}[v(\tau, R) - (-1)^\nu v(\tau, R')]\,d\tau$$
$$- \tfrac{1}{2}\int_{\sigma+\sigma_0}^\infty e^{-\gamma\tau}[v(\tau, R) - (-1)^\nu v(\tau, R')]\,d\tau\Big) + Z,$$
$$\text{in } D_2,$$

$$= \tfrac{1}{2}[-(2\pi)^{-1}]\Big(\int_{\rho+\rho_0}^\infty e^{-\gamma\tau}[v(\tau, R) + (-1)^\nu v(\tau, R')]\,d\tau$$
$$- \tfrac{1}{2}\int_{\sigma+\sigma_0}^\infty e^{-\gamma\tau}[v(\tau, R) - (-1)^\nu v(\tau, R)]\,d\tau\Big) + Z,$$
$$\text{in } D_3, \quad (25)$$

where

$$Z = -\frac{1}{2\pi}\Big[\int_{\rho+\sigma_0+d}^\infty e^{-\gamma\tau}\Big(\frac{\partial}{\partial y}\Big)^{2-\nu} w(\tau)\,d\tau + \int_{\sigma+\rho_0+d}^\infty e^{-\gamma\tau}\Big(\frac{\partial}{\partial y}\Big)^{2-\nu}\tilde{w}(\tau)\,d\tau\Big]$$

and where ρ, ρ_0 refer to the edge closest to P and σ, σ_0 to the other edge, w, $\tilde{w}$ denote the functions appearing in the upper, lower lines of (21).

IV. GREEN'S FUNCTION FOR THE STRIP

Let regions $\mathfrak{D}_i, \mathfrak{D}_r, \mathfrak{D}_1, \mathfrak{D}_2$ be defined again by (13), but with "strip" replacing "half-plane." Then the function ${}_0u_\nu(t, P)$, which is the solution for times t small enough that the two edges of the strip have not yet had time to interact ($t < d$), will again be given by (15) if we define ${}_0g_\nu(t - T, P, Q)$ this time by

$${}_0g_\nu(t - T, P, Q)$$
$$= -(2\pi)^{-1}v, \qquad \text{for } (T, Q) \in \mathfrak{D}_i,$$
$$= -(2\pi)^{-1}[v + (-1)^\nu v^*], \qquad \text{for } (T, Q) \in \mathfrak{D}_r,$$
$$= -(2\pi)^{-1}\tfrac{1}{2}[v + (-1)^\nu v^*], \qquad \text{for } (T, Q) \in \mathfrak{D}_1 \text{ or } \mathfrak{D}_2 \text{ but not in } \mathfrak{D}_1 \cap \mathfrak{D}_2,$$
$$= 0, \qquad \text{for } (T, Q) \in \mathfrak{D}_1 \cap \mathfrak{D}_2 \text{ with } P \text{ and } Q \text{ on the same side of the strip,}$$
$$= -(2\pi)^{-1}[v + (-1)^\nu v^*], \qquad \text{for } (T, Q) \in \mathfrak{D}_1 \cap \mathfrak{D}_2 \text{ with } P \text{ and } Q \text{ on opposite sides of the strip,} \quad (26)$$

in place of (14).

In this case one sees that, although ${}_0u_\nu$ satisfies the boundary condition for *all* times $t \geq 0$ (because ${}_0g_\nu$ does), it fails to satisfy the wave equation (1) in the plane of the strip. Of course, it is not required that (1) be satisfied on the strip itself, but it is required elsewhere in the plane of the strip and it is clear that ${}_0u_\nu$ will fail in this respect just as soon as $\mathfrak{D}_1 \cap \mathfrak{D}_2$ contains one of the edges of the strip, which occurs when $t = d$. However, if one recalls that ${}_1g_\nu(t - T, P, Q)$, defined by (21) and (23), also satisfies the (homogeneous) wave equation (16) everywhere except in the plane outside the strip and is zero and has zero derivatives on the strip itself, it is not difficult to see that the difference

$$g_\nu = {}_0g_\nu - {}_1g_\nu \quad (27)$$

is the desired Green's function for the strip; i.e., u_ν, defined by (4) with g_ν defined by (27), satisfies both the wave equation (1) and the boundary condition for *all* times $t \geq 0$.

The corresponding Green's function for the modified Helmholtz equation, found by taking Laplace transforms, is

$$G_\nu(P, Q)$$
$$= -(2\pi)^{-1}\Big(\int_R^\infty e^{-\gamma\tau}v\,d\tau + (-1)^\nu\int_{R'}^\infty e^{-\gamma\tau}v^*\,d\tau$$
$$- \tfrac{1}{2}\int_{\rho+\rho_0}^\infty e^{-\gamma\tau}[v + (-1)^\nu v^*]\,d\tau$$
$$- \tfrac{1}{2}\int_{\sigma+\sigma_0}^\infty e^{-\gamma\tau}[v + (-1)^\nu v^*]\,d\tau\Big) + Z, \quad \text{in } D_1,$$
$$= -(2\pi)^{-1}\Big(\int_R^\infty e^{-\gamma\tau}v\,d\tau$$
$$- \tfrac{1}{2}\int_{\rho+\rho_0}^\infty e^{-\gamma\tau}[v - (-1)^\nu v^*]\,d\tau$$
$$- \tfrac{1}{2}\epsilon\int_{\sigma+\sigma_0}^\infty e^{-\gamma\tau}[v + (-1)^\nu v^*]\,d\tau\Big) + Z, \quad \text{in } D_2,$$
$$= \tfrac{1}{2}[-(2\pi)^{-1}]\Big(\int_{\rho+\rho_0}^\infty e^{-\gamma\tau}[v + (-1)^\nu v^*]\,d\tau$$
$$+ \int_{\sigma+\sigma_0}^\infty e^{-\gamma\tau}[v + (-1)^\nu v^*]\,d\tau\Big) + Z, \quad \text{in } D_3, \quad (28)$$

where Z is the same as in Eq. (25) and where, as in (25), $v(\tau, R)$ and $v(\tau, R')$ have been abbreviated as just v and v^*, ρ and ρ_0 refer to the edge closest to P, ϵ is ± 1 according to whether P and Q are on the same or opposite sides of the strip, and D_1, D_2, D_3 are defined by (12) with "strip" in place of "half-plane."

V. DIFFRACTION OF PLANE WAVES

A. Half-Plane Diffraction

The usual Sommerfeld solution of the problem of diffraction of a time-harmonic plane wave incident upon a half-plane, when the direction of propagation lies in a plane perpendicular to the edge, can be obtained from (11) by a suitable limiting process. Of

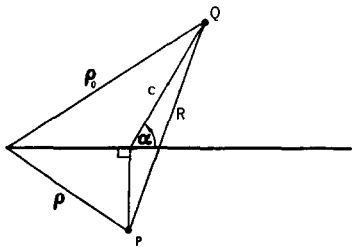


FIG. 4. A line source Q and observation point P are located at distances ρ_0 and ρ, respectively, from the edge of a half-plane, and are a distance R apart. The polar coordinates of Q with respect to the plane and the foot of the perpendicular from P are c and α, respectively.

course, the diffraction of such a plane wave by a slit plane or strip can also be obtained by the very same limiting process.

The first term in the first line on the right side of (11) is the "source" term, the one which becomes infinite as P approaches Q ($R \to 0$). This term, which could have been written as $K_0(\gamma R)$, since

$$K_0(\gamma R) = \int_R^\infty e^{-\gamma\tau}(\tau^2 - R^2)^{-\frac{1}{2}}\, d\tau, \qquad (29)$$

is, of course, the Green's function for (5) in free space. The asymptotic behavior of $K_0(ikR)$, k real, is

$$K_0(ikR) \sim (\pi/2ikR)^{\frac{1}{2}} e^{-ikR}, \quad \text{as} \quad R \to \infty. \qquad (30)$$

Consequently, to obtain the solution of the half-plane diffraction problem corresponding to a plane wave which at infinity had a direction of propagation making an angle α with the plane, we must put $\gamma = ik$, multiply (11) by a factor

$$A = (2ikc/\pi)^{\frac{1}{2}} e^{ikc},$$

where c is the distance of Q from the plane along the direction α, and then determine the limit as $c \to \infty$. For convenience in computation, we suppose Q has polar coordinates (c, α) relative to the foot of the perpendicular from P to the plane (see Fig. 4).

The first two terms in (11), after multiplication by A, tend to $e^{ik\tilde{R}}$ and $(-1)^\nu e^{ik\tilde{R}'}$, respectively, as $c \to \infty$, where

$$\tilde{R} = -y \sin\alpha \quad \text{and} \quad \tilde{R}' = y \sin\alpha.$$

To obtain Sommerfeld's result from the term

$$(-2\pi i)^{-\frac{1}{2}} \int_{\rho+\rho_0}^\infty e^{-ik(\tau-c)} \left(\frac{kc}{\tau^2 - R^2}\right)^{\frac{1}{2}} d\tau \qquad (31)$$

in the manner of Born and Wolf,[8] make the change of variable $\tau = R + \mu^2/|k|$:

$$(-2\pi i)^{-\frac{1}{2}} \int_{[|k|(\rho+\rho_0-R)]^{\frac{1}{2}}}^\infty \exp\left(-i\frac{k}{|k|}\mu^2 + ik(c-R)\right) \times \left(\frac{2kc}{R|k|}\right)^{\frac{1}{2}} \left(1 + \frac{\mu^2}{2R|k|}\right)^{-\frac{1}{2}} d\mu \qquad (32)$$

and then let $c \to \infty$. Since $c/R \to 1$ and

$$(1 + \mu^2/2R|k|)^{-\frac{1}{2}}$$

converges monotonically to unity for each fixed μ, the limit of (31) is

$$\left(\frac{-k}{\pi i|k|}\right)^{\frac{1}{2}} \exp(iky\sin\alpha) \times \int_{|k|(\rho+x\cos\alpha+y\sin\alpha)}^\infty \exp\left(-i\frac{k}{|k|}\mu^2\right) d\mu.$$

Alternatively, one can make the change of variable $\tau = \tau' + R - \tilde{R}$, $\tilde{R} = -y\sin\alpha$. Then (31) has the limit, as $c \to \infty$,

$$\frac{1}{2}\left(\frac{k}{\pi}\right)^{\frac{1}{2}} \int_{\rho+\tilde{\rho}_0}^\infty e^{-ik\tau'} \frac{d\tau'}{(\tau' - \tilde{R})^{\frac{1}{2}}},$$

with $\tilde{\rho}_0 = |x|\cos\alpha$, which is related to the usual Fresnel integral.[8]

Obviously, the terms involving R' can be treated in exactly the same way.

B. Diffraction by Slit or Strip

For the slit or strip, one sees that the terms in (25) and (28) are either of the same form as in the half-plane problem already treated or they involve the function denoted by w. To deal with these terms, write w in the form

$$w(\tau, x, y, X, Y) = \epsilon\pi^{-1} \iint [t'^2 - x'^2 - y^2]^{-\frac{1}{2}} \times [(t' - \tau)^2 - (x' + x - X)^2 - Y^2]^{-\frac{1}{2}}\, dt'\, dx',$$

by changing the integration variables. Put $\gamma = ik$, multiply the term in (25) by the factor A (Eq. 30), and make the change of variable $\tau = \tau' + R - \tilde{R}$ just as before, so that the required limit is the limit of

$$(kc)^{\frac{1}{2}} w(\tau' + R - \tilde{R}) = \epsilon\pi^{-1} \iint (t'^2 - x'^2 - y^2)^{-\frac{1}{2}} (kc)^{\frac{1}{2}} \times [(t' - \tau' - R + \tilde{R})^2 - (x' - c\cos\alpha)^2 - c^2\sin^2\alpha]^{-\frac{1}{2}}\, dt'\, dx',$$

where the integration is over the range

$$(x'^2 + y^2)^{\frac{1}{2}} < t' < \tau' + R - \tilde{R} - x' - x - \tfrac{1}{2}d - [(x + \tfrac{1}{2}d + c\cos\alpha)^2 + c^2\sin^2\alpha]^{\frac{1}{2}},$$
$$(x'^2 + y^2)^{\frac{1}{2}} < t' < x + x' - \tfrac{1}{2}d + [(x - \tfrac{1}{2}d)^2 + y^2]^{\frac{1}{2}}.$$

The desired limit is

$$\epsilon\pi^{-1}(\tfrac{1}{2}k)^{\frac{1}{2}} \iint (t'^2 - x'^2 - y^2)^{-\frac{1}{2}} \times (\tau' - t' + x'\cos\alpha)^{-\frac{1}{2}}\, dt'\, dx' \qquad (33)$$

over

$$(x'^2 + y^2)^{\frac{1}{2}} < t' < \tau' - x' - (x + \tfrac{1}{2}d)(1 + \cos\alpha),$$
$$(x'^2 + y^2)^{\frac{1}{2}} < t' < x + x' - \tfrac{1}{2}d + [(x - \tfrac{1}{2}d)^2 + y^2]^{\frac{1}{2}}.$$

One can then assemble the solutions for the perfectly reflecting slit plane for two-dimensional plane-wave excitation. The results may then be compared with those obtained by Karp and Russek[6] for the same problem attacked by a different method.

VI. ASYMPTOTIC BEHAVIOR OF DIFFRACTED TIME-HARMONIC WAVES

One wishes to consider the asymptotic behavior of the solutions of the various diffraction problems corresponding to time-harmonic waves—cylindrical or plane—as the wavenumber k becomes large. The case of a plane wave incident on a half-plane was treated by Sommerfeld[2] and the corresponding case for a plane with a slit has been treated by various authors.[6]

Upon examining the form of the solutions for the slit (25) and for the strip (28), one sees three types of integrals:

$$\int_R^\infty e^{-ik\tau} v(\tau, R)\, d\tau,$$

which is a Hankel function and poses no problem as $k \to \infty$,

$$\int_b^\infty e^{-ik\tau} \left(\frac{\partial}{\partial n}\right)^{2-\nu} w(\tau)\, d\tau, \tag{34}$$

which has no singularities in the domain of interest as $k \to \infty$ and so offers no difficulties, and

$$\int_a^\infty e^{-ik\tau}(\tau^2 - R^2)^{-\frac{1}{2}}\, d\tau, \quad \text{as} \quad k \to \infty. \tag{35}$$

Since the values of $a = \rho + \rho_0$ and R, corresponding to any point P, are equal if and only if the point P is on one of the "shadow lines" (i.e., on one of the half-lines extending from an edge in the direction away from the source Q or its image Q'), the asymptotic behavior of this last integral must be expected to be uniform with respect to a and R in any region bounded away from the half-line $R - a = 0$, but not otherwise. Therefore, it is necessary to consider this special case.

If $a - R \geq \delta > 0$, this integral, which is a Fourier integral of the form

$$\int_\alpha^\infty e^{ix\tau} \varphi(\tau)\, d\tau,$$

where $\varphi(\tau)$ is N times differentiable, admits the expansion[11]

$$\int_a^\infty e^{ik\tau}(\tau^2 - R^2)^{-\frac{1}{2}}\, d\tau = -\sum_{n=0}^{N-1} i^{n-1}\varphi^{(n)}(a)k^{-n-1}e^{ika} + o(k^{-N}), \tag{36}$$

where $\varphi(\tau) = (\tau^2 - R^2)^{-\frac{1}{2}}$. Of course this type of expansion also applies in the case of the integral (34) involving $w(\tau)$.

If, on the other hand, $a - R$ is too small for (36) to be practical, one can make a change of variable most useful for $a - R$ small. That is, make the change of variable $\tau = [R^2 + s^2(a^2 - R^2)]^{\frac{1}{2}}$ to obtain

$$\frac{1}{\pi R} e^{ikR} \int_1^\infty \frac{\exp -ik\{1 + s^2[(a^2 - R^2)/R^2]\}^{\frac{1}{2}}}{s(s^2 - 1)^{\frac{1}{2}}\{1 + s^2[(a^2 - R^2)/R^2]\}^{\frac{1}{2}}}\, ds.$$

The integral converges, for any fixed k, not necessarily large, to the value $\frac{1}{2}\pi$ as $a - R \to 0$.

[11] A. Erdélyi, *Asymptotic Expansions* (Dover Publications, Inc., New York, 1956).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Soluble Model of Condensation

M. D. GIRARDEAU
Institute of Theoretical Science and Department of Physics
University of Oregon, Eugene, Oregon

(Received 15 October 1968)

A soluble model of gas–liquid condensation is constructed by a quantum-mechanical extension of van Kampen's classical analysis. The Hamiltonian of a system of identical particles is separated into a "core" and a "tail" Hamiltonian. Following van Kampen, we simulate the actual core Hamiltonian by an effective Hamiltonian with the same free energy. The tail Hamiltonian is replaced by its diagonal part in a single-particle basis chosen self-consistently so that the thermodynamics is exactly soluble. The resultant free energy agrees with that of Lieb's quantum-mechanical extension of the van der Waals–Maxwell theory, without the necessity of taking the van der Waals limit of a suitably infinite-ranged and infinitely-weak attractive tail. However, the exchange contributions omitted from the model become negligible only in that limit.

1. INTRODUCTION

It has been shown by a number of authors[1–7] how the van der Waals–Maxwell theory of condensation of a classical gas follows from a model with hard cores plus a suitably infinitely-long-ranged and infinitely-weak attractive tail, and the analysis has been extended to the quantum-mechanical case by Lieb.[8] In this paper we shall show how an exactly-soluble model of condensation can be constructed by a suitable quantum-mechanical extension of van Kampen's classical analysis.[6] The mathematical method employed is Wentzel's "thermodynamically-equivalent Hamiltonian" (TEH) method,[9] which gives the exact free-energy density of the model in the thermodynamic limit (volume $\to \infty$ for fixed density or chemical potential), without the necessity of taking the van der Waals limit. The exchange contributions omitted from the model are, however, shown to be negligible only in that limit.

2. FORMULATION

The quantum-mechanical Hamiltonian of a system of identical particles interacting by a pair potential $v(|\mathbf{r} - \mathbf{r}'|)$ is

$$H = \int \psi^\dagger(\mathbf{r})\left(-\frac{\hbar^2}{2m}\nabla^2\right)\psi(\mathbf{r})\, d^3r + \tfrac{1}{2}\int \psi^\dagger(\mathbf{r})\psi^\dagger(\mathbf{r}')v(|\mathbf{r} - \mathbf{r}'|)\psi(\mathbf{r}')\psi(\mathbf{r})\, d^3r\, d^3r', \quad (1)$$

where $\psi(\mathbf{r})$ and $\psi^\dagger(\mathbf{r})$ are annihilation and creation operators for fermions or bosons at the point $\mathbf{r}$, satisfying canonical anticommutation or commutation relations

$$[\psi(\mathbf{r}), \psi^\dagger(\mathbf{r}')]_\pm = \delta(\mathbf{r} - \mathbf{r}'), \quad [\psi(\mathbf{r}), \psi(\mathbf{r}')]_\pm = 0. \quad (2)$$

The integrals in (1), as well as all subsequent integrals, run over the finite volume Ω in which the particles are contained.

Exact solution of the equilibrium statistical mechanics of the Hamiltonian (1) is, of course, out of the question. The manner in which (1) can be simplified so as to obtain an exactly soluble model is suggested by the TEH method,[9] which is closely related to the variational method.[10] The variational method involves replacement of the interaction Hamiltonian by its diagonal part in an appropriate representation. If the interaction v contains a hard core, this procedure leads to divergences. This difficulty can be circumvented by including only the attractive tail in the interaction Hamiltonian, simulating the hard cores by an effective hard-core Hamiltonian $H_{\text{h.c.}}$ suggested by van Kampen's classical theory.[6] Thus we replace the interaction Hamiltonian in (1) by $H_{\text{h.c.}} - W$, where

$$W = \tfrac{1}{2}\int \psi^\dagger(\mathbf{r})\psi^\dagger(\mathbf{r}')w(|\mathbf{r} - \mathbf{r}'|)\psi(\mathbf{r}')\psi(\mathbf{r})\, d^3r\, d^3r' \quad (3)$$

and $-w$ is the finite tail of the interaction, which we assume to be integrable and attractive on the average:

$$0 < w_0 \equiv \int w(r)\, d^3r < \infty. \quad (4)$$

Although finite, W still cannot be treated exactly, so that further simplifications must be made. We first reorder the annihilation and creation operators

[1] M. Kac, Phys. Fluids **2**, 8 (1959).
[2] G. Baker, Phys. Rev. **122**, 1477 (1961); **126**, 2071 (1962).
[3] M. Kac, G. Uhlenbeck, and P. Hemmer, J. Math. Phys. **4**, 216 (1963); **4**, 229 (1963); **5**, 60 (1964).
[4] P. Hemmer, J. Math. Phys. **5**, 75 (1964).
[5] E. Helfand, J. Math. Phys. **5**, 127 (1964).
[6] N. G. van Kampen, Phys. Rev. **135**, A362 (1964).
[7] J. L. Lebowitz and O. Penrose, J. Math. Phys. **7**, 98 (1966).
[8] E. Lieb, J. Math. Phys. **7**, 1016 (1966).
[9] G. Wentzel, Phys. Rev. **120**, 1572 (1960).
[10] M. Girardeau, J. Math. Phys. **3**, 131 (1962).

in W with the aid of (2), finding

$$W = -\tfrac{1}{2}w(0)N + \tfrac{1}{2}\int w(|\mathbf{r}-\mathbf{r}'|)\rho(\mathbf{r})\rho(\mathbf{r}')\,d^3r\,d^3r', \quad (5)$$

where ρ is the number-density operator

$$\rho(\mathbf{r}) = \psi^\dagger(\mathbf{r})\psi(\mathbf{r}), \quad (6)$$

and N the total particle-number operator:

$$N = \int \rho(\mathbf{r})\,d^3r. \quad (7)$$

Let $\{\varphi_\alpha(\mathbf{r})\}$ be any complete orthonormal set of single-particle states. Then

$$\psi(\mathbf{r}) = \sum_\alpha \varphi_\alpha(\mathbf{r})a_\alpha, \quad a_\alpha = \int \varphi_\alpha^*(\mathbf{r})\psi(\mathbf{r})\,d^3r. \quad (8)$$

The diagonal part of ρ in the a_α, $a_\alpha^\dagger$ representation is

$$\rho_{\text{diag.}}(\mathbf{r}) = \sum_\alpha |\varphi_\alpha(\mathbf{r})|^2\, a_\alpha^\dagger a_\alpha. \quad (9)$$

If $\rho(\mathbf{r})\rho(\mathbf{r}')$ is replaced by $\rho_{\text{diag.}}(\mathbf{r})\rho_{\text{diag.}}(\mathbf{r}')$ in W, then, as we shall presently see, the resultant model Hamiltonian is exactly soluble by the TEH method,[9] provided that the φ_α are chosen so that the single-particle part of the Hamiltonian is also diagonal.

A more physical justification for such a model is given by the variational method.[10] This method involves replacing the full Hamiltonian by its diagonal part in any representation. Taking this to be the a_α, $a_\alpha^\dagger$ representation, one finds that the diagonal part of the interaction term $\rho(\mathbf{r})\rho(\mathbf{r}')$ is

$$[\rho(\mathbf{r})\rho(\mathbf{r}')]_{\text{diag.}} = \rho_{\text{diag.}}(\mathbf{r})\rho_{\text{diag.}}(\mathbf{r}') + \sum_{\alpha\beta}{}' \varphi_\alpha^*(\mathbf{r})\varphi_\beta(\mathbf{r})\varphi_\beta^*(\mathbf{r}')\varphi_\alpha(\mathbf{r}')a_\alpha^\dagger a_\alpha a_\beta a_\beta^\dagger, \quad (10)$$

where the prime implies omission of terms with $\alpha = \beta$. The second term on the right of (10), when substituted into (5), gives rise to the exchange part of W:

$$W_{\text{ex.}} = -\tfrac{1}{2}\sum_{\alpha\beta}{}' J_{\alpha\beta}a_\alpha^\dagger a_\alpha a_\beta a_\beta^\dagger, \quad (11)$$

where

$$J_{\alpha\beta} = \int \varphi_\alpha^*(\mathbf{r})\varphi_\beta(\mathbf{r})w(|\mathbf{r}-\mathbf{r}'|)\varphi_\beta^*(\mathbf{r}')\varphi_\alpha(\mathbf{r}')\,d^3r\,d^3r'. \quad (12)$$

We shall see later that $J_{\alpha\beta}$ and hence the contribution of $W_{\text{ex.}}$ to the thermodynamics vanish in the van der Waals limit of infinitely-long-ranged w, due to the orthogonality of φ_α and φ_β for $\alpha \neq \beta$. Thus the variational method leads, in the van der Waals limit, to the replacement of $\rho(\mathbf{r})\rho(\mathbf{r}')$ by $\rho_{\text{diag.}}(\mathbf{r})\rho_{\text{diag.}}(\mathbf{r}')$.

Thus we adopt as our model Hamiltonian

$$H_{\text{mod.}} = \int\left[\psi^\dagger(\mathbf{r})\left(-\frac{\hbar^2}{2m}\nabla^2\right)\psi(\mathbf{r})\right]_{\text{diag.}} d^3r + H_{\text{h.c.}} + \tfrac{1}{2}w(0)N - \tfrac{1}{2}\int w(|\mathbf{r}-\mathbf{r}'|)\rho_{\text{diag.}}(\mathbf{r})\rho_{\text{diag.}}(\mathbf{r}')\,d^3r\,d^3r'. \quad (13)$$

Generalizing van Kampen's classical treatment of the hard cores, we choose[11]

$$H_{\text{h.c.}} = \int f_{\text{h.c.}}(\rho_{\text{diag.}}(\mathbf{r}))\,d^3r, \quad (14)$$

where $f_{\text{h.c.}}(n)$ is the hard-core contribution to the free energy per unit volume,[12] expressed as a function of the mean number density n. Actually, $f_{\text{h.c.}}(n)$ also depends on the hard-core diameter and the temperature, but these additional dependences will not be indicated explicitly since they are not important to our analysis. The detailed functional form of $f_{\text{h.c.}}(n)$ is unknown except at low densities, where it is known for both fermions and bosons. However, we need only certain qualitative properties of $f_{\text{h.c.}}$, in particular the property that it becomes infinite as $n \to n_p$ from below, where n_p is the number density at closest packing.

3. THERMODYNAMICALLY-EQUIVALENT HAMILTONIAN

The grand-canonical[13] thermodynamic potential of our model is

$$F = -\beta^{-1}\ln \operatorname{Tr} e^{-\beta(H_{\text{mod.}}-\mu N)}. \quad (15)$$

In order to evaluate (15), we decompose the Hamiltonian (13) in a manner similar to that used in the TEH method.[9] Thus we decompose the interaction term as follows:

$$\rho_{\text{diag.}}(\mathbf{r})\rho_{\text{diag.}}(\mathbf{r}') = [\rho_{\text{diag.}}(\mathbf{r}) - n(\mathbf{r})][\rho_{\text{diag.}}(\mathbf{r}') - n(\mathbf{r}')] + [n(\mathbf{r})\rho_{\text{diag.}}(\mathbf{r}') + n(\mathbf{r}')\rho_{\text{diag.}}(\mathbf{r})] - n(\mathbf{r})n(\mathbf{r}'), \quad (16)$$

where $n(\mathbf{r})$ is a c-number function to be determined. The operator $f_{\text{h.c.}}(\rho_{\text{diag.}}(\mathbf{r}))$ can be decomposed

[11] N. G. van Kampen, Phys. Rev. **135**, A362 (1964), Appendix I.

[12] More precisely, if $f^0(n)$ is the thermodynamic potential density of a system of hard spheres (no attractive tail) and $f_{\text{ideal}}(n)$ that of an ideal (Fermi or Bose) gas, then $f_{\text{h.c.}}$ is defined by the requirement that

$$f^0(n) = f_{\text{ideal}}(n) + f_{\text{h.c.}}(n).$$

[13] One might suppose that use of the grand ensemble implies that $f_{\text{h.c.}}$ should be the hard-core contribution to F/Ω. However, this is not the case. We shall see later that $f_{\text{h.c.}}$ must be chosen to be the hard-core contribution to the *Helmholtz* free-energy density.

similarly by expanding it in a Taylor series:

$$f_{\text{h.c.}}(\rho_{\text{diag.}}(\mathbf{r})) = f_{\text{h.c.}}(n(\mathbf{r})) + f'_{\text{h.c.}}(n(\mathbf{r}))[\rho_{\text{diag.}}(\mathbf{r}) - n(\mathbf{r})] + \sum_{l=2}^{\infty} (l!)^{-1} f^{(l)}_{\text{h.c.}}(n(\mathbf{r}))[\Delta\rho(\mathbf{r})]^l, \quad (17)$$

where $f'_{\text{h.c.}}$ is the first derivative and $f^{(l)}_{\text{h.c.}}$ the higher derivatives of $f_{\text{h.c.}}(n)$, and

$$\Delta\rho(\mathbf{r}) = \rho_{\text{diag.}}(\mathbf{r}) - n(\mathbf{r}). \quad (18)$$

Substitution of (17) and (18) into (13) leads to the decomposition

$$H_{\text{mod.}} = H_0 + H',$$

$$H_0 = E_0 + \int \left\{\psi^\dagger(\mathbf{r})\left[-\frac{\hbar^2}{2m}\nabla^2 - \bar{w}(\mathbf{r})\right]\psi(\mathbf{r})\right\}_{\text{diag.}} d^3r,$$

$$H' = \sum_{l=2}^{\infty} (l!)^{-1} \int f^{(l)}_{\text{h.c.}}(n(\mathbf{r}))[\Delta\rho(\mathbf{r})]^l \, d^3r \cdot - \tfrac{1}{2}\int w(|\mathbf{r} - \mathbf{r}'|)\Delta\rho(\mathbf{r})\Delta\rho(\mathbf{r}')\, d^3r\, d^3r', \quad (19)$$

with

$$E_0 = \int [f_{\text{h.c.}}(n(\mathbf{r})) - n(\mathbf{r})f'_{\text{h.c.}}(n(\mathbf{r}))]\, d^3r + \tfrac{1}{2}\int w(|\mathbf{r} - \mathbf{r}'|)n(\mathbf{r})n(\mathbf{r}')\, d^3r\, d^3r' \quad (20)$$

and

$$\bar{w}(\mathbf{r}) = -\tfrac{1}{2}w(0) + \int w(|\mathbf{r} - \mathbf{r}'|)n(\mathbf{r}')\, d^3r' - f'_{\text{h.c.}}(n(\mathbf{r})). \quad (21)$$

Let $\{\varphi_\alpha(\mathbf{r})\}$ be the complete orthonormal set of eigenstates of Schrödinger's equation in the potential $-\bar{w}$:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 - \bar{w}(\mathbf{r})\right]\varphi_\alpha(\mathbf{r}) = \epsilon_\alpha\varphi_\alpha(\mathbf{r}). \quad (22)$$

Then H_0 can be written as

$$H_0 = E_0 + \sum_\alpha \epsilon_\alpha a^\dagger_\alpha a_\alpha, \quad (23)$$

where the a_α are defined by (8). The thermodynamic potential of H_0 is then

$$F_0 = -\beta^{-1}\ln \text{Tr}\, e^{-\beta(H_0 - \mu N)} = E_0 \mp \beta^{-1}\sum_\alpha \ln[1 \pm e^{-\beta(\epsilon_\alpha - \mu)}], \quad (24)$$

where

$$N = \sum_\alpha a^\dagger_\alpha a_\alpha. \quad (25)$$

We wish to show that, with the appropriate choice of the function $n(\mathbf{r})$, H' becomes negligible in the thermodynamic limit in the sense that F differs from F_0 only by a thermodynamically negligible term. The "appropriate choice" referred to above is

$$n(\mathbf{r}) = \langle\rho(\mathbf{r})\rangle_0 = \langle\rho_{\text{diag.}}(\mathbf{r})\rangle_0 \equiv \frac{\text{Tr}\,[\rho_{\text{diag.}}(\mathbf{r})e^{-\beta(H_0 - \mu N)}]}{\text{Tr}\, e^{-\beta(H_0 - \mu N)}}, \quad (26)$$

so that all terms in H' are quadratic and higher in the operator $\Delta\rho(\mathbf{r})$ which represents the fluctuation of $\rho_{\text{diag.}}(\mathbf{r})$ about its mean value $n(\mathbf{r})$, which is the same as the mean value of $\rho(\mathbf{r})$. With (9), (22), and (26) one finds

$$n(\mathbf{r}) = \sum_\alpha n_\alpha |\varphi_\alpha(\mathbf{r})|^2, \quad (27)$$

where n_α is the Fermi–Dirac or Bose–Einstein distribution:

$$n_\alpha = [e^{\beta(\epsilon_\alpha - \mu)} \pm 1]^{-1}. \quad (28)$$

It is clear from (20), (21), and (23) that the n_α and φ_α depend implicitly on the function $n(\mathbf{r})$. Thus (27) is really a rather complicated functional equation.

The grand partition function Ξ of the model Hamiltonian (13) and that Ξ_0 of H_0 are given by

$$\Xi = e^{-\beta F} = \text{Tr}\, e^{-\beta(H_{\text{mod.}} - \mu N)}, \qquad \Xi_0 = e^{-\beta F_0} = \text{Tr}\, e^{-\beta(H_0 - \mu N)}. \quad (29)$$

Making use of (13), (14), (19), and (26) and noting from (9) that H_0 and H' commute, one can expand Ξ in terms of thermal averages of powers of H':

$$\Xi = \Xi_0 \sum_{j=0}^{\infty} \frac{(-\beta)^j}{j!}\langle(H')^j\rangle_0. \quad (30)$$

Upon substituting (19) into (30) and making use of the identity

$$\langle A\Delta\rho(\mathbf{r})B\rangle_0 = \langle A\rho_{\text{diag.}}(\mathbf{r})B\rangle'_0, \quad (31)$$

one sees that the various terms in (30) can be written as thermal averages $\langle\ \rangle'_0$ of various powers of $\rho_{\text{diag.}}$, where the prime implies omission of all terms in which any $\rho_{\text{diag.}}(\mathbf{r})$ is self-contracted in the expansion of $\langle\ \rangle_0$ according to Matsubara's theorem.[14] For example, the $j = 1$ term in (30) is

$$\Xi_1 = -\beta\Xi_0\left\{\sum_{l=2}^{\infty}(l!)^{-1}\int f^{(l)}_{\text{h.c.}}(n(\mathbf{r}))\langle[\rho_{\text{diag.}}(\mathbf{r})]^l\rangle'_0\, d^3r - \tfrac{1}{2}\int w(|\mathbf{r} - \mathbf{r}'|)\langle\rho_{\text{diag.}}(\mathbf{r})\rho_{\text{diag.}}(\mathbf{r}')\rangle'_0\, d^3r\, d^3r'\right\}. \quad (32)$$

Using (9), for the thermal averages occurring in the first line of (32), omitting self-contractions of any pair

[14] T. Matsubara, Progr. Theoret. Phys. (Kyoto) **14**, 351 (1955). A complete proof of the theorem, in the form we employ, is given by C. Bloch and C. DeDominicis, Nucl. Phys. **7**, 459 (1958).

$a_\alpha^\dagger a_\alpha$, one finds

$$\langle[\rho_{\text{diag.}}(\mathbf{r})]^2\rangle_0' = \sum_\alpha |\varphi_\alpha(\mathbf{r})|^4 n_\alpha(1 \mp n_\alpha),$$

$$\langle[\rho_{\text{diag.}}(\mathbf{r})]^3\rangle_0' = \sum_\alpha |\varphi_\alpha(\mathbf{r})|^6 [n_\alpha(1 \mp n_\alpha)^2 \mp n_\alpha^2(1 \mp n_\alpha)],$$

$$\cdot \quad \cdot \quad \cdot \tag{33}$$

where n_α is given by (28). We shall presently find that the single-particle functions φ_α are spatially extended and hence proportional, because of normalization, to $\Omega^{-\frac{1}{2}}$, where Ω is the volume of the system. We shall furthermore find that each $\sum_\alpha$ contributes a factor of the volume, since the φ_α are continuum states. Finally, the n_α are volume independent.[15] Hence the thermal average of $\rho_{\text{diag.}}^2$ is $O(\Omega^{-1})$ and that of $\rho_{\text{diag.}}^3$ is $O(\Omega^{-2})$. For $l = 4$, contractions become possible which allow two uncoupled summations over α and β, so that the thermal average of $\rho_{\text{diag.}}^4$ is also $O(\Omega^{-2})$. More generally, one finds that for $\nu \geq 2$, the thermal average of $\rho_{\text{diag.}}^{2\nu}$ is $O(\Omega^{-\nu})$ and that of $\rho_{\text{diag.}}^{2\nu+1}$ is $O(\Omega^{-\nu-1})$. Hence, on performing the volume integration, one sees that the terms involving the $f_{\text{h.c.}}^{(l)}$ give a net contribution of $O(1)$ (volume independent) to the curly bracket in (32). Similarly,

$$\langle\rho_{\text{diag.}}(\mathbf{r})\rho_{\text{diag.}}(\mathbf{r}')\rangle_0' = \sum_\alpha |\varphi_\alpha(\mathbf{r})|^2 |\varphi_\alpha(\mathbf{r}')|^2 n_\alpha(1 \mp n_\alpha) = O(\Omega^{-1}). \tag{34}$$

Since $w(|\mathbf{r} - \mathbf{r}'|)$ is volume independent and of finite range because of the integrability condition (4), one sees on performing the volume integrations that the net contribution of the term involving $w(|\mathbf{r} - \mathbf{r}'|)$ to the curly bracket in (32) is also $O(1)$. Hence the contribution of Ξ_1 to the free energy is negligible $[O(1)]$ in the thermodynamic limit. By straightforward extension of these considerations, one can show that the terms in the summation (30) with $j \geq 2$ are also $O(1)$.[16] Thus the free-energy density F/Ω [Eq. (15)] reduces to F_0/Ω [Eq. (24)] in the thermodynamic limit. In Wentzel's terminology, the Hamiltonian H_0 [Eq. (19)] is "thermodynamically equivalent" to the full model Hamiltonian $H_{\text{mod.}}$, provided that (26) is satisfied. The generalization of the TEH method[9] which we have derived here is similar to that previously carried out in connection with the theory of quasiparticles.[17]

[15] An exception may occur for a Bose–Einstein system below its condensation temperature T_0. In such a case we assume $T > T_0$.

[16] In order that this be true, it is essential that the l summation in (19) begins with $l = 2$, not $l = 1$, i.e., it is essential that the linear term in the expansion of $f_{\text{h.c.}}(\rho(\mathbf{r}))$ has been included in H_0.

[17] M. D. Girardeau, Phys. Rev. **140**, A1139 (1965).

4. SOLUTION IN THE ONE-PHASE REGION

To obtain more explicit results, it is necessary to solve the coupled equations (20), (21), (23), (27), and (28), so as to allow evaluation of E_0 and the ϵ_α. For one-phase states $n(\mathbf{r})$ is independent of $\mathbf{r}$:

$$n(\mathbf{r}) = \text{const} \equiv n \text{ (single phase)}. \tag{35}$$

Then (20) and (4) imply that

$$\bar{w}(\mathbf{r}) = -\tfrac{1}{2}w(0) + nw_0, \tag{36}$$

so that (21) becomes

$$-\frac{\hbar^2}{2m}\nabla^2\varphi_\alpha(\mathbf{r}) = [\lambda_\alpha - \tfrac{1}{2}w(0) + nw_0]\varphi_\alpha(\mathbf{r}). \tag{37}$$

The solutions are plane waves:

$$\alpha = \mathbf{k}, \quad \varphi_{\mathbf{k}}(\mathbf{r}) = \Omega^{-\frac{1}{2}}e^{i\mathbf{k}\cdot\mathbf{r}},$$
$$\lambda_{\mathbf{k}} - \tfrac{1}{2}w(0) + nw_0 = \hbar^2k^2/2m. \tag{38}$$

Then (23) shows that

$$\epsilon_{\mathbf{k}} = (\hbar^2k^2/2m) + \tfrac{1}{2}w(0) - nw_0 + f_{\text{h.c.}}'(n), \tag{39}$$

and (27) becomes, with (28),

$$n = \Omega^{-1}\sum_{\mathbf{k}} [e^{\beta(\epsilon_{\mathbf{k}}-\mu)} \pm 1]^{-1}. \tag{40}$$

In the thermodynamic limit the summation is converted into an integration in the usual way, yielding

$$n = \frac{1}{4\pi^2}\left(\frac{2m}{\beta\hbar^2}\right)^{\frac{3}{2}} I_\pm(\alpha), \tag{41}$$

where

$$I_\pm(\alpha) = \int_0^\infty \frac{y^{\frac{1}{2}}\,dy}{e^{y+\alpha} \pm 1} \tag{42}$$

and

$$\alpha = -\beta\mu',$$
$$\mu' = \mu - \tfrac{1}{2}w(0) + nw_0 - f_{\text{h.c.}}'(n). \tag{43}$$

Equation (41) is an implicit equation for μ as a function of n or n as a function of μ. From (24) and (20), the thermodynamic potential density f is

$$f = f_{\text{ideal}}(\beta, \mu') + f_{\text{h.c.}}(n) - nf_{\text{h.c.}}'(n) + \tfrac{1}{2}w_0n^2, \tag{44}$$

where $f_{\text{ideal}}(\beta, \mu')$ is the thermodynamic potential density of the ideal Fermi or Bose gas with chemical potential μ':

$$f_{\text{ideal}}(\beta, \mu') = \mp\frac{1}{4\pi^2\beta}\left(\frac{2m}{\beta\hbar^2}\right)^{\frac{3}{2}}\int_0^\infty y^{\frac{1}{2}}\ln(1 \pm e^{-y-\alpha})\,dy. \tag{45}$$

If one considers (41) as an equation for μ', given n, then it is clear from (43) that μ' is the chemical potential of an *ideal* (Fermi or Bose) gas of density n.

Thus, since f is minus the pressure p, one has

$$f_{\text{ideal}}(\beta, \mu') = -p_{\text{ideal}}(\beta, n), \tag{46}$$

where p_{ideal} is the pressure of the ideal gas of density n. The total pressure p is then

$$p(\beta, n) = p^0(\beta, n) - \tfrac{1}{2}w_0 n^2, \tag{47}$$

where

$$\begin{aligned} p^0(\beta, n) &\equiv p_{\text{ideal}}(\beta, n) + p_{\text{h.c.}}(\beta, n), \\ p_{\text{h.c.}}(\beta, n) &\equiv -f_{\text{h.c.}}(n) + n f'_{\text{h.c.}}(n). \end{aligned} \tag{48}$$

This is equivalent to Lieb's relationship[18]

$$a(\beta, n) = a^0(\beta, n) - \tfrac{1}{2}w_0 n^2, \tag{49}$$

provided that we choose

$$f_{\text{h.c.}}(n) = a_{\text{h.c.}}(\beta, n), \tag{50}$$

where $a_{\text{h.c.}}$ is the hard-core contribution to the *Helmholtz* free-energy density, i.e.,

$$a^0(\beta, n) = a_{\text{ideal}}(\beta, n) + a_{\text{h.c.}}(\beta, n) \tag{51}$$

with a^0 the Helmholtz free-energy density of a system of hard spheres (with no attractive tail) and a_{ideal} that of the corresponding ideal gas. In order to see this, recall that the pressure is given in terms of the Helmholtz free-energy density by

$$p(\beta, n) = n^2 \frac{\partial}{\partial n}[n^{-1}a(\beta, n)]. \tag{52}$$

Substituting (51) into (49) and differentiating in accordance with (52), one finds

$$\begin{aligned} p(\beta, n) = p_{\text{ideal}}(\beta, n) &- a_{\text{h.c.}}(\beta, n) \\ &+ n\frac{\partial}{\partial n} a_{\text{h.c.}}(\beta, n) - \tfrac{1}{2}w_0 n^2, \end{aligned} \tag{53}$$

which agrees with (47) and (48), provided that $f_{\text{h.c.}}$ is given by (50).

A few qualifying remarks should be made concerning the agreement between (47) and (49). Lieb's Eq. (49) is *exact* for the *exact* Hamiltonian in the van der Waals limit where the magnitude of $w(|\mathbf{r} - \mathbf{r}'|)$ vanishes and its range goes to infinity in such a way that w_0 [Eq. (4)] approaches a finite, positive limit. On the other hand, (47) was derived from the *approximate* model Hamiltonian (13) [with (14)], but is valid *without* the necessity of taking the van der Waals limit. Actually, however, the van der Waals limit is needed in order to justify our omission of the "exchange contribution" (11) from the model Hamiltonian. In the plane-wave basis (38), the "exchange matrix" (12) becomes

$$\begin{aligned} J_{\mathbf{k}\mathbf{k}'} &= \Omega^{-2}\int w(|\mathbf{r} - \mathbf{r}'|)e^{-i(\mathbf{k}-\mathbf{k}')\cdot(\mathbf{r}-\mathbf{r}')}\, d^3r\, d^3r' \\ &= \Omega^{-1}\tilde{J}(|\mathbf{k} - \mathbf{k}'|), \end{aligned} \tag{54}$$

where

$$\tilde{J}(k) = \int w(r)e^{-i\mathbf{k}\cdot\mathbf{r}}\, d^3r. \tag{55}$$

If $w(r)$ has range λ, then $\tilde{J}(k)$ has range λ^{-1}, i.e., $\tilde{J}(k)$ vanishes for $k \gg \lambda^{-1}$ and is equal to w_0 [Eq. (4)] for $k \ll \lambda^{-1}$. Thus as $\lambda \to \infty$ for fixed w_0 (the van der Waals limit), $\tilde{J}(k)$ goes to zero for all $k > 0$. Thus in the van der Waals limit, $W_{\text{ex.}}$ [Eq. (11)] contributes nothing. Without the van der Waals limit, $W_{\text{ex.}}$ may well be comparable to the interaction term in (13).

5. QUALITATIVE NATURE OF SOLUTIONS IN THE TWO-PHASE REGION

We have thus far restricted ourselves to a single homogeneous phase [Eq. (35)]. It is well known that this restriction leads to the physically unacceptable van der Waals loop in the two-phase region, which is removed by considering the possibility of two phases, thus leading in a natural way to the Maxwell equal-area construction on the curve $p(n^{-1})$, or the convex-envelope construction on the curve $a(n)$. The point is that although Eqs. (35)ff. are always a solution of the coupled equations (20), (21), (23), (27), and (28), they are not the solution of lowest free energy in the condensation region. To find the nature of the two-phase solutions, we shall sketch how the coupled equations might be solved by a self-consistent-field (iterative) procedure. Suppose that the total volume Ω is decomposed into two subregions Ω_g (occupied by the gas) and Ω_l (occupied by the liquid). These subregions are not necessarily connected; e.g., Ω_l may consist of a number of droplets; then Ω_g would be the remaining space $\Omega - \Omega_l$. Owing to quantum effects (nonzero deBroglie wavelength) and nonzero range of $w(|\mathbf{r} - \mathbf{r}'|)$, there is no solution in which $n(\mathbf{r})$ is completely constant in Ω_l and has a different constant value in Ω_g. However, it is clear physically that in the condensation region where the equations

$$p(\beta, n_l) = p(\beta, n_g), \quad n_l\Omega_l + n_g\Omega_g = n\Omega \tag{56}$$

[p given by (47) and (48)] have a solution $n_l(\beta)$, $n_g(\beta)$ with $n_l \neq n_g$, there will be solutions of the coupled equations in which $n(\mathbf{r})$ has the constant value n_l in the interior of Ω_l and the constant value n_g in the interior of Ω_g, changing continuously from n_l to n_g in a finite (Ω-independent) boundary layer separating Ω_l and Ω_g. One might find these by an iterative

[18] E. Lieb, J. Math. Phys. 7, 1016 (1966), Eq. (2.6), with the changes of notation $\rho \to n$, $\alpha \to -w_0$.

procedure, starting with the zero-order approximation[19]

$$n(\mathbf{r}) \approx n^{(0)}(\mathbf{r}) = n_l, \quad \mathbf{r} \in \Omega_l, \\ = n_g, \quad \mathbf{r} \in \Omega_g, \tag{57}$$

and hence[20]

$$\bar{w}(\mathbf{r}) \approx \bar{w}^{(0)}(\mathbf{r}) = -\tfrac{1}{2}w(0) + w_0 n_l, \quad \mathbf{r} \in \Omega_l, \\ = -\tfrac{1}{2}w(0) + w_0 n_g, \quad \mathbf{r} \in \Omega_g. \tag{58}$$

The first-order approximations $\varphi_\alpha^{(1)}$ and $\lambda_\alpha^{(1)}$ to the eigenfunctions φ_α and eigenvalues λ_α of (21) are then to be obtained by solving the Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\nabla^2 - \bar{w}^{(0)}(\mathbf{r})\right]\varphi_\alpha^{(1)}(\mathbf{r}) = \lambda_\alpha^{(1)}\varphi_\alpha^{(1)}(\mathbf{r}). \tag{59}$$

There are clearly two types of solutions. For

$$\tfrac{1}{2}w(0) - w_0 n_l < \lambda_\alpha^{(1)} < \tfrac{1}{2}w(0) - w_0 n_g,$$

there are bound states, bound in the square well of depth $w_0(n_l - n_g)$ occupying the region Ω_l. When substituted into (27) and (28) [with (23)], these will give rise to a contribution $n_b(\mathbf{r})$ to the density which, on physical grounds, one expects to be constant in the interior of Ω_l and to decay exponentially outside Ω_l (this can be verified explicitly in one dimension). For $\lambda_\alpha^{(1)} > \tfrac{1}{2}w(0) - w_0 n_g$, the solutions of (59) are unbound states, which are nonzero in both Ω_g and Ω_l. These will give a contribution $n_u(\mathbf{r})$ to the density which will have one constant value in the interior of Ω_l and a different constant value in the interior of Ω_g. Thus the density n_l of the liquid phase arises both from n_b and n_u, whereas the density n_g of the gas phase arises only from n_u. In first approximation, $n(\mathbf{r})$ has the constant values n_l in the interior of Ω_l and n_g in the interior of Ω_g, changing continuously from n_l to n_g in a boundary layer of thickness either λ_d or λ, whichever is larger, where λ_d is some mean de Broglie wavelength and λ is the range of $w(|\mathbf{r} - \mathbf{r}'|)$. If one continues the iteration, one expects that these qualitative features will be maintained in higher approximations, the readjustments taking place only in the thermodynamically-negligible boundary layer separating Ω_l and Ω_g.

6. COMPARISON WITH EMCH'S MODEL

Our theory of condensation is quite different from that discussed by Emch.[21] Emch works with a model Hamiltonian

$$H - \mu N = -\mu \sum_\alpha N_\alpha - \tfrac{1}{2}\sum_{\alpha\beta} w_{\alpha\beta} N_\alpha N_\beta, \tag{60}$$

where

$$N_\alpha = \sum_j a_\alpha^\dagger(j) a_\alpha(j) \tag{61}$$

and $a_\alpha(j)$, $a_\alpha^\dagger(j)$ annihilate and create a fermion at site j in cell α [Emch's Eqs. (56), (57)]. This differs from our model (13) in two essential ways. In the first place, the effects of the hard cores are not included explicitly, but only implicitly by use of a rigid lattice together with Fermi statistics. This difference is not serious so long as one does not attempt to apply the model to bosons, since the Fermi statistics prevent collapse of the system, as in the model of Andersen, Gartenhaus, and Stranahan.[22] A more important difference between Emch's model and ours is that the operators $a_\alpha(j)$ and $a_\alpha^\dagger(j)$ in terms of which his Hamiltonian is diagonal refer to *localized* particles, whereas the operators a_α and $a_\alpha^\dagger$ in terms of which our model Hamiltonian (13) is diagonal refer to particles in *spatially-extended* single-particle states $\varphi_\alpha(\mathbf{r})$. For the case of a single homogeneous phase, the φ_α are plane waves if one takes periodic boundary conditions, whereas we have shown that in the two-phase region some of the φ_α are states bound in the region occupied by the liquid phase and the remaining φ_α are unbound states extending throughout the container. In both cases (single-phase and two-phase), the states $\varphi_\alpha(\mathbf{r})$ are spatially extended and thus, by normalization, proportional to $\Omega^{-\frac{1}{2}}$. These factors $\Omega^{-\frac{1}{2}}$ are essential in making the model (13) exactly soluble in the thermodynamic limit, since the interaction $w(|\mathbf{r} - \mathbf{r}'|)$ is volume independent. The situation is quite different for Emch's model (60). In fact, the TEH method[9] shows that this model is soluble, via his Eqs. (58) and (59), only if the thermodynamic limit $\Omega \to \infty$ and the limit $w_{\alpha\alpha'} \to 0$ are *coupled* by the condition $w_{\alpha\alpha'} = O(\Omega^{-1})$, which is quite unphysical. On the other hand, our model is exactly soluble in the thermodynamic limit $\Omega \to \infty$, with $w(|\mathbf{r} - \mathbf{r}'|)$ *independent of volume*. If one wishes to take the van der Waals limit of infinite range and zero strength of $w(|\mathbf{r} - \mathbf{r}'|)$ subject to fixed w_0, in order to justify the neglect of exchange terms (11), then this limit is taken *after* the thermodynamic limit, as in previous work[1–8] on the van der Waals–Maxwell theory.

[19] Convergence of the iteration procedure would be improved by rounding off the discontinuity of $n^{(0)}(\mathbf{r})$ in a boundary layer of appropriate thickness. However, the simpler approximation (57) is sufficient to illustrate the general features of the solution.

[20] Strictly speaking, (58) follows exactly from (57) only in the case of the zero-range potential $w(|\mathbf{r} - \mathbf{r}'|) = w_0\delta(\mathbf{r} - \mathbf{r}')$. However, for any finite-range potential, (58) is correct in the *interior* of Ω_l and Ω_g. This is sufficient for our purposes.

[21] G. Emch, J. Math. Phys. **8**, 19 (1967), Sec. V.

[22] C. M. Andersen, S. Gartenhaus, and G. Stranahan, Phys. Rev. **146**, 101 (1966).

7. THEORY OF FREEZING

Finally, it should be pointed out that the model (13) may well lead not only to condensation, but, for sufficiently strong attractive interaction, also to freezing. In order to develop such a theory of freezing, one would look for solutions of the coupled equations in which $n(\mathbf{r})$, and hence $\bar{w}(\mathbf{r})$, is spatially periodic. In the two-phase region one would seek solutions for which $n(\mathbf{r})$ is periodic in the interior of the solid region Ω_s and constant in the interior of the fluid region Ω_f, with equal thermodynamic potential densities in the two phases. In this way it ought to be possible to develop a quantum-mechanical extension of previous classical self-consistent field theories of freezing.[23,24] The use of a realistic attractive interaction $w(|\mathbf{r}-\mathbf{r}'|)$ of fixed finite range is essential in such a theory. Thus the effects of the exchange terms (11) would have to be included, or else their neglect would have to be justified by some argument not involving the van der Waals limit.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation under grant GP-7689.

[23] J. G. Kirkwood and E. Monroe, J. Chem. Phys. **8**, 623 (1940).
[24] R. Brout, Physica **29**, 1041 (1963).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

N-Representability Theorem for Reduced Density Matrices*

HSIAO S. KIANG
Quantum Theory Project, University of Florida, Gainesville, Florida

(Received 15 July 1968)

A reduced density matrix representing a statistical mixture is shown to provide no definite information concerning its N-representability. A theorem about simultaneous N-representability and non-N-representability of a reduced density matrix is formulated and proved.

A problem of current interest has been to determine whether a given reduced density matrix associated with an identical particle subsystem corresponds to a total system with the desired statistics.[1,2] To find out the necessary and sufficient conditions imposed on the reduced density matrix by the statistics of the total system is termed the N-representability problem.[3,4] The N-representability problem has, in fact, raised the question of the proper definition of a reduced density matrix regarding the N-particle statistics. Reported difficulties in the search of a solution to this problem have appeared as stumbling stones in the application of the reduced density matrix theory.

This note shows that a reduced density matrix representing an interacting subsystem is always N-representable and at the same time non-N-representable. This is not surprising. Because the so-called N-representability constraints are not required by the physical arguments in the content of the density matrix theory, they were introduced in a sort of hope that a reduced density matrix always corresponds to some recognizable wavefunction; hence they would hopefully also facilitate the determination of a wavefunction with correct symmetry properties, given the reduced density matrix. It can, in fact, be shown that it is impossible to formulate constraints for a reduced density matrix so as to ensure N-representability; the statistical role that a reduced density matrix plays in the description of a subsystem is to ignore such constraints. Mathematically, such ignorance takes the form of the appearance of a set of independent parameters which are ineffective for the reduced density matrix, but, however, are effective in the wavefunction. These parameters determine the internal symmetry as well as other dynamical properties of the total system. Of course, the internal symmetry of a subsystem of identical particles requires that its reduced density matrix be defined in a symmetrized domain.[5]

* Research supported by the National Science Foundation.
[1] M. B. Ruskai and J. E. Harriman, Phys. Rev. **169**, 101 (1968).
[2] H. Kummer, J. Math. Phys. **8**, 2063 (1967).
[3] A. J. Coleman, Rev. Mod. Phys. **35**, 6681 (1963).
[4] For a complete list of authors contributing to this subject, see *Reduced Density Matrices with Applications to Physical and Chemical Systems*, A. J. Coleman and R. M. Erdahl, Eds. (Queen's University, Kingston, Ontario, Canada, 1968).
[5] H. S. Kiang, Phys. Rev. **167**, 116 (1968).

7. THEORY OF FREEZING

Finally, it should be pointed out that the model (13) may well lead not only to condensation, but, for sufficiently strong attractive interaction, also to freezing. In order to develop such a theory of freezing, one would look for solutions of the coupled equations in which $n(\mathbf{r})$, and hence $\bar{w}(\mathbf{r})$, is spatially periodic. In the two-phase region one would seek solutions for which $n(\mathbf{r})$ is periodic in the interior of the solid region Ω_s and constant in the interior of the fluid region Ω_f, with equal thermodynamic potential densities in the two phases. In this way it ought to be possible to develop a quantum-mechanical extension of previous classical self-consistent field theories of freezing.[23,24] The use of a realistic attractive interaction $w(|\mathbf{r}-\mathbf{r}'|)$ of fixed finite range is essential in such a theory. Thus the effects of the exchange terms (11) would have to be included, or else their neglect would have to be justified by some argument not involving the van der Waals limit.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation under grant GP-7689.

[23] J. G. Kirkwood and E. Monroe, J. Chem. Phys. **8**, 623 (1940).
[24] R. Brout, Physica **29**, 1041 (1963).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

N-Representability Theorem for Reduced Density Matrices*

HSIAO S. KIANG
Quantum Theory Project, University of Florida, Gainesville, Florida

(Received 15 July 1968)

A reduced density matrix representing a statistical mixture is shown to provide no definite information concerning its N-representability. A theorem about simultaneous N-representability and non-N-representability of a reduced density matrix is formulated and proved.

A problem of current interest has been to determine whether a given reduced density matrix associated with an identical particle subsystem corresponds to a total system with the desired statistics.[1,2] To find out the necessary and sufficient conditions imposed on the reduced density matrix by the statistics of the total system is termed the N-representability problem.[3,4] The N-representability problem has, in fact, raised the question of the proper definition of a reduced density matrix regarding the N-particle statistics. Reported difficulties in the search of a solution to this problem have appeared as stumbling stones in the application of the reduced density matrix theory.

This note shows that a reduced density matrix representing an interacting subsystem is always N-representable and at the same time non-N-representable. This is not surprising. Because the so-called N-representability constraints are not required by the physical arguments in the content of the density matrix theory, they were introduced in a sort of hope that a reduced density matrix always corresponds to some recognizable wavefunction; hence they would hopefully also facilitate the determination of a wavefunction with correct symmetry properties, given the reduced density matrix. It can, in fact, be shown that it is impossible to formulate constraints for a reduced density matrix so as to ensure N-representability; the statistical role that a reduced density matrix plays in the description of a subsystem is to ignore such constraints. Mathematically, such ignorance takes the form of the appearance of a set of independent parameters which are ineffective for the reduced density matrix, but, however, are effective in the wavefunction. These parameters determine the internal symmetry as well as other dynamical properties of the total system. Of course, the internal symmetry of a subsystem of identical particles requires that its reduced density matrix be defined in a symmetrized domain.[5]

* Research supported by the National Science Foundation.
[1] M. B. Ruskai and J. E. Harriman, Phys. Rev. **169**, 101 (1968).
[2] H. Kummer, J. Math. Phys. **8**, 2063 (1967).
[3] A. J. Coleman, Rev. Mod. Phys. **35**, 6681 (1963).
[4] For a complete list of authors contributing to this subject, see *Reduced Density Matrices with Applications to Physical and Chemical Systems*, A. J. Coleman and R. M. Erdahl, Eds. (Queen's University, Kingston, Ontario, Canada, 1968).
[5] H. S. Kiang, Phys. Rev. **167**, 116 (1968).

Now we prove the following theorem:

Theorem: A reduced density matrix representing a statistical mixture is simultaneously N-representable and non-N-representable.

Proof: Let ρ_{ij} be the matrix coefficients of a given positive Hermitian operator with finite trace, on the fixed symmetrized orthonormal basis $\{p_i(1, \cdots, m)\}$, $i = 1, \cdots, \alpha$. The space spanned by $\{p_i\}$ is an accessible space for an m-particle identical system. Let $N > m$. Consider a totally symmetrized wavefunction for a system of N identical particles of the form

$$\psi(1, \cdots, N) = \sum_{i,j=1}^{\alpha,\beta} \lambda_{ij} P_i(1, \cdots, m) q_j(m+1, \cdots, N),$$

where $\{q_j\}$, $j = 1, \cdots, \beta$ is some orthonormal symmetrized basis for the complementary $(N-m)$-particle subsystem. This wavefunction can be written alternatively

$$\psi(1, \cdots, N) = \sum_{i=1}^{\alpha} P_i(1, \cdots, m) \phi_i(m+1, \cdots, N),$$

where

$$\phi_i(m+1, \cdots, N) = \sum_{j=1}^{\beta} \lambda_{ij} q_j(m+1, \cdots, N).$$

Note that the number of independent parameters required to identify the ϕ_i's is less than

$$2\alpha\beta - 1 - [N! - m!\,(N-m)!].$$

In the conventional formalism of a composite system, ϕ_i is interpreted as the conditional wavefunction (unnormalized) of the $(N-m)$-particle subsystem. Specifying all the ϕ_i in one-to-one correspondence to the p_i implies specifying Ψ with the resulting m-particle reduced density matrix

$$\rho_{kl}^{\Psi} = \langle \phi_l \mid \phi_k \rangle = \sum_{s=1}^{\beta} \lambda_{ks} \lambda_{ls}^{\dagger}, \quad k, l = 1, \cdots, \alpha.$$

However, to specify ρ_{kl}^{Ψ} alone, one needs no complete knowledge about the ϕ_i's, provided the number of independent parameters in ϕ_i's is greater than $\alpha^2 - 1$. The latter is the maximal number of effective independent parameters in ρ_{kl}^{Ψ}.

Let us define the number of ineffective independent parameters in ρ_{kl}^{Ψ} as the difference between the number of independent parameters in Ψ and the maximal number of effective independent parameters in ρ_{kl}^{Ψ}. It is clear that there exists a family of Ψ with β satisfying the following inequality:

$$2\alpha\beta - 1 - [N! - m!\,(N-m)!] - (\alpha^2 - 1) \geq 0,$$

$$\beta \geq \frac{1}{2\alpha} [\alpha^2 + N! - m!\,(N-m)!].$$

This family consists of both unsymmetrized and totally symmetrized members which map to a given ρ_{ij}. Q.E.D.

The theorem obviously holds for $\alpha \to \infty$. In effect, it signifies the statistical characteristics of the reduced density matrix. It removes the conjecture that awkward constraints otherwise would be required in the evaluation of a reduced density matrix. On the other hand, the N-representability problem does arise in case one limits oneself to the consideration of N-particle wavefunctions involving insufficient independent parameters with respect to their m-particle fixed basis and, at the same time, at the sacrifice of the correct definition of a reduced density matrix. Discussion on the application of reduced density matrices to the energy eigenvalue problem for a many-particle system will be presented in a forthcoming paper.

ACKNOWLEDGMENTS

The author would like to express his thanks to Professor Per-Olov Löwdin who guided this work. It is a pleasure to thank Professor Harrison Shull, Indiana University, who first drew the author's attention to this subject. He wishes also to thank Professor Yngve Öhrn and Professor Darwin Smith for interesting discussions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Quantum Harmonic Oscillator with Time-Dependent Frequency*

S. SOLIMENO†, P. DI PORTO, AND B. CROSIGNANI
Fondazione Ugo Bordoni, Istituto Superiore Poste e Telecomunicazioni, Roma, Italy

(Received 28 February 1969)

The temporal evolution of the state vector relative to a harmonic oscillator with time-dependent frequency is examined. The Schrödinger equation is solved by choosing the instantaneous eigenstates of the Hamiltonian as the basis, thus getting an infinite set of coupled linear differential equations. This formulation is particularly suitable for studying the cases in which the Hamiltonian undergoes a very slow or a sudden variation from an initial constant value into a final one. A rigorous proof of adiabatic invariance to all orders in the slowness parameter is given for the transition probabilities. An application to the evolution of an initially coherent state is made.

1. INTRODUCTION

Dynamical systems whose Hamiltonian $H(t)$ is explicitly time-dependent play a fundamental role in many fields of physics. So long as no detailed prescription is given on the explicit form of $H(t)$, some general results can be obtained within the framework either of the adiabatic approximation, when the Hamiltonian changes very slowly with time, or of the sudden approximation, when $H(t)$ changes from a steady form to another over a very short time interval. This can be achieved through both a classical and a quantum-mechanical approach, leading to well-known adiabatic invariance theorems.[1,2]

In this paper, we wish to treat the case of the quantum-mechanical harmonic oscillator with a prescribed time-dependent frequency, in which the Hamiltonian specializes as $H(t) = \frac{1}{2}p^2 + \frac{1}{2}\omega^2(t)q^2$. We shall be able to furnish, besides a particularly simple set of equations for the temporal evolution of the state, the rigorous proof of the quantum-mechanical adiabatic theorem to all orders. This theorem, which is well established for a classical harmonic oscillator,[3] has been up to now rigorously stated only for quantum-mechanical systems with a finite number of energy levels.[2]

We approach our specific problem in a way closely related to the representation first adopted by Born and Fock[1] in order to describe very general systems with a time-varying Hamiltonian. This approach consists in looking for the solution of the time-dependent Schrödinger equation as an expansion in terms of the stationary-energy eigenfunctions of the instantaneous Hamiltonian, in which the time is regarded as a parameter. In Sec. 2, this method is outlined and applied to our specific problem, thus obtaining a very simple form for the infinite set of linear coupled differential equations which the coefficients of the expansion obey. We express the solution in terms of a unitary matrix **S**, which transforms the vector formed by the initial values of the expansion coefficients into the final one. This approach enables us to face in a straightforward manner the asymptotic expansion of **S**, corresponding to infinitely slow variations of the frequency. This is fully accomplished in Sec. 3, where **S** is given as an asymptotic power series for $\epsilon \to 0$. In a precise formulation, the smallness of the parameter ϵ is related to the slowness of frequency variation, so that $\epsilon \to 0$ implies the indefinite decrease in the rate of change of $\omega(t)$. Under suitable standard hypotheses on the frequency law, the off-diagonal elements of **S** are seen to vanish asymptotically to all orders in ϵ. Furthermore, the asymptotic expansions of the phase factors of diagonal elements are obtained, the moduli of which are shown to remain constant to all orders in ϵ. In this way the adiabatic theorem is stated in a complete and rigorous form for the harmonic oscillator. In Sec. 4 a sudden variation of the frequency is considered and the matrix **S** is shown to assume a particularly compact form. In Sec. 5 an application of these results is made to the evolution of an initially "coherent" state,[4] owing to the fundamental role that these states play in the framework of quantum optics and solid state physics.

Finally, we wish to note that an alternative method for quantum-mechanical systems whose Hamiltonian is explicitly time-dependent is furnished by the theory of explicitly time-dependent invariants. Along these lines, an interesting application has been recently made to the harmonic oscillator by Lewis and Riesenfeld.[5] This approach is very useful when one is interested in connecting the final state of the system to the initial one by disregarding the details of the

* This work was partially supported by C.N.R. (Italian National Council of Research).

† Present address: Istituto di Elettrotecnica, Università di Napoli, Napoli, Italy.

[1] M. Born and V. Fock, Z. Physik **51**, 165 (1928).

[2] See, for instance, A. Lenard, Ann. Phys. (N.Y.) **6**, 261 (1959).

[3] R. Kulsrud, Phys. Rev. **106**, 205 (1957).

[4] R. J. Glauber, Phys. Rev. **131**, 2766 (1963).

[5] H. R. Lewis, Jr., and W. B. Riesenfeld, J. Math. Phys. **10**, 1462 (1969).

evolution, but it is not particularly apt to furnish a complete description of the adiabatic evolution. However, a comparison between our conclusions and the results obtained in Ref. 5 will prove remarkably useful.

2. SCHRODINGER EQUATION FOR TIME-DEPENDENT HARMONIC OSCILLATOR

Let us consider a time-dependent Hamiltonian $H(t)$ and let us introduce the eigenkets $|n; t\rangle$, at each instant of time verifying the energy-eigenvalue equation

$$H(t)\,|n; t\rangle = E_n(t)\,|n; t\rangle. \tag{1}$$

Observe that, in order to obtain a uniquely defined set of orthogonal eigenvectors, one has to introduce the auxiliary condition[1] $\langle n; t|\,\partial/\partial t\,|n; t\rangle = 0$, which is identically satisfied in a nondegenerate case, as can be easily proved through perturbation theory.

Supposing the state vector $|\psi\rangle$ to be known at time t_0, we look for the solution of the Schrödinger equation

$$H(t)\,|\psi\rangle = i\hbar\frac{\partial}{\partial t}\,|\psi\rangle \tag{2}$$

at later times in the form of the expansion (see, e.g., Ref. 6):

$$|\psi\rangle = \sum_n a_n(t)\exp\left[-\frac{i}{\hbar}\int_{t_0}^{t}E_n(t')\,dt'\right]|n; t\rangle. \tag{3}$$

Inserting Eq. (3) into Eq. (2), we get, with the aid of Eq. (1),

$$\dot{a}_k = -\sum_n a_n \exp\left[\frac{i}{\hbar}\int_{t_0}^{t}(E_k - E_n)\,dt'\right] \times \langle k; t|\frac{\partial}{\partial t}\,|n; t\rangle. \tag{4}$$

The matrix elements on the right side of this last equation can be evaluated by exploiting the following identity[6]:

$$\langle k; t|\frac{\partial}{\partial t}\,|n; t\rangle = (E_n - E_k)^{-1}\langle k; t|\frac{\partial H}{\partial t}\,|n; t\rangle, \quad \text{for } k \neq n,$$
$$= 0, \quad \text{otherwise}, \tag{5}$$

thus getting for the harmonic oscillator (see, e.g., Ref. 6, p. 155):

$$\langle k; t|\,\partial/\partial t\,|n; t\rangle = [\dot{\omega}/\hbar(n-k)]\langle k; t|\,q^2\,|n; t\rangle = (\dot{\omega}/4\omega)[(k+2)^{\frac{1}{2}}(k+1)^{\frac{1}{2}}\delta_{k,n-2} - k^{\frac{1}{2}}(k-1)^{\frac{1}{2}}\delta_{k,n+2}], \tag{6}$$

$$\dot{a}_k = (\dot{\omega}/4\omega)[k^{\frac{1}{2}}(k-1)^{\frac{1}{2}}e^{2i\varphi(t,t_0)}a_{k-2} - (k+2)^{\frac{1}{2}}(k+1)^{\frac{1}{2}}e^{-2i\varphi(t,t_0)}a_{k+2}], \tag{7}$$

where $\varphi(t, t_0) \equiv \int_{t_0}^{t}\omega(t')\,dt'$. We wish to note that this last set of equations is particularly suited to be solved step by step, when the number of levels occupied at time t_0 is finite.

For our successive developments it is convenient to rewrite the set of Eqs. (7) in a more compact form. This can be achieved by introducing the matrix $\mathbf{G}$, defined by its components as

$$G_{\alpha\beta}(t) \equiv (\dot{\omega}/4\omega)[\alpha^{\frac{1}{2}}(\alpha-1)^{\frac{1}{2}}e^{2i\varphi(t,t_0)}\delta_{\alpha,\beta+2} - (\alpha+2)^{\frac{1}{2}}(\alpha+1)^{\frac{1}{2}}e^{-2i\varphi(t,t_0)}\delta_{\alpha,\beta-2}], \quad \alpha, \beta \geq 0, \tag{8}$$

and the vector $\mathbf{a}(t)$ defined by the components of $|\psi\rangle$ in the coordinate system

$$\{\exp[-i(n+\tfrac{1}{2})\varphi(t, t_0)\,|n; t\rangle\}$$

as

$$\mathbf{a}(t) \equiv (a_0, a_1, \cdots, a_n, \cdots). \tag{9}$$

Equations (7) then assume the form

$$\dot{\mathbf{a}}(t) = \mathbf{G}(t)\cdot\mathbf{a}(t). \tag{10}$$

The solution of Eq. (10) can be expressed conveniently in terms of a unitary matrix $\mathbf{S}$ which transforms $\mathbf{a}(t_0)$ into $\mathbf{a}(t)$ as

$$\mathbf{a}(t) \equiv \mathbf{S}(t, t_0)\mathbf{a}(t_0). \tag{11}$$

The components $S_{\alpha\beta}(t, t_0)$ coincide, safe for unessential phase factors, with the transition amplitudes $T(\beta \to \alpha) \equiv \langle\alpha; t\,|\,\beta; t_0\rangle$ from the eigenstate $|\beta; t_0\rangle$ of $H(t_0)$ to the eigenstate $|\alpha; t\rangle$ of $H(t)$. In addition, $\mathbf{S}$ is unitary, since the absolute magnitude of $\mathbf{a}(t)$ remains unchanged during the motion.

The matrix $\mathbf{S}$ can be obtained in the form of a series expansion by a generalization of the Liouville–Neumann method of successive substitutions,[7] which formally gives

$$\mathbf{S}(t, t_0) = \sum_{k=0}^{\infty}\mathbf{S}^{(k)}(t, t_0), \tag{12}$$

where $\mathbf{S}^{(0)}(t, t_0) \equiv \mathbf{1}$ and

$$\mathbf{S}^{(k)}(t, t_0) \equiv \int_{t_0}^{t}\mathbf{G}(t')\cdot\mathbf{S}^{(k-1)}(t', t_0)\,dt'. \tag{13}$$

Owing to the particularly simple form of the matrix $\mathbf{G}$, the components of the kth term of Eq. (12) can be written as

$$S^{(k)}_{\alpha\beta}(t, t_0) = \sum_{\sigma_2,\sigma_3,\cdots,\sigma_k}\int_{t_0}^{t}G_{\alpha,\alpha+\sigma_k}(t_k)\,dt_k \times \int_{t_0}^{t_k}G_{\alpha+\sigma_k,\alpha+\sigma_k+\sigma_{k-1}}(t_{k-1}) \times dt_{k-1}\cdots\int_{t_0}^{t_2}G_{\alpha+\sigma_k+\cdots+\sigma_2,\beta}(t_1)\,dt_1, \tag{14}$$

[6] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Co., New York, 1955), pp. 213–215.

[7] E. J. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge University Press, London, 1963), p. 221.

the summation running over the values $\sigma_i = \pm 2$. With the aid of Eq. (8), the generical term of this sum is given in a more explicit form by

$$\begin{aligned} &J^{(k)}(\boldsymbol{\sigma}^{(k)}; t, t_0) \\ &= g_{\alpha,\alpha+\sigma_k} g_{\alpha+\sigma_k,\alpha+\sigma_k+\sigma_{k-1}} \cdots g_{\alpha+\sigma_k+\sigma_{k-1}+\cdots+\sigma_2,\beta} \\ &\quad \times \int_{t_0}^{t} \frac{\dot{\omega}}{\omega} e^{-i\sigma_k\varphi}\, dt_k \int_{t_0}^{t_k} \frac{\dot{\omega}}{\omega} e^{-i\sigma_{k-1}\varphi} \\ &\quad \times dt_{k-1} \cdots \int_{t_0}^{t_2} \frac{\dot{\omega}}{\omega} e^{-i\sigma_1\varphi}\, dt_1 \\ &\equiv g_{\alpha,\alpha+\sigma_k} g_{\alpha+\sigma_k,\alpha+\sigma_k+\sigma_{k-1}} \cdots g_{\alpha+\sigma_k+\cdots+\sigma_2,\beta} I^{(k)}(\boldsymbol{\sigma}^{(k)}; t, t_0), \end{aligned} \tag{15}$$

where

$$\begin{aligned} \boldsymbol{\sigma}^{(k)} &\equiv (\sigma_1, \sigma_2, \cdots, \sigma_k), \\ g_{\gamma,\gamma-2} &\equiv \tfrac{1}{4}\gamma^{\frac{1}{2}}(\gamma-1)^{\frac{1}{2}}, \\ g_{\gamma,\gamma+2} &\equiv -\tfrac{1}{4}(\gamma+2)^{\frac{1}{2}}(\gamma+1)^{\frac{1}{2}}, \end{aligned} \tag{16}$$

and the relation $\sum_{i=1}^{k} \sigma_i = \beta - \alpha$ holds.

In order to avoid divergent integrals when we shall deal with the asymptotic behavior of $\mathbf{S}(t, -\infty)$ in the next section, we introduce the following quantities:

$$\begin{aligned} I^{(k)}(\boldsymbol{\sigma}^{(k)}; t, t', t_0) &\equiv e^{i(\beta-\alpha)\varphi(t,t_0)} I^{(k)}(\boldsymbol{\sigma}^{(k)}; t', t_0), \\ S^{(k)}_{\alpha\beta}(t, t_0) &\equiv e^{i(\beta-\alpha)\varphi(t,t_0)} S^{(k)}_{\alpha\beta}(t, t_0), \\ S_{\alpha\beta}(t, t_0) &\equiv e^{i(\beta-\alpha)\varphi(t,t_0)} S_{\alpha\beta}(t, t_0). \end{aligned} \tag{17}$$

It is interesting to note that the elements $S^{(k)}_{\alpha\beta}(t, t_0)$ vanish for $k + \frac{1}{2}|\alpha - \beta|$ not divisible by two and $k < \frac{1}{2}|\alpha - \beta|$, as it can be derived from the general structure of the matrix $\mathbf{G}$. In fact, if k^+ and k^- represent the number of terms $g_{\gamma,\delta}$ on the rhs of Eq. (15) with δ equal, respectively, to $\gamma + 2$ and $\gamma - 2$, one has $k^+ + k^- = k$ and $|k^+ - k^-| = \frac{1}{2}|\alpha - \beta|$. Then,

$$k + \tfrac{1}{2}|\alpha - \beta| = 2 \max(k^+, k^-) \geq |\alpha - \beta|,$$

which proves the assertion.

In connection with the series of Eq. (12), which formally represents the matrix $\mathbf{S}$, we must now face the problem of the convergence of the expansions

$$S_{\alpha\beta}(t, t_0) \equiv \sum_{k=0}^{\infty} S^{(k)}_{\alpha\beta}(t, t_0). \tag{18}$$

We wish to recall that, as long as the matrix $\mathbf{G}$ is of finite degree, the convergence of this series can be proved in a completely standard way. (For a specific application in the frame of the adiabatic theory, see, for instance, Ref. 8.)

As a first step, let us consider the case in which the phase factors appearing in Eq. (15) can be put identically equal to one. This actually happens in the case of sudden variation of the frequency at time t_0, that is, when the frequency settles into a final constant value ω_f in a time $t_f - t_0 \ll \omega^{-1}(t)$ for $t_0 \leq t \leq t_f$. In such a case, $I^{(k)}(\boldsymbol{\sigma}^{(k)}; t, t_0)$ is given by

$$I^{(k)}(\boldsymbol{\sigma}^{(k)}; t, t_0^-) = \left(\int_{t_0}^{t} \frac{\dot{\omega}}{\omega}\, dt'\right)^k (k!)^{-1} = \left(\ln \frac{\omega_f}{\omega_i}\right)^k (k!)^{-1}, \tag{19}$$

so that Eq. (18) specializes into

$$S_{\alpha\beta}(t, t_0^-) = \sum_{k=0} A^{(k)}_{\alpha\beta} (k!)^{-1} \left(\ln \frac{\omega_f}{\omega_i}\right)^k, \tag{20}$$

where

$$A^{(k)}_{\alpha\beta} \equiv \sum_{\sigma_2 \cdots \sigma_k} g_{\alpha,\alpha+\sigma_k} \cdots g_{\alpha+\sigma_k+\cdots+\sigma_2,\beta}$$

and $\omega_i = \omega(t_0-)$. On the other hand, $S_{\alpha\beta}(t, t_0)$ coincides with the overlap integral of the eigenstates of the Hamiltonian just before and after the discontinuity, as a consequence of the constancy of the state vector across the discontinuity. This leads to

$$\begin{aligned} S_{\alpha\beta}(t, t_0^-) &= \omega_i^{\frac{1}{2}} \omega_f^{\frac{1}{2}} [2^{(\alpha+\beta)} \pi \hbar (\alpha!)(\beta!)]^{-\frac{1}{2}} \\ &\quad \times \int_{-\infty}^{\infty} \exp[-(2\hbar)^{-1} x^2 (\omega_i + \omega_f)] \\ &\quad \times H_\alpha(x \omega_f^{\frac{1}{2}} \hbar^{-\frac{1}{2}}) H_\beta(x \omega_i^{\frac{1}{2}} \hbar^{-\frac{1}{2}})\, dx, \end{aligned} \tag{21}$$

where $H_\gamma(y)$ is the Hermite polynomial of order γ. The integral on the right-hand side coincides, for $z = \ln(\omega_f/\omega_i)$, with

$$\begin{aligned} &\int_{-\infty}^{\infty} \exp[-(2\hbar)^{-1} x^2 \omega_i (1 + e^z)] \\ &\qquad \times H_\alpha\left[x\left(\frac{\omega_i e^z}{\hbar}\right)^{\frac{1}{2}}\right] H_\beta\left[x\left(\frac{\omega_i}{\hbar}\right)^{\frac{1}{2}}\right] dx, \end{aligned} \tag{22}$$

which, regarded as a function of the complex variable z, converges only for $\operatorname{Re} e^z > -1$ or $\operatorname{Re} e^z = -1$ and $\operatorname{Im} e^z \neq 0$.[9] This in turn determines the radius of convergence ρ of the right-hand side of Eq. (20), regarded as a power series in z. A simple study of the equation $\operatorname{Re} e^z = -1$ furnishes the approximate value $\rho = 2.18$.

For any frequency law, one has

$$|I^{(k)}(\boldsymbol{\sigma}^{(k)}; t, t_0)| < \left(\int_{t_0}^{t} |\omega^{-1}\dot{\omega}|\, dt'\right)^k (k!)^{-1}, \tag{23}$$

so that

$$|S^{(k)}_{\alpha\beta}(t, t_0)| < |A^{(k)}_{\alpha\beta}| \left(\int_{t_0}^{t} |\omega^{-1}\dot{\omega}|\, dt'\right)^k (k!)^{-1}, \tag{24}$$

and the series on the right of Eq. (18) is absolutely convergent for $\int_{t_0}^{t} |\omega^{-1}\dot{\omega}|\, dt' < \rho$.

[8] S. Tamor, Ann. Phys. (N.Y.) 23, 161 (1963).

[9] A. Erdélyi, Ed., *Tables of Integral Transforms* (McGraw-Hill Book Co., New York, 1954), p. 291.

3. ADIABATIC CASE

The problem we now consider is the behavior of the harmonic oscillator when $\omega(t)$ varies adiabatically, that is, very slowly with time. In a precise formulation, we introduce a parameter ϵ running over the interval $(0, \epsilon_M)$ and the frequency law class defined by $\omega(t) \equiv \Omega(\epsilon t)$. Since the main task of the present paper is to demonstrate some adiabatic invariance theorems to all orders in the slowness parameter ϵ, we assume throughout $\Omega(\tau)$ to be a function infinitely differentiable for $\tau \in (-\infty, +\infty)$ and with absolutely integrable derivatives, which settles in the remote past and future in the constant values Ω_i and Ω_f. Along these lines, our aim will be reached by carrying out the asymptotic expansion of the transition matrix $\mathbf{S}(+\infty, -\infty)$, considered as a function of $\Omega(\epsilon t)$, for $\epsilon \to 0$. In turn, this will be accomplished by expressing the terms of the asymptotic expansion through the contributions originating from the single $\mathbf{S}^{(k)}$'s, whose typical dependence on ϵ is the one pertaining to $I^{(k)}(\boldsymbol{\sigma}^{(k)}; t, -\infty)$. Strictly speaking, this last integral has no meaning, since the phase factor $\varphi(t, t_0)$ appearing in Eq. (15) diverges as $t_0 \to -\infty$. Then, with the help of Eq. (17), we can carry out first the asymptotic expansion of $I^{(k)}(\boldsymbol{\sigma}^{(k)}; +\infty, +\infty, -\infty)$, and later the one relative to $I^{(k)}(\boldsymbol{\sigma}^{(k)}; +\infty, -\infty)$. To this end, we substitute t with $\epsilon^{-1}\tau$ and $\omega(t)$ with $\Omega(\tau)$. Thus, after some manipulations of Eqs. (15) and (17), we get

$$I^{(k)}(\boldsymbol{\sigma}^{(k)}; \epsilon; \tau, \tau', -\infty) = \int_{-\infty}^{\tau'} \Omega^{-1}\dot{\Omega} \exp(i\sigma_k\epsilon^{-1}\phi)\, d\tau_k \times \int_{-\infty}^{\tau_k} \Omega^{-1}\dot{\Omega} \exp(i\sigma_{k-1}\epsilon^{-1}\phi) \times d\tau_{k-1} \cdots \int_{-\infty}^{\tau_2} \Omega^{-1}\dot{\Omega} \exp(i\sigma_1\epsilon^{-1}\phi)\, d\tau_1, \quad (25)$$

where $\phi(\tau, \tau_i) = \int_{\tau_i}^{\tau} \Omega(\tau')\, d\tau'$ and the dot stands for derivation with respect to the argument.

Let us now assume that for a given k this last integral can be expressed as

$$I^{(k)}(\boldsymbol{\sigma}^{(k)}; \epsilon; \tau, \tau', -\infty) = \exp[iM\epsilon^{-1}\phi(\tau, \tau')] \sum_{h=0}^{H} A_h(\tau')\epsilon^{N+h} + \epsilon^{N+H} F(\boldsymbol{\sigma}^{(k)}; H; \epsilon; \tau, \tau'), \quad (26)$$

where H can assume any integer value, M and N are suitably chosen integer values, the A_h's are infinitely differentiable functions in the interval $(-\infty, +\infty)$, and $F(\boldsymbol{\sigma}^{(k)}; H; \epsilon; \tau, \tau')$ is a bounded function of $\boldsymbol{\sigma}^{(k)}$, ϵ, τ, and τ' for given k and H. On the other hand, it can be easily shown by successive integrations by parts that[10]

$$\int_{-\infty}^{\tau'} \Omega^{-1}(\tau')\dot{\Omega}(\tau')g(\epsilon; \tau, \tau') \exp[i\sigma\epsilon^{-1}\phi(\tau, \tau')]\, d\tau' = \exp[i\sigma\epsilon^{-1}\phi(\tau, \tau')] \sum_{k=0}^{K} g_k(\tau')(\epsilon\sigma^{-1})^{k+1} - i(\epsilon\sigma^{-1})^{K+1} \int_{-\infty}^{\tau'} g_{k+1}(\tau')\Omega(\tau') \times \exp[i\sigma\epsilon^{-1}\phi(\tau, \tau')]\, d\tau', \quad (27)$$

where

$$g_0(\tau') = i\Omega^{-2}(\tau')\dot{\Omega}(\tau')g(\epsilon; \tau, \tau'), \qquad g_{k+1}(\tau') = -i\Omega^{-1}(\tau')\dot{g}_k(\tau'), \quad (28)$$

and K can assume any integer value. Thus, with the help of Eqs. (25)–(27), we can write

$$\begin{aligned} I^{(k+1)}(\boldsymbol{\sigma}^{(k+1)}; \epsilon; \tau, \tau', -\infty) &= \int_{-\infty}^{\tau'} \Omega^{-1}(\tau')\dot{\Omega}(\tau')I^{(k)}(\boldsymbol{\sigma}^{(k)}; \epsilon; \tau, \tau'', -\infty) \times \exp[i\sigma_{k+1}\epsilon^{-1}\phi(\tau, \tau')]\, d\tau'' \\ &= \exp[i(M + \sigma_{k+1})\epsilon^{-1}\phi(\tau, \tau')] \times \sum_{h=0}^{H-1} \epsilon^{N+1+h} \sum_{l,m=0}^{h} g_{lm}(\tau')(M + \sigma_{k+1})^{l-1} \\ &\quad + \epsilon^{N+H}\Bigg(-i\sum_{m=0}^{H}(M + \sigma_{k+1})^{-H+m} \times \int_{-\infty}^{\tau'} g_{H-m,m}(\tau'')\Omega(\tau'') \times \exp i[(M + \sigma_{k+1})\epsilon^{-1}\phi(\tau, \tau')]\, d\tau'' \\ &\quad + \int_{-\infty}^{\tau'} \Omega^{-1}(\tau')\dot{\Omega}(\tau')F(\boldsymbol{\sigma}^{(k)}; H; \epsilon; \tau; \tau'') \times \exp[i\sigma_{k+1}\epsilon^{-1}\phi(\tau, \tau')]\, d\tau''\Bigg), \end{aligned} \quad (29)$$

where

$$g_{0m}(\tau') = i\Omega^{-2}(\tau')\dot{\Omega}(\tau')A_m(\tau'), \qquad g_{lm}(\tau') = -i\Omega^{-1}(\tau')\dot{g}_{l-1,m}(\tau'), \quad (30)$$

if $\sigma_{k+1} + M \neq 0$, and

$$I^{(k+1)}(\boldsymbol{\sigma}^{(k+1)}; \epsilon; \tau; \tau', -\infty) = \sum_{h=0}^{H} \epsilon^{N+h} \int_{-\infty}^{\tau'} \Omega^{-1}(\tau')\dot{\Omega}(\tau')A_h(\tau')\, d\tau'' + \epsilon^{N+H}\Bigg(\int_{-\infty}^{\tau'} \Omega^{-1}(\tau')\dot{\Omega}(\tau')F(\boldsymbol{\sigma}^{(k)}; H; \epsilon; \tau, \tau'') \times \exp[i\sigma_{k+1}\epsilon^{-1}\phi(\tau, \tau')]\, d\tau''\Bigg), \quad (31)$$

if $\sigma_{k+1} + M = 0$.

A simple inspection of Eqs. (29) and (31) allows us to affirm that the coefficients of ϵ^{N+1+h} are infinitely

[10] J. G. Van der Corput, "Asymptotic Expansions, Report No. 2" (Dept. of Mathematics, University of California, Berkeley, 1954), p. 18.

differentiable functions of $\Omega(\tau)$ and that the quantities in brackets are bounded functions of $\sigma^{(k+1)}$, ϵ, τ, and τ' for given k and H. Thus, since Eq. (26) is valid for $k = 1$, as can be proved by setting $g = 1$ in Eq. (27), it is valid for all $\bar{I}^{(k)}$'s. The importance of this result will be appreciated later, when we resort to Eq. (26) to get the asymptotic power expansion of $\mathbf{S}^{(k)}$ and $\mathbf{S}$ for $\epsilon \to 0$.

Now, it remains to calculate M and N, which can be readily achieved by inspecting Eqs. (29) and (31). We get in this way

$$M(\sigma^{(k)}) = \sigma_1 + \sigma_2 + \cdots + \sigma_k = \beta - \alpha,$$
$$N(\sigma^{(k)}) = k - f(\sigma^{(k)}), \tag{32}$$

where $f(\sigma^{(k)})$ represents the number of times the relation $\sigma_1 + \sigma_2 + \cdots + \sigma_l = 0$, $l = 2, 3, \cdots, k$, can be satisfied with the components of a given vector $\sigma^{(k)}$. Since $f(\sigma^{(k)})$ reaches its upper bound when $\sigma_i + \sigma_{i+1} = 0$ for all $i < k$, we can easily evaluate the lower bound $N^{(k)}$ of $N(\sigma^{(k)})$, which reads

$$N^{(k)} = \tfrac{1}{2}k, \qquad \text{for } k \text{ even},$$
$$N^{(k)} = \tfrac{1}{2}(k+1), \quad \text{for } k \text{ odd}. \tag{33}$$

Putting $H = 0$, with the aid of Eqs. (32), Eq. (26) specializes into

$$\bar{I}^{(k)}(\sigma^{(k)}; \epsilon; \tau, \tau', -\infty) = \{\exp[i(\beta - \alpha)\epsilon^{-1}\phi(\tau, \tau')]A_0(\tau') \times F(\sigma^{(k)}; 0; \epsilon; \tau, \tau')\}\epsilon^{N(\sigma^{(k)})}, \tag{34}$$

thus giving rise to the following inequalities:

$$|\bar{I}^{(k)}(\sigma^{(k)}; \epsilon; \tau, \tau', -\infty)| < Q^{(k)}\epsilon^{N^{(k)}}, \tag{35}$$

where $Q^{(k)}$ is independent of $\sigma^{(k)}$, ϵ, τ, and τ', and, with the help of the condition $k = \frac{1}{2}|\beta - \alpha| + 2n$ established in Sec. 2,

$$|S^{(k)}_{\alpha\beta}(\infty, -\infty)| < |A^{(k)}_{\alpha\beta}|\, Q^{(k)}\epsilon^{N^{(k)}}, \tag{36}$$

$$\lim_{\epsilon \to 0} S^{(k+2)}_{\alpha\beta}(\infty, -\infty)/S^{(k)}_{\alpha\beta}(+\infty, -\infty) = 0. \tag{37}$$

This last equation ensures that the sequence of functions $\{S^{(k)}_{\alpha\beta}\}$ is an asymptotic sequence[11] for $\epsilon \to 0$. Accordingly, the solution of Eq. (11) defined by the right-hand side of Eq. (12) has to be considered as a perturbative expansion for $\epsilon \to 0$.

In turn, Eqs. (35) and (36), together with Eqs. (14) and (15), yield

$$|\bar{I}^{(k+2n)}(\sigma^{(k+2n)}; \epsilon; +\infty, -\infty)| \leq Q^{(k)}\epsilon^{N^{(k)}}[(2n)!]^{-1}\left(\int_{-\infty}^{\infty} |\Omega^{-1}\dot{\Omega}|\, d\tau\right)^{2n} \equiv Q^{(k)}\epsilon^{N^{(k)}}[(2n)!]^{-1}s^{2n}, \tag{38}$$

$$|\bar{S}^{(k+2n)}_{\alpha\beta}(+\infty, -\infty)| < |A^{(k+2n)}_{\alpha\beta}|\, Q^{(k)}\epsilon^{N^{(k)}}[(2n)!]^{-1}s^{2n}, \tag{39}$$

$$\left|\bar{S}_{\alpha\beta}(+\infty, -\infty) - \sum_{n=0}^{N-1} \bar{S}^{(\frac{1}{2}|\alpha-\beta|+2n)}_{\alpha\beta}(+\infty, -\infty)\right| < Q^{(\frac{1}{2}|\alpha-\beta|+2N)}\epsilon^{N(\frac{1}{2}|\alpha-\beta|+2N)} \times \sum_{l=0}^{\infty} |A^{(\frac{1}{2}|\alpha-\beta|+2N+2l)}_{\alpha\beta}|\, [(2l)!]^{-1}s^{2l}. \tag{40}$$

In accordance with the conclusions of Sec. 2, the series on the right of Eq. (40) converges, provided $s < \rho$ and the following order relation[11] holds:

$$\bar{S}_{\alpha\beta}(+\infty, -\infty) = \sum_{n=0}^{N-1} \bar{S}^{(\frac{1}{2}|\alpha-\beta|+2n)}_{\alpha\beta}(+\infty, -\infty) + O(\epsilon^{N(\frac{1}{2}|\alpha-\beta|+2N)}), \tag{41}$$

where O stands for the well-known Landau–Beckman order symbol.[11]

It is worth now remembering that the first and major item of our program involves evaluating the asymptotic expansion of $S(+\infty, -\infty)$ for $\epsilon \to 0$. To this end, indeed, we have already obtained some useful relations represented by the set of Eqs. (26), (31), and (41). It remains yet to calculate the coefficients in the asymptotic power series of $S_{\alpha\beta}(+\infty, -\infty)$ and $S^{(k)}_{\alpha\beta}(+\infty, -\infty)$. To this end, let us focus our attention on some interesting features of the $I^{(k)}$'s for $\tau = \tau' \to +\infty$. Taking into account that the functions $g_{lm}(\tau)$ vanish identically as $\tau \to \infty$, Eqs. (29), (31), and (32) yield

$$|\bar{I}^{(k)}(\sigma^{(k)}; \epsilon; +\infty, -\infty)| = \epsilon^{N(\sigma^{(k)})+H}\,|F(\sigma^{(k)}; H; \epsilon)|, \quad \text{for } \beta - \alpha \neq 0, \tag{42}$$

$$\bar{I}^{(k)}(\sigma^{(k)}; \epsilon; +\infty, -\infty) = \sum_{h=0}^{H} \epsilon^{N(\sigma^{(k)})+h} \int_{-\infty}^{\infty} \Omega^{-1}\dot{\Omega}A_h\, d\tau + \epsilon^{N(\sigma^{(k)})+H}F(\sigma^{(k)}; H; \epsilon), \quad \text{for } \beta - \alpha = 0, \tag{43}$$

where H can assume any integer value and F is a bounded function of $\sigma^{(k)}$ and ϵ for given k and H. In other words, Eq. (42) means that $\bar{I}^{(k)}$ and consequently $\bar{S}^{(k)}_{\alpha\beta}(+\infty, -\infty)$ are adiabatic-invariant to all orders in the parameter ϵ, when $\beta - \alpha \neq 0$. This feature, together with Eq. (41), implies in turn that $\bar{S}_{\alpha\beta}(+\infty, -\infty)$ and consequently $S_{\alpha\beta}(+\infty, -\infty)$ vanish asymptotically to all orders in ϵ, thus proving the main property of the asymptotic behavior of $S_{\alpha\beta}(+\infty, -\infty)$. In addition, Eq. (43), with the help of Eq. (17), assures that the asymptotic power expansion

[11] A. Erdélyi, *Asympotic Expansions* (Dover Publications, Inc., New York, 1956).

of $S_{\alpha\alpha}^{(k)}(+\infty, -\infty)$ exists. Therefore, Eq. (41) yields

$$S_{\alpha\alpha}^{[N]} = \sum_{n=0}^{N} S_{\alpha\alpha}^{(2n)\,[N]}, \tag{44}$$

where $S_{\alpha\alpha}^{[N]}$ and $S_{\alpha\alpha}^{(2n)[N]}$ represent, respectively, the coefficients of ϵ^N in the asymptotic power series of the matrix components $S_{\alpha\alpha}$ and $S_{\alpha\alpha}^{(2n)}$ considered as functions of ϵ.

Since the matrix S is unitary and its off-diagonal components vanish asymptotically to all orders, most likely the moduli of the diagonal components must remain asymptotically equal to one. This conjecture can be rigorously proved, as we will do below. Although this problem can be solved in principle by means of Eq. (44), we prefer for the sake of simplicity that it hinge upon some results of Ref. 5. In that paper the following expression is obtained, to within a phase factor:

$$\begin{aligned} S_{\alpha\beta}(+\infty, -\infty) = \exp(i\chi)(\beta!)^{-\frac{1}{2}}(\cosh\delta + 1)^{-\frac{1}{4}}2^{\frac{1}{4}} \\ \times \sum_{r=0}^{\infty}\left(\frac{-\zeta}{2\eta}\right)^{r}[(2r)!]^{\frac{1}{2}}(r!)^{-1} \\ \times \langle u_{\alpha,+\infty}|\,(\zeta^* b + \eta^* b^{\dagger})^{\beta}\,|u_{2r,+\infty}\rangle, \end{aligned} \tag{45}$$

where the constants δ, η, and ζ can be determined in terms of the frequency law $\omega(t)$ and obey the relations

$$\begin{aligned} |\eta|^2 - |\zeta|^2 &= 1, \\ |\eta| &= (1 + \cosh\delta)^{\frac{1}{2}}2^{-\frac{1}{2}}. \end{aligned} \tag{46}$$

For $\beta = 0$ and $\alpha \neq 0$, Eq. (45) yields

$$\begin{aligned} |S_{\alpha 0}(+\infty, -\infty)|^2 = \alpha!\, 2^{-\alpha+\frac{1}{2}}[(\alpha 2^{-1})!]^{-2} \\ \times (\cosh\delta - 1)^{\frac{1}{2}\alpha}(\cosh\delta + 1)^{-\frac{1}{2}(\alpha+1)}. \end{aligned} \tag{47}$$

Since $S_{\alpha 0}$ vanishes asymptotically to all orders in ϵ, this equation gives

$$\begin{aligned} \lim_{\epsilon\to 0} \epsilon^n |S_{\alpha 0}(+\infty, -\infty)| \\ = \lim_{\epsilon\to 0} \epsilon^n \delta^{\frac{1}{2}\alpha} \times \lim_{\delta\to 0} \delta^{-\frac{1}{2}\alpha} |S_{\alpha 0}(+\infty, -\infty)| \\ = (\alpha!)^{\frac{1}{2}}2^{-\alpha}[(2^{-1}\alpha)!]^{-1} \lim_{\epsilon\to 0} \epsilon^n \delta^{\frac{1}{2}\alpha} = 0, \end{aligned} \tag{48}$$

for all integers n, which, together with Eqs. (46), implies that $|\eta| - 1$, ζ, and δ vanish asymptotically to all orders in ϵ. Thus, with the help of Eq. (45) and the theorem on the asymptotic expansion of the product of two functions (see, e.g., Ref. 11, p. 18), we can write

$$\begin{aligned} S_{\alpha\alpha}(+\infty, -\infty) &\sim \exp[(i\chi)\eta^{*\alpha}] \\ &\sim \exp[i(\chi + \alpha\varphi_{\eta^*})] \equiv \exp(i\Psi), \\ |S_{\alpha\alpha}(+\infty, -\infty)| &\sim 1, \end{aligned} \tag{49}$$

where the symbol $\sim$ means that the two sides of the equivalence relation are asymptotically equal with respect to $\{\epsilon^n\}$.

4. THE SUDDEN CASE

In this section we wish to take into account again the case of sudden variation of the frequency, which has been already treated in Sec. 2. Our aim is to put the matrix S in the form of an operator $\hat{\mathbf{S}}$ particularly suited to study the evolution of an initially coherent state, as we show in the next section.

Let us remember that, according to Eq. (20),

$$S_{\alpha\beta}^{(k)}(t, t_{0^-}) = A_{\alpha\beta}^{(k)} s^k/k!, \tag{50}$$

where $s = \ln \omega_f/\omega_i$. Let us introduce the operators $\hat{A}^{(k)}$ defined by the relations

$$\langle\alpha; +\infty|\,\hat{A}^{(k)}\,|\beta; -\infty\rangle = A_{\alpha\beta}^{(k)}. \tag{51}$$

They allow to pass from the Hilbert space spanned by the initial eigenkets $|\beta; -\infty\rangle$ to the Hilbert space relative to the $|\alpha; +\infty\rangle$'s. From the structure of $A^{(k)}$ it is easy to obtain

$$\hat{A}^{(k)} = \sum_{\sigma_1,\sigma_2,\cdots,\sigma_k} \hat{a}^{\sigma_k}\hat{a}^{\sigma_{k-1}}\cdots\hat{a}^{\sigma_1} = (\hat{a}^{+2} + \hat{a}^{-2})^k, \tag{52}$$

where

$$\begin{aligned} \hat{a}^{\sigma}\,|\beta; -\infty\rangle &= -(\beta+1)^{\frac{1}{2}}(\beta+2)^{\frac{1}{2}}4^{-1}\,|\beta+2; +\infty\rangle, \\ &\qquad \text{if} \quad \sigma = 2, \\ &= \beta^{\frac{1}{2}}(\beta-1)^{\frac{1}{2}}4^{-1}\,|\beta-2; +\infty\rangle, \\ &\qquad \text{if} \quad \sigma = -2. \end{aligned} \tag{53}$$

Observe that $\hat{a}^{+2} = (2^{-1}a^{\dagger})^2$ and $\hat{a}^{-2} = (2^{-1}a)^2$, $a^{\dagger}$ and a being an obvious generalization of the usual creation and annihilation operators.

The previous results immediately lead, through Eq. (50), to the definition of a linear operator $\hat{\mathbf{S}}$:

$$\begin{aligned} \hat{\mathbf{S}} = \sum_{k=0}^{\infty} \frac{\hat{A}^{(k)}s^k}{k!} = \sum_{n=0}^{\infty} \frac{\frac{1}{4}(a^{+2} + a^2)^n s^n}{n!} \\ = \exp[\tfrac{1}{4}(a^{+2} + a^2)s], \end{aligned} \tag{54}$$

which furnishes the expression of the final state by operating on the initial one.

5. APPLICATIONS TO THE COHERENT STATES

The coherent states of a time-independent harmonic oscillator of frequency ω_0 are defined[4] as the eigenkets of the annihilation operator a with complex eigenvalues

$$a\,|\alpha\rangle_c = \alpha\,|\alpha\rangle_c. \tag{55}$$

The expression of the $|\alpha\rangle_c$'s is usually given in the Heisenberg picture and reads in terms of the energy

eigenstates $|n\rangle$ as

$$|\alpha\rangle_c = \exp(-\tfrac{1}{2}|\alpha|^2)\sum_{n=0}^{\infty}\alpha^n(n!)^{-\frac{1}{2}}|n\rangle. \quad (56)$$

In the Schrödinger picture and in our notation, they read

$$|\alpha(t)\rangle_c$$
$$= \exp(-\tfrac{1}{2}|\alpha|^2)\sum_{n=0}^{\infty}\{\alpha\exp[-i\omega_0(t-t_0)]^n\}(n!)^{-\frac{1}{2}}|n;t_0\rangle$$
$$= |\alpha\exp[-i\omega_0(t-t_0)]\rangle_c. \quad (57)$$

Thus, the coherent state can be represented in a c plane as a vector of constant absolute magnitude rotating at constant rate. On the other hand, by anticipating the conclusions of this section, in the case of time-dependent frequency we can affirm that our state remains coherent in the adiabatic limit and evolves in such a way as to maintain its modulus constant to all orders in ϵ, while it rotates with instantaneous angular velocity equal on the average to $\omega(t)$ to the first order in ϵ.

As a matter of fact, assuming

$$a_n(-\infty) = \alpha^n(n!)^{-\frac{1}{2}}\exp(-\tfrac{1}{2}|\alpha|^2) \quad (58)$$

and applying Eq. (45), we get, to within an unessential phase factor,

$$a_n(+\infty) \sim \exp(-\tfrac{1}{2}|\alpha|^2)\alpha^n(n!)^{\frac{1}{2}}\exp(i\eta\varphi_{\eta^*}). \quad (59)$$

Then, the final expression of the initially coherent states reads

$$|\alpha(+\infty)\rangle_c \sim \exp(-\tfrac{1}{2}|\alpha|^2)\sum_{n=0}^{\infty}\alpha_f^n(n!)^{-\frac{1}{2}}|n;+\infty\rangle, \quad (60)$$

where $\alpha_f = \alpha\exp(-i\int_{-\infty}^{+\infty}\omega(t)\,dt + i\varphi_{\eta^*})$. This formally proves our assertion that the state remains coherent. On the other hand, since $\lim_{\epsilon\to 0}\varphi_{\eta^*} = 0$, as can be easily proved by observing that the leading term of the asymptotic expansion of $S_{\alpha\alpha}(+\infty, -\infty)$ is equal to unity and resorting to Eq. (49), we can further write

$$|\alpha(+\infty)\rangle_c = \left|\alpha\exp\left(-i\int_{-\infty}^{\infty}\omega(t)\,dt\right)\right\rangle_c + O(\epsilon). \quad (61)$$

Let us finally apply the results of Sec. 4 to the evolution of a coherent state. In the sudden case this is described by Eq. (54) through

$$|\alpha(t_{0+})\rangle = \hat{S}\,|\alpha(t_{0-})\rangle_c$$
$$= \exp[\tfrac{1}{4}(a^{+2} + a^2)s]\,|\alpha(t_{0-})\rangle_c. \quad (62)$$

In the case of $s \ll 1$, this assumes the particularly compact form

$$|\alpha(t_{0+})\rangle \cong \left(1 + \tfrac{1}{4}s\alpha_0^2 - \tfrac{1}{4}s\exp(-\tfrac{1}{2}|\alpha_0|^2)\frac{\partial^2}{\partial\alpha_0^2}\right.$$
$$\left.\times\exp(\tfrac{1}{2}|\alpha_0|^2)\right)|\alpha(t_{0-})\rangle_c, \quad (63)$$

having introduced $\alpha_0 \equiv \alpha(t_{0-})$ and where use has been made of Eq. (55) and of the relation[12]

$$a^{+2}|\alpha_0\rangle_c = \exp(-\tfrac{1}{2}|\alpha_0|^2)\frac{\partial^2}{\partial\alpha_0^2}\exp(\tfrac{1}{2}|\alpha_0|^2)\,|\alpha_0\rangle_c. \quad (64)$$

Equations (61) and (62) summarize the behavior of an initially coherent state in the two opposite cases of adiabatic and sudden variation of the frequency.

ACKNOWLEDGMENTS

We wish to thank Dr. F. Engelmann for many helpful suggestions and H. R. Lewis and W. B. Riesenfeld for letting us have a preprint of their work on the explicitly time-dependent invariants.

[12] See, for example, C. L. Mehta, P. Chaud, E. C. G. Sudarshan, and R. Vedam, Phys. Rev. 157, 1198 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Alternative Dynamics for Classical Relativistic Particles

AMNON KATZ
Physics Department, Weizmann Institute, Rehovot, Israel

(Received 19 February 1969)

Equations of motion for interacting classical relativistic particles (different from Van Dam–Wigner equations) are derived from an action principle. In the classical limit they reduce to Newton's equations with an interparticle potential depending on distance only.

1. INTRODUCTION

In this paper we propose a dynamical theory for interacting classical relativistic particles that is different from the one proposed by Van Dam and Wigner.[1]

We derive our equations of motion via the variation principle from a manifestly invariant action [Eq. (2.1)] in Sec. 2. In Sec. 3, we exhibit a conserved momentum–energy vector. Section 4 deals with the nonrelativistic limit of the theory. In the limit, Newtonian equations of motion are recovered with an interparticle potential that depends only on the interparticle distance. The relation between the classical potential that appears in this limit and the arbitrary function characterizing the interaction in the relativistic equations is studied in Sec. 5.

The original Van Dam–Wigner[1] equations can also be derived from an action (which is a natural generalization of the Tetrode–Fokker[2,3] action leading to electrodynamics with time-symmetric interactions). This is done in Appendix A. The nonrelativistic limit of the Van Dam–Wigner equations is identical to that of the equations of motion proposed in the present paper; both theories are generalizations of Newtonian mechanics. Therefore, the treatment of the limit and of the classical potential that appears here is relevant also to the original Van Dam–Wigner theory. This treatment is not available in the original paper.[4] Conversely, many of the remarks of the original paper, concerning the Cauchy problem and approximation schemes for collision processes, are relevant to our equations of motion; they are not repeated here.

A third admissible form of the action is investigated in Appendix B. It is shown that no new equations of motion result.

[1] H. Van Dam and E. P. Wigner, Phys. Rev. **138**, B1576 (1965).
[2] H. Tetrode, Z. Physik **10**, 317 (1922); A. D. Fokker, *ibid.* **58**, 386 (1929); Physica **9**, 33 (1929); **12**, 145 (1932).
[3] J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. **21**, 425 (1949).
[4] The conclusion of Van Dam and Wigner, "The interaction ... reduces to the gravitational interaction in the nonrelativistic limit and if the distances between the particles ... are always large," is rather empty since it refers to an expansion of the classical potential in $1/r$. The full classical potential is discussed here in Secs. 4 and 5.

2. EQUATIONS OF MOTION

We postulate the action integral

$$A = m_1c^2\int(\dot{x}_1^2)^{\frac{1}{2}}\,d\tau_1 + m_2c^2\int(\dot{x}_2^2)^{\frac{1}{2}}\,d\tau_2 + \tfrac{1}{2}\iint\chi((x_1-x_2)^2)(\dot{x}_1^2)^{\frac{1}{2}}(\dot{x}_2^2)^{\frac{1}{2}}\,d\tau_1\,d\tau_2. \quad (2.1)$$

x_1 and x_2 are the four-vector positions of the particles labeled 1 and 2; they depend on the parameters τ_1 and τ_2, respectively; $\dot{x}_i$ is the derivative of x_i with respect to τ_i. The generalization to any number of particles is immediate. (For alternative interaction terms, see Appendices A and B.)

There is a great amount of arbitrariness in the choice of a parameter in the parametric description of the world line $x_i(\tau_i)$. We limit this arbitrariness by imposing the condition

$$\dot{x}_1^2 = c^2, \quad (2.2)$$

which makes τ_i the proper time of particle i. (Our metric is $g_{00} = c^2$, $g_{rs} = \delta_{rs}$, $r, s = 1, 2, 3$, $g_{0r} = 0$.) Once the variation of the action is performed, Eq. (2.2) will be substituted in all expressions.

The function χ which describes the interaction is required to satisfy

$$\chi(p) = 0, \quad \text{when} \quad p > 0. \quad (2.3)$$

The interaction occurs only at spacelike separations, which is the relativistic generalization of instantaneous interaction.[1]

The variation of A with respect to the world lines of the particles gives rise to the following equations of motion:

$$\left(m_1 + \frac{1}{2c}\int\chi((x_1-x_2)^2)\,d\tau_2\right)\ddot{x}_1^\mu = \left(\delta_\nu^\mu - \frac{\dot{x}_1^\mu\dot{x}_{1\nu}}{c^2}\right)\int\chi'((x_1-x_2)^2)(x_1^\nu - x_2^\nu)c\,d\tau_2 \quad (2.4)$$

and the corresponding equation for $\ddot{x}_2$, obtainable from Eq. (2.4) by exchanging the indices 1 and 2. The τ_2 integration is along the whole world line of particle 2. χ' is the derivative of χ with respect to its argument.

The summation convention is used. Consistency of the equation of motion (2.4) with the condition (2.2) is apparent.

3. ENERGY AND MOMENTUM

The total energy and momentum of the system may be described by the four-vector:

$$P(\tau_1, \tau_2) = \left(m_1 + \frac{1}{2c}\int \chi\, d\tau_2\right)\dot{x}_1 + \left(m_2 + \frac{1}{2c}\int \chi\, d\tau_1\right)\dot{x}_2 - \left(\int_{-\infty}^{\tau_1}\int_{\tau_2}^{\infty} - \int_{\tau_1}^{\infty}\int_{-\infty}^{\tau_2}\right) c(x_1 - x_2)\chi'\, d\tau_1'\, d\tau_2'. \tag{3.1}$$

The argument of χ and χ' in the last expression is still $(x_1 - x_2)^2$; x_1, x_2 in the integrand in the last term depend on τ_1', τ_2' rather than τ_1, τ_2. It is straightforward to verify that, by virtue of the equations of motion (2.4),

$$\frac{d}{d\tau_1} P(\tau_1, \tau_2) = \frac{d}{d\tau_2} P(\tau_1, \tau_2) = 0. \tag{3.2}$$

$P(\tau_1, \tau_2)$ is thus a constant of the motion, independent of τ_1 and τ_2.

4. NONRELATIVISTIC LIMIT

The nonrelativistic limit may be reached by letting $c \to \infty$. Dropping terms that obviously tend to zero in Eq. (2.4), we are left with

$$m_1\ddot{x}_1 = \int \chi' c\, d\tau_2 (x_1 - x_2) = \frac{\partial}{\partial x_1}\frac{1}{2}\int \chi c\, d\tau_2. \tag{4.1}$$

In the nonrelativistic limit, the proper time τ_i becomes identical with the ordinary time x_i^0 up to order $1/c^2$:

$$x_i^0 = \tau_i + f_i(\tau_i)/c^2. \tag{4.2}$$

The integral on the right-hand side of Eq. (4.1) becomes

$$\tfrac{1}{2}\int \chi\{c^2(\tau_1 - \tau_2)^2 + 2(\tau_1 - \tau_2)(f_1 - f_2) - [\mathbf{x}_1(\tau_1) - \mathbf{x}_2(\tau_2)]^2\} c\, d\tau_2. \tag{4.3}$$

We now introduce as a new variable of integration

$$\theta = c(\tau_2 - \tau_1), \tag{4.4}$$

which turns the integral (4.3) into

$$\tfrac{1}{2}\int \chi\left\{\theta^2 + 2\frac{\theta}{c}(f_1 - f_2) - \left[\mathbf{x}_1(\tau_1) - \mathbf{x}_2\left(\tau_1 + \frac{\theta}{c}\right)\right]^2\right\} d\theta. \tag{4.5}$$

As the limiting process $c \to \infty$ is completed, this is reduced to

$$\tfrac{1}{2}\int \chi\{\theta^2 - [\mathbf{x}_1(\tau_1) - \mathbf{x}_2(\tau_1)]^2\}\, d\theta. \tag{4.6}$$

The space components of Eq. (4.1) may now be rewritten as

$$m_1\ddot{\mathbf{x}}_1 = -\frac{\partial}{\partial \mathbf{x}_1} V(|\mathbf{x}_1 - \mathbf{x}_2|), \tag{4.7}$$

where both $\mathbf{x}_1$ and $\mathbf{x}_2$ depend on the same time parameter and the potential V is defined by

$$V(r) \equiv \tfrac{1}{2}\int \chi(\theta^2 - r^2)\, d\theta. \tag{4.8}$$

The time component is satisfied identically, since both sides vanish.

As for the energy and momentum, we find in the limit that

$$P^0(\tau, \tau)c^2 - m_1c^2 - m_2c^2 \to \tfrac{1}{2}m_1\dot{\mathbf{x}}_1^2 + \tfrac{1}{2}m_2\dot{\mathbf{x}}_2^2 + V(|\mathbf{x}_1 - \mathbf{x}_2|), \tag{4.9}$$

$$\mathbf{P}(\tau, \tau) \to m_1\dot{\mathbf{x}}_1 + m_2\dot{\mathbf{x}}_2. \tag{4.10}$$

Newtonian equations of motion, governed by an interparticle potential depending on interparticle distance only, are thus regained in the nonrelativistic limit. The same is true for the Wigner–Van Dam dynamics (see Appendix A). The nonrelativistic limit cannot distinguish between the two forms of relativistic dynamics which are both generalizations of Newton's equations.

5. THE NEWTONIAN AND THE RELATIVISTIC POTENTIAL

Equation (4.8), which expresses the Newtonian potential $V(r)$ in terms of the "relativistic potential" χ, may be recast in the form

$$V(r) = \frac{1}{2}\int_0^{r^2} \frac{\chi(-p)\, dp}{(r^2 - p)^{\frac{1}{2}}}. \tag{5.1}$$

This relation may be inverted into[5]

$$\chi(-p) = -\frac{1}{\pi}\int_0^{p^{\frac{1}{2}}} \frac{V(r) - V(p^{\frac{1}{2}})}{(p - r^2)^{\frac{3}{2}}}\, 2r\, dr + \frac{2}{\pi}\frac{V(p^{\frac{1}{2}})}{p^{\frac{1}{2}}}. \tag{5.2}$$

The two potentials determine one another uniquely through the linear relations (5.1) and (5.2).

Note Added in Proof: The condition (2.3) could be relaxed to $X(p) \to 0$ as $p \to +\infty$ fast enough for all integrals used to converge. The author is indebted to P. Havas for pointing out the relevant work of Havas,[6] and Dettman and Schild.[7]

[5] I am indebted to M. Peshkin for help in deriving Eq. (5.2).
[6] P. Havas, Phys. Rev. **87**, 309 (1951).
[7] J. W. Dettman and A. Schild, Phys. Rev. **95**, 1057 (1954).

ACKNOWLEDGMENT

The author thanks Professor H. J. Lipkin for fruitful comments.

APPENDIX A: THE VAN DAM–WIGNER RELATIVISTIC DYNAMICS

The Van Dam–Wigner equations[1] may also be derived from an action integral similar to (2.1), namely

$$A = m_1c^2\int(\dot{x}_1^2)^{\frac{1}{2}}\,d\tau_1 + m_2c^2\int(\dot{x}_2^2)^{\frac{1}{2}}\,d\tau_2 + \tfrac{1}{2}\iint\chi((x_1-x_2)^2)\dot{x}_1\cdot\dot{x}_2\,d\tau_1\,d\tau_2. \quad \text{(A1)}$$

The expression $(\dot{x}_1^2)^{\frac{1}{2}}(\dot{x}_2^2)^{\frac{1}{2}}$ in (2.1) is here replaced by $\dot{x}_1\cdot\dot{x}_2$. The equation of motion that follows is

$$m_1\ddot{x}_1 = \int\chi' c\,d\tau_2\left\{\frac{\dot{x}_1}{c}\cdot\frac{\dot{x}_2}{c}(x_1-x_2) - (x_1-x_2)\cdot\frac{\dot{x}_1}{c}\frac{\dot{x}_2}{c}\right\}. \quad \text{(A2)}$$

This is the Van Dam–Wigner equation of motion. The correspondence between our χ and their ϕ is

$$\phi = \chi'. \quad \text{(A3)}$$

For the particular choice $\chi(p) = 2e_1e_2\delta(p)$, Eq. (A1) becomes the Tetrode–Fokker action,[2,3] and the corresponding theory is electrodynamics in which all interactions are time-symmetric (half-retarded, half-advanced) and there is no self-interaction.

The classical limit of Eq. (A2) may be derived analogously and is identical to that of Eq. (2.4) studied in Sec. 4.

The energy–momentum vector in the present case is[3]

$$P(\tau_1,\tau_2) = m_1\dot{x}_1 + m_2\dot{x}_2 + \left(\int_{-\infty}^{\tau_1}\int_{\tau_2}^{\infty} - \int_{\tau_1}^{\infty}\int_{-\infty}^{\tau_2}\right)\chi'\,d\tau_1'\,d\tau_2' \times\{(x_1-x_2)\dot{x}_1\cdot\dot{x}_2 - (x_1-x_2)\cdot\dot{x}_1\dot{x}_2 - (x_1-x_2)\cdot\dot{x}_2\dot{x}_1\}. \quad \text{(A4)}$$

The two kinds of interaction could coexist. One should then use an action integral with two interaction terms, one of the form of (A1) and the other of the form of (2.1), with two different functions χ. The classical potentials V in the nonrelativistic limit would simply be added.

APPENDIX B: ANOTHER INTERACTION TERM

In this appendix we examine the possibility of using another simple invariant form for the interaction term in the action integral, namely

$$\int\chi((x_1-x_2)^2)(x_1-x_2)\cdot\dot{x}_1(x_1-x_2)\cdot\dot{x}_2\,d\tau_1\,d\tau_2. \quad \text{(B1)}$$

An integration by parts on τ_1 transforms (B1) into

$$-\tfrac{1}{2}\int\chi^{(-1)}((x_1-x_2)^2)\dot{x}_1\cdot\dot{x}_2\,d\tau_1\,d\tau_2 + \left[\int\chi^{(-1)}((x_1-x_2)^2)(x_1-x_2)\cdot x_2\,d\tau_2\right]_{\tau_1=-\infty}^{\tau_1=\infty} = -\tfrac{1}{2}\int\chi^{(-1)}((x_1-x_2)^2)\dot{x}_1\cdot\dot{x}_2\,d\tau_1\,d\tau_2 + \left[\left[\chi^{(-2)}((x_1-x_2)^2)\right]_{\tau_2=-\infty}^{\tau_2=\infty}\right]_{\tau_1=-\infty}^{\tau_1=\infty}. \quad \text{(B2)}$$

This differs from the interaction term in Eq. (A1) by a quantity that is invariant to the variation of the world lines of the two particles. The theory resulting from use of (B1) as the interaction term is therefore identical to the Van Dam–Wigner theory with $\chi^{(-1)}$ replacing χ. [$\chi^{(-1)}$ is a primitive of χ—a function whose derivative is χ. This determines $\chi^{(-1)}$ only up to a constant. We adjust the constant so as to insure $\chi^{(-1)}(p) = \chi(p) = 0$ when $p > 0$.]

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

New Methods for Reduction of Group Representations. IV*

J. R. GABRIEL
Applied Mathematics Division, Argonne National Laboratory, Argonne, Illinois

(Received 31 March 1969)

Some of the methods given in an earlier paper are extended; in particular, it is shown how to construct the central from the matrix representatives of the group generators and how to apply the techniques given for the central to simplify the reduction of the representation.

1. INTRODUCTION

Previous papers[1–3] (these will be called I, II, and III in this paper) have shown how to use the matrix representatives in a space S of the generators of a group G to construct matrices which span the representation in S of $\mathcal{C}$, the commutator algebra of G, and how to use the representation of $\mathcal{C}$ to reduce S into irreducible invariant subspaces with respect to G.

The methods for constructing the representation of $\mathcal{C}$ (for the rest of this paper, we shall not distinguish between $\mathcal{C}$ and its representation, or G and its representation) work very well in practice, but the use of $\mathcal{C}$ to reduce S in the manner suggested in III proves inconvenient because of the difficulty in guaranteeing that all eigenvectors belonging to a given multiple eigenvalue have been found, since if S is reducible but not fully reducible, elements of $\mathcal{C}$ do not necessarily have linear elementary divisors.[4]

In this paper we shall give methods for constructing $\mathcal{E}$ (the central of $\mathcal{C}$ which contains all the elements of $\mathcal{C}$ which commute with every element of $\mathcal{C}$), the explicit construction of primitive idempotents for $\mathcal{E}$ without the use of eigenvectors, and the use of related techniques to reduce $\mathcal{C}$.

2. CONSTRUCTION OF THE CENTRAL

Suppose that the techniques of Ref. 3 (hereafter called III) have been used to construct a set of matrices u_i which span $\mathcal{C}$. To construct $\mathcal{E}$, we need to find every linear combination

$$\zeta = \sum_i x_i u_i, \tag{2.1}$$

such that

$$\zeta u_j = u_j \zeta, \quad \text{for all } j. \tag{2.2}$$

Because $\mathcal{C}$ is an algebra,

$$u_i u_j = \sum_k \sigma_{ij}^k u_k, \tag{2.3}$$

where the σ_{ij}^k are numbers. Since we may assume that the u_i are orthogonal in the sense of III, i.e.,

$$\mathrm{Tr}\,(u_i^\dagger u_j) = \delta_{ij},$$

the σ_{ij}^k may be readily calculated as

$$\sigma_{ij}^k = \mathrm{Tr}\,(u_k^\dagger u_i u_j). \tag{2.4}$$

Equation (2.2) in terms of the σ's becomes

$$\sum_i x_i(\sigma_{ij}^k - \sigma_{ji}^k) = 0, \quad \text{for all } j, k. \tag{2.5}$$

That is, it becomes a set of homogeneous, simultaneous equations for the x_i whose solutions define matrices

$$v_j = \sum_i x_i^j u_j, \tag{2.6}$$

where the vector x_i^j, $i = 1, 2, \cdots, n_c$, is the jth linearly independent solution of (2.5).

3. THE REDUCTION OF S WITH RESPECT TO $\mathcal{E}$

Suppose that the n_e matrices v_i span $\mathcal{E}$, and they have been made orthonormal in the sense that

$$\mathrm{Tr}\,(v_i^\dagger v_j) = \delta_{ij}. \tag{3.1}$$

Because the elements of $\mathcal{E}$ commute with each other and every element of $\mathcal{C}$, they must have linear elementary divisors. It follows that $\mathcal{E}$ has n_e primitive idempotents e_j and

$$v_i = \sum_{j=1}^{n_e} \lambda_{ij} e_j. \tag{3.2}$$

The e_j are themselves members of $\mathcal{E}$, i.e.,

$$e_j = \sum_{i=1}^{n_e} \mu_{ji} v_j \tag{3.3}$$

and our task is to find the coefficients μ_{ji}.

Let

$$v_i v_j = \sum_k \sigma_{ij}^k v_k. \tag{3.4}$$

By (3.1), we have

$$\sigma_{ij}^k = \mathrm{Tr}\,(v_k^\dagger v_i v_j). \tag{3.5}$$

Because the e_i are primitive idempotents,

$$e_i e_j = \delta_{ij} e_i \tag{3.6}$$

* Work performed under the auspices of the U.S. Atomic Energy Commission.

[1] J. R. Gabriel, J. Math. Phys. **5**, 494 (1964).
[2] J. R. Gabriel, J. Math. Phys. **9**, 973 (1968).
[3] J. R. Gabriel, J. Math. Phys. **10**, 1789 (1969).
[4] J. H. Wilkinson, *The Algebraic Eigenvalue Problem* (Oxford University Press, London, 1967), p. 12.

and, therefore,

$$v_i e_j = \lambda_{ij} e_j . \tag{3.7}$$

By (3.3) and (3.7),

$$v_i \sum_k \mu_{jk} v_k = \lambda_{ij} \sum_l \mu_{jl} v_l . \tag{3.8}$$

Using (3.1) and (3.4),

$$\sum_k \mu_{jk} \sigma^l_{ik} = \lambda_{ij} \mu_{jl} . \tag{3.9}$$

Writing

$$\sigma^l_{ik} = A^i_{lk}, \quad \mu_{jk} = x^j_k , \tag{3.10}$$

and $\lambda_{ij} = \lambda^i_j$, Eq. (3.9) becomes

$$\sum_k A^i_{lk} x^i_k = \lambda^i_j x^j_l , \tag{3.11}$$

i.e., x is an eigenvector of A^i belonging to the eigenvalue λ^i_j. Thus, if A^i has n_e distinct eigenvalues, the eigenvectors define the idempotents e_i.

Since the x^j must be simultaneous eigenvectors of all the A^i, any degeneracy for one i value may be resolved by examining another of the A^i.

Because the eigenvalues of the A^i are the same as the coefficients λ_{ij} in the expansion (3.2) of the v_i, the algebra of the matrices A^i is isomorphic with the algebra $\mathfrak{E}$ of the v_i.

4. THE REDUCTION OF S WITH RESPECT TO $\mathfrak{C}$ AND G

The primitive idempotents e_i of $\mathfrak{E}$ project subspaces S_i out of S. By Schur's lemma and the arguments of Lemma 2 of I or the arguments of Weyl,[5] it follows that each S_i is invariant under both $\mathfrak{C}$ and G and that S_i affords one or more equivalent irreducible representations of $\mathfrak{C}$ and one or more equivalent irreducible representations of G. If the dimension of the typical irreducible representation of G in S_i is n_g and the dimension of the typical irreducible representation of $\mathfrak{C}$ in S is n_c, then the dimension of S_i is $n_c \times n_g$, i.e., the representation Γ^i_g of G occurs n_c times and the representation Γ^i_c of $\mathfrak{C}$ occurs n_g times.

If S_i is fully reducible, then Γ_c is isomorphic to $GL(n_c)$ and Γ_g is isomorphic to $GL(n_g)$. Let $\mathfrak{C}^i$ be the representation of $\mathfrak{C}$ in S_i.

Then if the matrices u_j span $\mathfrak{C}$ in S, the matrices $e_i u_j$ span $\mathfrak{C}^i$ in S_i. They are not linearly independent, but a set of linearly independent matrices which spans $\mathfrak{C}^i$ can be constructed from them. Call these matrices

$$u^i_k = \sum_j t_{kj} e_i u_j , \tag{4.1}$$

where the t_{kj} are numbers.

For the rest of this section, we shall deal with the space S_i alone so that the label i can be dropped. Thus we have a set of matrices u_k which spans $\mathfrak{C}$ and operates in a space S. The irreducible components Γ_c of $\mathfrak{C}$ are equivalent and occur n_g times.

Augment $\mathfrak{C}$ by its Hermitian conjugate $\mathfrak{C}^\dagger$, let $\mathfrak{C}_H$ be the Hermitian subalgebra of $\mathfrak{C} \cup \mathfrak{C}^\dagger$, and let u be an arbitrary element of $\mathfrak{C}_H$. The techniques of Sec. 2 may then be used to construct the subalgebra $\mathcal{A}$ of $\mathfrak{C}_H$ which commutes with u. If u has n_c distinct eigenvalues, then u and all the elements of $\mathcal{A}$ may be made simultaneously diagonal by the same change of basis. If u has less than n_c distinct eigenvalues, then $\mathcal{A}$ has more than n_c elements and they do not all commute. In this case, choose an element of $\mathcal{A}$ which does not commute with every element of $\mathcal{A}$; call it u', and restrict $\mathcal{A}$ to those elements that commute with u'. This process may be repeated until $\mathcal{A}$ contains just n_c elements, all of which commute.

Having found the $\mathcal{A}$, the techniques of Sec. 3 may be used to construct the primitive idempotents e_k which divide S into irreducible subspaces with respect to $\mathfrak{C}$. Let the space projected out of S by e_k be S_k. Then we know that each space S_k affords the same irreducible representation of $\mathfrak{C}$. If we choose S_1 to afford the "standard form," then the bases in S_2, $S_3, \cdots, S_{n_g}$ must be changed to put them also in standard form. This can be done by examining the matrices

$$V_{ij} = e_1 u_i e_j . \tag{4.2}$$

Schur's lemma asserts that there exists a change of basis in S_j which displays V_{ij} as a multiple of the $n_c \times n_c$ unit submatrix in the $1,j$ block, and zeros elsewhere. If the multiplier is zero, then $V_{ij} = 0$; if not, Gauss's condensation will give the necessary transformation.

5. SUMMARY

This seems to be an appropriate place to summarize the results and evaluate the significance of Papers I–IV in this series, since it is the last of the set and completes the task set in I, namely, the reduction of a group representation using the absolute minimum of information so that no unnecessary matrices are constructed or computations done.

The work was started with the aim of computation of antisymmetrized molecular wavefunctions.[6] The first attempt[6] proved unsatisfactory because of the impracticability of computing all $f!$ matrices R for substitution in

$$e^\alpha_{ij} = \sum \Gamma^\alpha_{ij}(R^{-1})R, \tag{5.1}$$

but a conversation with Powell in 1960[7] turned the

[5] H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, N.J., 1946), Chap. III.

[6] J. R. Gabriel, Proc. Cambridge Phil. Soc. 57, 330 (1961).

[7] See Ref. 1, p. 504.

author's attention to the group generators[1] as the absolute minimum of information needed to define a representation. A first attempt, stated in terms of a physical discussion of the set of all possible Hamiltonians that show a given symmetry, was submitted for publication and finally withdrawn. Reflection on the mathematical content of that paper led to the conclusion that the commutator algebra had to be the central theme of any line of attack using the generators, and this resulted in Paper I.[1] How to construct the commutator algebra and use it for the reduction was not clearly understood in I, but Secs. 1 and 3, together with Lemmas 1 and 2 of Sec. 2 of I, are still valid and useful.

Paper II was an unsuccessful attempt to improve the unsatisfactory parts of I. Although formally valid, it does not address the real problems.

Paper III is the first real success in constructing the commutator algebra $\mathfrak{C}$ with a minimum of labor, and the use of $\mathfrak{C}$ to reduce the space in III is valid, but not very satisfactory from a computational point of view. This is because it is oriented towards using eigenvectors of matrices which have many repeated eigenvalues. Whenever a repeated eigenvalue occurs, it is a mistake to search for eigenvectors. Instead one should be calculating projection operators for the associated spaces, since they are unique, while the vectors are not.[8] Trying to find projection operators instead of vectors led to the results of Secs. 3 and 4 of this paper where the smallest matrix algebra isomorphic to $\mathfrak{C}$ is analyzed instead of $\mathfrak{C}$. This means that eigenvalues which appear repeatedly in $\mathfrak{C}$ only appear once in the analysis, and projection operators appear instead of eigenvectors. Since the eigenvalue calculation for an $n \times n$ matrix takes a time proportional to at least n^3, this compression is very significant for practical purposes.

Finally, an outline of future plans: This paper essentially completes the method proposed in Sec. 1 of I. The application to antisymmetrized molecular wavefunctions is discussed in another paper.[9,10] Now, all that remains is to write the computer programs to implement the method. This is being done at Argonne National Laboratory.

ACKNOWLEDGMENTS

Many of my colleagues at Argonne have listened while I tried to clarify my thoughts on papers III and IV, and I am grateful to them all. But I would particularly like to thank my wife Marian, who has listened more than anyone else, and Dr. Wallace Givens who has not only been technically helpful on the algebraic eigenvalue problem, but also encouraged this work as one of the projects of the Mathematics Division in computer programming research.

[8] T. Kato, Progr. Theoret. Phys. (Kyoto) 4, 514 (1949).

[9] J. R. Gabriel, J. Chem. Phys. 51, 1533 (1969).

[10] J. R. Gabriel, J. Chem. Phys. (to be published).

Relations between the Elements of the Phase Matrix for Scattering*

K. D. ABHYANKAR† AND A. L. FYMAT
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

(Received 3 February 1969)

The process of scattering of radiation is usually characterized by a 4 × 4 transformation matrix which relates the radiation field vector incident on the scatterer to the scattered field vector. The nine relations between the 16 elements of this phase matrix for scattering are derived explicitly for the three most commonly used representations of the intensity vector, viz., Wolf's coherency matrix formalism, Chandrasekhar's and Stokes's representations. The invariance of these relations under the action of any optical train containing one or more elements characterized by their Jones representation is demonstrated. These relations should be useful in the theory of polarization optics. The same relations are also shown to hold after rotation of the axes of reference for the electric vectors in the incident and scattered beams. Since such a transformation is required in the formulation of the theory of radiative transfer, the relations derived here may find use in multiple scattering problems as well.

I. INTRODUCTION

The process of scattering of electromagnetic radiation is characterized by a transformation matrix which relates the radiation field vector incident on the scatterer to the scattered field vector. If the incident and scattered electric vectors $\mathbf{E}^{(i)}$ and $\mathbf{E}^{(s)}$ are represented by the column matrices

$$\mathbf{E}^{(i)} = \begin{bmatrix} E_l \\ E_r \end{bmatrix}^{(i)} \quad \text{and} \quad \mathbf{E}^{(s)} = \begin{bmatrix} E_l \\ E_r \end{bmatrix}^{(s)}, \tag{1}$$

where the subscripts l and r refer to directions parallel and perpendicular to the plane of scattering, respectively, we have the relation

$$\mathbf{E}^{(s)} = \mathsf{A} \cdot \mathbf{E}^{(i)} \tag{2}$$

where A is the amplitude (scattering) matrix,

$$\mathsf{A} = \begin{bmatrix} A_{ll} & A_{lr} \\ A_{rl} & A_{rr} \end{bmatrix} = \begin{bmatrix} A_2 & A_3 \\ A_4 & A_1 \end{bmatrix} \tag{3}$$

with

$$A_j = \alpha_j e^{i\theta_j}, \quad j = 1, \cdots, 4. \tag{4}$$

On the other hand, one is usually concerned with the intensity and state of polarization of the fields, which can be represented by four vectors derived from the corresponding electric vectors. If $\mathbf{I}^{(i)}$ and $\mathbf{I}^{(s)}$ are these intensity vectors for the incident and scattered beams, respectively, then, within a constant multiplying factor, in place of Eq. (2) we have the relation

$$\mathbf{I}^{(s)} = \mathsf{P} \cdot \mathbf{I}^{(i)}, \tag{5}$$

where P is known as the phase matrix for scattering. The 16 elements P_{ij}, $i,j = 1, \cdots, 4$, of P are functions of the elements of A. Since the four complex elements of A contain only seven independent constants a_1, a_2, a_3, a_4, $(\theta_2 - \theta_1)$, $(\theta_3 - \theta_2)$, and $(\theta_4 - \theta_3)$, it is obvious that there must be nine relations between the 16 elements of P.[1] On account of their intrinsic interest in scattering theory, these relations are derived here for the most commonly used representations of the intensity vector.

II. RELATIONS BETWEEN THE ELEMENTS OF THE COHERENCY PHASE MATRIX $^{j}\mathsf{P}$

In optics one uses the Stokes representation of the intensity vector, whereas, in the theory of radiative transfer, Chandrasekhar's[2] representation is found more advantageous. It has been shown by Marathay[3] and Sekera[4] that both representations can be derived from the more fundamental Wolf's[5] coherency matrix formalism.

Defining the column vector (hereafter called coherency vector) $\mathbf{J}$ by

$$\mathbf{J} = \langle \mathbf{E} \times \mathbf{E}^* \rangle = \begin{bmatrix} J_{ll} \\ J_{lr} \\ J_{rl} \\ J_{rr} \end{bmatrix},$$

where $\times$ represents Kronecker product, * denotes complex conjugate, and $\langle \cdots \rangle$ stands for time average, we obtain, from Eq. (2),

$$\mathbf{J}^{(s)} = {}^{j}\mathsf{P} \cdot \mathbf{J}^{(i)}, \tag{6}$$

* This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. NAS 7-100, sponsored by the National Aeronautics and Space Administration.

† Senior Postdoctoral Resident Research Associate of NRC–NASA. On leave of absence from Osmania University, Hyderabad, India.

[1] H. C. Van de Hulst, *Light Scattering by Small Particles* (John Wiley & Sons, Inc., New York, 1957), Chap. 5.

[2] S. Chandrasekhar, *Radiative Transfer* (Clarendon Press, Oxford, England, 1950), Chap. 1.

[3] A. S. Marathay, J. Opt. Soc. Am. **55**, 969 (1965); see also E. L. O'Neill, *Introduction to Statistical Optics* (Addison-Wesley Publ. Co., Reading, Mass., 1963), Chap. 9.

[4] Z. Sekera, J. Opt. Soc. Am. **56**, 1732 (1966).

[5] E. Wolf, Proc. Roy. Soc. **A230**, 246 (1955); see also M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Inc., New York, 1959), Chap. 10.

where

$$^{j}\mathsf{P} = \mathsf{A} \times \mathsf{A}^{*} \tag{7}$$

is the coherence phase matrix for scattering. Substituting Eqs. (3) and (4) in Eq. (7), we get

$$\begin{bmatrix} {}^{j}P_{11} & {}^{j}P_{12} & {}^{j}P_{13} & {}^{j}P_{14} \\ {}^{j}P_{21} & {}^{j}P_{22} & {}^{j}P_{23} & {}^{j}P_{24} \\ {}^{j}P_{31} & {}^{j}P_{32} & {}^{j}P_{33} & {}^{j}P_{34} \\ {}^{j}P_{41} & {}^{j}P_{42} & {}^{j}P_{43} & {}^{j}P_{44} \end{bmatrix} = \begin{bmatrix} \alpha_2^2 & \alpha_2\alpha_3 e^{-i(\theta_3-\theta_2)} & \alpha_2\alpha_3 e^{i(\theta_3-\theta_2)} & \alpha_3^2 \\ \alpha_2\alpha_4 e^{-i(\theta_4-\theta_2)} & \alpha_1\alpha_2 e^{i(\theta_2-\theta_1)} & \alpha_3\alpha_4 e^{-i(\theta_4-\theta_3)} & \alpha_1\alpha_3 e^{i(\theta_3-\theta_1)} \\ \alpha_2\alpha_4 e^{i(\theta_4-\theta_2)} & \alpha_3\alpha_4 e^{i(\theta_4-\theta_3)} & \alpha_1\alpha_2 e^{-i(\theta_2-\theta_1)} & \alpha_1\alpha_3 e^{-i(\theta_3-\theta_1)} \\ \alpha_4^2 & \alpha_1\alpha_4 e^{i(\theta_4-\theta_1)} & \alpha_1\alpha_4 e^{-i(\theta_4-\theta_1)} & \alpha_1^2 \end{bmatrix}, \tag{8}$$

where it is seen that four of the 16 elements of $^{j}\mathsf{P}$ are real quantities, while the remaining 12 elements form six complex conjugate pairs as follows:

$$\begin{aligned} &{}^{j}P_{13} = {}^{j}P_{12}^{*}, \quad {}^{j}P_{31} = {}^{j}P_{21}^{*}, \quad {}^{j}P_{42} = {}^{j}P_{43}^{*}, \\ &{}^{j}P_{34} = {}^{j}P_{24}^{*}, \quad {}^{j}P_{33} = {}^{j}P_{22}^{*}, \quad {}^{j}P_{23} = {}^{j}P_{32}^{*}. \end{aligned} \tag{9}$$

We note that, in general, the product of the end elements of a column, a row, or a diagonal of $^{j}\mathsf{P}$ is equal to the product of the corresponding intermediate elements. There are ten such relations out of which only eight are independent; they are

$$^{j}P_{11}\,{}^{j}P_{14} = {}^{j}P_{12}\,{}^{j}P_{13}, \tag{10a}$$

$$^{j}P_{41}\,{}^{j}P_{44} = {}^{j}P_{42}\,{}^{j}P_{43}, \tag{11a}$$

$$^{j}P_{11}\,{}^{j}P_{41} = {}^{j}P_{21}\,{}^{j}P_{31}, \tag{12a}$$

$$^{j}P_{14}\,{}^{j}P_{44} = {}^{j}P_{24}\,{}^{j}P_{34}, \tag{13a}$$

$$^{j}P_{11}\,{}^{j}P_{44} = {}^{j}P_{22}\,{}^{j}P_{33}, \tag{14a}$$

$$^{j}P_{14}\,{}^{j}P_{41} = {}^{j}P_{23}\,{}^{j}P_{32}, \tag{15a}$$

$$^{j}P_{12}\,{}^{j}P_{42} = {}^{j}P_{22}\,{}^{j}P_{32}, \tag{16}$$

$$^{j}P_{21}\,{}^{j}P_{24} = {}^{j}P_{22}\,{}^{j}P_{23}. \tag{17}$$

The other two relations of this type are obtained by taking the complex conjugates of Eqs. (16) and (17). The ninth independent relation between the 16 elements of $^{j}\mathsf{P}$, which is of a different kind, is

$$^{j}P_{12}\,{}^{j}P_{43} = {}^{j}P_{21}\,{}^{j}P_{34}. \tag{18}$$

The relations (10a)–(15a) involve only real quantities, whereas the relations (16)–(18) equate complex expressions. The latter can be reduced to real equations by means of Eq. (9) as

$$^{j}P_{12}\,{}^{j}P_{42}\,{}^{j}P_{33}\,{}^{j}P_{23} = {}^{j}P_{22}\,{}^{j}P_{32}\,{}^{j}P_{13}\,{}^{j}P_{43}, \tag{19a}$$

$$^{j}P_{21}\,{}^{j}P_{24}\,{}^{j}P_{33}\,{}^{j}P_{32} = {}^{j}P_{22}\,{}^{j}P_{23}\,{}^{j}P_{31}\,{}^{j}P_{34}, \tag{20a}$$

$$^{j}P_{12}\,{}^{j}P_{43}\,{}^{j}P_{31}\,{}^{j}P_{24} = {}^{j}P_{21}\,{}^{j}P_{34}\,{}^{j}P_{13}\,{}^{j}P_{42}. \tag{21a}$$

The relations corresponding to Eqs. (10a)–(21a) for the other representations are derived in the next section.

III. RELATIONS BETWEEN THE ELEMENTS OF THE CHANDRASEKHAR AND STOKES PHASE MATRICES $^{c}\mathsf{P}$ AND $^{s}\mathsf{P}$

In his theory of radiative transfer, Chandrasekhar[2] uses the following components of the intensity vector **C**:

$$\begin{aligned} I_l &= \langle E_l E_l^{*} \rangle, \\ I_r &= \langle E_r E_r^{*} \rangle, \\ U &= \langle E_l E_r^{*} \rangle + \langle E_r E_l^{*} \rangle, \\ V &= -i(\langle E_l E_r^{*} \rangle - \langle E_r E_l^{*} \rangle), \end{aligned} \tag{22}$$

while the components of the original Stokes vector **S** are

$$\begin{aligned} I &= I_l + I_r = \langle E_l E_l^{*} \rangle + \langle E_r E_r^{*} \rangle, \\ Q &= I_l - I_r = \langle E_l E_l^{*} \rangle - \langle E_r E_r^{*} \rangle, \\ U &= \langle E_l E_r^{*} \rangle + \langle E_r E_l^{*} \rangle, \\ V &= -i(\langle E_l E_r^{*} \rangle - \langle E_r E_l^{*} \rangle). \end{aligned} \tag{23}$$

They can be obtained from **J** by the transformations[3,4]

$$\mathbf{C} = \mathsf{T}_c \cdot \mathbf{J}, \tag{24}$$

$$\mathbf{S} = \mathsf{T}_s \cdot \mathbf{J}, \tag{25}$$

where

$$\mathsf{T}_c = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{bmatrix} \tag{26}$$

and

$$\mathsf{T}_s = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{bmatrix}. \tag{27}$$

Then, from Eq. (6), we find

$$^{j}\mathsf{P} = \mathsf{T}_c^{-1} \cdot {}^{c}\mathsf{P} \cdot \mathsf{T}_c = \mathsf{T}_s^{-1} \cdot {}^{s}\mathsf{P} \cdot \mathsf{T}_s, \tag{28}$$

or

$$\begin{bmatrix} {}^jP_{11} & {}^jP_{12} & {}^jP_{13} & {}^jP_{14} \\ {}^jP_{21} & {}^jP_{22} & {}^jP_{23} & {}^jP_{24} \\ {}^jP_{31} & {}^jP_{32} & {}^jP_{33} & {}^jP_{34} \\ {}^jP_{41} & {}^jP_{42} & {}^jP_{43} & {}^jP_{44} \end{bmatrix}$$

$$= \begin{bmatrix} {}^cP_{11} & {}^cP_{13}-i{}^cP_{14} & {}^cP_{13}+i{}^cP_{14} & {}^cP_{12} \\ \frac{1}{2}({}^cP_{31}+i{}^cP_{41}) & \frac{1}{2}({}^cP_{33}+{}^cP_{44}-i{}^cP_{34}+i{}^cP_{43}) & \frac{1}{2}({}^cP_{33}-{}^cP_{44}+i{}^cP_{34}+i{}^cP_{43}) & \frac{1}{2}({}^cP_{32}+i{}^cP_{42}) \\ \frac{1}{2}({}^cP_{31}-i{}^cP_{41}) & \frac{1}{2}({}^cP_{33}-{}^cP_{44}-i{}^cP_{34}-i{}^cP_{43}) & \frac{1}{2}({}^cP_{33}+{}^cP_{44}+i{}^cP_{34}-i{}^cP_{43}) & \frac{1}{2}({}^cP_{32}-i{}^cP_{42}) \\ {}^cP_{21} & {}^cP_{23}-i{}^cP_{24} & {}^cP_{23}+i{}^cP_{24} & {}^cP_{22} \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} {}^sP_{11}+{}^sP_{12}+{}^sP_{21}+{}^sP_{22} & {}^sP_{13}+{}^sP_{23}-i({}^sP_{14}+{}^sP_{24}) & {}^sP_{13}+{}^sP_{23}+i({}^sP_{14}+{}^sP_{24}) & {}^sP_{11}-{}^sP_{12}+{}^sP_{21}-{}^sP_{22} \\ {}^sP_{31}+{}^sP_{32}+i({}^sP_{41}+{}^sP_{42}) & {}^sP_{33}+{}^sP_{44}-i({}^sP_{34}-{}^sP_{43}) & {}^sP_{33}-{}^sP_{44}+i({}^sP_{34}+{}^sP_{43}) & {}^sP_{31}-{}^sP_{32}+i({}^sP_{41}-{}^sP_{42}) \\ {}^sP_{31}+{}^sP_{32}-i({}^sP_{41}+{}^sP_{42}) & {}^sP_{33}-{}^sP_{44}-i({}^sP_{34}+{}^sP_{43}) & {}^sP_{33}+{}^sP_{44}+i({}^sP_{34}-{}^sP_{43}) & {}^sP_{31}-{}^sP_{32}-i({}^sP_{41}-{}^sP_{42}) \\ {}^sP_{11}+{}^sP_{12}-{}^sP_{21}-{}^sP_{22} & {}^sP_{13}-{}^sP_{23}-i({}^sP_{14}-{}^sP_{24}) & {}^sP_{13}-{}^sP_{23}+i({}^sP_{14}-{}^sP_{24}) & {}^sP_{11}-{}^sP_{12}-{}^sP_{21}+{}^sP_{22} \end{bmatrix}. \quad (29)$$

Substituting the values of ${}^jP_{ij}$ from Eq. (29) into Eqs. (10a)–(21a), we obtain the desired relations between the elements of either ${}^c\mathsf{P}$ or ${}^s\mathsf{P}$. They are

$${}^cP_{11}{}^cP_{12} = {}^cP_{13}^2 + {}^cP_{14}^2, \quad (10b)$$

$${}^cP_{21}{}^cP_{22} = {}^cP_{23}^2 + {}^cP_{24}^2, \quad (11b)$$

$$4\,{}^cP_{11}{}^cP_{21} = {}^cP_{31}^2 + {}^cP_{41}^2, \quad (12b)$$

$$4\,{}^cP_{12}{}^cP_{22} = {}^cP_{32}^2 + {}^cP_{42}^2, \quad (13b)$$

$$4\,{}^cP_{11}{}^cP_{22} = ({}^cP_{33} + {}^cP_{44})^2 + ({}^cP_{34} - {}^cP_{43})^2, \quad (14b)$$

$$4\,{}^cP_{12}{}^cP_{21} = ({}^cP_{33} - {}^cP_{44})^2 + ({}^cP_{34} + {}^cP_{43})^2, \quad (15b)$$

$$2({}^cP_{13}{}^cP_{23} - {}^cP_{14}{}^cP_{24})({}^cP_{33}{}^cP_{34} + {}^cP_{44}{}^cP_{43}) = ({}^cP_{13}{}^cP_{24} + {}^cP_{14}{}^cP_{23})({}^cP_{33}^2 - {}^cP_{44}^2 + {}^cP_{43}^2 - {}^cP_{34}^2), \quad (19b)$$

$$2({}^cP_{31}{}^cP_{32} - {}^cP_{41}{}^cP_{42})({}^cP_{33}{}^cP_{43} + {}^cP_{44}{}^cP_{34}) = ({}^cP_{31}{}^cP_{42} + {}^cP_{41}{}^cP_{32})({}^cP_{33}^2 - {}^cP_{44}^2 + {}^cP_{34}^2 - {}^cP_{43}^2), \quad (20b)$$

$$({}^cP_{14}{}^cP_{23} - {}^cP_{13}{}^cP_{24})({}^cP_{31}{}^cP_{32} + {}^cP_{41}{}^cP_{42}) = ({}^cP_{42}{}^cP_{31} - {}^cP_{41}{}^cP_{32})({}^cP_{13}{}^cP_{23} + {}^cP_{14}{}^cP_{24}), \quad (21b)$$

and

$$({}^sP_{11} + {}^sP_{21})^2 - ({}^sP_{12} + {}^sP_{22})^2 = ({}^sP_{13} + {}^sP_{23})^2 + ({}^sP_{14} + {}^sP_{24})^2, \quad (10c)$$

$$({}^sP_{11} - {}^sP_{21})^2 - ({}^sP_{12} - {}^sP_{22})^2 = ({}^sP_{13} - {}^sP_{23})^2 + ({}^sP_{14} - {}^sP_{24})^2, \quad (11c)$$

$$({}^sP_{11} + {}^sP_{12})^2 - ({}^sP_{21} + {}^sP_{22})^2 = ({}^sP_{31} + {}^sP_{32})^2 + ({}^sP_{41} + {}^sP_{42})^2, \quad (12c)$$

$$({}^sP_{11} - {}^sP_{12})^2 - ({}^sP_{21} - {}^sP_{22})^2 = ({}^sP_{31} - {}^sP_{32})^2 + ({}^sP_{41} - {}^sP_{42})^2, \quad (13c)$$

$$({}^sP_{11} + {}^sP_{22})^2 - ({}^sP_{12} + {}^sP_{21})^2 = ({}^sP_{33} + {}^sP_{44})^2 + ({}^sP_{34} - {}^sP_{43})^2, \quad (14c)$$

$$({}^sP_{11} - {}^sP_{22})^2 - ({}^sP_{12} - {}^sP_{21})^2 = ({}^sP_{33} - {}^sP_{44})^2 + ({}^sP_{34} + {}^sP_{43})^2, \quad (15c)$$

$$({}^sP_{13}{}^sP_{14} - {}^sP_{23}{}^sP_{24})({}^sP_{33}^2 - {}^sP_{34}^2 + {}^sP_{43}^2 - {}^sP_{44}^2) = ({}^sP_{33}{}^sP_{34} + {}^sP_{43}{}^sP_{44})({}^sP_{13}^2 - {}^sP_{14}^2 - {}^sP_{23}^2 + {}^sP_{24}^2), \quad (19c)$$

$$({}^sP_{31}{}^sP_{41} - {}^sP_{32}{}^sP_{42})({}^sP_{33}^2 - {}^sP_{43}^2 + {}^sP_{34}^2 - {}^sP_{44}^2) = ({}^sP_{33}{}^sP_{43} + {}^sP_{34}{}^sP_{44})({}^sP_{31}^2 - {}^sP_{41}^2 - {}^sP_{32}^2 + {}^sP_{42}^2), \quad (20c)$$

$$({}^sP_{14}{}^sP_{23} - {}^sP_{13}{}^sP_{24})({}^sP_{31}^2 - {}^sP_{32}^2 + {}^sP_{41}^2 - {}^sP_{42}^2) = ({}^sP_{42}{}^sP_{31} - {}^sP_{41}{}^sP_{32})({}^sP_{14}^2 - {}^sP_{24}^2 + {}^sP_{13}^2 - {}^sP_{23}^2). \quad (21c)$$

If we partition the Chandrasekhar phase matrix in the following manner,

$$\left[\begin{array}{cc:cc} {}^cP_{11} & {}^cP_{12} & {}^cP_{13} & {}^cP_{14} \\ {}^cP_{21} & {}^cP_{22} & {}^cP_{23} & {}^cP_{24} \\ \hdashline {}^cP_{31} & {}^cP_{32} & {}^cP_{33} & {}^cP_{34} \\ {}^cP_{41} & {}^cP_{42} & {}^cP_{43} & {}^cP_{44} \end{array}\right] = \left[\begin{array}{c:c} \mathsf{C}_1 & \mathsf{C}_2 \\ \hdashline \mathsf{C}_3 & \mathsf{C}_4 \end{array}\right],$$

then we see that:

(i) Equations (10b) and (11b) relate the elements of C_1 to those of C_2.

(ii) Equations (12b) and (13b) relate the elements of C_1 to those of C_3.

(iii) Equations (14b) and (15b) relate the elements of C_1 to those of C_4. In particular, subtracting Eq. (15b) from Eq. (14b), we get Det of C_1 = Det of C_4.

(iv) Equation (21b) relates the elements of C_2 to those of C_3.

(v) Equation (19b) relates the elements of C_2 to those of C_4.

(vi) Equation (20b) relates the elements of C_3 to those of C_4.

Similar remarks hold in the case of the Stokes phase matrix ${}^{s}\mathsf{P}$.

IV. SOME REMARKS ON THE INVARIANCE OF THE RELATIONS

From the definition of ${}^{j}\mathsf{P}$ by Eq. (7), and from the method of derivation of the relations Eqs. (10a)–(21a), it is obvious that these relations will be valid for any phase matrix

$$\mathsf{P} = \mathsf{K} \times \mathsf{K}^*, \tag{30}$$

where K is an arbitrary 2×2 matrix of complex elements. Consequently, since

$$\mathsf{P}' = \mathsf{P}_1 \cdot \mathsf{P}_2 \cdot \mathsf{P}_3 \cdots \mathsf{P}_n = \mathsf{K}' \times \mathsf{K}'^*, \tag{31}$$

where

$$\mathsf{K}' = \mathsf{K}_1 \cdot \mathsf{K}_2 \cdot \mathsf{K}_3 \cdots \mathsf{K}_n, \tag{32}$$

the elements of P' will also satisfy the same relations.

Now, the action of any single optical device can be described by a matrix of the form of Eq. (30), where K is the Jones matrix[6] of the device; similarly, the matrix representing any optical train is of the form of Eq. (31), where K_n is the Jones matrix of the nth element in the train. Hence, the elements of the phase matrix describing the transmission of light through such devices will always satisfy the relations (10a)–(21a). Further, since the rotation of the axes of reference for the electric vector in either the incident or the scattered beam is equivalent to the action of a rotator, the relations between the elements of the phase matrix will remain invariant after such rotation. If $\mathsf{R}(\phi)$ and $\mathsf{R}(\psi)$ are the matrices representing rotations through angles ϕ and ψ for the incident and scattered beams, respectively, the final phase matrix P'' will be

$$\mathsf{P}'' = [\mathsf{R}(\psi) \times \mathsf{R}(\psi)] \cdot \mathsf{P} \cdot [\mathsf{R}(\phi) \times \mathsf{R}(\phi)]. \tag{33}$$

In the equation of radiative transfer, $\phi = -i_1$ and $\psi = \pi - i_2$, where i_1 and i_2 are the angles between the scattering plane and the meridian planes passing through the incident and scattered beams, respectively.[2] Hence, both P and P'' satisfy the relations (10a)–(21a).

The phase matrix P can be written in the form given by Eq. (30) only in the coherency representation [Eq. (7)]. But the invariance properties of the relations between the elements of P will also hold in other representations. These invariant relations are given by Eqs. (10b)–(21b) in the Chandrasekhar representation and by Eqs. (10c)–(21c) in the Stokes representation. For example, in the Chandrasekhar representation, the phase matrix for Rayleigh scattering is

$${}^{c}\mathsf{P}(\Theta) = \tfrac{3}{2}\begin{bmatrix} \cos^2\Theta & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\Theta & 0 \\ 0 & 0 & 0 & \cos\Theta \end{bmatrix}, \tag{34}$$

where Θ is the scattering angle, and the rotation matrix is given by

$$\mathsf{L}(\phi) = \mathsf{T}_c \cdot [\mathsf{R}(\phi) \times \mathsf{R}(\phi)] \cdot \mathsf{T}_c^{-1} = \begin{bmatrix} \cos^2\phi & \sin^2\phi & \tfrac{1}{2}\sin 2\phi & 0 \\ \sin^2\phi & \cos^2\phi & -\tfrac{1}{2}\sin 2\phi & 0 \\ -\sin 2\phi & \sin 2\phi & \cos 2\phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{35}$$

where ϕ is measured positive clockwise. Then, it is easily verified that ${}^{c}\mathsf{P}(\Theta)$, $\mathsf{L}(\phi)$, and ${}^{c}\mathsf{P}''(\Theta) = \mathsf{L}(\pi - i_2) \cdot {}^{c}\mathsf{P}(\Theta) \cdot \mathsf{L}(-i_1)$ all satisfy the relations (10b)–(21b). Similarly, in the Stokes representation, it can also be verified that the elements of the Mueller matrix of any optical device[6] satisfy the relations (10c)–(21c).

In conclusion, we may note that the relations derived here between the elements of the phase matrix will find use not only in the theory of scattering but also in polarization optics.

[6] W. A. Shurcliff, *Polarized Light* (Harvard University Press, Cambridge, Mass., 1962), Chap. 8 and Appendix 2.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Conformal-Invariant Scattering Amplitude for Spinless Particles*

NAREN F. BALI, DARRYL D. COON, AND AMNON KATZ†
Department of Physics, University of Washington, Seattle, Washington

(Received 31 March 1969)

The implications of exact conformal symmetry for the S matrix of spinless massless particles are studied. Explicit forms for the scattering amplitudes are obtained in configuration space. Differential conditions for conformal-invariant amplitudes are derived in momentum space. These are solved for the case of elastic two-body scattering and lead to an integral representation of the general two-body amplitude in terms of Lorentz scalars. The integral representation involves two arbitrary functions of two variables and four explicit quadratures. Further study of the four quadratures appears in appendices.

1. INTRODUCTION

The conformal group[1]—the group of all coordinate transformations that conserve the light cone—has often been the subject of investigation by physicists.[2] Maxwell's equations, the electrodynamics of massless particles, scalar ϕ^3 theory,[4] and other field theories of massless particles, are conformal-invariant. As dilatations form a subgroup of the conformal group, all masses in conformal-invariant theories must vanish or continuous mass spectra must exist. Therefore, unbroken conformal invariance is in contradiction with observation; but broken or approximate conformal invariance might prove useful, as does broken SU_3. In particular, it has been speculated[3] that conformal invariance might be relevant to processes at high-energy and high-momentum transfer in which the masses of the particles involved are negligible in comparison.

This paper is concerned with the implications of exact and unbroken conformal invariance for the S matrix of scalar particles. We concentrate on two-body scattering and arrive (in Sec. 9) at an integral representation for the two-body scattering amplitude.

We deal with scalar massless particles, but make no specific assumptions on the dynamics governing them. We only assume that the equations determining the scattering matrix are conformal-invariant and give rise to a conformally scalar S matrix. (The possibility of spontaneously broken conformal invariance is thus excluded.)

Our requirements of conformal invariance are off the mass shell as well as on it. Off the mass shell, these requirements are an arbitrary but natural choice. This remark is particularly important, as it turns out that the integral representation of Sec. 9 [Eq. (9.6)] appears to have structure only off the mass shell and vanish on it.[5] As in physical applications the symmetry will not be exact and the mass-shell condition will have to be broken, the off-mass-shell information may not prove devoid of interest. Our condition of conformal invariance off and on the mass shell is the same that results from study of Feynman diagrams applied to the ϕ^4 theory. We do not know whether the trivial mass-shell behavior persists for production amplitudes or scattering of particles with spin.

We briefly introduce the conformal group in Sec. 2 and the scalar massless field in Sec. 3. For full discussion of the conformal group, its representations, and its Lie algebra, the reader is referred elsewhere.[1] In Sec. 4, we determine the requirements of conformal invariance for the S matrix. It is here that the mathematical problem we treat is set. The S matrix is initially introduced in coordinate space, where the requirements of conformal invariance are readily solved. To be of physical use, it must be transformed into momentum space. This transformation is discussed in Sec. 5.

For two-body scattering, the Fourier transform involves a 16-fold integration. Of these, four integrations are immediate (by virtue of translation invariance) and two involve an arbitrary function that figures in the general scattering amplitude. A tenfold

* This paper was supported in part by the Atomic Energy Commission, AT(45-1)-1388.

† On leave from the Weizmann Institute, Rehovoth, Israel, 1967–68.

[1] Y. Murai, Progr. Theoret. Phys. (Kyoto) **9**, 147 (1953); A. Esteve and P. G. Sona, Nuovo Cimento **32**, 473 (1964); L. Gross, J. Math. Phys. **5**, 687.(1964); R. Raçzka, N. Limić, and J. Niederle, J. Math. Phys. **7**, 1861, 2026 (1966); **8**, 1079 (1967); T. Fulton, F. Rohrlich, and L. Witten, Rev. Mod. Phys. **34**, 442 (1962); T. Yao, J. Math. Phys. **8**, 1931 (1967).

[2] E. Cunningham, Proc. London Math. Soc. **8**, 77 (1910); H. Bateman, *ibid.* **8**, 223 (1910); P. A. M. Dirac, Ann. Math. **37**, 429 (1935); J. Wess, Nuovo Cimento **18**, 1086 (1960); H. A. Kastrup, Ann. Physik **9**, 388 (1962); H. A. Kastrup, Phys. Rev. **142**, 1060 (1965); L. Castell, Nucl. Phys. **B4**, 343 (1968).

[3] See, for example, H. A. Kastrup, Nucl. Phys. **58**, 561 (1964).

[4] However, the four-fermion interaction and the ϕ^3 interaction are not conformal invariant even for massless particles.

[5] This does not mean that the scattering amplitude itself vanishes on the mass shell, because the integral representation of Sec. 9 leaves out certain possibilities including that of a scattering amplitude that is everywhere a constant. In fact, all the examples of conformal invariant amplitudes that we know become constant on the mass shell. We do not know whether a different behavior on the mass shell is possible.

explicit integral is still to be done. The main contribution of the present paper is what amounts to the reduction of this tenfold integral to only a fourfold integral in Eq. (9.6). To achieve this, we reformulate the requirements of conformal invariance into the language of momentum space. This is done in Secs. 6 and 7, and the conditions emerge as a dilatation property and three partial differential equations in terms of Lorentz scalars that the momentum-space-invariant scattering amplitude must satisfy. The differential equations are solved in Sec. 8 via a Fourier transform involving Lorentz scalars rather than vectors. The final integral representation for the invariant amplitude is presented in Sec. 9. The integrals appearing in it are interpreted as generalized integrals and remarks concerning them appear in Sec. 10. Section 11 contains concluding remarks.

As mentioned above, the final result in Sec. 9 still features a fourfold explicit integral. This integral should be treated further, and Eq. (9.6) should not be regarded as the end of the way. However, the present paper represents the end of our collaboration and we feel that the reduction of the tenfold integral to a fourfold integral is a nontrivial step which should be on record.

We have tried our hand at doing the remaining four integrations. We can do two of the four integrations explicitly. In one of the terms of Eq. (9.6) we can do what amounts to three of the four integrations. These attempts at integration are presented in Appendix A. For particular values of the parameters appearing in the arbitrary functions, we have been able to carry out all four quadratures in terms of elementary functions. Such elementary solutions of the conformal-invariance conditions are presented in Appendix B. These have been checked to satisfy the differential equations.

In general we know (Appendix A) that the four quadratures cannot all be performed in terms of elementary functions. We do not know whether they can all be performed in terms of standard special functions, or how this might best be attempted. We have therefore not included our attempts regarding the four remaining quadratures in the text of the paper. We do present them in the appendices, as readers might find some of our devices useful.

2. THE CONFORMAL GROUP

Under a general coordinate transformation $y^\mu = y^\mu(x)$, the metric of space–time transforms according to

$$g'_{\mu\nu}(y) = \frac{\partial x^\alpha}{\partial y^\mu} g_{\alpha\beta}(x) \frac{\partial x^\beta}{\partial y^\nu}. \tag{2.1}$$

The transformations of the Poincaré group are distinguished by the property that they conserve the Minkowski metric; i.e., if we put $g_{\alpha\beta}(x) = \eta_{\alpha\beta}$ ($\eta_{\alpha\beta}$ being the Minkowski metric), we find $g'_{\mu\nu}(y) = \eta_{\mu\nu}$. The conformal group consists of all transformations that, starting with $g_{\alpha\beta}(x) = \eta_{\alpha\beta}$ as above, yield $g'_{\mu\nu}(y) = f(y)\eta_{\mu\nu}$. By considering the determinant of the two sides of Eq. (2.1), we find

$$f(y) = [\det(\partial x/\partial y)]^{\frac{1}{2}}.$$

Thus, for conformal transformations

$$\frac{\partial x^\alpha}{\partial y^\mu} \eta_{\alpha\beta} \frac{\partial x^\beta}{\partial y^\nu} = \left[\det\left(\frac{\partial x}{\partial y}\right)\right]^{\frac{1}{2}} \eta_{\mu\nu}, \tag{2.2}$$

by definition.

By conserving the Minkowski form $\eta_{\mu\nu}$, Poincaré transformations leave physical equations constructed with this space–time metric unchanged, hence the role of the Poincaré group as the symmetry group of physical theories. Some equations in physics (notably the equations governing the electromagnetic field) remain unchanged when $\eta_{\mu\nu}$ is replaced by $f(x)\eta_{\mu\nu}$. These equations admit the conformal group as their symmetry group.

Obviously, the Poincaré group is a subgroup of the conformal group. The full conformal group may be obtained by adjoining to the Poincaré group two additional operations: the operation of dilatation

$$x^\mu \to y^\mu = ax^\mu \tag{2.3}$$

(depending on the continuous parameter a) and the discrete operation

$$x^\mu \to y^\mu = x^\mu/x^2. \tag{2.4}$$

This operation is called the inverse radius transformation[1] and is not contained in the proper conformal group. Here and in the following, x^2 denotes $\eta_{\alpha\beta}x^\alpha x^\beta$.

To be a conformal scalar, a quantity has to be a Poincaré scalar, it has to be dilatation invariant (i.e., dimensionless), and it has to be invariant under the inverse radius transformation. The determinant of the inverse radius transformation (2.4) is

$$\det\left(\frac{\partial x}{\partial y}\right) = -x^8. \tag{2.5}$$

We use x^{2n} as notation for $(x^2)^n$.

The Poincaré scalar $(x_1 - x_2)^2$ transforms under (2.3) into $a^2(x_1 - x_2)^2$, and under (2.4) into $(x_1 - x_2)^2/x_1^2x_2^2$. It is therefore a scalar density (of weight $-\frac{1}{4}$) rather than a scalar under the conformal group. It takes four points to make a true conformal scalar. Given four points, two independent scalars can be

constructed as can be easily seen using the six-dimensional notation of Dirac.[2] They can be chosen to be

$$X = \frac{(x_1 - x_2)^2(x_3 - x_4)^2}{(x_1 - x_3)^2(x_2 - x_4)^2}, \tag{2.6}$$

$$Y = \frac{(x_1 - x_2)^2(x_3 - x_4)^2}{(x_1 - x_4)^2(x_2 - x_3)^2}. \tag{2.7}$$

Notice that x^μ is not a conformal vector. Notice also that, since $\eta_{\mu\nu}$ is a tensor density rather than a tensor, it turns tensors into tensor densities, when used for raising or lowering indices.

3. THE MASSLESS SCALAR FIELD

The free massless scalar field (by definition) satisfies the massless Klein–Gordon equation

$$\Box\, \phi(x) = 0. \tag{3.1}$$

It is easily verified that if $\phi(x)$ satisfies this equation, so does $(1/x^2)\phi(x/x^2)$. This suggests that we consider the "scalar" (meaning Poincaré scalar) field to be a conformal scalar density of weight $\frac{1}{4}$.[6] With this transformation law, the free-field equations are conformally covariant.

The interacting field equations will also be conformally covariant if we choose a conformally invariant interaction. The ϕ^4 interaction, for one, satisfies this criterion, since $[\phi(x)]^4$ is a scalar density of weight 1 and its integral is, therefore, a scalar.

In the following we do not actually specify the interaction, nor do we commit ourselves to any field-theoretical model. We only assume that the equations determining the scattering, whatever they are, are conformally covariant and that they lead to a conformally scalar S matrix.

4. THE SCATTERING MATRIX

We assume that the S matrix can be expanded in the free fields:

$$S = 1 + \sum_{n=1}^{\infty} \int S_n(x_1, \cdots, x_n) \times :\phi(x_1) \cdots \phi(x_n): d^4x_1 \cdots d^4x_n. \tag{4.1}$$

Each term in this sum must be a conformal scalar. A sufficient condition for this to be true is that $S_n(x_1, \cdots, x_n)$ be a scalar density of weight $\frac{3}{4}$ in each of its variables. Since $\phi(x)$ is a scalar density of weight $\frac{1}{4}$, this would make the integrand a scalar density (of weight 1) in each variable, and the integral a scalar.

If the fields $\phi(x)$ at different space–time points x were all linearly independent, the above sufficient condition would also be necessary. However, $\phi(x)$ in the neighborhood of any spacelike surface determines $\phi(x)$ everywhere through the Klein–Gordon equation (3.1). Thus, it is not quite *necessary* to have $S_n(x_1, \cdots, x_n)$ a scalar density of weight $\frac{3}{4}$, but it is *possible*. We *choose* to enforce this condition. By making this choice we are in effect defining the scattering matrix off the mass shell.[7] This remark is particularly pertinent in the present paper, where we find that the S matrix for two-body scattering possesses structure only off the mass shell. By the above choice, S_n satisfies

$$\frac{S_n(x_1/x_1^2, \cdots, x_n/x_n^2)}{x_1^6 \cdots x_n^6} = S_n(x_1, \cdots, x_n). \tag{4.2}$$

The most general scalar density may be expressed as a product of a special scalar density and the most general scalar. This allows us to express S_n as

$$S_n(x_1, \cdots, x_n) = \frac{F_n(x_1, \cdots, x_n)}{(x_1 - x_2)^3(x_2 - x_3)^3 \cdots (x_{n-1} - x_n)^3(x_n - x_1)^3}, \tag{4.3}$$

where F_n is a scalar.[8] The denominator contains n brackets. For even indices some simplification is possible. We may write S_{2n} as

$$S_{2n}(x_1, \cdots, x_{2n}) = \frac{f_{2n}(x_1, \cdots, x_{2n})}{(x_1 - x_2)^6(x_3 - x_4)^6, \cdots, (x_{2n-1} - x_{2n})^6}, \tag{4.4}$$

where f_{2n} is a scalar. The denominator in the last equation contains n (rather than $2n$) brackets.

For S_4 (which corresponds to two-body scattering), Eq. (4.4) reduces to

$$S_4(x_1, x_2, x_3, x_4) = \frac{f_4(x_1, x_2, x_3, x_4)}{(x_1 - x_2)^6(x_3 - x_4)^6}. \tag{4.5}$$

Conformal invariance thus restricts S_4 to the above form, with f_4 being the most general conformal scalar that can be constructed from the coordinates of the points x_1 through x_4. The considerations of Sec. 2 in turn indicate that f_4 is a function of the quantities X and Y of Eqs. (2.6) and (2.7), rather than of the individual x_i $(i = 1, \cdots, 4)$. The requirement of conformal invariance thereby reduces S_4 to the form

$$S_4(x_1, x_2, x_3, x_4) = \frac{f(X, Y)}{(x_1 - x_2)^6(x_3 - x_4)^6}, \tag{4.6}$$

[6] J. Mathews and R. Walker, *Mathematical Methods of Physics* (W. A. Benjamin, Inc., New York, 1964), p. 385.

[7] This off-mass-shell continuation is the same as results from Feynman diagrams.

[8] The δ function $\delta^4(x_i - x_j)$ is also a scalar density. Generalized function densities of this type are not included in Eq. (4.3), but may be readily added separately.

where f is an arbitrary function of two variables. This last explicit form is mentioned in the next section but will not be used in the derivation of our final result which appears in Sec. 9.

5. THE SCATTERING AMPLITUDE IN MOMENTUM SPACE

Experimental data on scattering come in terms of momentum rather than coordinate variables. Therefore, we are not so much interested in $S_n(x_1, \cdots, x_n)$ as in its Fourier transform

$$\tilde{S}_n(p_1, \cdots, p_n) = \int d^4x_1 \cdot \cdots d^4x_n [e^{i(p_1 \cdot x_1 + \cdots + p_n \cdot x_n)} S_n(x_1 \cdots x_n)]. \tag{5.1}$$

Of the $4n$ integrations in Eq. (5.1), four can be easily performed by using translation invariance. By defining new variables of integration

$$x_i' = x_i - x_n, \quad i = 1, \cdots, n-1, \tag{5.2}$$

and invoking the translation invariance of S_n, we transform the integral in Eq. (5.1) into

$$\begin{aligned}\tilde{S}_n(p_1, \cdots, p_n) &= \int d^4x_1' \cdots d^4x_{n-1}' \, d^4x_n e^{i(p_1 + \cdots + p_n) \cdot x_n} \\ &\quad \times e^{i(p_1 \cdot x_1' + \cdots + p_{n-1} \cdot x_{n-1}')} S_n(x_1', \cdots, x_{n-1}', 0) \\ &= (2\pi)^4 \delta^4(p_1 + \cdots + p_n) M_n(p_1, \cdots, p_{n-1}),\end{aligned} \tag{5.3}$$

where

$$M_n(p_1, \cdots, p_{n-1}) = \int d^4x_1 \cdots d^4x_{n-1} e^{i(p_1 \cdot x_1 + \cdots + p_{n-1} \cdot x_{n-1})} \times S_n(x_1, \cdots, x_{n-1}, 0). \tag{5.4}$$

The presence of the momentum-conserving delta function in Eq. (5.4) was to be expected. The dynamical information is in the function $M_n(p_1, \cdots, p_{n-1})$, which is essentially the usual invariant amplitude of scattering theory.

For $n = 4$ (elastic two-body scattering), we have[9]

$$M_4(p_1, p_2, p_3) = \int d^4x_1 \, d^4x_2 \, d^4x_3 e^{i(p_1 \cdot x_1 + p_2 \cdot x_2 + p_3 \cdot x_3)} \frac{f(X', Y')}{(x_1 - x_2)^6 x_3^6}. \tag{5.5}$$

Here we have used Eq. (4.6); X', Y' stand for the quantities X, Y of Eqs. (2.6) and (2.7) in which x_4 has been set equal to zero. The twelve integrations in Eq. (5.5) cannot all be carried out because f is arbitrary.

A natural way to proceed with Eq. (5.5) would be to change variables from the 12 components of x_1, x_2, x_3, to X', Y', and ten other variables. Ten integrations would then be performed explicitly, leaving $M_4(p_1, p_2, p_3)$ as a double integral over an integrand involving an arbitrary function of two variables. However, the ten integrations in question proved to be beyond the capabilities of the present authors. We therefore chose a roundabout way that leads in (Sec. 9) to an integral representation of M involving only four explicit quadratures (as well as a double integral over an arbitrary function of two variables).

6. DIFFERENTIAL EQUATIONS FOR M

The inverse radius transformation is a discrete operation that cannot be continuously connected to the unit (do nothing) operation. However, an inverse radius transformation followed by a Poincaré operation and another inverse radius transformation constitutes a conformal transformation which depends on a continuous parameter (the parameter of the Poincaré operation) and is continuously connected with the identity operation. This can be nontrivially done only for those Poincaré operations that do not commute with the inverse radius transformation, i.e., translations. An inverse radius transformation followed by a translation and another inverse radius transformation constitutes a "special conformal transformation"

$$x \to \frac{x + ax^2}{1 + a^2x^2 + 2a \cdot x}, \tag{6.1}$$

where a is a parameter of the translation $x \to x + a$. The infinitesimal generator of this operation is

$$K_\mu(x) = -i(x^2 \delta_\mu^\nu - 2x_\mu x^\nu) \frac{\partial}{\partial x^\nu}. \tag{6.2}$$

For a quantity $f_n(x_1, \cdots, x_n)$ to be a conformal scalar, it is necessary that

$$\sum_{i=1}^{n} K_\mu(x_i) f_n(x_1, \cdots, x_n) = 0 \tag{6.3}$$

as well as

$$\sum_{i=1}^{n} P_\mu(x_i) f_n(x_1, \cdots, x_n) = 0, \tag{6.4}$$

where $P_\mu(x)$ is the generator of translations

$$P_\mu(x) = -i \frac{\partial}{\partial x^\mu}. \tag{6.5}$$

It turns out that conditions (6.3) and (6.4) are also sufficient for invariance under the whole connected

[9] Except for singular contributions; see Ref. 8.

part of the conformal group. This is so because of the commutation relation

$$[K_\mu(x), P_\nu(x)] = -2i(D\eta_{\mu\nu} + M_{\mu\nu}), \quad (6.6)$$

where D and $M_{\mu\nu}$ are the infinitesimal generators of dilatations and Lorentz rotations, respectively.[10]

For a scalar density $f_n^w(x_1, \cdots, x_n)$ of weight w, Eq. (6.3) is replaced by

$$\sum_{i=1}^{n} \{K_\mu(x_i) - 8wx_{i\mu}\} f_n^w(x_1, \cdots, x_n) = 0. \quad (6.7)$$

The equations defining the scalar density $S_n(x_1, \cdots, x_n)$ are thus

$$\sum_{i=1}^{n} \frac{\partial}{\partial x_i^\mu} S_n(x_1, \cdots, x_n) = 0, \quad (6.8)$$

$$\sum_{i=1}^{n} \left[x_i^2 \frac{\partial}{\partial x_i^\mu} - 2x_{i\mu} x_i^\nu \frac{\partial}{\partial x_i^\nu} - 6x_{i\mu} \right] S_n(x_1, \cdots, x_n) = 0. \quad (6.9)$$

But Eq. (5.4) for M_n uses $S_n(x_1, \cdots, x_{n-1}, 0)$ rather than $S_n(x_1, \cdots, x_n)$. The substitution of $x_n = 0$ destroys translation invariance and makes Eq. (6.8) useless. The substitution can be effected in Eq. (6.9) because the coefficient of $\partial/\partial x_n$ is quadratic in x_n. Thus we find

$$\sum_{i=1}^{n-1} \left[x_i^2 \frac{\partial}{\partial x_i^\mu} - 2x_{i\mu} x_i^\nu \frac{\partial}{\partial x_i^\nu} - 6x_{i\mu} \right] \times S_n(x_1, \cdots, x_{n-1}, 0) = 0 \quad (6.10)$$

as a necessary condition on $S_n(x_1, \cdots, x_{n-1}, 0)$.

Given $S_n(x_1, \cdots, x_{n-1}, 0)$ as a function of $x_1, \cdots, x_{n-1}$, it is easy to restore translation invariance by putting

$$S_n(x_1, \cdots, x_n) = S_n(x_1 - x_n, \cdots, x_{n-1} - x_n, 0). \quad (6.11)$$

The $S_n(x_1, \cdots, x_n)$ so constructed will automatically satisfy Eq. (6.8), but not necessarily Eq. (6.9). Equation (6.10), applying to $S_n(x_1, \cdots, x_{n-1}, 0)$, is not sufficient to insure that $S_n(x_1, \cdots, x_n)$ constructed by Eq. (6.11) satisfies Eq. (6.9). It turns out, however, that Eq. (6.10) together with the dilatation property of S_n and its being a Lorentz scalar are sufficient for this purpose.

A conformal scalar must be dilatation-invariant, i.e., dimensionless. Being a scalar density of weight $\frac{3}{4}$ in each of its variables, $S_n(x_1, \cdots, x_n)$ must satisfy

$$S_n(\lambda x_1, \cdots, \lambda x_n) = \lambda^{-3n} S_n(x_1, \cdots, x_n). \quad (6.12)$$

This equation is a consequence of Eqs. (6.8) and (6.9).

[10] See, for example, H. A. Kastrup, Ref. 2.

We now show that if $S_n(x_1, \cdots, x_{n-1}, 0)$ satisfies

$$S_n(\lambda x_1, \cdots, \lambda x_{n-1}, 0) = \lambda^{-3n} S_n(x_1, \cdots, x_{n-1}, 0) \quad (6.13)$$

as well as Eq. (6.10) and is scalar under Lorentz rotations, then $S_n(x_1, \cdots, x_n)$ of Eq. (6.11) satisfies Eq. (6.9).

Proof: A direct application of Eq. (6.10) yields

$$\sum_{i=1}^{n-1} \left[(x_i - x_n)^2 \frac{\partial}{\partial x_i^\mu} - 2(x_{i\mu} - x_{n\mu}) \times (x_i^\nu - x_n^\nu) \frac{\partial}{\partial x_i^\nu} - 6(x_{i\mu} - x_{n\mu}) \right] \times S_n(x_1, \cdots, x_n) = 0. \quad (6.14)$$

Simple algebraic manipulations and the use of Eq. (6.8) bring the last equation into the form

$$\left[\sum_{i=1}^{n} \left(x_i^2 \frac{\partial}{\partial x_i^\mu} - 2x_{i\mu} x_i^\nu \frac{\partial}{\partial x_i^\nu} - 6x_{i\mu} \right) + 2x_n^\nu \sum_{i=1}^{n} \left(x_{i\mu} \frac{\partial}{\partial x_i^\nu} - x_{i\nu} \frac{\partial}{\partial x_i^\mu} \right) + 2x_{n\mu} \left(\sum_{i=1}^{n} x_i^\nu \frac{\partial}{\partial x_i^\nu} + 3n \right) \right] S_n(x_1, \cdots, x_n) = 0. \quad (6.15)$$

The last term in the square bracket vanishes because of the dilatation condition Eq. (6.13). The second term in the square bracket is made of generators of Lorentz rotations and vanishes because S_n is a Lorentz scalar by assumption. This leaves us with Eq. (6.9). Q.E.D.

We have so far shown that $S_n(x_1, \cdots, x_{n-1}, 0)$ is completely characterized by the conditions:

1. It is a Lorentz scalar;
2. It satisfies the dilatation property (6.13);
3. It satisfies Eq. (6.10).

These three conditions are readily translated into equivalent conditions on $M_n(p_1, \cdots, p_{n-1})$, the Fourier transform of $S_n(x_1, \cdots, x_{n-1}, 0)$. The conditions on M_n are:

1. M_n is a Lorentz scalar;
2. M_n satisfies the dilatation property

$$M_n(\lambda p_1, \cdots, \lambda p_{n-1}) = \lambda^{4-n} M_n(p_1, \cdots, p_{n-1}); \quad (6.16)$$

3. M_n satisfies the differential equation

$$\sum_{i=1}^{n-1} \left[p_{i\mu} \left(\frac{\partial}{\partial n} \right)^2 - 2p_{i\nu} \frac{\partial}{\partial p_{i\nu}} \frac{\partial}{\partial p_i^\mu} - 2 \frac{\partial}{\partial p_i^\mu} \right] \times M_n(p_1, \cdots, p_{n-1}) = 0. \quad (6.17)$$

In the following sections, we use these conditions to obtain an integral representation of $M_4(p_1, p_2, p_3)$.

7. M_4 EXPRESSED IN TERMS OF LORENTZ SCALARS

By virtue of condition 1 of the previous section, the scattering amplitude $M_4(p_1, p_2, p_3)$ does not depend on the twelve components of p_1, p_2, p_3 separately, but rather on the six Lorentz scalars

$$u_i = p_i^2, \quad i = 1, 2, 3, \tag{7.1}$$

$$v_i = p_j \cdot p_k, \tag{7.2}$$

where i, j, k, here and wherever they appear together, are any permutation of 1, 2, 3. We therefore write

$$M_4(p_1, p_2, p_3) = M(u; v) = M(u_1, u_2, u_3; v_1, v_2, v_3). \tag{7.3}$$

Condition 1 is now satisfied. We turn to condition 3 and substitute $M(u; v)$ for M_4 in Eq. (6.17), using

$$\frac{\partial}{\partial p_{i\mu}} M_4(p_1, p_2, p_3) = 2p_i^\mu M_{u_i} + p_j^\mu M_{v_k} + p_k^\mu M_{v_j}. \tag{7.4}$$

In the last expression, derivatives of $M(u; v)$ are denoted by subscripts

$$M_{u_i} \equiv \frac{\partial}{\partial u_i} M(u; v), \quad M_{v_i^2} \equiv \frac{\partial}{\partial v_i^2} M(u; v), \quad \text{etc.} \tag{7.5}$$

Equation (6.17) now takes on the form

$$p_{1\mu}T_1 + p_{2\mu}T_2 + p_{3\mu}T_3 = 0, \tag{7.6}$$

where

$$\begin{aligned} T_i = {} & -4u_i M_{u_i^2} + (u_k - 2v_j)M_{v_j^2} + (u_j - 2v_k)M_{v_k^2} \\ & - 2v_i M_{v_i v_j} - 2v_i M_{v_i v_k} + 2v_i M_{v_j v_k} - 4u_j M_{u_j v_k} \\ & - 4u_k M_{u_k v_j} - 2M_{v_j} - 2M_{v_k}. \end{aligned} \tag{7.7}$$

Since the vectors p_1, p_2, p_3 are, in general, independent, Eq. (7.6) entails

$$T_i = 0, \quad i = 1, 2, 3. \tag{7.8}$$

These are three partial differential equations for $M(u; v)$, which are equivalent to conditions 1 and 3 of Sec. 6.

8. FOURIER TRANSFORM SOLUTION

The differential equations (7.8) are of second order, and we are not able to solve them directly. We therefore introduce a Fourier transformation that turns them into first-order partial differential equations in the transformed variables. We put

$$\begin{aligned} M(u; v) = {} & \int dU_1\, dU_2\, dU_3\, dV_1\, dV_2\, dV_3 \\ & \times \exp\left[i(u_1U_1 + u_2U_2 + u_3U_3 + v_1V_1 + v_2V_2 + v_3V_3)\right] \\ & \times \tilde{M}(U; V); \end{aligned} \tag{8.1}$$

Eqs. (7.7) and (7.8) for M are equivalent to the following equations for $\tilde{M}$:

$$\begin{aligned} & 4U_i^2 \tilde{M}_{U_i} + (4U_jV_k - V_k^2)\tilde{M}_{U_j} + (4U_kV_j - V_j^2)\tilde{M}_{U_k} \\ & \quad + 2(V_iV_j - V_jV_k + V_kV_i)\tilde{M}_{V_i} + 2V_j^2\tilde{M}_{V_j} + 2V_k^2\tilde{M}_{V_k} \\ & \qquad = -8(U_i + V_j + V_k)\tilde{M}. \end{aligned} \tag{8.2}$$

The above expression is invariant under the exchange of the indices j and k. When all permutations of the three indices are used, (8.2) gives rise to three distinct equations, obtainable from each other by a cyclic permutation of the indices.

For any given choice of indices, Eq. (8.2) is a first-order partial differential equation that can be readily solved. The most general solution is a product of a special solution with the general solution of the equation in which the right-hand side has been set equal to zero.[11] This in turn is an arbitrary function of five particular functions of the six variables. These five forms may be obtained by the method of characteristic equations and characteristic curves.[12]

Our task is to satisfy all three equations. It turns out that when the solution to one equation is substituted into the second equation, the arbitrary function is restricted to depend on only four functions of the original five forms, i.e., on only four particular forms in the six variables. By the time the third equation is taken into account, the arbitrary function may depend on only three forms in the six variables.

We have followed the procedure outlined above. The general solution to the three equations contained in (8.2) came out as

$$\begin{aligned} & \tilde{M}(U; V) \\ & = (V_1 - V_3)^2(2U_3V_3 - V_1V_2)^{-2} \\ & \quad \times (2U_1V_1 + 2U_2V_2 - 4U_1U_2 - V_1V_3 - V_2V_3 + V_3^2)^{-2} \\ & \quad \times H(\omega, \chi, \lambda), \end{aligned} \tag{8.3}$$

[11] This is not quite the most general solution as some generalized function solutions are not included. One of the solutions so excluded is $\delta^3(U)\delta^3(V)$, which satisfies all three equations in (8.2).

[12] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York, 1962), Vol. II, p. 62.

where

$$\omega = \tfrac{1}{2}(2V_1 - V_2 - V_3) \times \left[\frac{V_2 - V_3}{V_2(2U_2V_2 - V_3V_1)} + \frac{V_1 - V_3}{V_1(2U_1V_1 - V_2V_3)} - \frac{1}{V_1V_2}\right], \tag{8.4}$$

$$\chi = \tfrac{1}{2}(2V_1 - V_2 - V_3) \times \left[\frac{V_3 - V_2}{V_3(2U_3V_3 - V_1V_2)} + \frac{V_1 - V_2}{V_1(2U_1V_1 - V_2V_3)} - \frac{1}{V_3V_1}\right], \tag{8.5}$$

$$\lambda = \frac{2V_1 - V_2 - V_3}{V_3 - V_2}, \tag{8.6}$$

and where H is an arbitrary function of three variables.

The above solution is not manifestly symmetric under a permutation of the indices 1, 2, 3, but it is, of course, symmetric. A permuted solution with some particular choice of the arbitrary function H equals the original solution with some other appropriate choice of H.

The quantity λ is dimensionless, but ω and χ have a dimension of length to the -2 power. When condition 2 of Sec. 6 (the dilatation property which has not been used so far) is applied to $M(u; v) = M_4$, it requires $M(u; v)$ to be dimensionless. This in turn implies that the arbitrary function H in Eq. (8.3) must also be dimensionless. Consequently, H can depend only on the ratio of ω and χ. It may still, however, depend on the signs of ω and χ separately. This restricts H to the form

$$H(\omega, \chi, \lambda) = G(\omega/|\omega|, \chi/\omega, \lambda) = f_1(\chi/\omega, \lambda) + \epsilon(\omega) f_2(\chi/\omega, \lambda), \tag{8.7}$$

where f_1 and f_2 are arbitrary functions of two variables, and $\epsilon(\omega)$ is the sign function:

$$\epsilon(\omega) \equiv 1, \quad \text{when} \quad \omega > 0, \qquad \equiv -1, \quad \text{when} \quad \omega < 0. \tag{8.8}$$

When H in Eq. (8.3) is restricted to the form (8.7) and $\tilde{M}(U; V)$ is substituted from Eq. (8.3) into Eq. (8.1), we recover $M(u; v)$. This form of $M(u; v)$ is the most general which is allowed by conformal invariance except for excluded contributions of singular $\tilde{M}(U; V)$ (see Footnote 8). The possible solution $\tilde{M}(U; V) = \delta^3(U)\delta^3(V)$ gives rise to $M(u; v) = 1$, which satisfies all our conditions for a conformally invariant amplitude.

9. CHANGE OF VARIABLES

To separate the integrations involving arbitrary functions from others that can, in principle, be carried out, we introduce a change of variables into the Fourier integral of Eq. (8.1). Naturally we choose λ as one of the new variables and

$$\delta \equiv \chi/\omega \tag{9.1}$$

as another. For the other four variables we choose

$$\alpha \equiv \tfrac{1}{2}(V_2 - V_3), \tag{9.2}$$

$$\xi \equiv (V_2 + V_3)/(V_2 - V_3), \tag{9.3}$$

$$\nu \equiv (V_2 - V_3)^3/8V_1(2U_1V_1 - V_2V_3), \tag{9.4}$$

and

$$\tau \equiv \alpha\lambda\omega. \tag{9.5}$$

When the integral in Eq. (8.1) is re-expressed in terms of these new variables and $\tilde{M}(U; V)$, or Eqs. (8.3) through (8.7) are used, we find

$$\begin{aligned} M(u; v) &= \int d\delta\, d\lambda \frac{d\alpha}{\alpha}\frac{d\tau}{\tau} d\xi\, d\nu \\ &\times \exp\left\{i\alpha\left[v_1(\xi - \lambda) + v_2(\xi + 1) + v_3(\xi - 1)\right.\right. \\ &+ \tfrac{1}{2}u_1\frac{\xi^2 - 1}{\xi - \lambda} + \tfrac{1}{2}u_2(\xi - \lambda)\frac{\xi - 1}{\xi + 1} \\ &+ \tfrac{1}{2}u_3(\xi - \lambda)\frac{\xi + 1}{\xi - 1} + \tfrac{1}{2}u_1\frac{1}{(\xi - \lambda)^2\nu} \\ &+ \frac{u_2}{(\xi + 1)^2}\left[\tau + \nu(\lambda - 1) + \frac{1}{(\xi - \lambda)(\xi + 1)}\right]^{-1} \\ &\left.\left. + \frac{u_3}{(\xi - 1)^2}\left[\delta\tau - \nu(\lambda + 1) - \frac{1}{(\xi - \lambda)(\xi - 1)}\right]^{-1}\right]\right\} \\ &\times [\epsilon(\alpha\tau) f_1\delta(, \lambda) + f_2(\delta, \lambda)]. \end{aligned} \tag{9.6}$$

The range of all variables of integration is the real axis from $-\infty$ to $+\infty$. The variables U_i, V_i expressed in terms of the new variables of integration appear in Eq. (9.6) as the coefficients of the corresponding u_i, v_i in the exponent. The λ and δ integrations involve the arbitrary functions; the other integrations are explicit.

10. REMARKS

$M(u; v)$ of Eq. (9.6) is not quite the most general conformal-invariant scattering amplitude, because singular $\tilde{M}(U; V)$ were excluded in Sec. 8. Among the possibilities thus excluded is $\tilde{M}(U; V) = \delta^3(U)\delta^3(V)$, which leads to $M(u; v) = 1$. Therefore, a constant may always be added to Eq. (9.6).

$M(u; v)$ of Eq. (9.6) may be written as

$$M(u; v) = \int d\delta\, d\lambda [M_1(u; v \mid \delta, \lambda) f_1(\delta, \lambda) + M_2(u; v \mid \delta, \lambda) f_2(\delta, \lambda)]. \tag{10.1}$$

M_1 and M_2 are obtained by performing the α, τ, ν, and ξ integrations in the coefficients of f_1 and f_2 in Eq. (9.6) [see also Eqs. (A2.2) and (A2.3) in Appendix A]. M_1 and M_2 are themselves possible scattering amplitudes for any values of the parameters δ and λ as are any linear combinations. In fact, Eq. (10.1) expresses $M(u; v)$ as the most general linear combination of M_1 and M_2. The "arbitrary functions" f_1 and f_2 need not actually be functions; they could be generalized functions, e.g., δ functions giving rise to M_1 or M_2 themselves or derivatives of δ functions giving rise to derivatives of M_1 and M_2 with respect to the parameters δ and λ.

The α, τ, ν, and ξ integrations in Eq. (9.6) are not convergent in the ordinary sense. We understand these integrations as generalized integrations, and regularize them accordingly.[13] The final test of M_1 and M_2 so obtained is that they satisfy the differential equations (7.7), (7.8), and be homogeneous of order zero in the u, v variables.

Here are a few remarks about the interpretation of the divergent integrals. In terms of ν, the integrand tends to a constant at infinity. This can be regularized by introducing $1/\nu$ as a new variable and interpreting the double pole at the origin according to Ref. 13 [p. 52, Eq. (7)]. Equivalently, one could close the contour of the ν integration at infinity above and below the real axis and take the average. In terms of τ, the integrand behaves at infinity as $1/\tau$ in M_2 and as $1/|\tau|$ in M_1. The first of these is most easily interpreted by deforming the contour into a half-circle at infinity that contributes πi (see Appendix A6). The α and ξ integrations at infinity are interpreted by closing the contour of integration on that side of the real axis on which the half-circle at infinity contributes nothing to the integral.

The pole at $\tau = 0$ may be interpreted as any linear combination of $1/(\tau + i0)$ and $1/(\tau - i0)$ with coefficients adding to one. This is so because it is a remnant of the explicit double pole in the Fourier transform (8.1) with $\tilde{M}(U; V)$ substituted from Eq. (8.3). However, singular functions that result from change of variable are interpreted to be consistent with the original integrations. This leads to the interpretation of singular functions as real generalized functions. Accordingly, the $1/\alpha$ is interpreted as a principal value.

A special remark is due regarding the interpretation of $1/|\alpha|$ and $1/|\tau|$.[14] There exists no interpretation of these singular functions that conserves their formal homogeneity. The homogeneity of the $d\alpha/|\alpha|$ integration is what insures the homogeneity of M_1 in the variables u, v. Therefore, the M_1 resulting from the generalized integration might not be homogeneous. However, the singularity at $\alpha = 0$ may not actually be present after the τ, ν, and ξ integrations have been performed. In fact, when $\alpha = 0$ is substituted in the integrand, the generalized integrals on τ, ν, and ξ give zero. Also, the explicit special example of M_1 in Appendix B is homogeneous. Even if M_1 is not homogeneous, it should not be discarded, because for some class of functions f_1 the resulting $M(u; v)$ may still be homogeneous. For example, the singularity at $\alpha = 0$ and with it any possible inhomogeneity disappear in the derivative of M_1 with respect to δ (see Appendix A5).

On the mass shell we have $u_1 = u_2 = u_3 = v_1 + v_2 + v_3 = 0$, and the integrand in Eq. (9.6) becomes independent of ν and ξ. The ν and ξ integrals then diverge as ordinary integrals and vanish as generalized integrals.[15] Equation (9.6) thus vanishes upon substitution of the mass shell. We are not sure whether a different mass-shell result cannot sometimes be obtained by taking the limit of approaching the mass shell after the quadratures are carried out. This does not happen in any of the explicit forms of Eq. (9.6) that we know (Appendix B). If, in fact, Eq. (9.6) does vanish in the mass-shell limit, the on-mass-shell scattering amplitude is made up of a constant and possibly other contributions of singular $\tilde{M}$. All the examples we know are constant on the mass shell.

The ultimate test that M_1 and M_2 must pass is that they be homogeneous of order zero in the u, v and satisfy the differential equations (7.7) and (7.8). Ideally we should have carried out the four integrations and tested the resulting M_1 and M_2. However, it is shown at the end Appendix A6 that the four integrations cannot, in general, be carried out in terms of elementary functions. The authors were not able to carry out explicitly all these integrations in terms of standard special functions, although this might be possible. We were, however, able to carry all integrations out for special values of δ and λ. A list of M_1 and M_2 for certain special values of δ and λ is given in Appendix B. These special M_1 and M_2 are elementary functions of u, v (homogeneous of order zero) that have been checked to satisfy the differential equations.

For general δ and λ, we have been able to do the α and τ integrations. We have also been able to carry out explicitly the α, τ, and ν integrations in the δ derivative of M_2 as well as in M_2 at $\delta = \infty$. These two

[13] I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic Press Inc., New York, 1964), Vol. 1.

[14] See p. 25 of Ref. 13.

[15] See p. 68 of Ref. 13.

between themselves contain the same information as the original M_2. These explicit integrations are described in Appendix A.

From any one function of u, v that solves the differential equations, others may be generated, in general, by permutations of the indices on the u, v. It is also permissible to interchange p_k with the p_4 that was eliminated by the δ function in Eq. (5.3). This interchange is effected by the substitutions

$$u_i \to u_i, \tag{10.2}$$

$$u_j \to u_j, \tag{10.3}$$

$$u_k \to u_1 + u_2 + u_3 + 2v_1 + 2v_2 + 2v_3, \tag{10.4}$$

$$v_i \to -v_i - v_k - u_j, \tag{10.5}$$

$$v_j \to -v_j - v_k - u_i, \tag{10.6}$$

$$v_k \to v_k. \tag{10.7}$$

11. CONCLUSION

Our final result, Eq. (9.6), is the most general conformal invariant two-body scattering amplitude, except for the contribution of singular $\tilde{M}(U; V)$ excluded in Sec. 8. One such possible contribution is $M = 1$.

If the mass-shell limit of Eq. (9.6) vanishes (as seems to be the case), it may still assume physical significance in conjunction with some scheme for breaking the symmetry and the mass-shell conditions. This is outside our scope here.

Equation (9.6) has been obtained on the basis of conformal symmetry alone. All dynamical information is contained in the arbitrary functions $f_1(\delta, \lambda)$ and $f_2(\delta, \lambda)$. Constraints on f_1 and f_2 are expected to follow from the conditions of unitarity and statistics. Confrontation with experiment of a symmetry-breaking scheme involving our result will consist of checking whether experimental data can be reproduced by some choice of f_1 and f_2.

Our result applies to the elastic two-body spinless case. Production amplitudes with three particles in the final state can probably be treated quite similarly. Other production amplitudes involve much greater complication in the procedure of Sec. 7.

ACKNOWLEDGMENT

The authors gratefully acknowledge useful conversations with Professor Marshall Baker.

APPENDIX A: EXPLICIT INTEGRATIONS

1. Introduction

Four of the integrations in Eq. (9.6), the α, τ, ν, and ξ integrations, are explicit. These quadratures cannot all be carried out in terms of elementary functions (see Appendix A6). We were not able to carry them all out in terms of standard special functions. However, the α and τ integrations can be carried out in terms of logarithms and dilogarithms; this is done in Secs. A3, A4. In Secs. A5 and A6, an alternative approach is applied to $M_2(u; v \mid \delta, \lambda)$, which allows us to perform the ν as well as α and τ integrations. The result is an elliptic integral.

2. The Integrals

$M(u; v)$ of Eq. (9.6) may be written as

$$M(u; v) = \int d\delta\, d\lambda [M_1(u; v \mid \delta, \lambda) f_1(\delta, \lambda) + M_2(u; v \mid \delta, \lambda) f_2(\delta, \lambda)], \tag{A2.1}$$

where

$$M_1(u; v \mid \delta, \lambda) = \int \frac{d\alpha}{|\alpha|} \frac{d\tau}{|\tau|} d\nu\, d\xi \exp\left\{ i\alpha \left[R + \frac{a}{\nu} + \frac{b}{\tau + \nu(\lambda - 1) + B} + \frac{c}{\delta\tau - \nu(\lambda + 1) - C} \right] \right\}, \tag{A2.2}$$

and

$$M_2(u; v \mid \delta, \lambda) = \int \frac{d\alpha}{\alpha} \frac{d\tau}{\tau} d\nu\, d\xi \exp\left\{ i\alpha \left[R + \frac{a}{\nu} + \frac{b}{\tau + \nu(\lambda - 1) + B} + \frac{c}{\delta\tau - \nu(\lambda + 1) - C} \right] \right\}. \tag{A2.3}$$

All integrations are along the real axis from $-\infty$ to $+\infty$. In the last two integrals we used the abbreviations

$$R = v_1(\xi - \lambda) + v_2(\xi + 1) + v_3(\xi - 1) + \tfrac{1}{2}u_1 \frac{\xi^2 - 1}{\xi - \lambda} + \tfrac{1}{2}u_2(\xi - \lambda)\frac{\xi - 1}{\xi + 1} + \tfrac{1}{2}u_3(\xi - \lambda)\frac{\xi + 1}{\xi - 1}, \tag{A2.4}$$

$$a = \tfrac{1}{2}[u_1/(\xi - \lambda)^2], \tag{A2.5}$$

$$b = u_2/(\xi + 1)^2, \tag{A2.6}$$

$$c = u_3/(\xi - 1)^2, \tag{A2.7}$$

$$B = 1/(\xi - \lambda)(\xi + 1), \tag{A2.8}$$

$$C = 1/(\xi - \lambda)(\xi - 1). \tag{A2.9}$$

Notice that all dependence on the variables of integration α, τ, and ν, and on the parameter δ in Eqs. (A2.2) and (A2.3) is still explicit.

3. The α and τ Integrations in $M_2(u; v \mid \lambda, \delta)$

The α integration in Eq. (A2.3) is the Fourier transform of $1/\alpha$, which is $i\pi$ times the sign of the Fourier transform variable.[16] Therefore,

$$M_2(u; v \mid \delta, \lambda) = \int \frac{d\tau}{\tau} dv\, d\xi i\pi\epsilon\left(R + \frac{a}{\nu} + \frac{b}{\tau + \nu(\lambda - 1) + B} + \frac{c}{\delta\tau - \nu(\lambda + 1) - C}\right), \tag{A3.1}$$

where

$$\epsilon(t) = +1, \quad \text{when} \quad t > 0,$$
$$= -1, \quad \text{when} \quad t < 0. \tag{A3.2}$$

This can be alternatively written as

$$M_2(u; v \mid \delta, \lambda) = \int dv\, d\xi \epsilon(R) \times \int \frac{d\tau}{\tau} \epsilon[(\tau - \gamma_1)(\tau - \gamma_2)(\tau - \gamma_3)(\tau - \gamma_4)], \tag{A3.3}$$

where $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ are the two values

$$-\nu(\lambda - 1) - B, \quad [\nu(\lambda + 1) + C]\delta^{-1}, \tag{A3.4}$$

and the two roots of the quadratic equation

$$(R + a/\nu)[\tau + \nu(\lambda - 1) + B][\delta\tau - \nu(\lambda + 1) - C] + b[\delta\tau - \nu(\lambda + 1) - C] + c[\tau + \nu(\lambda - 1) + B] = 0 \tag{A3.5}$$

are ordered so that γ_1 and γ_2 are real and

$$\gamma_2 > \gamma_1, \tag{A3.6}$$

and if γ_3 and γ_4 are also real, then also

$$\gamma_4 > \gamma_3 > \gamma_2. \tag{A3.7}$$

The $1/\tau$ in Eq. (A3.3) may be interpreted as either $1/(\tau + i0)$ or $1/(\tau - i0)$ or an appropriate linear combination. However, $1/(\tau \pm i0) = 1/\tau \mp i\pi\delta(\tau)$, where $1/\tau$ now stands for the principal value. As the integration over the δ function is easy, we concentrate on $1/\tau$ interpreted as the principal value.

Notice that, as a generalized integral,[17]

$$\int_{-\infty}^{+\infty} \frac{d\tau}{\tau} = 0. \tag{A3.8}$$

We subtract a multiple of Eq. (A3.8) from Eq. (A3.3) and make use of

$$\epsilon(t) - 1 = -2\theta(-t), \tag{A3.9}$$

where $\theta(t)$ is the step function

$$\theta(t) = 1, \quad \text{when} \quad t > 0,$$
$$= 0, \quad \text{when} \quad t < 0. \tag{A3.10}$$

This turns Eq. (A3.3) into

$$M_2(u; v \mid \delta, \lambda) = -2\int dv\, d\xi \epsilon(R) \times \int \frac{d\tau}{\tau} \theta[-(\tau - \gamma_1)(\tau - \gamma_2)(\tau - \gamma_3)(\tau - \gamma_4)]$$
$$= \begin{cases} -2\int dv\, d\xi \epsilon(R) \log \dfrac{|\gamma_2\gamma_4|}{|\gamma_1\gamma_3|}, \\ \quad \text{when} \quad \gamma_i, \quad i = 1, 2, 3, 4 \text{ are all real}, \\ -2\int dv\, d\xi \epsilon(R) \log \dfrac{|\gamma_2|}{|\gamma_1|}, \\ \quad \text{when } \gamma_3 \text{ and } \gamma_4 \text{ are complex.} \end{cases} \tag{A3.11}$$

The dependence of the last result on ν and ξ is rather complicated.

4. The α and τ Integrations in $M_1(u; v \mid \delta, \lambda)$

The α integration in Eq. (A2.2) is the Fourier transform of $1/|\alpha|$, which is $\log|\sigma|$,[18] where σ is the Fourier transform variable. This result is not homogeneous of order zero in σ as formally suggested by Eq. (A2.2). $M(u; v)$ must be homogeneous of order zero in the variables u, v as part of the requirements of the conformal group. Therefore, if the inhomogeneity does not disappear through the τ, ν, or ξ integrations, the arbitrary function $f_1(\delta, \lambda)$ must be restricted to a class of functions that make the inhomogeneity disappear through the δ and λ integrations.

Using the form of the Fourier transform, Eq. (A2.2) becomes

$$M_1(u; v \mid \delta, \lambda) = \int dv\, d\xi \frac{d\tau}{|\tau|} \log\left| R\frac{(\tau - \mu_1)(\tau - \mu_2)}{(\tau - \rho_1)(\tau - \rho_2)}\right|, \tag{A4.1}$$

where ρ_1 and ρ_2 are the two quantities (A3.4), and μ_1 and μ_3 are the two roots of the quadratic equation (A3.5).

The τ integral is divergent at both $\tau = 0$ and $\tau = \infty$. The interpretation of this integral as a generalized integral cannot conserve the homogeneity of $d\tau/|\tau|$; it

[16] See Ref. 13, p. 360, entry 19.
[17] See Ref. 13, p. 52.
[18] See Ref. 13, p. 361, entry 27 for $m = 0$. This Fourier transform involves an arbitrary additive constant. The choice of Ref. 13 differs from ours.

involves four arbitrary constants and is[14]

$$\int_{-\infty}^{+\infty}\frac{d\tau}{|\tau|}\log\left|R\frac{(\tau-\mu_1)(\tau-\mu_2)}{(\tau-\rho_1)(\tau-\rho_2)}\right|$$
$$=\int_{-\infty}^{-M}\frac{d\tau}{|\tau|}\log\left|\frac{(\tau-\mu_1)(\tau-\mu_2)}{(\tau-\rho_1)(\tau-\rho_2)}\right|$$
$$+\int_{-M}^{-N}\frac{d\tau}{|\tau|}\log\left|R\frac{(\tau-\mu_1)(\tau-\mu_2)}{(\tau-\rho_1)(\tau-\rho_2)}\right|$$
$$+\int_{-N}^{P}\frac{d\tau}{|\tau|}\log\left|\frac{(\tau-\mu_1)(\tau-\mu_2)\rho_1\rho_2}{(\tau-\rho_1)(\tau-\rho_2)\mu_1\mu_2}\right|$$
$$+\int_{P}^{Q}\frac{d\tau}{\tau}\log\left|R\frac{(\tau-\mu_1)(\tau-\mu_2)}{(\tau-\rho_1)(\tau-\rho_2)}\right|$$
$$+\int_{Q}^{\infty}\frac{d\tau}{\tau}\log\left|\frac{(\tau-\mu_1)(\tau-\mu_2)}{(\tau-\rho_1)(\tau-\rho_2)}\right|, \quad \text{(A4.2)}$$

where M, N, P, Q are the arbitrary constants. They are subject only to the conditions

$$M > N > 0, \quad \text{(A4.3)}$$

$$Q > P > 0. \quad \text{(A4.4)}$$

The integrals in Eq. (A4.2) may be expressed in terms of the dilogarithm L_2.[19]

5. Integration by Parts on δ—An Alternative Approach Applied to $M_2(u; v \mid \delta, \lambda)$

We return to Eq. (A2.1) and perform an integration by parts on δ. This yields

$$M(u;v) = -\int d\delta\, d\lambda\{M_{1\delta}(u; v \mid \delta, \lambda)g_1(\delta, \lambda)$$
$$+ M_{2\delta}(u; v \mid \delta, \lambda)g_2(\delta, \lambda)\}$$
$$+\int\{M_1(u; v \mid \infty, \lambda)g_1(\infty, \lambda)$$
$$+ M_2(u; v \mid \infty, \lambda)g_2(\infty, \lambda)\}\, d\lambda, \quad \text{(A5.1)}$$

where

$$g_i(\delta, \lambda) = \int_{-\infty}^{\delta} dt f_i(t, \lambda), \quad i = 1, 2, \quad \text{(A5.2)}$$

and

$$M_{i\delta}(u; v \mid \delta, \lambda) \equiv \left(\frac{\partial}{\partial\delta}\right)M_i(u; v \mid \delta, \lambda), \quad i = 1, 2. \quad \text{(A5.3)}$$

Following Eq. (A5.1), we shall work on the original $M_i(u; v \mid \delta, \lambda)$ only for $\delta = \infty$ and for general δ we shall treat $M_{i\delta}(u; v \mid \delta, \lambda)$ instead. The

$$M_{1\delta}(u; v \mid \delta, \lambda)$$
$$= -\int d\alpha\, d\tau\, dv\, d\xi\epsilon(\alpha\tau)\frac{ic}{(\delta\tau - \nu(\lambda+1) - C)^2}$$
$$\times \exp\left[i\alpha\left(R + \frac{a}{\nu} + \frac{b}{\tau + \nu(\lambda-1) + B}\right.\right.$$
$$\left.\left. + \frac{c}{\delta\tau - \nu(\lambda+1) - C}\right)\right], \quad \text{(A5.4)}$$

$$M_{2\delta}(u; v \mid \delta, \lambda)$$
$$= -\int d\alpha\, d\tau\, dv\, d\xi\frac{ic}{(\delta\tau - \nu(\lambda+1) - C)^2}$$
$$\times \exp\left[i\alpha\left(R + \frac{a}{\nu} + \frac{b}{\tau + \nu(\lambda-1) + B}\right.\right.$$
$$\left.\left. + \frac{c}{\delta\tau - \nu(\lambda+1) - C}\right)\right] \quad \text{(A5.5)}$$

turn out more tractable than the original M_i. Notice that the singularities at $\alpha = 0$ and $\tau = 0$ and the ambiguities that go with them have disappeared. The double pole in Eqs. (A5.4) and (A5.5) is to be interpreted as[20]

$$\frac{1}{2}\left(\frac{1}{(z+i0)^2} + \frac{1}{(z-i0)^2}\right).$$

We now concentrate on $M_2(u; v \mid \delta, \lambda)$ and on Eq. (A5.5), and introduce a new variable of integration,

$$\psi = \frac{\delta}{\delta\tau - \nu(\lambda+1) - C} + \frac{\delta}{\nu\Lambda + C + \delta B}, \quad \text{(A5.6)}$$

where

$$\Lambda \equiv \lambda + 1 + \delta(\lambda - 1). \quad \text{(A5.7)}$$

In terms of the new variable, Eq. (A5.5) becomes

$$M_{2\delta}(u; v \mid \delta, \lambda) = -i\delta^{-2}\int d\alpha\, d\psi\, dv\, d\xi C$$
$$\times \exp\left\{i\alpha\left[R + \frac{a}{\nu} + \frac{\delta b - c}{\nu\Lambda + C + \delta B}\right.\right.$$
$$\left.\left. + \frac{c}{\delta}\psi - \frac{\delta^2 b}{(\nu\Lambda + C + \delta B)^2}\frac{1}{\psi}\right]\right\}. \quad \text{(A5.8)}$$

The integrand has essential singularities in ψ at the origin and at infinity. The real axis integration on ψ may be interpreted as a contour integration in the complex ψ plane, with the contour passing above (below) the origin for $\alpha b > 0$ ($\alpha b < 0$) and closing at infinity above (below) the real axis for $\alpha c/\delta > 0$ (< 0). The integral vanishes unless the signs of b and

[19] A. Erdélyi, Ed., *Higher Transcendental Functions* (McGraw-Hill Book Co., New York, 1953), Vol. 1, p. 31.

[20] See Ref. 13, p. 60 and Eq. (7), p. 52.

c/δ are opposite. When they are, the integral is $2\pi i$ times the -1 coefficient in the Laurent expansion of the integrand in powers of ψ. This expansion is in terms of Bessel functions,[21]

$$\exp\left[i\alpha\left(s\psi - \frac{t}{\psi}\right)\right] = \sum_{n=-\infty}^{+\infty} \psi^n \left(\frac{s}{t}\right)^{\frac{1}{2}n} J_n(2i\alpha[ts]^{\frac{1}{2}}). \tag{A5.9}$$

The -1 Laurent coefficient involves J_{-1}. However, $J_{-1}(x) = -J_1(x)$, so that

$$M_{2\delta}(u; v \mid \delta, \lambda) = 2\pi\delta^{-\frac{1}{2}} \int d\alpha\, dv\, d\xi \frac{(bc)^{\frac{1}{2}}}{v\Lambda + C + \delta B} \times J_1\left(2i\alpha \frac{(\delta bc)^{\frac{1}{2}}}{v\Lambda + C + \delta B}\right) \epsilon(\alpha\delta c)\theta(-\delta bc) \times \exp\left[i\alpha\left(R + \frac{a}{v} + \frac{\delta b - c}{v\Lambda + C + \delta B}\right)\right]. \tag{A5.10}$$

Because of the θ function, (A5.10) vanishes unless $\delta bc < 0$. For negative δbc, the coefficient of α in the Bessel function is real. When the $\epsilon(\alpha\delta c)$ is taken into account, the α integration becomes the cosine Fourier transform of J_1 which can be found in tables.[22] When this is substituted, Eq. (A5.10) becomes

$$M_{2\delta}(u; v \mid \delta, \lambda) = -\frac{2\pi i}{\delta} \int dv\, d\xi \epsilon(-b)\theta(-\delta bc) \times [1 - [|X|/(X^2 - 1)^{\frac{1}{2}}]\theta(X^2 - 1)], \tag{A5.11}$$

where

$$X = Fv + G + H/v, \tag{A5.12}$$

and

$$F = \tfrac{1}{2}(\Lambda R)(-\delta bc)^{-\frac{1}{2}}, \tag{A5.13}$$

$$G = \tfrac{1}{2}[R(C + \delta B) + a\Lambda + \delta b - c](-\delta bc)^{-\frac{1}{2}}, \tag{A5.14}$$

$$H = \tfrac{1}{2}a(C + \delta B)(-\delta bc)^{-\frac{1}{2}}. \tag{A5.15}$$

Our next step is to introduce X as a new variable of integration. This process requires some attention, since v in terms of X is double valued

$$v = (1/2F)\{X - G \pm [(X - G)^2 - 4FH]^{\frac{1}{2}}\}, \tag{A5.16}$$

and both branches must be used to cover all values of v. Figure 1(a) shows this relation for $FH > 0$ and Fig. 1(b) shows it for $FH < 0$. Inspection of Fig. 1(b) shows that, for $FH < 0$, the sum of the two branches must be used. This leads to the substitution

$$dv = dX/|F|, \tag{A5.17}$$

which turns the v integration in Eq. (A5.11) into an elementary integral that is readily carried out and is found to vanish. For $FH > 0$, inspection of Fig. 1(a) shows that the difference of the two branches should be used, and the substitution

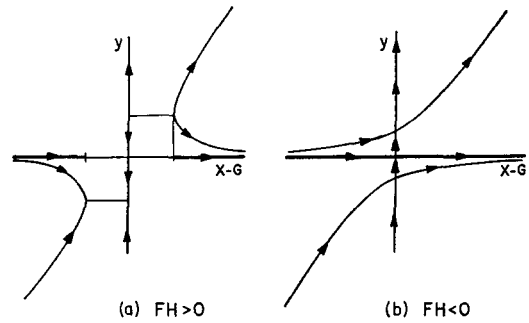


FIG. 1. The relation between v and $X - G$ [Eq. (A5.16)] for $F > 0$ is plotted for $FH > 0$ (a) and $FH < 0$ (b). In case (b), as X runs from $-\infty$ to $+\infty$, one branch of v goes from 0 to ∞ while the other goes from $-\infty$ to 0. The sum of both branches covers all values of v in the positive direction. In case (a), as X ranges from $B + 2(FH)^{\frac{1}{2}}$ to ∞, one branch of v goes from B to ∞ while the other goes from B to 0. The difference of the two branches covers all positive values of v in the positive direction. Similarly, as X ranges from $-\infty$ to $B - 2(FH)^{\frac{1}{2}}$, the opposite difference of the two branches covers all negative values of v in the positive direction.

$$dv = \frac{|X - G|\, dX}{|F|\,[(X - G)^2 - 4FH]^{\frac{1}{2}}} \theta[(X - G)^2 - 4FH] \tag{A5.18}$$

follows. Equation (A5.11) is, therefore, turned into

$$M_{2\delta}(u; v \mid \delta, \lambda) = 4\pi i \int d\xi \epsilon(-b)\theta(-\delta bc)\theta(FH)(1/|F|)Y, \tag{A5.19}$$

where

$$Y = \int dX \frac{|X - G|}{[(X - G)^2 - 4FH]^{\frac{1}{2}}} \theta[(X - G)^2 - 4FH] \times \left[1 - \frac{|X|}{(X^2 - 1)^{\frac{1}{2}}}\theta(X^2 - 1)\right] = Y_1 + Y_2 + Y_3 + Y_4, \tag{A5.20}$$

and where

$$Y_1 = \int dX \left(1 - \frac{4FH}{(X - G)^2}\right)^{-\frac{1}{2}} [1 - (1 - X^{-2})^{-\frac{1}{2}}], \tag{A5.21}$$

$$Y_2 = -\int dX \left(1 - \frac{4FH}{(X - G)^2}\right)^{-\frac{1}{2}} \theta[4FH - (X - G)^2], \tag{A5.22}$$

$$Y_3 = \int dX \left(1 - \frac{4FH}{(X - G)^2}\right)^{-\frac{1}{2}} (1 - X^{-2})^{-\frac{1}{2}} \times \{\theta[(X - G)^2 - 4FH]\theta(1 - X^2) + \theta[4FH - (X - G)^2]\theta(X^2 - 1)\}, \tag{A5.23}$$

$$Y_4 = \int dX \left(1 - \frac{4FH}{(X - G)^2}\right)^{-\frac{1}{2}} (1 - X^{-2})^{-\frac{1}{2}} \times \theta[4FH - (X - G)^2]\theta(1 - X^2). \tag{A5.24}$$

[21] Ref. 19, Vol. 2, p. 7, Eq. (25).

[22] A. Erdélyi, Ed., *Tables of Integral Transforms* (McGraw-Hill Book Co., New York, 1954), Vol. 1, p. 43, entry 3.

Because the integrand is an analytic function that vanishes at infinity, Y_1 vanishes (in all directions in the complex plane) as X^{-2} and has no singularity off the real axis. The contour of integration may be closed at infinity and encloses no singularity. Y_2 and Y_3 are imaginary; as the original Y as well as Y_4 are real, they must cancel out. This leaves us with

$$Y = Y_4 = -\int dX \frac{|X-G|}{[4FH-(X-G)^2]^{\frac{1}{2}}} \frac{|X|}{(1-X^2)^{\frac{1}{2}}} \times \theta[4FH-(X-G)^2]\theta(1-X^2). \tag{A5.25}$$

This is an elliptic integral that can be readily expressed in terms of the standard elliptic integrals.[23]

6. Treatment of $M_2(u; v \mid \infty, \lambda)$

In this section we treat $M_2(u; v \mid \infty, \lambda)$ which appears in the second term in Eq. (A5.1). At $\delta = \infty$, the last term in the exponent in Eq. (A2.3) disappears. We interpret $1/\tau$ as $1/(\tau \pm i0)$ and write

$$M_2(u; v \mid \infty, \lambda) = \int \frac{d\alpha}{\alpha} \frac{d\tau}{\tau \pm i0} dv\, d\xi \times \exp\left\{i\alpha\left[R + \frac{a}{v} + \frac{b}{\tau + \nu(\lambda-1) + B}\right]\right\}. \tag{A6.1}$$

The τ integration may be interpreted as a contour integration. The contour bypasses the pole at $\tau = 0$ above or below the real axis as indicated by the sign in $1/(\tau \pm i0)$; it bypasses the essential singularity at $\tau = -\nu(\lambda-1) - B$ below (above) the real axis when αb is positive (negative). We now deform this contour (Fig. 2) away from the essential singularity into a half-circle at infinity. If the pole is in the way, we leave a small circle around the pole. The half-circle at infinity contributes

$$\pi i \int \frac{d\alpha}{\alpha} dv\, d\xi \exp\left[i\alpha\left(R + \frac{a}{v}\right)\right] \tag{A6.2}$$

when it is below the real axis, and minus that when it is above the real axis. Putting all this together we find

$$M_2(u; v \mid \infty, \lambda) = Z_1 + Z_2, \tag{A6.3}$$

with

$$Z_1 = i\pi \int \frac{d\alpha}{|\alpha|} dv\, d\xi \exp\left[i\alpha\left(R + \frac{a}{v}\right)\right]\epsilon(b), \tag{A6.4}$$

$$Z_2 = \mp 2\pi i \int \frac{d\alpha}{\alpha} dv\, d\xi \theta(\pm\alpha b) \times \exp\left[i\alpha\left(R + \frac{a}{v} + \frac{b}{\nu(\lambda-1)+B}\right)\right]. \tag{A6.5}$$

[23] W. Grobner and N. Hofreiter, *Integraltafel* (Springer-Verlag, Vienna, 1958), Vol. 1, Eq. (8a6), p. 85, and the preceding tables.

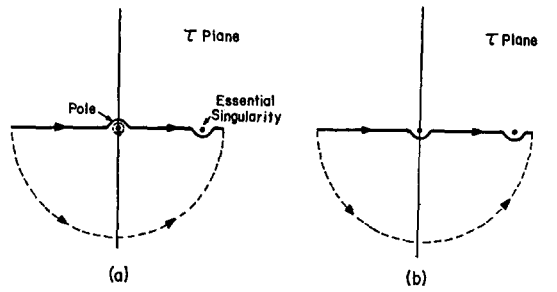


FIG. 2. The contour of the τ integration in Eq. (A6.1) is shown (heavy line). It bypasses a pole (at the origin) and an essential singularity. We deform the contour away from the essential singularity into a half-circle at infinity (dotted). In case (a), where the pole is in the way, we leave a small circle around the pole in the process of deformation. If the pole and the essential singularity are on the same side of the contour, as in case (b), this is not necessary.

The α integration in Z_1 is the Fourier transform of $1/|\alpha|$.[18] Performing it we find

$$Z_1 = \pi i \int dv\, d\xi \epsilon(b) \log|R + a/v|. \tag{A6.6}$$

Similarly, the α integration in Z_2 is the Fourier transform of

$$\theta(\pm\alpha b)\frac{1}{\alpha} = \frac{1}{2\alpha}[1 + \epsilon(\pm\alpha b)] = \frac{1}{2\alpha} \pm \epsilon(b)\frac{1}{2|\alpha|}. \tag{A6.7}$$

When the Fourier transforms of the last two expressions are substituted in Eq. (A6.5), we find

$$Z_2 = Z_{21} + Z_{22}, \tag{A6.8}$$

where

$$Z_{21} = \pm\pi^2 \int dv\, d\xi \epsilon\left(R + \frac{a}{v} + \frac{b}{\nu(\lambda-1)+B}\right), \tag{A6.9}$$

$$Z_{22} = -\pi i \int dv\, d\xi \epsilon(b) \log\left|R + \frac{a}{v} + \frac{b}{\nu(\lambda-1)+B}\right|. \tag{A6.10}$$

We now recombine Z_{22} with Z_1:

$$Z_1 + Z_{22} = \pi i \int dv\, d\xi \epsilon(b) \times \log\left|\left[R + \frac{a}{v}\right]\left[R + \frac{a}{v} + \frac{b}{\nu(\lambda-1)+B}\right]^{-1}\right| = 0. \tag{A6.11}$$

The v integration in (A6.11) is elementary. (Some care must be exercised at the infinite limits.) This leaves us with Z_{21}; the v integration in this expression diverges

as an ordinary integration. As a generalized integral,

$$M_2(u; v \mid \infty, \lambda) = Z_{21} = \pm\pi^2 \int d\xi \int_{-\infty}^{+\infty} dv \times \left[\epsilon\left(R + \frac{a}{\nu} + \frac{b}{\nu(\lambda-1)+B}\right) - \epsilon(R)\right] = \pm 2\pi^2 \int d\xi \epsilon(b) \frac{1}{|R(\lambda-1)|} \times \{[RB + a(\lambda-1) + b]^2 - 4RBa(\lambda-1)\}^{\frac{1}{2}}. \quad \text{(A6.12)}$$

The remaining ξ integration in Eq. (A6.12) is elliptic. This shows that $M_2(u; v \mid \delta, \lambda)$ cannot in general be expressed in terms of elementary functions.

APPENDIX B: SPECIAL SOLUTIONS

In a few special cases, special values or limits of δ and λ, we have been able to carry out the quadratures involved in M_1 and M_2 [see Eqs. (A2.2) and (A2.3)]. The resulting M_1 and M_2 are given below:

$$M_2(u; v \mid \infty, 1) = -i\pi^3 \frac{|u_1|}{[(s + v_1 + v_3)^2 - su_4]^{\frac{1}{2}}} \times \left|\frac{(v_1+v_3)^2}{s^2} + \frac{v_1 - u_2 + v_3}{2s}\right|, \quad \text{when} \quad u_2u_4 > 0,$$

$$= \tfrac{1}{2}i\pi^3 \frac{|u_1|}{s^2}(u_2 - u_4), \quad \text{when} \quad u_2u_4 < 0; \quad \text{(B1)}$$

$$\lim_{\lambda\to\infty} \lambda^2 M_1(u; v \mid \delta, \lambda) = -2\pi^2 i K_1 \theta(u_1u_4)\left\{1 - \frac{|u_1 + v_2 + v_3|}{[(u_1+v_2+v_3)^2 - u_1u_4]^{\frac{1}{2}}} \times \theta[(u_1 + v_2 + v_3)^2 - u_1u_4]\right\}; \quad \text{(B2)}$$

$$\lim_{\lambda\to\infty} \lambda^2 M_2(u; v \mid \delta, \lambda) = -2\pi^2 i K_2 \theta(u_1u_4) \frac{u_1 + v_2 + v_3}{[u_1u_4 - (u_1+v_2+v_3)^2]^{\frac{1}{2}}} \times \theta[u_1u_4 - (u_1 + v_2 + v_3)^2]. \quad \text{(B3)}$$

The notations

$$s = u_1 + 2v_2 + u_3 \quad \text{(B4)}$$

and

$$u_4 = u_1 + u_2 + u_3 + 2v_1 + 2v_2 + 2v_3 \quad \text{(B5)}$$

have been used. (In terms of p_1, p_2, p_3, p_4, the momenta of the four particles, $s = (p_1 + p_3)^2$ and $u_4 = p_4^2$. The K_1 in Eq. (B2) is the arbitrary result of the regularized integral $\int_{-\infty}^{+\infty} d\tau/|\tau|$. The K_2 in Eq. (B3) stands for $\int_{-\infty}^{+\infty} d\tau/\tau$; this integral is interpreted as an arbitrary linear combination of

$$\int_{-\infty}^{+\infty} d\tau/(\tau + i0) = -\pi i$$

and

$$\int_{-\infty}^{+\infty} d\tau/(\tau - i0) = \pi i,$$

thus K_2 is also arbitrary. The amplitudes in Eqs. (B1)–(B3) have been checked to satisfy the differential equations. They all vanish as the mass shell is approached.

Here are a few more simple functions of the variables $u; v$ that satisfy the conditions on conformal invariant scattering amplitudes:

$$v_2/(v_2^2 - u_1u_3)^{\frac{1}{2}}, \quad \text{(B6)}$$

$$u_1u_2u_3u_4/s^4, \quad \text{(B7)}$$

$$u_1u_2/s^2. \quad \text{(B8)}$$

Each of the examples in this Appendix can generate several others by permutations of the indices 1, 2, 3 or the substitutions of Eqs. (10.2)–(10.7). All the explicit examples we have come across feature trivial mass-shell behavior.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Effective-Range Formula for Two-Body Scattering Coupled to a Closed Secondary Channel

MARK W. KERMODE*
Department of Physics, McMaster University, Hamilton, Ontario

(Received 3 September 1968)

The multichannel effective-range formulas are extended to the case of a closed secondary two-body channel in which there is a repulsive Coulomb potential.

I. INTRODUCTION

The results of low-energy scattering experiments can best be expressed in terms of the partial-wave phase shifts δ_l. The energy dependence of a phase shift in a given energy region may provide useful information about the scattering potential even though the potential itself cannot be determined uniquely. In particular, one of the elementary results of scattering theory is that, for the single-channel S-wave scattering of uncharged particles, $k \cot \delta_0$ is a slowly varying function of energy. An elementary or composite particle incident on a nucleus can give rise to a large number of intermediate states and hence a complicated energy dependence of the phase shift. It would be very useful if an effective-range-type formula expressing a function of the phase shift in terms of slowly varying functions of energy could be derived for this case.

Effective-range formulas for the scattering of a projectile b_i by a target c_i in the presence of $n-1$ secondary *two-body* channels containing particles b_j and c_j $(j=1,\cdots,n; j\neq i)$ have been derived and it was found that they can be written in a convenient determinantal form.[1] One interesting physical example is α–α elastic scattering in which one of the α-particles becomes excited. The formulas given by Kermode[1] for a repulsive Coulomb field in a closed secondary channel, e.g.,

$$\begin{array}{ccc} \alpha+\alpha & \longrightarrow & \alpha+\alpha, \\ & \searrow \quad \nearrow & \\ & \alpha+\alpha^* & \end{array}$$

can only be used if we introduce a hard-core potential in the secondary channel. The main purpose of the present paper is to remove this limitation. To extend the previous formulas to the case of scattering by a repulsive Coulomb field and a short-range potential in one or more of the secondary channels we have to study the logarithmic derivative of the Whittaker function. We shall consider a two-channel reaction, since the extension to more channels is straightforward.[1]

The effective-range formula for the case of a hard core of radius a in the second channel can be written as

$$\begin{vmatrix} M_{11}(E)-y(k_1,\delta_{l_1}) & (m_1/m_2)^{\frac{1}{2}}M_{12}(E) \\ (m_2/m_1)^{\frac{1}{2}}M_{12}(E) & M_{22}(E)-z(k_2) \end{vmatrix} \equiv 0, \quad (1)$$

where the functions y and z are given by

$$y(k_1,\delta_{l_1}) = \frac{(1+\eta_1^2)\cdots(l_1^2+\eta_1^2)}{(l_1!)^2} \times 2\eta_1 k_1^{2l_1+1}\left[\frac{\pi\cot\delta_{l_1}}{\exp(2\pi\eta_1)-1}+h^+(\eta_1)\right]$$

and

$$z(k_2) = \frac{[(2l_2-1)!!]^2}{a^{2l_2}}\, k_2 \frac{W'_{l_2}(\eta_2, k_2a)}{W_{l_2}(\eta_2, k_2a)}, \quad (2)$$

with

$$\eta_i = Z_{1i}Z_{2i}e^2m_i/\hbar^2k_i, \quad i=1,2,$$

and

$$h^+(\eta) = \mathrm{Re}\,\psi(1+i\eta) - \log\eta. \quad (3)$$

E is the energy matrix

$$\begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}.$$

The suffix on the relative momentum

$$k_i = (2m_i|E_i|/\hbar^2)^{\frac{1}{2}},$$

the angular momentum l_i, the reduced mass m_i, and the two charges Z_{1i} and Z_{2i} refer to the first or second channel. The elements of the M matrix are slowly varying functions of energy and can be written as a series in k^2. Usually one considers only the first two or three terms. $W_l(\eta,\rho)$ is the Whittaker function and the prime denotes differentiation with respect to $\rho = ka$. The phase shift δ_{l_1} is necessarily real.

In principle, the extension of (1) to the case of $a=0$ is straightforward: One takes the limit of $z(k_2)$ [Eq. (2)] as the hard-core radius tends to zero. Unfortunately, Eq. (2) contains a series in k^2 with coefficients that become infinite. One argues that this

* Present address: Department of Applied Mathematics, Liverpool University, U.K.

[1] M. W. Kermode, Nucl. Phys. A99, 605 (1967).

series can be absorbed by $M_{22}(E)$ since (1) is an identity and that only the renormalized series is important. This is very similar to the open channel case and we note that the renormalized expansion of Eq. (2) with the irregular Coulomb wavefunction in place of the Whittaker function has already been reported.[2] In the following we consider the special cases $l_2 = 0$, 1, and 2 and from the results we can write down the expression for any l_2.

II. THE LOGARITHMIC DERIVATIVE OF THE WHITTAKER FUNCTION

The integral representation of the Whittaker function is[3,4]

$$W_l(\eta, \rho) = \frac{e^{-\rho-\eta \log 2\rho}}{\Gamma(l+1+\eta)} \int_0^\infty t^{\eta+l} e^{-t} \left(1 + \frac{t}{2\rho}\right)^{l-\eta} dt.$$

Differentiating $W_l(\eta, \rho)$ with respect to ρ, we have

$$\frac{W_l'(\eta, \rho)}{W_l(\eta, \rho)} = -1 - \frac{l}{\rho} + \frac{(\eta - l)}{\rho} J_l(\eta, \rho), \tag{4}$$

where

$$J_l(\eta, \rho) = I_l(\eta + 1, \rho)/I_l(\eta, \rho) - 1, \tag{5}$$

and

$$I_l(\eta, \rho) = (2\rho)^{-2l-1} e^{2\rho} \int_0^\infty t^{\eta+l} e^{-t} (2\rho + t)^{l-\eta}\, dt$$
$$= \int_1^\infty z^{2l} \left(1 - \frac{1}{z}\right)^{l+\eta} e^{-2\rho z}\, dz.$$

The function $I_l(\eta, \rho)$ can be written

$$I_l(\eta, \rho) = \int_1^\infty z^{2l} \left(1 - \frac{(\eta + l)}{z} + \frac{(\eta + l)(\eta + l - 1)}{2!\, z^2} - \cdots + \frac{(\eta + l) \cdots (\eta + 1 - l)}{(2l)!\, z^{2l}}\right) e^{-2\rho z}\, dz$$
$$- \frac{(\eta + l) \cdots (\eta - l)}{(2l + 1)!} \int_1^\infty \frac{e^{-2\rho z}}{z}\, dz + K_l(\eta, \rho), \tag{6}$$

where

$$K_q(\eta, \rho) = \int_1^\infty \left(\frac{(\eta + q) \cdots (\eta - q - 1)}{(2q + 2)!\, z^2} - \frac{(\eta + q) \cdots (\eta - q - 2)}{(2q + 3)!\, z^3} + \cdots\right) e^{-2\rho z}\, dz. \tag{7}$$

In Eq. (7), q can be integer or half-integer. The function $K_q(\eta, \rho)$ obeys the equation

$$K_q(\eta + 1, \rho) - K_q(\eta, \rho) = K_{q-\frac{1}{2}}(\eta + \tfrac{1}{2}, \rho).$$

In the following, we shall require the value of $K_q(\eta, \rho = 0)$. This can be written as

$$K_q(\eta, 0) = (-1)^{2q} \int_1^\infty z^{2q} \left[\left(1 - \frac{1}{z}\right)^{\eta+q} - 1 + \frac{(\eta + q)}{z} - \frac{(\eta + q)(\eta + q - 1)}{2!\, z^2} + \cdots + (-1)^{2q} \frac{(\eta + q) \cdots (\eta - q)}{(2q + 1)!\, z^{2q+1}}\right] dz. \tag{8}$$

From Eq. (8) we find the useful recurrence relation

$$K_q(\eta, 0) = \frac{(-1)^{2q}}{2q + 1} \left(1 - (\eta + q) + \frac{(\eta + q)(\eta + q - 1)}{2!} - \cdots (-1)^{2q} \frac{(\eta + q) \cdots (\eta - q)}{(2q + 1)!}\right) + \frac{(\eta + q)}{2q + 1} K_{q-\frac{1}{2}}(\eta - \tfrac{1}{2}, 0),$$

with $K_{-\frac{1}{2}}(\eta - q - \frac{1}{2}, 0) = C + \psi(\eta - q)$. Euler's constant C and the digamma function ψ are well known.[5]

The evaluation of the first integral on the right-hand side of Eq. (6) is straightforward. The second integral and the $K_l(\eta, \rho)$ function can be written in terms of the exponential integrals $E_n(2\rho)$ $(n = 1, 2, \cdots)$.[5] From the results of these integrations we obtain the power-series expansion of $J_l(\eta, \rho)$ [Eq. (5)].

A. *S*-Wave Result

The evaluation of the integrals for zero angular momentum is simple and one finds that

$$I_0(\eta, \rho) = (2\rho)^{-1} e^{-2\rho} - \eta E_1(2\rho) + K_0(\eta, \rho).$$

Substituting this result into Eqs. (4) and (5), we have

$$\frac{W_0'(\eta, \rho)}{W_0(\eta, \rho)} = -1 + \frac{2\eta e^{-2\rho}[-E_1(2\rho) + K_0(\eta + 1, \rho) - K_0(\eta, \rho)]}{1 + 2\rho e^{2\rho}[-\eta E_1(2\rho) + K_0(\eta, \rho)]}. \tag{9}$$

For small ρ, Eq. (9) becomes

$$\frac{W_0'(\eta, \rho)}{W_0(\eta, \rho)} = -1 + 2\eta[-E_1(2\rho) + K_{-\frac{1}{2}}(\eta + \tfrac{1}{2}, 0)] + \mathcal{O}(\rho \log \rho). \tag{10}$$

[2] M. W. Kermode, Nucl. Phys. **A104**, 49 (1967).
[3] M. H. Hull and G. Breit, S. Flugge, Ed., *Handbuch der Physik* (Springer-Verlag, Berlin, 1959), Vol. 41, Pt. 1, p. 408.
[4] A. M. Lane and R. G. Thomas, Rev. Mod. Phys. **30**, 257 (1958).
[5] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions* (National Bureau of Standards, Washington, D.C., 1965).

In the limit as ρ tends to zero, Eqs. (2) and (10) give

$$z(k) = -k + 2\eta k h^-(\eta) + 2\eta k \times [2C + \log(2\eta k) + \log a], \quad (11)$$

where

$$h^-(\eta) = \psi(1+\eta) - \log\eta \quad (12)$$

and the logarithmic singularity is shown explicitly.[6] The last three terms on the right-hand side of Eq. (11) are independent of k and, consequently, they also occur for positive energies.[2] We absorb these terms in the matrix element $M_{22}(E)$. Since the effective-range formula is, in fact, an identity, we assume that $M_{22}(E)$ has a singularity of the form $2\eta k \log a$ so that $M_{22}(E) - z(k_2)$ remains finite. We can then define the renormalized logarithmic derivative by

$$z^R(k) = k[-1 + 2\eta h^-(\eta)], \quad l = 0.$$

B. *P*-Wave Result

For *P*-waves, Eq. (6) becomes

$$I_1(\eta,\rho) = \frac{e^{-2\rho}}{4\rho}\left(\frac{1}{\rho^2} - \frac{(\eta-1)}{\rho} + \eta(\eta-1)\right) - \frac{(\eta+1)\eta(\eta-1)}{3!}E_1(2\rho) + K_1(\eta,\rho). \quad (13)$$

From Eqs. (5) and (13), we find that, for small ρ,

$$J_1(\eta,\rho) = \frac{\rho\{-1 + 2\eta\rho + 4\rho^2[-\frac{1}{2}(\eta+1)\eta E_1(2\rho) + K_1(\eta+1,0) - K_1(\eta,0)]\}}{1 + (1-\eta)\rho - \eta(1-\eta)\rho^2} + O(\rho^4\log\rho)$$
$$= -\rho + \rho^2(\eta+1) + \rho^3[(\eta-1)(2\eta+1) - 2(\eta+1)\eta E_1(2\rho) + 4K_{\frac{1}{2}}(\eta+\tfrac{1}{2},0)] + O(\rho^4\log\rho). \quad (14)$$

Equations (2), (4), and (14) give, as $\rho \to 0$,

$$z(k) = (\eta^2-1)k^3[-1 + 2\eta h^-(\eta)] + A_1 + B_1k^2, \quad l = 1,$$

where

$$A_1 = -\frac{1}{a^3} - \frac{\eta k}{a^2} + \frac{\eta^2k^2}{a} - \eta^3k^3 + 2\eta^3k^3[2C + \log(2\eta k) + \log a]$$

and

$$B_1 = -a^{-1} + \eta k - 2\eta k[2C + \log(2\eta k) + \log a].$$

Again the "infinite" coefficients A_1 and B_1 are absorbed into the M-matrix element and we are left with the renormalized logarithmic derivative

$$z^R(k) = (\eta^2-1)k^3[-1 + 2\eta h^-(\eta)], \quad l = 1.$$

C. *D*-Wave Result

The evaluation of the function $I_2(\eta,\rho)$ and the insertion of the result into Eq. (5) give, for small ρ,

$$J_2(\eta,\rho) = \left[3 + \sum_{i=1}^{5} a_i(\eta+1)\rho^i\right] \Big/ \left[3 + \sum_{i=1}^{5} a_i(\eta)\rho^i\right] - 1 + O(\rho^6\log\rho), \quad (15)$$

where

$$a_1(\eta) = 6 - \tfrac{3}{2}(\eta+2),$$
$$a_2(\eta) = 6 - 3(\eta+2) + \tfrac{1}{2}(\eta+2)(\eta+1),$$
$$a_3(\eta) = 4 - 3(\eta+2) + (\eta+2)(\eta+1) - (1/3!)(\eta+2)(\eta+1)\eta,$$
$$a_4(\eta) = 2[1 - (\eta+2) + \tfrac{1}{2}(\eta+2)(\eta+1) - (1/3!)(\eta+2)(\eta+1)\eta + (1/4!) \times (\eta+2)(\eta+1)\eta(\eta-1)],$$
$$a_5(\eta) = 4K_2(\eta,0) - (4/5!)(\eta+2)(\eta+1) \times \eta(\eta-1)(\eta-2)E_1(2\rho).$$

The expansion of the right-hand side of Eq. (15) as a power series in ρ gives rise to straightforward but tedious algebra. The coefficients of ρ^4 and ρ^5 consist of the sum of many products of brackets of the form $(\eta+i)$. These products were calculated and the coefficients of similar powers of η collected together by computer. The resulting power-series expansion for $J_2(\eta,\rho)$ was found to be

$$J_2(\eta,\rho) = -\tfrac{1}{2}\rho + \frac{(\eta+2)\rho^2}{12}\left(1 - \frac{\eta\rho}{2} + \frac{(8-5\eta^2)\rho^2}{12} - \frac{(\eta^2-1)}{3}\{1 - 2\eta[C + \psi(1+\eta) - E_1(2\rho)]\}\rho^3 + \frac{\eta(16-19\eta^2)}{72}\rho^3\right) + O(\rho^6\log\rho). \quad (16)$$

As $\rho \to 0$, Eqs. (2), (4), and (16) give

$$z(k) = \frac{(\eta^2-4)(\eta^2-1)}{4}k^5[-1 + 2\eta h^-(\eta)] + A_2 + B_2k^2 + C_2k^4, \quad l = 2,$$

where the coefficients A_2, B_2, and C_2 are infinite for zero a. The renormalized logarithmic derivative is then

$$z^R(k) = \frac{(\eta^2-4)(\eta^2-1)}{4}k^5[-1 + 2\eta h^-(\eta)], \quad l = 2.$$

[6] Equations (3) and (12) can be used to define the function $h(z) = \mathrm{Re}[\psi(1+z) - \log z]$, where $z = \eta$ and $z = i\eta$ for negative and positive energies, respectively.

D. Extension of Results to Arbitrary l

The renormalized logarithmic derivative

$$z^R(k) = \frac{(\eta^2 - l^2)(\eta^2 - (l-1)^2)\cdots(\eta^2 - 1)}{(l!)^2} \times k^{2l+1}[-1 + 2\eta h^-(\eta)] \quad (17)$$

was derived for the S-, P-, and D-wave cases and, although it seems to hold for arbitrary l, no general proof has been found. We show below that the second term in Eq. (17) is true for any angular-momentum value and point out the difficulty in confirming, or otherwise, that the first term is also true.

The three contributions to the second part of the function z^R come from

(i)

$$\frac{(2l)!}{(2\rho)^{2l+1}}, \quad \text{which is part of the integral } \int_1^\infty z^{2l} e^{-2\rho z}\, dz,$$

(ii)

$$-\frac{(\eta + l)(\eta + l - 1)\cdots(\eta - l)}{(2l+1)!}\int_1^\infty \frac{e^{-2\rho z}}{z}\, dz,$$

and (iii)

$$\frac{(\eta + l)(\eta + l - 1)\cdots(\eta - l)}{(2l+1)!}[C + \psi(1 + \eta)].$$

Part (iii) arises from (a) the reduction of the function $K_l(\eta, 0)$ to the form $K_{-\frac{1}{2}}(\eta - l - \frac{1}{2}, 0)$ and (b) the repeated use of the important property of the digamma function, viz., $\psi(1 + z) = \psi(z) + z^{-1}$.

From Eqs. (2), (4), (5), and (6), we find the contribution

$$-[(2l-1)!!]^2 k^{2l+1} \frac{(\eta + l)(\eta + l - 1)\cdots(\eta - l)}{[(2l)!]^2} \times 2^{2l+1}[E_1(2\rho) - C - \psi(1 + \eta)],$$

which, on rearranging and subtracting a series of the form

$$A_l + B_l k^2 + \cdots + C_l k^{2l},$$

becomes

$$k^{2l+1} \frac{(\eta^2 - l^2)(\eta^2 - (l-1)^2)\cdots(\eta^2 - 1)}{(l!)^2} 2\eta h^-(\eta).$$

This term is, in fact, part of the contribution to the coefficient of ρ^{2l+1} arising from

$$b_0^{-2}(b_0 a_{2l+1} - a_0 b_{2l+1})$$

in the expansion of

$$\sum_{i=0}^{2l+1} a_i \rho^i \Big/ \sum_{i=0}^{2l+1} b_i \rho^i.$$

The other contributions to z^R come from the reduction procedure mentioned in (iii) and from the coefficient of ρ^{2l+1} in the series expansion of

$$I_l^{(1)}(\eta + 1, \rho)/I_l^{(1)}(\eta, \rho)$$

[$I_l^{(1)}(\eta, \rho)$ is the first integral in Eq. (6)]. The calculation of the latter term for arbitrary l has proved too difficult to perform and consequently we have been unable to give a rigorous derivation of Eq. (17).

III. EFFECTIVE-RANGE FORMULAS

The two-channel effective-range formula can be written

$$\begin{vmatrix} M_{11}(E) - y(k_1, \delta_{l_1}) & (m_1/m_2)^{\frac{1}{2}} M_{12}(E) \\ (m_2/m_1)^{\frac{1}{2}} M_{12}(E) & M_{22}(E) - z(k_2) \end{vmatrix} \equiv 0,$$

where

$$y(k_1, \delta_{l_1}) = \frac{(\eta_1^2 + 1)\cdots(\eta_1^2 + l_1^2)}{(l_1!)^2} \times k_1^{2l_1+1}\left[\frac{2\pi\eta_1 \cot \delta_{l_1}}{\exp(2\pi\eta_1) - 1} + 2\eta_1 h^+(\eta_1)\right]$$

and

$$z(k_2) = \frac{(\eta_2^2 - 1)\cdots(\eta_2^2 - l_2^2)}{(l_2!)^2} \times k_2^{2l_2+1}[-1 + 2\eta_2 h^-(\eta_2)], \quad E_2 < 0,$$

or

$$z(k_2) = \frac{(\eta_2^2 + 1)\cdots(\eta_2^2 + l_2^2)}{(l_2!)^2} \times k_2^{2l_2+1}\left[\frac{2\pi\eta_2 i}{\exp(2\pi\eta_2) - 1} + 2\eta_2 h^+(\eta_2)\right], \quad E_2 > 0.$$

The effective-range formulas for three or more channels are given in Ref. 1. These formulas are generalized to the case of no hard core and the Coulomb potential by including Eq. (17) in Table 3 of that paper.

Equation (17) is also useful for the single-channel scattering of two charged particles which have a bound state near threshold. The effective-range parameters can be used to obtain the energy of this state in the same way that the bound-state energy of the deuteron can be calculated from the S-wave n–p scattering data.[7]

ACKNOWLEDGMENTS

It is a pleasure to thank Dr. A. Perkins for his help in checking some of the algebra. The award of a postdoctoral fellowship at McMaster University is gratefully acknowledged.

[7] See, e.g., M. A. Preston, *Physics of the Nucleus* (Addison-Wesley Publ. Co., Inc., Reading, Mass., 1962), p. 24.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

On the Uncoupling of Wiener–Hopf Equations for Half-Plane Diffraction Problems in a Gyrotropic Medium

H. S. TAN* AND D. K. CHENG
Department of Electrical Engineering, Syracuse University, Syracuse, New York

(Received 24 April 1969)

Integral equations of the Wiener–Hopf type appear in the study of electromagnetic diffraction problems by a half-plane. When the environment is gyroelectric and characterized by an arbitrary permittivity tensor and a scalar permeability, the equations become coupled and no general method of solution is available unless a certain specific criterion is met. Hurd formulated this criterion and examined a number of solvable cases. The present paper extends Hurd's work to include oblique incidence on a half-plane in a gyrotropic medium. Media which are characterized by an arbitrary permittivity tensor and a diagonal permeability tensor are considered. An analogous but more general class of situations is obtained which permits the uncoupling of the Wiener–Hopf equations.

I. INTRODUCTION

The problem of electromagnetic diffraction by a half-plane has been the subject of numerous studies in the past. It constitutes the fundamental diffraction problem by an edge, and has been the basis of various approximation theories in diffraction by more complex geometries of plane and nonplanar screens. Examples such as the geometrical diffraction theory,[1] the edge current theories,[2] and other high-frequency asymptotic theories[3] may be mentioned. Solutions to the half-plane problem in isotropic media with Dirichlet, Neumann, and impedance boundary conditions and with various types of incident waves have been reported. Recent interest in this problem has been extended to diffraction in anisotropic plasma media. Seshadri and Rajagopal[4] and Jull[5] considered a gyroelectric medium with axis parallel to the half-plane edge, while Felsen[6] and Rulf[7] were concerned with a uniaxial medium with axis perpendicular to the edge. Williams[8] has shown that the case of general plane-wave incidence on a half-plane in a uniaxial medium may be formally expressed in terms of solutions of the isotropic half-plane problem. More recently, Rosenbaum[9] has reported formal solutions of the general anisotropic case, using plane-wave modal superposition methods, with the modal amplitudes obtained from a Wiener–Hopf procedure. These previous studies have been concerned only with electrical anisotropy; i.e., the medium has been assumed to be described by a permittivity tensor ϵ and a scalar constant permeability μ, with the exceptions of Felsen[6] and Przezdziecki,[10] who considered uniaxial electric and magnetic anisotropy.

Recently, Hurd[11] developed a method for uncoupling the simultaneous Wiener–Hopf equations obtained from a formulation of the half-plane diffraction problem in a general electrically anisotropic medium. He found that under certain conditions the kernel matrix $\mathbf{G}$ of these equations can be cast into the following form:

$$\mathbf{G} = \Gamma_1(\mathbf{A} + \Gamma_2\mathbf{B}), \tag{1}$$

where Γ_1 and Γ_2 are scalar nonpolynomials and $\mathbf{A}$ and $\mathbf{B}$ are polynomial matrices. In such cases the equations can be uncoupled if $\mathbf{A}$ and $\mathbf{B}$ satisfy certain restrictions contained in an equation which he called "the criterion." Application of this criterion tnen yields a list of different forms of the permittivity tensor ϵ for which the half-plane problem can, in principle, be solved by the Wiener–Hopf technique.

The purpose of this paper is to show that Hurd's method can be generalized to include media characterized by a general permittivity tensor and a biaxial permeability tensor.[12] An analogous but more general class of solvable problems is obtained and shown to contain various interesting possibilities. These problems reduce to those considered by Hurd as special cases. The factorization process in the actual Wiener–Hopf technique remains difficult, but, in a few of the simpler cases, this appears to be possible. Most of the procedures in this paper are parallel to those of Hurd, although the amount of manipulation and algebra

* On leave from the University of Malaya, Kuala Lumpur, Malaysia.

[1] J. B. Keller, J. Appl. Phys. **28**, 426 (1957).
[2] R. F. Millar, IEEE Monograph 152R, 1955.
[3] D. S. Jones, *The Theory of Electromagnetism* (The Macmillan Co., New York, 1964).
[4] S. R. Seshadri and A. K. Rajagopal, IEEE Trans. Antennas Propagation **AP-11**, 497 (1963).
[5] E. V. Jull, Can. J. Phys. **42**, 1455 (1964).
[6] L. B. Felsen, Proc. IEEE **111**, 445 (1964).
[7] B. Rulf, SIAM J. Appl. Math. **15**, 120 (1967).
[8] W. E. Williams, J. Inst. Math. Appl. **2**, 186 (1966).
[9] S. Rosenbaum, Can. J. Phys. **45**, 3479 (1967).
[10] S. Przezdziecki, J. Appl. Phys. **37**, 2768 (1966).
[11] R. A. Hurd, Can. J. Phys. **45**, 2561 (1967).
[12] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Inc., New York, 1964), p. 678.

involved is even more tedious. Only the essential steps leading to the criterion and the various solvable combinations will be given.

II. FORMULATION OF DIFFRACTION PROBLEM

We consider a monochromatic plane wave with angular frequency ω,

$$\mathbf{E}^{\text{inc}} = \mathbf{E}^0 \exp i(\alpha_0 x + \beta y + \gamma_0 z - \omega t),$$
$$\mathbf{H}^{\text{inc}} = \mathbf{H}^0 \exp i(\alpha_0 x + \beta y + \gamma_0 z - \omega t), \quad (2)$$

which travels from the region $z < 0$ and impinges obliquely on a perfectly conducting, thin half-plane occupying $z = 0$, $x \geq 0$, in an infinite medium characterized by an arbitrary permittivity tensor

$$\boldsymbol{\epsilon} = \epsilon_0 \begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{21} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{31} & \epsilon_{32} & \epsilon_{33} \end{pmatrix} \quad (3)$$

and a diagonal permeability tensor

$$\boldsymbol{\mu} = \mu_0 \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}. \quad (4)$$

For a general plane wave with field amplitudes $(\mathbf{E}_j, \mathbf{H}_j)$ and propagation constants α, β, and γ_j, the Maxwell curl equations can be written as

$$\mathbf{K} \cdot \mathbf{E}_j = \omega \boldsymbol{\mu} \cdot \mathbf{H}_j, \quad (5)$$

$$\mathbf{K} \cdot \mathbf{H}_j = -\omega \boldsymbol{\epsilon} \cdot \mathbf{E}_j, \quad (6)$$

where

$$\mathbf{K} = \begin{pmatrix} 0 & -\gamma_j & \beta \\ \gamma_j & 0 & -\alpha \\ -\beta & \alpha & 0 \end{pmatrix}. \quad (7)$$

Elimination of $\mathbf{H}_j$ from (5) and (6) yields a wave equation for $\mathbf{E}_j$:

$$\boldsymbol{\Phi} \cdot \mathbf{E}_j = (\omega^2 \boldsymbol{\epsilon} + \mathbf{K} \cdot \boldsymbol{\mu}^{-1} \cdot \mathbf{K}) \cdot \mathbf{E}_j = 0. \quad (8)$$

The existence of nontrivial solutions demands that

$$\det(\boldsymbol{\Phi}) = \begin{vmatrix} e_{11} - \frac{\beta^2}{c} - \frac{\gamma_j^2}{b} & e_{12} + \frac{\alpha\beta}{c} & e_{13} + \frac{\alpha\gamma_j}{b} \\ e_{21} + \frac{\alpha\beta}{c} & e_{22} - \frac{\gamma_j^2}{a} - \frac{\alpha^2}{c} & e_{23} + \frac{\beta\gamma_j}{a} \\ e_{31} + \frac{\alpha\gamma_j}{b} & e_{32} + \frac{\beta\gamma_j}{a} & e_{33} - \frac{\alpha^2}{b} - \frac{\beta^2}{a} \end{vmatrix} = 0, \quad (9)$$

where

$$e_{mn} = \omega^2 \mu_0 \epsilon_0 \epsilon_{mn}, \quad m, n = 1, 2, 3. \quad (10)$$

Equation (9) is a fourth-degree dispersion equation in γ_j, from which four roots γ_1, γ_2, γ_3, and γ_4 can be determined. Two of the roots, say γ_1 and γ_2, are appropriate for waves traveling in the $+z$ direction; and the other two, γ_3 and γ_4, are for waves in the $-z$ direction.

We next represent the scattered electric and magnetic fields in terms of superpositions of plane waves as Fourier integrals. We write

$$\mathbf{E}^s(z > 0) = \int_{-\infty}^{\infty} [\mathbf{E}_1(\alpha) e^{i\gamma_1 z} + \mathbf{E}_2(\alpha) e^{i\gamma_2 z}] e^{i\alpha x} d\alpha, \quad (11)$$

$$\mathbf{E}^s(z < 0) = \int_{-\infty}^{\infty} [\mathbf{E}_3(\alpha) e^{i\gamma_3 z} + \mathbf{E}_4(\alpha) e^{i\gamma_4 z}] e^{i\alpha x} d\alpha, \quad (12)$$

$$\mathbf{H}^s(z > 0) = \frac{1}{\omega} \boldsymbol{\mu}^{-1} \int_{-\infty}^{\infty} [\mathbf{K}(\gamma_1) \cdot \mathbf{E}_1(\alpha) e^{i\gamma_1 z} + \mathbf{K}(\gamma_2) \cdot \mathbf{E}_2(\alpha) e^{i\gamma_2 z}] e^{i\alpha x} d\alpha, \quad (13)$$

$$\mathbf{H}^s(z < 0) = \frac{1}{\omega} \boldsymbol{\mu}^{-1} \int_{-\infty}^{\infty} [\mathbf{K}(\gamma_3) \cdot \mathbf{E}_3(\alpha) e^{i\gamma_3 z} + \mathbf{K}(\gamma_4) \cdot \mathbf{E}_4(\alpha) e^{i\gamma_4 z}] e^{i\alpha x} d\alpha, \quad (14)$$

where the common factor $\exp[i(\beta y - \omega t)]$ has been suppressed and the $\mathbf{E}_j(\alpha)$ are the unknowns to be determined. It is noted that the x, y, z components of $\mathbf{E}_j$ are not independent, but are related through (8) to give

$$E_{ji} = D_{ij} E_{jz} / D_{zj}, \quad i = x, y, \quad j = 1, 2, 3, 4, \quad (15)$$

where the D_{ij} are defined in terms of the elements of the $\boldsymbol{\Phi}$ matrix as

$$D_{xj} = \begin{vmatrix} \phi_{12} & \phi_{13} \\ \phi_{22} & \phi_{23} \end{vmatrix}, \quad D_{yj} = \begin{vmatrix} \phi_{13} & \phi_{11} \\ \phi_{23} & \phi_{21} \end{vmatrix}, \quad D_{zj} = \begin{vmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{vmatrix}. \quad (16)$$

The boundary conditions to be satisfied are

$$x \geq 0: E_i(z = 0) = 0, \quad (18)$$

$$x < 0: E_i(z = 0+) = E_i(z = 0-), \quad (17)$$

$$H_i(z = 0+) = H_i(z = 0-), \quad i = x, y. \quad (19)$$

Substitution of the incident and scattered field expressions in (17)–(19) yields the following relations:

$$E_{1i} + E_{2i} = U_i - E_{ji}^0 / 2\pi i(\alpha - \alpha_0) = V_i, \quad (20)$$

$$E_{1i} + E_{2i} = E_{3i} + E_{4i}, \quad (21)$$

$$[\mathbf{K}(\gamma_1) \cdot \mathbf{E}_1(\alpha) + \mathbf{K}(\gamma_2) \cdot \mathbf{E}_2(\alpha) - \mathbf{K}(\gamma_3) \cdot \mathbf{E}_3(\alpha) - \mathbf{K}(\gamma_4) \cdot \mathbf{E}_4(\alpha)]_i = L_i(\alpha), \quad (22)$$

where U_i and L_i are analytic functions of α in the upper and lower half-planes, respectively, and are of algebraic growth in their respective half-planes. α_0 is assumed to lie above the real axis and E^0_{ji} is given by

$$E^0_{ji} = D^0_{ij}E^0_{jz}/D^0_{zj}, \quad i = x, y, \quad j = 1, 2. \tag{23}$$

The D^0_{ij} in (23) are defined in (16) except that α and γ_j are replaced by α_0 and γ_0. Equations (15) and (23) can now be used to eliminate the x and y components of $\mathbf{E}_j$ in (20)–(22). After much algebra and the eventual elimination of E_{jz}, we arrive at a pair of coupled simultaneous Wiener-Hopf equations in the unknowns V_i and L_i, namely,

$$G_{11}V_x + G_{12}V_y = L_x, \tag{24}$$

$$G_{12}V_x + G_{21}V_y = L_y. \tag{25}$$

In (24) and (25) the components of the kernel matrix **G** are defined as

$$G_{11} = \left[\frac{D_{y2}}{\Delta_1}(\beta D_{z1} - \gamma_1 D_{y1}) - \frac{D_{y1}}{\Delta_1}(\beta D_{z2} - \gamma_2 D_{y2}) - \frac{D_{y4}}{\Delta_2}(\beta D_{z3} - \gamma_3 D_{y3}) + \frac{D_{y3}}{\Delta_2}(\beta D_{z4} - \gamma_4 D_{y4})\right], \tag{26}$$

$$G_{12} = \left[-\frac{D_{x2}}{\Delta_1}(\beta D_{z1} - \gamma_1 D_{y1}) + \frac{D_{x1}}{\Delta_1}(\beta D_{z2} - \gamma_2 D_{y2}) + \frac{D_{x4}}{\Delta_2}(\beta D_{z3} - \gamma_3 D_{y3}) - \frac{D_{x3}}{\Delta_2}(\beta D_{z4} - \gamma_4 D_{y4})\right], \tag{27}$$

$$G_{21} = \left[-\frac{D_{y2}}{\Delta_1}(\alpha D_{z1} - \gamma_1 D_{x1}) + \frac{D_{y1}}{\Delta_1}(\alpha D_{z2} - \gamma_2 D_{x2}) + \frac{D_{y4}}{\Delta_2}(\alpha D_{z3} - \gamma_3 D_{x3}) - \frac{D_{y3}}{\Delta_2}(\alpha D_{z4} - \gamma_4 D_{x4})\right], \tag{28}$$

$$G_{22} = \left[\frac{D_{x2}}{\Delta_1}(\alpha D_{z1} - \gamma_1 D_{x1}) - \frac{D_{x1}}{\Delta_1}(\alpha D_{z2} - \gamma_2 D_{x2}) - \frac{D_{x4}}{\Delta_2}(\alpha D_{z3} - \gamma_3 D_{x3}) + \frac{D_{x3}}{\Delta_2}(\alpha D_{z4} - \gamma_4 D_{x4})\right], \tag{29}$$

where

$$\Delta_1 = (D_{x1}D_{y2} - D_{x2}D_{y1}), \tag{30}$$

$$\Delta_2 = (D_{x3}D_{y4} - D_{y3}D_{x4}). \tag{31}$$

Equations (26)–(29) can be simplified when D_{zj} is eliminated by using (8):

$$D_{zj} = \frac{-1}{\phi_{33}}[D_{xj}(e_{31} + \alpha\gamma_j/b) + D_{yj}(e_{32} + \beta\gamma_j/a)]. \tag{32}$$

The elements of the kernel matrix **G** now become

$$G_{11} = \frac{1}{\Delta_1\phi_{33}}\left[\left(e_{33} - \frac{\alpha^2}{b}\right)(\gamma_2 - \gamma_1)D_{y1}D_{y2} + \frac{\alpha\beta}{b}(\gamma_2 D_{x2}D_{y1} - \gamma_1 D_{x1}D_{y2})\right] - \frac{1}{\Delta_2\phi_{33}}\left[\left(e_{33} - \frac{\alpha^2}{b}\right)(\gamma_4 - \gamma_3)D_{y3}D_{y4} + \frac{\alpha\beta}{b}(\gamma_4 D_{x4}D_{y3} - \gamma_3 D_{x3}D_{y4})\right], \tag{33}$$

$$G_{12} = \frac{1}{\Delta_1\phi_{33}}\left[\frac{\alpha\beta}{b}(\gamma_1 - \gamma_2)D_{x1}D_{x2} + \left(e_{33} - \frac{\alpha^2}{b}\right)(\gamma_1 D_{x2}D_{y1} - \gamma_2 D_{x1}D_{y2})\right] - \frac{1}{\Delta_2\phi_{33}}\left[\frac{\alpha\beta}{b}(\gamma_3 - \gamma_4)D_{x3}D_{x4} + \left(e_{33} - \frac{\alpha^2}{b}\right)(\gamma_3 D_{x4}D_{y3} - \gamma_4 D_{x3}D_{y4})\right], \tag{34}$$

$G_{21} = -G_{12}$ (interchange x and y, 1 and 2, 3 and 4, a and b, α and β; ϕ_{33}, Δ_1, and Δ_2 unchanged), (35)

$G_{22} = -G_{11}$ (interchange x and y, 1 and 2, 3 and 4, a and b, α and β; ϕ_{33}, Δ_1, and Δ_2 unchanged). (36)

III. TRANSFORMATION OF KERNEL MATRIX

Our object now is to express the kernel matrix **G** in the form of (1). In general, it is probably not possible to do this; but Hurd has shown that this can be done if we assume that the propagation constants γ_j have the relations

$$\gamma_3 = -\gamma_1, \quad \gamma_4 = -\gamma_2. \tag{37}$$

Equation (9) then reduces to the form

$$\gamma^4 - \lambda\gamma^2 - \zeta = 0, \tag{38}$$

where

$$\lambda = \gamma_1^2 + \gamma_2^2 \tag{39}$$

and

$$\zeta = -\gamma_1^2\gamma_2^2. \tag{40}$$

Detailed expressions for λ and ζ are given in (54) and (55). With this assumption, we first transform **G** to **G′** by a premultiplication by a polynomial $\mathbf{N}_1$ given by

$$\mathbf{N}_1 = \begin{pmatrix} \alpha\beta/a, & e_{33} - \alpha^2/b \\ e_{33} - \beta^2/a, & \alpha\beta/b \end{pmatrix}. \tag{41}$$

We have

$$G'_{11} = (e_{33}/\Delta_1)[\gamma_1 D_{x1} D_{y2} - \gamma_2 D_{x2} D_{y1}] - (e_{33}/\Delta_2)[\gamma_3 D_{x3} D_{y4} - \gamma_4 D_{x4} D_{y3}], \quad (42)$$

$$G'_{12} = (e_{33}/\Delta_1)(\gamma_2 - \gamma_1) D_{x1} D_{x2} - (e_{33}/\Delta_2)(\gamma_4 - \gamma_3) D_{x3} D_{x4}, \quad (43)$$

$$G'_{21} = (e_{33}/\Delta_1)(\gamma_2 - \gamma_1) D_{y1} D_{y2} - (e_{33}/\Delta_2)(\gamma_4 - \gamma_3) D_{y3} D_{y4}, \quad (44)$$

$$G'_{22} = (e_{33}/\Delta_1)(\gamma_1 D_{x2} D_{y1} - \gamma_2 D_{x1} D_{y2}) - (e_{33}/\Delta_2)(\gamma_3 D_{x4} D_{y3} - \gamma_4 D_{x3} D_{y4}). \quad (45)$$

After much tedious algebra and a considerable amount of manipulation, (42)–(45) can be rewritten as

$$G'_{11} = -\rho[SI_{11} - \zeta R I_{21} - U I_{31} + \zeta T I_{41} + \gamma_1\gamma_2(R I_{11} + S I_{21} - T I_{31} - U I_{41})], \quad (46)$$

$$G'_{12} = \rho[SJ_{12} - \zeta R J_{22} - U J_{32} + \zeta T J_{42} + \gamma_1\gamma_2(R J_{12} + S J_{22} - T J_{32} - U J_{42})], \quad (47)$$

$$G'_{21} = \rho[SJ_{11} - \zeta R J_{21} - U J_{31} + \zeta T J_{41} + \gamma_1\gamma_2(R J_{11} + S J_{21} - T J_{31} - U J_{41})], \quad (48)$$

$$G'_{22} = -\rho[SI_{12} - \zeta R I_{22} - U I_{32} + \zeta T I_{42} + \gamma_1\gamma_2(R I_{12} + S I_{22} - T I_{32} - U I_{42})], \quad (49)$$

where the expressions for R, S, T, U, ρ, the I's, and the J's are defined below:

$$R = \alpha\nu_2/ab - (\beta/ab)\nu_1 + \eta_2 e_{13} - \eta_1 e_{23}, \quad (50)$$

$$S = b\eta_1\nu_2 - a\eta_2\nu_1 + \left(\frac{\beta e_{13}}{a^2 b} - \frac{\alpha e_{23}}{ab^2}\right)\zeta + \left(\frac{\alpha\nu_2}{ab} - \frac{\beta\nu_1}{ab}\right)\lambda, \quad (51)$$

$$T = \alpha\eta_2/b - (\beta/a)\eta_1, \quad (52)$$

$$U = e_{13}\nu_2/a - e_{23}\nu_1/b, \quad (53)$$

$$\lambda = -\frac{ab}{e_{33}}\left[\alpha\eta_1 + \beta\eta_2 + \frac{e_{13}e_{31}}{a} + \frac{e_{23}e_{32}}{b} - \chi\phi_{33}\right], \quad (54)$$

$$\zeta = -(ab/e_{33})[e_{31}\nu_1 + e_{32}\nu_2 + \phi_{33}\nu_3], \quad (55)$$

$$\nu_1 = (e_{12} + \alpha\beta/c)e_{23} - (e_{22} - \alpha^2/c)e_{13}, \quad (56)$$

$$\nu_2 = (e_{21} + \alpha\beta/c)e_{13} - (e_{11} - \beta^2/c)e_{23}, \quad (57)$$

$$\nu_3 = \left(e_{11} - \frac{\beta^2}{c}\right)\left(e_{22} - \frac{\alpha^2}{c}\right) - \left(e_{12} + \frac{\alpha\beta}{c}\right)\left(e_{21} + \frac{\alpha\beta}{c}\right), \quad (58)$$

$$\eta_1 = \frac{\beta}{ab}\left(e_{12} + \frac{\alpha\beta}{c}\right) - \frac{\alpha}{b^2}\left(e_{22} - \frac{\alpha^2}{c}\right), \quad (59)$$

$$\eta_2 = \frac{\alpha}{ab}\left(e_{21} + \frac{\alpha\beta}{c}\right) - \frac{\beta}{a^2}\left(e_{11} - \frac{\beta^2}{c}\right), \quad (60)$$

$$\chi = \frac{e_{11}}{a} + \frac{e_{22}}{b} - \frac{\alpha^2}{bc} - \frac{\beta^2}{ac}, \quad (61)$$

$$\phi_{33} = e_{33} - \alpha^2/b - \beta^2/a, \quad (62)$$

$$\rho = 2(\gamma_1 - \gamma_2)^2(\gamma_1 + \gamma_2)e_{33}/\Delta_1\Delta_2, \quad (63)$$

$$I_{11} = b\eta_1\nu_2 + \alpha\nu_2\lambda/ab - \alpha e_{23}\zeta/ab^2, \quad (64)$$

$$I_{21} = \eta_2 e_{13} - \beta\nu_1/ab, \quad (65)$$

$$I_{31} = \nu_1\nu_2 - (\zeta/ab)e_{13}e_{23} + \nu_2(e_{13}\lambda/a) - \zeta T, \quad (66)$$

$$I_{41} = ab\eta_1\eta_2 - \alpha\beta\zeta/a^2b^2 + (\alpha\lambda/b)\eta_2 + U, \quad (67)$$

$$I_{12} = a\eta_2\nu_1 + \beta\nu_1\lambda/ab - \beta e_{13}\zeta/a^2 b, \quad (68)$$

$$I_{22} = \eta_1 e_{23} - \alpha\nu_2/ab, \quad (69)$$

$$I_{32} = \nu_1\nu_2 - \zeta e_{13}e_{23}/ab + \nu_1 e_{23}\lambda/b + \zeta T, \quad (70)$$

$$I_{42} = ab\eta_1\eta_2 - (\alpha\beta/a^2b^2)\zeta + (\beta/a)\lambda\eta_1 - U, \quad (71)$$

$$J_{11} = a\eta_2\nu_2 + \beta\nu_2\lambda/ab - \beta e_{23}\zeta/ab^2, \quad (72)$$

$$J_{21} = (a/b)\eta_2 e_{23} - \beta\nu_2/ab, \quad (73)$$

$$J_{31} = \nu_2^2 + \nu_2 e_{23}\lambda/b - e_{23}^2\zeta/b^2, \quad (74)$$

$$J_{41} = a^2\eta_2^2 + \eta_2\beta\lambda/b - \beta^2\zeta/a^2b^2, \quad (75)$$

$$J_{12} = b\eta_1\nu_1 + \alpha\nu_1\lambda/ab - \alpha e_{13}\zeta/a^2 b, \quad (76)$$

$$J_{22} = (b/a)\eta_1 e_{13} - \alpha\nu_1/ab, \quad (77)$$

$$J_{32} = \nu_1^2 + \nu_1 e_{13}\lambda/a - e_{13}^2\zeta/a^2, \quad (78)$$

$$J_{42} = b^2\eta_1^2 + \alpha\eta_1\lambda/a - \alpha^2\zeta/a^2b^2. \quad (79)$$

These are extremely complicated expressions, but certain symmetry exists in the structure of the new symbols relative to their subscripts.

The expressions $(RJ_{12} + SJ_{22} - TJ_{32} - UJ_{42})$ in G'_{12} and $(RJ_{11} + SJ_{21} - TJ_{31} - UJ_{41})$ in G'_{21}, appearing in (47) and (48), respectively, are both identically zero, but have been inserted to facilitate a further transformation of $\mathbf{G}'$. Premultiplying $\mathbf{G}'$ by $\mathbf{N}_2$ and postmultiplying by matrix $\mathbf{M}_1$, where

$$\mathbf{N}_2 = \begin{pmatrix} \beta/ab & \alpha/ab \\ e_{23}/b, & e_{13}/a \end{pmatrix}, \quad (80)$$

$$\mathbf{M}_1 = \begin{pmatrix} \nu_1 & b\eta_1 \\ \nu_1 & a\eta_2 \end{pmatrix}, \quad (81)$$

we obtain a new transformed kernel matrix $\mathbf{G}''$:

$$\mathbf{G}'' = \mathbf{N}_2 \cdot \mathbf{G}' \cdot \mathbf{M}_1 = \rho(\mathbf{A}'' + \gamma_1\gamma_2\mathbf{B}''), \quad (82)$$

where

$$A''_{11} = (\zeta C/ab)[U^2 + ab\nu_3 T^2 + 4ab\nu_3 BC - \nu_3\chi C^2 - (B^2/ab)\chi], \quad (83)$$

$$A''_{12} = (C/ab)\{(ab\nu_3 + \zeta)[\nu_3 CT + (B/ab)U] - (\nu_3\lambda + \zeta\chi)(BT + CU)\}, \quad (84)$$

$$A''_{21} = (\zeta C/ab)\{[\nu_3 + (\zeta/ab)](BT + CU) + (\lambda - ab\chi)[\nu_3 CT + (B/ab)U]\}, \quad (85)$$

$$A''_{22} = (C/ab)[ab\nu_3 U^2 + \zeta^2 T^2 - ab\chi\zeta TU + SD(\zeta - ab\nu_3) - S\lambda C\nu_3], \quad (86)$$

$$B''_{11} = WB/ab, \quad (87)$$

$$B''_{12} = WT, \quad (88)$$

$$B''_{21} = WU, \quad (89)$$

$$B''_{22} = WD, \quad (90)$$

with

$$B = \alpha\nu_2 - \beta\nu_1, \quad (91)$$

$$C = (\beta/a)e_{13} - (\alpha/b)e_{23}, \quad (92)$$

$$D = \eta_2 e_{13} - \eta_1 e_{23}, \quad (93)$$

$$W = \zeta T^2 - U^2 - \lambda TU + RS. \quad (94)$$

The other symbols have been defined previously. The negative sign before the RS term in the expression for W in Hurd's paper [our Eq. (94)] appears to be a typographical error.

$\mathbf{G}''$ in (82) is now of the desired form of (1) and is to be diagonalized in order to uncouple the Wiener–Hopf equations (24) and (25). The Appendix contains a brief derivation of the "criterion" which must be satisfied by the elements of $\mathbf{G}''$ for the uncoupling of (24) and (25). A formal solution by the Wiener–Hopf technique is also sketched there.

IV. RESULTS AND DISCUSSION

It remains now to apply simultaneously the restriction implied in (38) and the criterion in (A6) in order to determine the form of the ϵ and μ tensors for which the half-plane diffraction problem will be solvable. Expanding the determinant in (9) and comparing it with (38), we find that the following four conditions must hold:

$$e_{13} + e_{31} = 0, \quad (95)$$

$$\beta(e_{23} + e_{32}) = 0, \quad (96)$$

$$e_{21}e_{32} + e_{12}e_{23} = 0, \quad (97)$$

$$\beta(e_{21}e_{13} + e_{12}e_{31}) = 0. \quad (98)$$

The criterion in (A6) is written in a more explicit form by substituting A''_{ij} and B''_{ij} from (82)–(90). After some more tedious work, we obtain

$$r^2 = \nu_3^2(C^4W^2/a^2b^2) \times [(S^2 - \zeta R^2 - \lambda U^2 + \lambda\zeta T^2 + 4\zeta TU)^2 + 4\zeta W^2] = \text{perfect square in } \alpha. \quad (99)$$

The quantity in the brackets on the right-hand side of (99) is to be made a perfect square. Our task is thus reduced to finding the elements of ϵ and μ which will simultaneously satisfy (95)–(99). Despite the complexity of the algebra involved, there appears to be a large number of possibilities, most of which impose rather severe restrictions on the elements of ϵ and μ, and are therefore descriptions of unusual media which probably cannot be realized physically. We list below some of the more interesting types of media which are parallel to those considered by Hurd,[11] but which represent extensions to include biaxial gyromagnetic cases.

Case I: Biaxial gyrotropic media with coincident axes ($e_{ij} = 0$, $i \neq j$). When both the permittivity tensor ϵ and the permeability tensor μ are diagonal, we have a general form of a biaxial gyrotropic medium[12] with coincident principal axes, and Eqs. (95)–(98) are satisfied identically. Also, since $R = S = U = 0$, the criterion in (99) reduces to

$$\lambda^2 + 4\zeta = \text{perfect square.} \quad (100)$$

Two possible ways of satisfying (100) are

$$\frac{\beta^2}{a} e_{33}\left(\frac{e_{33}}{c} - \frac{e_{22}}{b}\right)\left(\frac{e_{22}}{b} - \frac{e_{11}}{a}\right) = 0, \quad (101)$$

$$e_{11}/a - e_{33}/c = 0, \quad \beta \text{ arbitrary.} \quad (102)$$

Condition (101) gives three distinct possibilities: (i) arbitrary incidence in a biaxial gyrotropic medium with principal axes in the x direction, i.e., in the plane of the half-plane and perpendicular to its edge; (ii) arbitrary incidence in a biaxial gyrotropic medium with principal axes in the z direction, i.e., perpendicular to the plane and the edge of the half-plane; (iii) normal incidence in a biaxial gyrotropic medium. Condition (102) gives rise to (iv) arbitrary incidence in a biaxial gyrotropic medium with principal axes parallel to the half-plane edge. There exist other special situations which also satisfy (100), but they lead to rather unusual media.

Case II: Biaxial gyrotropic media with noncoincident axes. This is a generalization of Case I above. The medium may be assumed to have a biaxial μ with axes in the (x, y, z) directions and a biaxial ϵ with axes

in the (x', y', z') directions, where x and y' axes coincide, and z' makes an angle θ with z. A rotation about the x axis through θ and another rotation about the z axis through 90° enable the ϵ tensor to be written as

$$\epsilon = \epsilon_0 \begin{pmatrix} \epsilon_2 & 0 & 0 \\ 0 & \epsilon_1 \cos^2\theta + \epsilon_3 \sin^2\theta & (\epsilon_3 - \epsilon_1)\cos\theta \sin\theta \\ 0 & (\epsilon_3 - \epsilon_1)\cos\theta \sin\theta & \epsilon_1 \sin^2\theta + \epsilon_3 \cos^2\theta \end{pmatrix}, \tag{103}$$

$$\mu = \mu_0 \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}. \tag{104}$$

From (16) and (33)–(37), it is evident that $G_{11} = G_{22} = 0$, so that the original kernel matrix is diagonal, and the subsequent transformations are unnecessary.

Case III: Normal plane of Incidence ($\beta = 0$). Two general cases arise. (v) $e_{12} = e_{21} = e_{23} = e_{32} = 0$, $e_{13} = e_{31}$. Equations (95)–(98) are satisfied identically, while (99) reduces to $r^2 = 0$. A short calculation shows that this is admissible, so that we have here a gyrotropic medium with a biaxial magnetic permeability. Various special cases are of course possible. For example, with $e_{11} = e_{33}$ and $a = c$, the gyroelectric and gyromagnetic axes are both parallel to the half-plane; or, with $e_{11} = e_{33}$ and $a = b$, the two axes are now perpendicular to each other. (vi) $e_{12} = e_{21} = e_{13} = e_{31} = 0$. In this case, (95)–(99) are satisfied identically, so that we have a general gyrotropic medium with a biaxial magnetic permeability. Various special cases are again possible.

Case IV: $e_{12} = e_{21}$, $e_{13} = -e_{31}$, $e_{23} = -e_{32}$. Under these conditions, (95)–(98) are satisfied, but (99) is satisfied only by rather restrictive types of media. So far no new worthwhile solutions for this case have been obtained.

Lastly we mention an especially simple case in which the uncoupling of Wiener–Hopf equations can be carried out without satisfying the criterion. If we write

$$V_i = U_i - P_i \tag{105}$$

in (20), Eqs. (24) and (25) become

$$\mathbf{G} \cdot \mathbf{U} = \mathbf{L} + \mathbf{G} \cdot \mathbf{P}. \tag{106}$$

Suppose the kernel matrix $\mathbf{G}$ can be put in the form of (82) with $\gamma_1\gamma_2$ a polynomial in α; $(\mathbf{A} + \gamma_1\gamma_2\mathbf{B}) \cdot \mathbf{U}$ will then be a matrix function of α which is regular in the upper half-plane and, consequently, the Wiener–Hopf procedure can be carried out directly. An example of such a case is a biaxial gyrotropic medium with coincident axes under the conditions of $\beta = 0$ and $e_{22}/b = e_{33}/c$.

V. CONCLUSIONS

It has been shown that the half-plane diffraction problem can be solved by the Wiener–Hopf technique for a number of types of gyroelectric and gyromagnetic media. These include the types of gyroelectric media mentioned by Hurd as special cases when the permeability is a scalar constant. As might be expected, the addition of a μ tensor complicates the problem somewhat, but, fortunately, the criteria contained in (95)–(99) can be satisfied, in most cases, by restrictions on the elements of ϵ alone, as in cases (v) and (vi) mentioned above. In other cases, the restrictions apply to e_{11}/a, e_{22}/b, e_{33}/c, and their combinations, so that relatively simple types of realizable media will satisfy the criteria.

In general, the Wiener–Hopf factorization in (A9) remains a difficult process, although there are formal methods, and, failing these, approximate methods[13] for dealing with this problem. Examples of the simpler cases where the complete solution has been worked out have been reported for the half-plane in gyroelectric media. It has also been shown[9] that information about the scattering process may be deduced from Eqs. (11)–(14) and (20)–(22) without going through the complete Wiener–Hopf procedure.

ACKNOWLEDGMENT

This work was supported in part by the U.S. Army Research Office, Durham.

APPENDIX: DIAGONALIZATION OF KERNEL MATRIX AND FORMAL SOLUTION

With the kernel matrix $\mathbf{G}''$ written in the form of (82), it is easily seen that $\mathbf{N} \cdot \mathbf{G}'' \cdot \mathbf{M}$ is diagonal if $\mathbf{M}$ and $\mathbf{N}$ are two constant nonsingular polynomial matrices and if

$$(N_{11}A''_{11} + N_{12}A''_{21})M_{12} + (N_{11}A''_{12} + N_{12}A''_{22})M_{22} = 0, \tag{A1}$$

$$(N_{21}A''_{11} + N_{22}A''_{21})M_{11} + (N_{21}A''_{12} + N_{21}A''_{22})M_{21} = 0, \tag{A2}$$

together with two similar equations for the elements of $\mathbf{B}''$. Solving these equations for the elements of $\mathbf{M}$

[13] B. Noble, *Methods Based on the Wiener–Hopf Technique for the Solution of Partial Differential Equations* (Pergamon Press, Inc., New York, 1958).

and $\mathbf{N}$, we have

$$r_2 x_1^2 + r_1 x_1 + r_0 = 0, \tag{A3}$$

$$r_2 x_2^2 + r_1 x_2 + r_0 = 0, \tag{A4}$$

where

$$x_1 = N_{11}/N_{12}, \quad x_2 = N_{21}/N_{22},$$
$$r_2 = (A''_{11}B''_{12} - A''_{12}B''_{11}),$$
$$r_1 = (A''_{11}B''_{22} + A''_{21}B''_{12} - B''_{11}A''_{22} - B''_{21}A''_{12}),$$
$$r_0 = (A''_{21}B''_{22} - B''_{21}A''_{22}).$$

The discriminant in the solution of (A3) and (A4) must be a perfect square in α, since x_1 and x_2 cannot be allowed branch points. This is the restriction on the elements of G'' which Hurd calls the "criterion." We require

$$r^2 = r_1^2 - 4r_0 r_2 = \text{perfect square in } \alpha \tag{A5}$$

or

$$(A''_{11}B''_{22} + A''_{22}B''_{11} - A''_{12}B''_{21} - A''_{21}B''_{12})^2 - 4 \det(\mathbf{A}'') \cdot \det(\mathbf{B}'') = \text{perfect square in } \alpha. \tag{A6}$$

Once the elements of $\mathbf{N}$ are found, those of $\mathbf{M}$ follow from (A1) and (A2). The diagonalized matrix $\mathbf{N} \cdot \mathbf{G}'' \cdot \mathbf{M}$ then becomes

$$\mathbf{N} \cdot \mathbf{G}'' \cdot \mathbf{M} = -\rho \det(\mathbf{M}) \det(\mathbf{M}_1) \times \begin{pmatrix} 1 + \gamma_1\gamma_2 \dfrac{x_1 B''_{11} + B''_{21}}{x_1 A''_{11} + A''_{21}} & 0 \\ 0 & -1 - \gamma_1\gamma_2 \dfrac{x_2 B''_{11} + B''_{21}}{x_2 A''_{11} + A''_{21}} \end{pmatrix}. \tag{A7}$$

The uncoupled equations are now

$$\mathbf{N} \cdot \mathbf{G}'' \cdot \mathbf{M} \cdot \mathbf{U}' = \det(\mathbf{M}) \det(\mathbf{M}_1)[\mathbf{L}' + \mathbf{N} \cdot \mathbf{N}_2 \cdot \mathbf{N}_1 \cdot \mathbf{G} \cdot \mathbf{P}], \tag{A8}$$

where $\mathbf{P}$ is the known vector $E^0_{ji}/2\pi i(\alpha - \alpha_0)$ in (20), $\mathbf{L}' = \mathbf{N} \cdot \mathbf{N}_2 \cdot \mathbf{N}_1 \cdot \mathbf{L}$, and

$$\mathbf{U}' = \det(\mathbf{M}) \det(\mathbf{M}_1) \mathbf{M}^{-1} \cdot \mathbf{M}_1^{-1} \cdot \mathbf{U}.$$

In the usual Wiener–Hopf procedure, the elements of the matrix $\mathbf{N} \cdot \mathbf{G}'' \cdot \mathbf{M}$ have to be factorized into terms which are separately regular in the upper and lower half-planes of α. Writing (A7) as

$$\mathbf{N} \cdot \mathbf{G}'' \cdot \mathbf{M} = \det(\mathbf{M}) \det(\mathbf{M}_1) \begin{pmatrix} K_1 & 0 \\ 0 & K_2 \end{pmatrix} \tag{A9}$$

and letting $K_i = K_i^+/K_i^-$, where the $+$ and $-$ signs denote regularity in the upper and lower half-planes, respectively, we see that the uncoupled equations (A8) are finally reduced to

$$K_i^+ U_i' - [K_i^-(\mathbf{N} \cdot \mathbf{N}_2 \cdot \mathbf{N}_1 \cdot \mathbf{G} \cdot \mathbf{P})_i]^+ = Q_i, \tag{A10}$$

$$K_i^- L_i' + [K_i^-(\mathbf{N} \cdot \mathbf{N}_2 \cdot \mathbf{N}_1 \cdot \mathbf{G} \cdot \mathbf{P})_i]^- = Q_i, \tag{A11}$$

where the Q_i are polynomials in α which may be uniquely determined from the analytic behavior of $\mathbf{L}$, $\mathbf{L}'$, $\mathbf{U}$, and $\mathbf{U}'$. The formal solution is now effected, since $\mathbf{L}$ and $\mathbf{U}$ are now found from $\mathbf{L}'$ and $\mathbf{U}'$ in terms of the polynomials Q_i.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Numerical Solution of Integral Equations Using Legendre Polynomials

H. KSCHWENDT
Euratom, 21020 *Ispra (Va)*, *Italy*

(Received 25 November 1968)

This paper deals with the solution of certain integral equations of displacement type. A previously developed theory, which uses an expansion of the solution in Legendre polynomials, is applied to two special cases. One is a singular equation encountered in polymer physics. The results of the Legendre expansion are compared with calculations based on two other recently proposed methods, and the merits of an expansion in Gegenbauer polynomials are discussed. For the second case, which arises in electrodynamics, a very accurate solution is given for a large range of parameter values.

1. INTRODUCTION

Various problems in mathematical physics lead to integral equations of the displacement type

$$\varphi(x) = f(x) + \lambda \int_{-1}^{1} K(s\,|x-y|)\varphi(y)\,dy, \quad (1.1)$$

where s is a parameter. Among them there is a well-known problem in polymer physics[1,2] where the kernel is a singular one and has the form $|x-y|^{-\alpha}$, $0 < \alpha < 1$. Another problem which arises in electrodynamics leads to equations where the kernels themselves are somewhat complicated integrals.[3,4]

The polymer equation was treated in two recent papers,[5,6] where the special case

$$\varphi(x) = x^2 - \lambda \int_{-1}^{1} |x-y|^{-\frac{1}{2}}\,\varphi(y)\,dy \quad (1.2)$$

was considered numerically. Ullmann and Ullmann[5] proposed two methods: One is a Gaussian quadrature procedure, where the singularity of (1.2) was eliminated by an approximation; the second uses an expansion of the solution of (1.2) in a Fourier series. The singularity was thereby removed in a natural way. The Gaussian quadrature procedure gave better results than the Fourier expansion, probably because the Fourier coefficients of the kernel and the inhomogeneous term did not drop off very rapidly. Ullmann and Ullmann[5] stated explicitly that the expansion of (1.2) in Legendre polynomials would probably lead to a more accurate solution, "but unfortunately the Fourier–Legendre coefficients of $|x-y|^{-\frac{1}{2}}$ are not easily expressed in terms of known functions." We shall show that these coefficients can be calculated in an elementary manner.

Schlitt[6] proposed a different method in which he added and subtracted a certain term on the right of (1.2), thus eliminating the singularity. The resultant equation was then again treated by a Gaussian quadrature formula and gave better results than the Gaussian procedure used by Ullmann and Ullmann.[5] The reason seems to be that Schlitt treated the singularity analytically.

Another way of solving (1.2) was suggested by Auer and Gardner.[7] They observed that $|x-y|^{-\frac{1}{2}}$ possesses an expansion which involves only diagonal terms if certain Gegenbauer polynomials are used. However, we find that this method is impracticable for the numerical solution of Eq. (1.2) because no simple way seems to be known to calculate the corresponding matrix elements.

The previous cited electrodynamical problem arises when the resistance of a plasma slab between juxtaposed circular disk electrodes is determined. It leads to equations[3] whose kernels can be written in displacement form as

$$F(x) = 1 + \frac{2s}{\pi}\int_{-1}^{1} F(y)\,dy \int_{0}^{\infty} \frac{\cos s(x-y)z}{\exp z + 1}\,dz, \quad (1.3)$$

$$G(x) = x - \frac{2s}{\pi}\int_{-1}^{1} G(y)\,dy \int_{0}^{\infty} \frac{\cos s(x-y)z - 1}{\exp z - 1}\,dz, \quad (1.4)$$

where s is the ratio of disk radius to slab thickness. Laporte and Fowler[4] expanded the solution of (1.3) in Legendre polynomials and calculated the lowest matrix elements directly, but stated "it has not been possible to derive a type form."

The purpose of the present paper is to demonstrate that it is relatively simple to obtain general expressions for the matrix elements by using some recent results in connection with the Legendre expansion of arbitrary integral equations of the displacement type.[8–10] This

[1] J. G. Kirkwood and J. Riseman, J. Chem. Phys. **16**, 565 (1948).
[2] R. Ullmann, J. Chem. Phys. **40**, 2193 (1964).
[3] O. Laporte and R. G. Fowler, J. Math. Phys. **8**, 518 (1967).
[4] O. Laporte and R. G. Fowler, Phys. Rev. **148**, 170 (1966).
[5] N. Ullmann and R. Ullmann, J. Math. Phys. **7**, 1743 (1966).
[6] D. W. Schlitt, J. Math. Phys. **9**, 436 (1968).
[7] P. L. Auer and C. S. Gardner, J. Chem. Phys. **23**, 1545, 1546 (1955).
[8] H. Kschwendt, Z. Angew. Math. Mech. **48**, T20 (1968).
[9] H. Kschwendt, "Legendre Expansion and Integral Equations of Displacement Type," J. Comp. Phys. (to be published).
[10] H. Kschwendt, "The Fourier Transform of a Product of Two Spherical Bessel Functions," Atomkernenergie (to be published).

expansion is of particular interest because, in many cases, the resultant system of algebraic equations can be truncated at a relatively low order. This has been shown in Refs. 8 and 9 and was mentioned for (1.3) in Ref. 4. Another advantage of the use of the Legendre polynomials (or any other set of orthogonal functions) over the Gaussian quadrature formalism is the fact that the solution of the equation of interest can be easily evaluated for *any* value of x without the need for an interpolation.

2. LEGENDRE EXPANSION

In this section, the main results of Refs. 8, 9, and 10 are shortly repeated. The kernel of (1.1) is expanded in Legendre polynomials as

$$K(s\,|x-y|) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}(2m+1)(2n+1)T_{m,n}(s)P_m(x)P_n(y). \quad (2.1)$$

Then the solution of (1.1) is given by

$$\varphi(x) = \sum_{m=0}^{\infty}(2m+1)X_m P_m(x), \quad (2.2)$$

where X_m obeys the linear system of algebraic equations

$$X_m = f_m + 2\lambda\sum_{n=0}^{\infty}(2n+1)T_{m,n}(s)X_n, \qquad m = 0, 1, 2, \cdots. \quad (2.3)$$

The f_m's are the Fourier–Legendre coefficients of $f(x)$. Equations (2.1) and (2.2) both converge in the mean, i.e., for almost all s, x, and y.

One of the results of Refs. 8 and 9 is the reduction of the usual double integral for the determination of the matrix elements,

$$T_{m,n}(s) = \tfrac{1}{4}\int_{-1}^{1}dx\int_{-1}^{1}dy K(s\,|x-y|)P_m(x)P_n(y),$$

to a single integral, *independently* of the actual kernel under consideration. Thus,

$$T_{m,n}(s) = \int_0^1 Q_{m,n}(q)K(2sq)\,dq, \quad m+n \text{ even}, \qquad = 0, \quad m+n \text{ odd}. \quad (2.4)$$

The function

$$Q_{m,n}(q) = \sum_{\nu=0}^{m+n+1} c_\nu^{m,n} q^\nu, \quad (2.5)$$

where

$$c_0^{m,n} = \frac{2\delta_{m,n}}{(2m+1)}$$

and

$$c_\nu^{m,n} = \frac{2(-1)^\nu}{\nu!\,(\nu-1)!}\prod_{\kappa=1}^{\nu-1}(m+n+1+\nu-2\kappa) \times (|m-n|+\nu-2\kappa)$$

depends in *no way* on the equation to be solved.[8–10] At the same time, the possible discontinuity of the derivative of the kernel is removed. Here and later, empty products are to be replaced by unity and empty sums by zero.

Since $T_{m,n}(s)$ vanishes for $m+n$ odd, the matrix equation (2.3) degenerates in two *uncoupled* systems. Concerning the special cases to be considered here, for Eqs. (1.2) and (1.3), we have to determine only those matrix elements where m and n are both even, and, for Eq. (1.4), only those where m and n are both odd.

There exists a simple recurrence relation[10] between the four Q's:

$$(2n+1)[Q_{m+2,n}(q) - Q_{m,n}(q)] = (2m+3)[Q_{m+1,n-1}(q) - Q_{m+1,n+1}(q)]; \quad (2.6)$$

and, due to Eq. (2.4), the same relation holds for $T_{m,n}(s)$ *independently* of the actual equation of interest.

If the kernel in (2.4) possesses a series expansion with respect to q, the determination of $T_{m,n}(s)$ is reduced to integrals of the form[9]

$$\psi_\nu^{m,n} = \frac{(\nu+1)(\nu+2)}{2}\int_0^1 q^\nu Q_{m,n}(q)\,dq = (-1)^m\prod_{\kappa=1}^{\tau}\frac{(\nu+3-2\kappa)}{(\nu+1+2\kappa)}\prod_{\sigma=1}^{\rho}\frac{(\nu+2-2\sigma)}{(\nu+2+2\sigma)}, \quad m+n \text{ even}, \quad (2.7)$$

where throughout the paper $\tau = \frac{1}{2}|m-n|$ and $\rho = \frac{1}{2}(m+n)$. The latter results holds not only for integral values of ν, but also for any $\nu > -1$.[10]

3. THE SINGULAR EQUATION

The determination of the matrix elements for the Legendre expansion of Eq. (1.2) is now extremely simple. Substitution of (2.7) into (2.4) yields

$$T_{m,n} = \frac{2^{\frac{5}{2}}}{3}\prod_{\kappa=1}^{\tau}\frac{(5-4\kappa)}{(1+4\kappa)}\prod_{\sigma=1}^{\rho}\frac{(3-4\sigma)}{(3+4\sigma)}. \quad (3.1)$$

This form is well suited for machine computation. Similar simple results are obtained for any kernel of the form $|x-y|^\beta$, $\beta > -1$.

The elements (3.1) do not drop off very rapidly with increasing m and n. Therefore, in (2.3), one has to use a relatively high truncation order to get reasonable results. But in the polymer problem, which is considered here, the solution $\varphi(x)$ itself is not actually needed. Only the integral of $\varphi(x)$ over x in $(-1, 1)$ is required. It is a remarkable fact that, because of the orthogonality of the Legendre polynomials, only the term X_0 in Eq. (2.2) contributes to this integral. We

show later that this quantity can be determined sufficiently accurately from relatively small matrices except when λ is large.

For the solution of the matrix equation (2.3) the ordinary Gaussian algorithm with full pivotal condensation was performed using double precision arithmetic. Substitution of the result in the original matrix shows an accuracy of 14 digits. The execution time to compute the solution for one value of λ from an 80 × 80 matrix was about 60 seconds on an IBM 360/65 machine. The Legendre polynomials were calculated from the trigonometric expansion because in this representation the coefficients are consecutively decreasing. Even for $P_{160}(x)$, the deviation from the correct result was, in double precision, smaller than 10^{-13}.

To provide a comparison, some results of the Legendre expansion are compared in Table I with results obtained by Ullmann and Ullmann[5] and Schlitt[6] from the Gaussian quadrature procedure. All these results are calculated from 80 × 80 matrices. We shall present only a few numerical values, but they are sufficient to indicate the trend between the different solutions. In particular, we give results for $\lambda = 0.5$ and $\lambda = 200$. Since the solution of (1.2) can be obtained so simply by the Legendre expansion, there is no need to publish too much numerical data.

For small values of λ and any x and large values of λ and small or intermediate x, the results from the Legendre expansion are very close to the results of Schlitt. He concluded[6] that his results are probably better than those of Ullmann and Ullmann[5] and, therefore, the same holds for our values. For large values of λ and values of x near to unity, the Legendre results differ from the others by, at most, 4%. The discrepancy probably comes from the fact that the asymptotic solution[5-7]

$$\lim_{\lambda\to\infty} \lambda\varphi(x) = 2^{\frac{1}{2}}(4x^2 - 1)(1 - x^2)^{-\frac{1}{4}}/3\pi \quad (3.2)$$

is singular at $x = 1$ and, consequently, the numerical values of X_m in (2.3) do not drop off fast enough.

TABLE I. Comparison of Legendre results with two other solutions of Eq. (1.2).

x	Reference	$\lambda = 0.5$	$\lambda = 200$
0.019511	This paper	−0.081004	−0.00074416
	Schlitt[6]	−0.081002	−0.00074540
	Ullmann[5]	−0.081611	−0.00075671
0.502804	This paper	0.062542	0.0000123556
	Schlitt[6]	0.062541	0.0000111018
	Ullmann[5]	0.062666	0.0000079162
0.999554	This paper	0.67447	0.012566
	Schlitt[6]	0.67454	0.012984
	Ullmann[5]	0.67470	0.013045

TABLE II. Integral of $\varphi(x)$ and size N of matrices.

λ	X_0	N	$2\lambda X_0$
0.5	0.12383	10	0.12383
5.0	0.019807	25	0.19807
20.0	0.0052689	35	0.21076
100.0	0.0010735	40	0.21469
200.0	0.00053803	50	0.21521
500.0	0.00021553	60	0.21553
1000.0	0.00010782	80	0.21563
∞			0.21574

In Table II, we give numerical values for X_0 and show the size N of the matrices which must be used to obtain five-digit accuracy. The quantity $2\lambda X_0$ is the integral of $\lambda\varphi(x)$ over x. The value for $\lambda = \infty$ is calculated from (3.2).

4. GEGENBAUER EXPANSION

Auer and Gardner[7] suggested solving (1.2) by an expansion in Gegenbauer polynomials $P_m^{(\gamma)}(x)$, which are defined by

$$P_m^{(\gamma)}(x) = \frac{(-1)^m\Gamma(\gamma+1)}{2^m\Gamma(m+\gamma+1)}(1-x^2)^{-\gamma}\frac{d^m}{dx^m}(1-x^2)^{m+\gamma}.$$

($\gamma = 0$ are the Legendre polynomials.) These polynomials are normalized according to

$$\int_{-1}^{1}(1-x^2)^{\gamma}P_m^{(\gamma)}(x)P_n^{(\gamma)}(x)\,dx = N_m\delta_{m,n},$$

$$N_m = \frac{2^{2\gamma+1}[\Gamma(\gamma+1)]^2 m!}{(2m+2\gamma+1)\Gamma(m+2\gamma+1)}.$$

The reason for this proposal is the fact that $|x-y|^{-\alpha}$ has a *diagonal* representation

$$|x-y|^{-\alpha} = \sum_{m=0}^{\infty} a_m P_m^{(\gamma)}(x)P_m^{(\gamma)}(y),$$

$$\text{for}\quad \gamma = (\alpha-1)/2,\quad 0<\alpha<1. \quad (4.1)$$

In the special case, Eq. (1.2), which is considered here, $\gamma = -\frac{1}{4}$. The coefficients a_m are given by[7]

$$a_m = \frac{-\pi}{2^{2\gamma+1}\sin\pi\gamma}\left[\frac{\Gamma(m+2\gamma+1)}{\Gamma(\gamma+1)m!}\right]^2\frac{(2m+2\gamma+1)}{\Gamma(2\gamma+1)}.$$

Substitution of (4.1) into (1.2) gives, after some calculation,

$$\varphi(x) = \sum_{m=0}^{\infty} Z_m P_m^{(\gamma)}(x). \quad (4.2)$$

Since the $P_m^{(\gamma)}(x)$ are *not* the eigenfunctions of (1.2), the coefficients Z_m have to be determined from a linear system of algebraic equations

$$Z_m = h_m - \lambda a_m \sum_{n=0}^{\infty} V_{m,n}Z_n, \quad (4.3)$$

with

$$N_m h_m = \int_{-1}^{1} x^2(1-x^2)^\gamma P_m^{(\gamma)}(x)\,dx.$$

For the integral of $\lambda\varphi(x)$, we get

$$\lambda\int_{-1}^{1}\varphi(x)\,dx = (h_0 - Z_0)/a_0.$$

The matrix elements of (4.3) are defined by

$$V_{m,n} = \int_{-1}^{1} P_m^{(\gamma)}(x)P_n^{(\gamma)}(x)\,dx. \tag{4.4}$$

If one compares the Legendre with the Gegenbauer expansion, one might expect similar results because both methods consist in a power-series expansion of $\varphi(x)$. Both methods describe the inhomogeneous term of (1.2) exactly and the expansion coefficients are, in both cases, obtained from a system of linear equations. Also, the integral over $\varphi(x)$ depends, in both methods, only on the lowest expansion coefficient. From the practical point of view, the only difference thus lies in the possibility of the determination of the respective matrix elements. The integral defining $T_{m,n}$ led to the simple result (3.1), where each element is calculated from a previous one by multiplication with a certain constant. In the case of Eq. (4.4), no way was found to solve the integral for arbitrary m and n. Moreover, direct calculation of an element with relatively low indices like $V_{4,4}$ already produced a very complicated expression. So long as there is no simple way to evaluate $V_{m,n}$, the use of the Gegenbauer expansion for the numerical solution of (1.2) cannot be recommended.

5. THE ELECTRODYNAMICAL PROBLEM

A rough solution of Eq. (1.3) was obtained by Laporte and Fowler[4] by expansion in Legendre polynomials and direct calculation of the matrix elements up to $m, n = 4$. To obtain the form of these elements for arbitrary indices, we use the representation (2.4):

$$T_{m,n}(s) = 2s\int_0^1 dq\,Q_{m,n}(q)\int_0^\infty \frac{\cos 2sqz}{\exp z + 1}\,dz. \tag{5.1}$$

Expanding the denominator in a series of $\exp(-kz)$ and integrating over z yields

$$T_{m,n}(s) = \sum_{k=1}^{\infty}(-1)^{k+1}Y_{m,n}\left(\frac{2s}{k}\right), \tag{5.2}$$

with

$$Y_{m,n}(\vartheta) = \vartheta\int_0^1 Q_{m,n}(q)(1+\vartheta^2q^2)^{-1}\,dq. \tag{5.3}$$

Substitution of (2.5) into (5.3) gives

$$Y_{m,n}(\vartheta) = \ln(1+\vartheta^2)\sum_{\nu=0}^{\rho}\frac{(-1)^\nu c_{2\nu+1}^{m,n}}{2\vartheta^{2\nu+1}} + \tan^{-1}\vartheta\sum_{\nu=0}^{\tau}\frac{(-1)^\nu c_{2\nu}^{m,n}}{\vartheta^{2\nu}} - \sum_{\kappa=1}^{\rho}\frac{(-1)^\kappa}{\vartheta^{2\kappa-1}}\sum_{\nu=2\kappa}^{m+n+1}\frac{c_\nu^{m,n}}{(\nu+1-2\kappa)}, \tag{5.4}$$

where again $\rho = \frac{1}{2}(m+n)$ and $\tau = \frac{1}{2}|m-n|$.

For small values of ϑ, Eq. (5.4) is useless, because negative powers of ϑ appear in the result. Expanding (5.3) with respect to ϑ and using (2.7), we have

$$T_{m,n}(s) = \sum_{k=1}^{J-1}(-1)^{k+1}Y_{m,n}\left(\frac{2s}{k}\right) + \sum_{\nu=\rho}^{\infty}\sum_{k=J}^{\infty}\frac{(-1)^{k+\nu+1}}{k^{2\nu+1}(\nu+1)(2\nu+1)}\psi_{2\nu}^{m,n}(2s)^{2\nu+1}, \qquad J = 1, 2, 3, \cdots. \tag{5.5}$$

The convergence of (5.5) is assured if $J > 2s$, since, for any $\nu \geq 0$,

$$\left|\sum_{k=J}^{\infty}\frac{(-1)^{k+\nu+1}}{k^{2\nu+1}}\right| \leq J^{-2\nu-1}$$

and

$$|\psi_\nu^{m,n}| \leq 1.$$

In the same way the matrix elements of (1.4) are obtained, the only difference being the disappearance of the sign factor $(-1)^{k+1}$ in (5.2) and (5.5).

The series (5.5) was found to be very convenient for numerical calculations. Moreover, the recurrence relation (2.6) also holds for $\psi_\nu^{m,n}$ and $Y_{m,n}(\vartheta)$. Thus, only those elements with $n = 0$ and those with $n = m$ have to be calculated directly from Eq. (5.5).

Following the theory of Laporte and Fowler,[4] the resistance function R, which is described by (1.3), is proportional to the inverse of the lowest expansion coefficient of (2.2):

$$R = 1/X_0. \tag{5.6}$$

In Table III, we give results for $R(s)$ up to $s = 50$,

TABLE III. Resistance obtained from Eq. (5.6) and size N of matrices.

$2s$	R Eq. (5.6)	R Ref. 4	N	$\frac{1}{2}\pi sR$
0.1	0.95597		1	0.07508
0.5	0.79016	0.788	2	0.31030
1.0	0.62665	0.615	2	0.49217
5.0	0.21510		3	0.84469
10.0	0.11680		3	0.91732
50.0	0.025021		6	0.98258
100.0	0.012621		8	0.99123
∞				1.00000

together with the size N of the matrices, which were used to obtain five-digit accuracy. Two values, which were computed by Laporte and Fowler,[4] are also included.

For values of $s > 50$, one needs matrices larger than 10×10 to obtain the resistance exact to five digits. But, for $m, n \geq 20$, the determination of $Y_{m,n}(\vartheta)$ becomes inaccurate, even with double precision arithmetic.

To study the asymptotic behavior of $R(s)$, we perform in Eq. (5.1) a q integration by parts:

$$T_{m,n}(s) = -\int_0^1 dq Q'_{m,n}(q) \int_0^\infty \frac{\sin 2sqz}{z(\exp z + 1)} dz.$$

Here we have used the property[10] $Q_{m,n}(1) = 0$. Since

$$\lim_{s\to\infty} \frac{\sin s(x-y)}{\pi(x-y)} = \delta(x-y)$$

is Dirac's δ function, we obtain

$$\lim_{s\to\infty} 2(2n+1)\dot{T}_{m,n}(s) = \pi\delta_{m,n}.$$

From the matrix equation (2.3) we see that all terms tend to zero in the limit $s \to \infty$ and, consequently, also $R(s)$ vanishes. Numerical calculations show that the expression $2sR(s)$ approaches the value $4/\pi$ for very large s and not unity, which was assumed by Laporte and Fowler.[4] In Table III, values for $\frac{1}{2}\pi sR(s)$ are also included.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Representation Functions for $SU(3)$

JOUKO MICKELSSON
Research Institute for Theoretical Physics, University of Helsinki, Helsinki, Finland

(Received 27 December 1968)

"Euler" decomposition for $SU(3)$ has been found by analytic continuation from the "Euler" decomposition of $SU(2, 1)$, which is based on the factorization $SU(2, 1) = SU(2) \times A_u \times A \times SU(2)$ derived by Hillion [$A_u \sim U(1)$ and A is one-parameter subgroup of the V-spin subgroup.] For the degenerate case the $SU(3)$ representation functions can be found by transforming to a new basis where the V-spin subgroup is diagonal. The result is a sum of products, each product containing two D-functions, one d-function, and an exponential function exp [$im\chi$]. As is well known, every representation $D(p, q)$ can be reduced from the direct product $D(p, 0) \times D(q, q)$. The nondegenerate representation functions can then be directly written down using the $SU(3)$ CG coefficients, which are relatively simple for the above-mentioned case.

REPRESENTATION FUNCTIONS FOR $SU(3)$

An arbitrary group element of $SU(2)$ can be factorized according to

$$g = e^{iK_3\phi}e^{iK_2\psi}e^{iK_3\theta}, \quad g \in SU(2). \tag{1}$$

ϕ, ψ, and θ are the usual Euler angles and K_i ($i = 1, 2, 3$) are the generators of the Lie algebra $SU(2)$.

Our first problem is to find similar factorization for the group $SU(3)$.

We start with the decomposition[1]

$$SU(2, 1) = SU(2) \times A_u \times A \times SU(2) \tag{2}$$

for the $SU(2, 1)$. "$\times$" means the semidirect product. A_u is the center of $S[U(2) \times U(1)]$, which is isomorphic to $U(1)$. A is the same A as in the Iwasawa decomposition

$$SU(2, 1) = KAN \tag{3}$$

and its Lie algebra is a Cartan subalgebra of $\bar{P}$ in the decomposition

$$su(2, 1) = \bar{K} + \bar{P}, \tag{4}$$

where $\bar{K}$ is the maximal compact subalgebra of $su(2, 1)$. $su(2, 1)$ has the following basis[2]:

$$\begin{aligned} K_i &= \tfrac{1}{2}\lambda_i \quad (i = 1, 2, 3), \\ K_j &= \tfrac{1}{2}i\lambda_j \quad (j = 4, 5, 6, 7), \\ K_8 &= (1/2\sqrt{3})\lambda_8. \end{aligned} \tag{5}$$

$\bar{K}$ is now spanned by K_i ($i = 1, 2, 3, 8$) and $\bar{P}$ by K_i ($i = 4, 5, 6, 7$). The subgroup $SU(2)$ is generated by K_1, K_2, and K_3.

The Lie algebra of A is one-dimensional and its basis can be chosen to be $\{K_5\}$. Therefore, an arbitrary

[1] P. Hillion, Nuovo Cimento 51A, 674 (1967).

[2] L. C. Biedenharn, "Cargese Lectures 1965, Application of Mathematics to Problems in Theoretical Pysics," F. Lurçat, Ed. (Gordon & Breach, Science Publishers, Inc., New York, 1967).

together with the size N of the matrices, which were used to obtain five-digit accuracy. Two values, which were computed by Laporte and Fowler,[4] are also included.

For values of $s > 50$, one needs matrices larger than 10×10 to obtain the resistance exact to five digits. But, for $m, n \geq 20$, the determination of $Y_{m,n}(\vartheta)$ becomes inaccurate, even with double precision arithmetic.

To study the asymptotic behavior of $R(s)$, we perform in Eq. (5.1) a q integration by parts:

$$T_{m,n}(s) = -\int_0^1 dq Q'_{m,n}(q) \int_0^\infty \frac{\sin 2sqz}{z(\exp z + 1)} dz.$$

Here we have used the property[10] $Q_{m,n}(1) = 0$. Since

$$\lim_{s\to\infty} \frac{\sin s(x-y)}{\pi(x-y)} = \delta(x-y)$$

is Dirac's δ function, we obtain

$$\lim_{s\to\infty} 2(2n+1)\dot{T}_{m,n}(s) = \pi\delta_{m,n}.$$

From the matrix equation (2.3) we see that all terms tend to zero in the limit $s \to \infty$ and, consequently, also $R(s)$ vanishes. Numerical calculations show that the expression $2sR(s)$ approaches the value $4/\pi$ for very large s and not unity, which was assumed by Laporte and Fowler.[4] In Table III, values for $\frac{1}{2}\pi sR(s)$ are also included.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 10 OCTOBER 1969

Representation Functions for $SU(3)$

JOUKO MICKELSSON
Research Institute for Theoretical Physics, University of Helsinki, Helsinki, Finland

(Received 27 December 1968)

"Euler" decomposition for $SU(3)$ has been found by analytic continuation from the "Euler" decomposition of $SU(2, 1)$, which is based on the factorization $SU(2, 1) = SU(2) \times A_u \times A \times SU(2)$ derived by Hillion [$A_u \sim U(1)$ and A is one-parameter subgroup of the V-spin subgroup.] For the degenerate case the $SU(3)$ representation functions can be found by transforming to a new basis where the V-spin subgroup is diagonal. The result is a sum of products, each product containing two D-functions, one d-function, and an exponential function exp [$im\chi$]. As is well known, every representation $D(p, q)$ can be reduced from the direct product $D(p, 0) \times D(q, q)$. The nondegenerate representation functions can then be directly written down using the $SU(3)$ CG coefficients, which are relatively simple for the above-mentioned case.

REPRESENTATION FUNCTIONS FOR $SU(3)$

An arbitrary group element of $SU(2)$ can be factorized according to

$$g = e^{iK_3\phi}e^{iK_2\psi}e^{iK_3\theta}, \quad g \in SU(2). \tag{1}$$

ϕ, ψ, and θ are the usual Euler angles and K_i ($i = 1, 2, 3$) are the generators of the Lie algebra $SU(2)$.

Our first problem is to find similar factorization for the group $SU(3)$.

We start with the decomposition[1]

$$SU(2, 1) = SU(2) \times A_u \times A \times SU(2) \tag{2}$$

for the $SU(2, 1)$. "$\times$" means the semidirect product. A_u is the center of $S[U(2) \times U(1)]$, which is isomorphic to $U(1)$. A is the same A as in the Iwasawa decomposition

$$SU(2, 1) = KAN \tag{3}$$

and its Lie algebra is a Cartan subalgebra of $\bar{P}$ in the decomposition

$$su(2, 1) = \bar{K} + \bar{P}, \tag{4}$$

where $\bar{K}$ is the maximal compact subalgebra of $su(2, 1)$. $su(2, 1)$ has the following basis[2]:

$$\begin{aligned} K_i &= \tfrac{1}{2}\lambda_i \quad (i = 1, 2, 3), \\ K_j &= \tfrac{1}{2}i\lambda_j \quad (j = 4, 5, 6, 7), \\ K_8 &= (1/2\sqrt{3})\lambda_8. \end{aligned} \tag{5}$$

$\bar{K}$ is now spanned by K_i ($i = 1, 2, 3, 8$) and $\bar{P}$ by K_i ($i = 4, 5, 6, 7$). The subgroup $SU(2)$ is generated by K_1, K_2, and K_3.

The Lie algebra of A is one-dimensional and its basis can be chosen to be $\{K_5\}$. Therefore, an arbitrary

[1] P. Hillion, Nuovo Cimento 51A, 674 (1967).

[2] L. C. Biedenharn, "Cargese Lectures 1965, Application of Mathematics to Problems in Theoretical Pysics," F. Lurçat, Ed. (Gordon & Breach, Science Publishers, Inc., New York, 1967).

group element of $SU(2, 1)$ can be written in the form

$$g = e^{iK_3\phi}e^{iK_2\psi}e^{iK_3\theta}e^{iK_8\chi}e^{iK_5\omega}e^{iK_3\alpha}e^{iK_2\beta}e^{iK_3\gamma}, \quad g \in SU(2, 1), \quad (6)$$

corresponding to (2).

Now we can "continue (6) analytically" to $SU(3)$, substituting $\omega \to -i\omega$ and defining $\bar{K}_5 = -iK_5 = \frac{1}{2}\lambda_5$. We get

$$g = e^{iK_3\phi}e^{iK_2\psi}e^{iK_3\theta}e^{iK_8\chi}e^{i\bar{K}_5\omega}e^{iK_3\alpha}e^{iK_2\beta}e^{iK_3\gamma}, \quad g \in SU(3). \quad (7)$$

Take now the $\langle pq\lambda'\mu'm'| \cdots |pq\lambda\mu m\rangle$ matrix element [where λ and μ are the isotopic spin variables, $\mu = -\lambda, -\lambda + 1, \cdots, \lambda$. $m = \frac{1}{2}Y$ (Y is the hypercharge)] of the representation operator of (7) in an unitary irreducible representation (labeled by pq):

$$\begin{aligned} D^{pq}_{\lambda'\mu'm',\lambda\mu m}(\phi, \psi, \theta, \chi, \omega, \alpha, \beta, \gamma) &\equiv \langle\lambda'\mu'm'| e^{iK_3\phi}\cdots e^{iK_3\gamma}|\lambda\mu m\rangle \\ &= \sum_{\bar{\lambda}\bar{\mu}\bar{m}}\sum_{\tilde{\lambda}\tilde{\mu}\tilde{m}} \langle\lambda'\mu'm'| e^{iK_3\phi}e^{iK_2\psi}e^{iK_3\theta}|\bar{\lambda}\bar{\mu}\bar{m}\rangle \\ &\quad \times \langle\bar{\lambda}\bar{\mu}\bar{m}| e^{iK_8\chi}e^{i\bar{K}_5\omega}|\tilde{\lambda}\tilde{\mu}\tilde{m}\rangle \\ &\quad \times \langle\tilde{\lambda}\tilde{\mu}\tilde{m}| e^{iK_3\alpha}e^{iK_2\beta}e^{iK_3\gamma}|\lambda\mu m\rangle \\ &= \sum_{\bar{\mu}\tilde{\mu}} D^{\lambda'}_{\mu'\bar{\mu}}(\phi, \psi, \theta)e^{im'\chi}\langle\lambda'\bar{\mu}m'| e^{i\bar{K}_5\omega}|\lambda\tilde{\mu}m\rangle \\ &\quad \times D^{\lambda}_{\tilde{\mu}\mu}(\alpha, \beta, \gamma). \end{aligned} \quad (8)$$

In (8) we have suppressed the labels pq and used the completeness of states. The D^{λ}'s are usual $SU(2)$ rotation functions. All difficulties are contained in the matrix element

$$d^{pq}_{\lambda'\mu'm',\lambda\mu m}(\omega) \equiv \langle\lambda'\mu'm'| e^{i\bar{K}_5\omega}|\lambda\mu m\rangle. \quad (9)$$

We transform now to a new basis where the subgroup $U(2)$ generated by $\bar{K}_3$, $\bar{K}_4$, $\bar{K}_5$, $\bar{K}_8$ is diagonal. The $\bar{K}_i$'s ($i = 3, 4, 5, 8$) are defined in the defining representation in following way:

$$\begin{aligned} \bar{K}_3 &= \tfrac{1}{2}\,\mathrm{diag}\,(1, 0, -1) = \tfrac{1}{2}(K_3 + 3K_8), \\ \bar{K}_4 &= -iK_4 = \tfrac{1}{2}\lambda_4, \\ \bar{K}_5 &= -iK_5 = \tfrac{1}{2}\lambda_5, \\ \bar{K}_8 &= \tfrac{1}{6}\,\mathrm{diag}\,(1, -2, 1) = \tfrac{1}{2}(K_3 - K_8). \end{aligned} \quad (10)$$

We denote the quantum numbers corresponding to $(\lambda\mu m)$ in the old basis by (JMn). From (10) we see that

$$M = \tfrac{1}{2}(\mu + 3m), \quad n = \tfrac{1}{2}(\mu - m), \quad \bar{K}_3^2 + \bar{K}_4^2 + \bar{K}_5^2 \to J(J + 1). \quad (11)$$

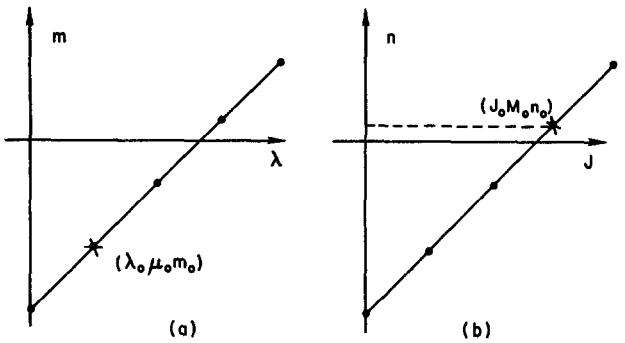


FIG. 1. The representation D^{+} ($q = 0$). The dotted line $n = \frac{1}{2}(\mu_0 - m_0)$ [Eq. (11)] intersects the net of figure (b) only in one point. Therefore every basis vector $|\lambda\mu m\rangle$ can be expressed using only one basis vector $|JMn\rangle$.

Now (9) can be written as follows:

$$\begin{aligned} &\langle\lambda'\mu'm'| e^{i\bar{K}_5\omega}|\lambda\mu m\rangle \\ &= \sum_J \langle\lambda'\mu'm' \,|\, JM'n'\rangle\langle JM'n'| e^{i\bar{K}_5\omega}|JMn\rangle\langle JMn \,|\, \lambda\mu m\rangle \\ &= \sum_J \langle\lambda'\mu'm' \,|\, JM'n'\rangle\, d^{J}_{M'M}(\omega)\langle JMn \,|\, \lambda\mu m\rangle\delta^{n'}_{n}. \end{aligned} \quad (12)$$

From (11) and (12) it can be seen that (12) is different from zero only if $\mu - m = \mu' - m'$. Up to this, we have only shifted the difficulties from (9) to the transformation coefficients $\langle JMn \,|\, \lambda\mu m\rangle$. In case of degenerate representations ($q = 0$ or $q = p$; see Ref. 3) only one of the coefficients $\langle JMn \,|\, \lambda\mu m\rangle$ is different from zero (see Figs. 1 and 2).

Using the matrix elements for the $SU(3)$ generators, which are well known,[3] and operating with $\bar{K}_3^2 + \bar{K}_4^2 + \bar{K}_5^2$ to a state $|\lambda\mu m\rangle$ in a degenerate representation we get, for D^{+},

$$\begin{aligned} J(J + 1) = \frac{1}{27}\frac{\lambda + \mu + 1}{(2\lambda + \lambda)(2\lambda + 2)}&(2p - 3\lambda - 3m) \\ &\times (p + 3\lambda + 3m + 3)(p + 3m + 3\lambda + 6) \\ &+ \tfrac{1}{4}(\mu + 3m)(\mu + 3m + 2) \end{aligned} \quad (13a)$$

and for D^{-},

$$\begin{aligned} J(J + 1) = \frac{1}{27}\frac{\lambda - \mu}{2\lambda(2\lambda + 1)}&(p + 3\lambda - 3m + 3) \\ &\times (2p - 3\lambda + 3m + 3)(p + 3\lambda - 3m) \\ &+ \tfrac{1}{4}(\mu + 3m)(\mu + 3m + 2). \end{aligned} \quad (13b)$$

From (8) and (12) it follows that

$$D^{pq}_{\lambda'\mu'm',\lambda\mu m}(\phi, \psi, \theta, \chi, \omega, \alpha, \beta, \gamma) \times \sum_{\bar{\mu}} D^{\lambda'}_{\mu'\bar{\mu}}(\phi, \psi, \theta)e^{im\chi}\, d^{J}_{M'M}(\omega)D^{\lambda}_{\bar{\mu}\mu}(\alpha, \beta, \gamma), \quad (14)$$

[3] G. E. Baird and L. C. Biedenharn, J. Math. Phys. 4, 1449 (1963).

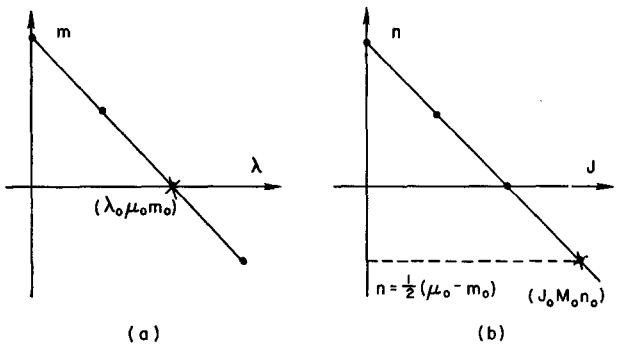


FIG. 2. The representation D^- $(q = p)$. Again the dotted line intersects the net of figure (b) in only one point, so that

$$|\lambda_0\mu_0 m_0\rangle = |J_0 M_0 n_0\rangle.$$

for the degenerate case only, where

$$M' = \tfrac{1}{2}(\tilde{\mu} + 3m'), \quad M = \tfrac{1}{2}(\tilde{\mu} + 3m),$$
$$\tilde{\mu} = \bar{\mu} - m' + m, \tag{15}$$
$$J = -\frac{1}{2} + \Big[\frac{1}{4} + \frac{1}{27}\frac{\lambda + \tilde{\mu} + 1}{(2\lambda + 1)(2\lambda + 2)} \times (2p - 3\lambda - 3m)(p + 3\lambda + 3m + 3) \times (p + 3m + 3\lambda + 6) + \tfrac{1}{4}(\tilde{\mu} + 3m)(\tilde{\mu} + 3m + 2)\Big]^{\frac{1}{2}},$$

for D^+, and

$$J = -\frac{1}{2} + \Big[\frac{1}{4} + \frac{1}{27}\frac{\lambda - \tilde{\mu}}{2\lambda(2\lambda + 1)}(p + 3\lambda - 2m + 3) \times (p + 3\lambda - 3m)(2p - 3\lambda + 3m + 3) + \tfrac{1}{4}(\tilde{\mu} + 3m)(\tilde{\mu} + 3m + 2)\Big]^{\frac{1}{2}},$$

for D^-.

For the nondegenerate case we use the following fact: Every nondegenerate representation (p, q) can be reduced from direct product of the type

$$D^-(q, q) \times D^+(p, 0). \tag{16}$$

[Note that the representation space of $D(p, q)$ is spanned by tensors with $p - q$ upper indices and q lower indices and which are symmetric in upper and lower indices and traceless.] Therefore, we can write

$$D^{pq}_{\lambda'\mu'm',\lambda\mu m} = \sum_{\lambda_1\mu_1 m_1} \sum_{\lambda_1'\mu_1' m_1'} \sum_{\lambda_2\mu_2 m_2} \sum_{\lambda_2'\mu_2' m_2'} \times C[qq; \lambda_1\mu_1 m_1 | p0; \lambda_2\mu_2 m_2 | pq; \lambda\mu m] \times C^*[qq; \lambda_1'\mu_1' m_1' | p0; \lambda_2'\mu_2' m_2' | pq; \lambda'\mu' m'] \times \delta^{\mu}_{\mu_1+\mu_2}\delta^{\mu'}_{\mu_1'+\mu_2'}\delta^{m}_{m_1+m_2}\delta^{m'}_{m_1'+m_2'} \times D^{qq}_{\lambda_1'\mu_1' m_1',\lambda_1\mu_1 m_1} D^{p0}_{\lambda_2'\mu_2' m_2',\lambda_2\mu_2 m_2}.$$

The $C[\cdots]$'s are $SU(3)$ Clebsch–Gordan coefficients and * means complex conjugation. For the case (16) they are relatively simple, and they can be found in Ref. 4 as follows:

$$C[qq; \lambda_1\mu_1 m_1 | p0; \lambda_2\mu_2 m_2 | pq; \lambda\mu m] = \{[(p + 1)!]^{-1} \times [\tfrac{1}{3}(2p - q) - m + \lambda + 1]! \times [m + \lambda + 1 + \tfrac{1}{3}(p + q)]! \times [\tfrac{1}{3}(2p - q) - m - \lambda]! \times [\tfrac{1}{3}(p + q) + m - \lambda]!\}^{\frac{1}{2}} \times \{[\tfrac{1}{3}(p - q) + 2m_2 - m + \lambda + 1]! \times [\tfrac{1}{3}(p - q) + 2m_2 - m - \lambda]! \times [\tfrac{2}{3}q - 2m_2 + 2m]!\,[\tfrac{1}{3}p - 2m_2]!\}^{-\frac{1}{2}} \times \left\langle \begin{matrix} \lambda_1, \lambda_2 \\ \mu_1\ \mu_2 \end{matrix} \middle| \begin{matrix} \lambda \\ \mu \end{matrix} \right\rangle.$$

SUMMARY

For the degenerate case, the representation functions can be found from (14). Doing the summation, note that J, M', M, and $\tilde{\mu}$ are functions of $\bar{\mu}$. For nondegenerate case the representation functions can be found from (17).

ACKNOWLEDGMENTS

The author is thankful to Prof. Pekka Tarjanne for encouragement and critical reading of the manuscript.

[4] R. T. Sharp and H. von Baeyer, J. Math. Phys. **7**, 1105 (1966).