JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Analyticity of Green's Functions of Dilute Quantum Gases

DAVID RUELLE
I.H.E.S. 91, Bures-sur-Yvette, France
and
Institute for Advanced Study, Princeton, New Jersey

(Received 13 July 1970)

Previous results on reduced density matrices of quantum gases are used to obtain theorems on the existence and analyticity of Green's functions.

In this article we point out that Ginibre's results on the reduced density matrices of quantum gases[1] have immediate implications for the existence and analyticity of Green's functions. If H_Λ is the Hamiltonian in the bounded region Λ, we define Green's functions by

$$G_\Lambda(\mathbf{x}_1, \cdots, \mathbf{x}_m; \zeta_1, \cdots, \zeta_m) = Z^{-1} \operatorname{Tr} (A_1(\mathbf{x}_1)e^{-(\zeta_2-\zeta_1)H_\Lambda} \times A_2(\mathbf{x}_2) \cdots e^{-(\zeta_m-\zeta_{m-1})H_\Lambda}A_m(\mathbf{x}_m)e^{-(\beta+\zeta_1-\zeta_m)H_\Lambda}),$$

where $Z = \operatorname{Tr} e^{-\beta H_\Lambda}$ and

$$A_k(\mathbf{x}_k) = a^*(x'_{k1}) \cdots a^*(x'_{kp(k)})a(x''_{k1}) \cdots a(x''_{kq(k)}).$$

Let $\varphi_k \in L^2((\mathbb{R}^\nu)^{p(k)+q(k)})$ for Fermi statistics, or $\varphi_k = \varphi'_k\varphi''_k$ with $\varphi'_k \in L^2(\mathbb{R}^{\nu p(k)})$, $\varphi''_k \in L^2(\mathbb{R}^{\nu q(k)})$ for Bose statistics; we write

$$G^\varphi_\Lambda(\zeta_1, \cdots, \zeta_m) = \int_{\Lambda^{p(1)+q(1)}} d\mathbf{x}_1 \cdots \int_{\Lambda^{p(m)+q(m)}} d\mathbf{x}_m \varphi_1(\mathbf{x}_1) \cdots \varphi(\mathbf{x}_m) \times G_\Lambda(\mathbf{x}_1, \cdots, \mathbf{x}_m; \zeta_1, \cdots, \zeta_m).$$

In the case of a system of particles interacting through a suitable pair potential Φ, and for small activity, the operator $e^{-\lambda H_\Lambda}$, with $\lambda > 0$, may be defined in terms of Wiener integrals and is of trace class. The operators $e^{-\lambda H_\Lambda}A_k(\mathbf{x}_k)e^{-\lambda H_\Lambda}$ can also be expressed in terms of Wiener integrals and are of trace class.

When λ is complex and $\operatorname{Re} \lambda > 0$, $e^{-\lambda H_\Lambda}$ is defined and analytic; therefore G_Λ is an analytic function of the complex variables $\zeta_k = \beta_k - it_k$ in the domain

$$\mathfrak{D} = \{(\zeta_1, \cdots, \zeta_m): \beta_1 < \cdots < \beta_m < \beta_1 + \beta\}.$$

If $t_1 = \cdots = t_m$, and $\beta_1 < \cdots < \beta_m < \beta_1 + \beta$, G_Λ can be expressed in terms of Wiener integrals and it follows from Ginibre's analysis[2] that, when $\Lambda \to \infty$ (e.g., Λ is a sphere centered at the origin and with radius tending to infinity),

$$G_\Lambda(\mathbf{x}_1, \cdots, \mathbf{x}_m; \beta_1, \cdots, \beta_m) \to G(\mathbf{x}_1, \cdots, \mathbf{x}_m; \beta_1, \cdots, \beta_m) \quad (1)$$

uniformly on compacts with respect to $\mathbf{x}_1, \cdots, \mathbf{x}_m$, and

$$G^\varphi_\Lambda(\beta_1, \cdots, \beta_m) \to G^\varphi(\beta_1, \cdots, \beta_m) = \int d\mathbf{x}_1 \cdots d\mathbf{x}_m \varphi_1(\mathbf{x}_1) \cdots \varphi_m(\mathbf{x}_m) \times G(\mathbf{x}_1, \cdots, \mathbf{x}_m; \beta_1, \cdots, \beta_m). \quad (2)$$

Proposition 1: There exists a function

$$G(\mathbf{x}_1, \cdots, \mathbf{x}_m; \zeta_1, \cdots, \zeta_m),$$

analytic with respect to $(\zeta_1, \cdots, \zeta_m) \in \mathfrak{D}$, and such that

$$\lim_{\Lambda\to\infty} G_\Lambda(\mathbf{x}_1, \cdots, \mathbf{x}_m; \zeta_1, \cdots, \zeta_m) = G(\mathbf{x}_1, \cdots, \mathbf{x}_m; \zeta_1, \cdots, \zeta_m)$$

uniformly on compacts with respect to $\mathbf{x}_1, \cdots, \mathbf{x}_m$, $\zeta_1, \cdots, \zeta_m$. Furthermore, if $(\zeta_1, \cdots, \zeta_m) \in \mathfrak{D}$,

$$\lim_{\Lambda\to\infty} G^\varphi_\Lambda(\zeta_1, \cdots, \zeta_m) = G^\varphi(\zeta_1, \cdots, \zeta_m) = \int d\mathbf{x}_1 \cdots d\mathbf{x}_m \varphi_1(\mathbf{x}_1) \cdots \varphi_m(\mathbf{x}_m) \times G(\mathbf{x}_1, \cdots, \mathbf{x}_m; \zeta_1, \cdots, \zeta_m).$$

We notice first that $\mathfrak{D}$ is the union (over $n > 0$) of the sets

$$\mathfrak{K}_n = \{(\zeta_1, \cdots, \zeta_m) \in \mathfrak{D}: \beta_2 - \beta_1 \geq \beta/2n, \cdots, \beta_m - \beta_{m-1} \geq \beta/2n, \beta + \beta_1 - \beta_m \geq \beta/2n\}.$$

If $(\zeta_1, \cdots, \zeta_m) \in \mathfrak{K}_n$, we may express G_Λ in terms of operators $e^{[it_k-(1/4n)\beta]H_\Lambda}A_k(\mathbf{x}_k)e^{-(it_k+(1/4n)\beta)H_\Lambda}$, and $e^{-\lambda H_\Lambda}$ with $0 < \lambda < \beta$. Using Hölder's inequality[3] we find an upper bound for $|G_\Lambda|$ in terms of the expressions

$$Z^{-1} \operatorname{Tr} \{[e^{-(1/4n)\beta H_\Lambda}A_k(\mathbf{x}_k)^* \times e^{-(1/2n)\beta H_\Lambda}A_k(\mathbf{x}_k)e^{-(1/4n)\beta H_\Lambda}]^n\},$$

which are known by (1) to have a limit when $\Lambda \to \infty$. We may thus assume that G_Λ is bounded on each $\mathfrak{K}_n$ uniformly with respect to Λ and $\mathbf{x}_1, \cdots, \mathbf{x}_m$ in a compact, and the convergence of G_Λ on the real points

of $\mathfrak{D}$ implies its uniform convergence on the compacts of $\mathfrak{D}$. The uniformity of the convergence with respect to $\mathbf{x}_1, \cdots, \mathbf{x}_m$ on compacts follows from the uniformity of (1).

The proof of the convergence of G_Λ^φ proceeds like the proof of the convergence of G_Λ and shows in particular that the limit of G_Λ^φ is a bounded multilinear functional of $\varphi_1, \cdots, \varphi_m$ (Fermi) or $\varphi_1', \cdots, \varphi_m''$ (Bose) on the product of the relevant L^2 spaces, identification of the limit follows from (2), taking $\varphi_1, \cdots, \varphi_m$ with compact supports.

Proposition 2: Let $m = 2$ and let the pair potential $\Phi \in L^1(\mathbb{R}^\nu) \cap L^2(\mathbb{R}^\nu)$. Then G^φ extends to a bounded continuous function on $\overline{\mathfrak{D}}$ such that

$$\lim_{\Lambda \to \infty} G_\Lambda^\varphi(\zeta_1, \zeta_2) = G^\varphi(\zeta_1, \zeta_2)$$

uniformly on the compacts of the closure $\overline{\mathfrak{D}}$ of $\mathfrak{D}$.

The operators $A_k(\varphi_k)e^{-\lambda H_\Lambda}$ are of trace class [consider $e^{-\lambda H_\Lambda}A_k(\varphi_k)^* A_k(\varphi_k)e^{-\lambda H_\Lambda}$] and, if $\beta_2 = \beta_1$ or $\beta_2 = \beta_1 + \beta$, we have

$$|G_\Lambda^\varphi(\zeta_1, \zeta_2)| \leq [G_{1\Lambda} G_{2\Lambda}]^{\frac{1}{2}}, \tag{3}$$

$$G_{k\Lambda} = Z^{-1} \operatorname{Tr} \{[A_k(\varphi_k)^* A_k(\varphi_k) + A_k(\varphi_k) A_k(\varphi_k)^*] e^{-\beta H_\Lambda}\}. \tag{4}$$

We assume first Bose statistics. The reduced density matrices are integral kernels of bounded operators in L^2. When (Λ_n) tends to infinity, these bounded operators form a bounded sequence converging in the strong operator topology.[1] Therefore, there exists $C > 0$ such that

$$G_{k\Lambda_n} \leq C \int d\mathbf{x}_k \, |\varphi_k(\mathbf{x}_k)|^2 \tag{5}$$

for all n. Since G_Λ^φ is analytic and bounded, (3) holds for all $(\zeta_1, \cdots, \zeta_2)$ in $\overline{\mathfrak{D}}$ and, using (5), this gives

$$|G_{\Lambda_n}^\varphi(\zeta_1, \zeta_2)| \leq C \, \|\varphi_1\|_2 \, \|\varphi_2\|_2 \tag{6}$$

for all n, and $(\zeta_1, \zeta_2) \in \overline{\mathfrak{D}}$. In the case of Fermi statistics, (6) holds again (with $C = 1$) because

$$\|A_k(\varphi_k)\| \leq \|\varphi_k\|_2 .$$

In view of (6), it suffices to prove the proposition when $A_k(\varphi_k)$ is of the form

$$A_k(\varphi_k) = a^*(\psi_{k'}') \cdots a^*[\psi_{kp(k)}'] a(\psi_{k'}'') \cdots a[\psi_{kq(k)}''],$$

where $\psi_k', \cdots, \psi_{kq(k)}''$ are of class C^2 with compact support.

We have

$$\frac{d}{d\zeta_2} G_\Lambda^\varphi = Z^{-1} \operatorname{Tr} \{A_1(\varphi_1) \times e^{-(\zeta_2 - \zeta_1)H_\Lambda}[A_2(\varphi_2), H_\Lambda] e^{-(\beta + \zeta_1 - \zeta_2)H_\Lambda}\},$$

and therefore

$$\left| \frac{d}{d\zeta_2} G_\Lambda^\varphi \right| \leq [G_{1\Lambda} G_{2\Lambda}']^{\frac{1}{2}},$$

where $G_{2\Lambda}$ is given by (4) with $A_k(\varphi_k)$ replaced by $[A_2(\varphi_2), H_\Lambda]$. Since $\psi_{21}', \cdots, \psi_{2q(2)}''$ are of class C^2 with compact support, the commutator of $A_2(\varphi_2)$ with the kinetic energy part of H_Λ is again of the form $A(\varphi)$. In view of this, $G_{2\Lambda}'$ is a sum of integrals of reduced density matrices $G_\Lambda'(x_1, \cdots, x_r)$ multiplied by continuous functions $\psi(x_i)$, with compact support, and the pair potential $\Phi(x_k - x_j)$. The pair potential appears as factor 0, 1, or 2 times; if $\Phi(x_k - x_j)$ appears, there also appears a factor $\psi(x_j)$ or $\psi(x_k)$; for each variable x_i in $G_\Lambda'(x_1, \cdots, x_r)$ which does not appear in a factor $\Phi(x_j - x_i)$, there is a factor $\psi(x_i)$. Using the condition $\Phi \in L^1(\mathbb{R}^\nu) \cap L^2(\mathbb{R}^\nu)$ and the fact that the reduced density matrices G_Λ' are bounded functions uniformly in Λ[1], we obtain a bound on $G_{2\Lambda}'$ which is independent of Λ. Therefore

$$\left| \frac{d}{d\zeta_2} G_\Lambda^\varphi \right| = \left| \frac{d}{d\zeta_1} G_\Lambda^\varphi \right|$$

is bounded on $\mathfrak{D}$ uniformly in Λ. The convergence of G_Λ^φ in $\mathfrak{D}$ implies then its uniform convergence on the compacts of $\overline{\mathfrak{D}}$.

Remark 1: Let $m = 3$, $\Phi \in L^1(\mathbb{R}^\nu) \cap L^2(\mathbb{R}^\nu)$. In the case of Fermi statistics, G^φ extends to a bounded continuous function on $\overline{\mathfrak{D}}$ such that

$$\lim_{\Lambda \to \infty} G_\Lambda^\varphi(\zeta_1, \zeta_2, \zeta_3) = G^\varphi(\zeta_1, \zeta_2, \zeta_3)$$

uniformly on the compacts of the closure $\overline{\mathfrak{D}}$ of $\mathfrak{D}$.

To estimate $|dG_\Lambda^\varphi / d\zeta_i|$ it suffices to consider the expression

$$Z^{-1} \operatorname{Tr} \{e^{(-\beta + it)H_\Lambda}[A_i, H_\Lambda] e^{it'H_\Lambda} A_j e^{it''H_\Lambda} A_k\}$$

and similar ones, where $[A_i, H_\Lambda]$ and A_j, A_k are circularly permuted. If $[A_i, H_\Lambda]$ occupies the middle position, we rewrite the expression in terms of $[A_j, H_\Lambda]$ and $[A_k, H_\Lambda]$. The rest of the argument goes as for $m = 2$ (using the boundedness of A_j).

Proposition 3: In the case of Fermi statistics, introduce the operators

$$A_k(\varphi_k, f_k) = \int dt f_k(t) e^{itH_\Lambda} A_k(\varphi_k) e^{-itH_\Lambda},$$

where $f_k \in L^1(\mathbb{R})$, then the limit

$$\lim_{\Lambda \to \infty} Z^{-1} \operatorname{Tr} (A_1(\varphi_1, f_1) \cdots A_m(\varphi_m, f_m) e^{-\beta H_\Lambda})$$

exists.

It is sufficient to prove this for f_k of class C^1 with compact support. We construct Green's functions

with the operators $A_k(\varphi_k, f_k)$ instead of $A_k(\varphi_k)$, and use the fact that the derivatives of these functions are bounded in $\mathfrak{D}$ uniformly with respect to Λ.

Remark 2: Proposition 3 has obvious implications for the description of time evolution of a dilute Fermi gas. It does not, however, exhibit this time evolution as a group of automorphisms of the C^* algebra of the anticommutation relations. Streater[4] and Hepp[5] have shown that such a group of automorphisms exists for some nonlocal interactions.

I am indebted to J. Ginibre for helpful discussions; I also wish to thank K. Hepp for explaining to me his results before publication.

[1] See J. Ginibre, J. Math. Phys. **6**, 238, 252, 1432 (1965); in *Statistical Mechanics*, edited by T. Bak (Benjamin, New York, 1967), p. 148.
[2] This result is contained in C. Gruber, thesis, Princeton University, 1968 (unpublished); see also J. Ginibre and C. Gruber, Commun. Math. Phys. **11**, 198 (1969).
[3] See N. Dunford and J. Schwartz, *Linear Operators* (Interscience, New York, 1963), p. 1105, Lemma XI. 9–20.
[4] R. F. Streater, Commun. Math. Phys. **7**, 93 (1968).
[5] K. Hepp (unpublished).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Relaxation to Quantum-Statistical Equilibrium of the Wigner–Weisskopf Atom in a One-Dimensional Radiation Field. III. The Quantum-Mechanical Solution

RUSSELL DAVIDSON
Center for Statistical Mechanics and Thermodynamics, The University of Texas at Austin Austin, Texas 78712

AND

JOHN J. KOZAK
Department of Chemistry, The University of Notre Dame, Notre Dame, Indiana 46556

(Received 23 September 1970)

A study is undertaken to cast light on difficulties, which arose in the first two papers of this series, pertaining to the occurrence of negative probabilities in the weak-coupling solution of the generalized Prigogine–Résibois master equation for the model of the Wigner–Weisskopf atom in a one-dimensional radiation field. The Schrödinger equation is solved exactly for the model with the initial condition for spontaneous emission, and then the weak-coupling approximations to the solution, both for an infinite and for a finite system, are derived as inverse Laplace transform integrals. An extensive analysis, theoretical and numerical, of these is undertaken, and comparison is made with the corresponding results based on the master equation. In particular, quantitative estimates of the Poincaré recurrence times for finite systems are made. It is found that considerable differences exist between the statistical-mechanical and quantum-mechanical results, but that both manifest nonanalyticity in the coupling parameter as it tends to zero. Suggestions are given for further work toward the resolution of these discrepancies and a better understanding of the weak-coupling limit.

I. INTRODUCTION

The work of this paper is a continuation of that of two previous papers[1,2] by the authors, hereafter referred to respectively as I and II. Certain problems were considered there whose existence in general had been pointed out by Zwanzig[3] and which arose in particular in the treatment, by the methods of non-equilibrium statistical mechanics, of the Wigner–Weisskopf atom, a model which consists of a two-level quantum system in interaction with a massless boson field, always taken as one-dimensional. In I, the possibility of nonexponential behavior was discussed for the problem of spontaneous emission—the Prigogine–Résibois master equation[4] for the diagonal elements of the density matrix was solved in the approximation of weak coupling and it was seen to give rise to an exponential decay overlaid by a (considerably smaller numerically) sequence of slowly damped oscillations. The separate question was treated in II of the behavior of finite systems, that is, those for which the usual "thermodynamic limit" of statistical mechanics, with its assumption of unlimited

with the operators $A_k(\varphi_k, f_k)$ instead of $A_k(\varphi_k)$, and use the fact that the derivatives of these functions are bounded in $\mathfrak{D}$ uniformly with respect to Λ.

Remark 2: Proposition 3 has obvious implications for the description of time evolution of a dilute Fermi gas. It does not, however, exhibit this time evolution as a group of automorphisms of the C^* algebra of the anticommutation relations. Streater[4] and Hepp[5] have shown that such a group of automorphisms exists for some nonlocal interactions.

I am indebted to J. Ginibre for helpful discussions; I also wish to thank K. Hepp for explaining to me his results before publication.

[1] See J. Ginibre, J. Math. Phys. **6**, 238, 252, 1432 (1965); in *Statistical Mechanics*, edited by T. Bak (Benjamin, New York, 1967), p. 148.
[2] This result is contained in C. Gruber, thesis, Princeton University, 1968 (unpublished); see also J. Ginibre and C. Gruber, Commun. Math. Phys. **11**, 198 (1969).
[3] See N. Dunford and J. Schwartz, *Linear Operators* (Interscience, New York, 1963), p. 1105, Lemma XI. 9–20.
[4] R. F. Streater, Commun. Math. Phys. **7**, 93 (1968).
[5] K. Hepp (unpublished).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Relaxation to Quantum-Statistical Equilibrium of the Wigner–Weisskopf Atom in a One-Dimensional Radiation Field. III. The Quantum-Mechanical Solution

RUSSELL DAVIDSON
Center for Statistical Mechanics and Thermodynamics, The University of Texas at Austin Austin, Texas 78712

AND

JOHN J. KOZAK
Department of Chemistry, The University of Notre Dame, Notre Dame, Indiana 46556

(Received 23 September 1970)

A study is undertaken to cast light on difficulties, which arose in the first two papers of this series, pertaining to the occurrence of negative probabilities in the weak-coupling solution of the generalized Prigogine–Résibois master equation for the model of the Wigner–Weisskopf atom in a one-dimensional radiation field. The Schrödinger equation is solved exactly for the model with the initial condition for spontaneous emission, and then the weak-coupling approximations to the solution, both for an infinite and for a finite system, are derived as inverse Laplace transform integrals. An extensive analysis, theoretical and numerical, of these is undertaken, and comparison is made with the corresponding results based on the master equation. In particular, quantitative estimates of the Poincaré recurrence times for finite systems are made. It is found that considerable differences exist between the statistical-mechanical and quantum-mechanical results, but that both manifest nonanalyticity in the coupling parameter as it tends to zero. Suggestions are given for further work toward the resolution of these discrepancies and a better understanding of the weak-coupling limit.

I. INTRODUCTION

The work of this paper is a continuation of that of two previous papers[1,2] by the authors, hereafter referred to respectively as I and II. Certain problems were considered there whose existence in general had been pointed out by Zwanzig[3] and which arose in particular in the treatment, by the methods of non-equilibrium statistical mechanics, of the Wigner–Weisskopf atom, a model which consists of a two-level quantum system in interaction with a massless boson field, always taken as one-dimensional. In I, the possibility of nonexponential behavior was discussed for the problem of spontaneous emission—the Prigogine–Résibois master equation[4] for the diagonal elements of the density matrix was solved in the approximation of weak coupling and it was seen to give rise to an exponential decay overlaid by a (considerably smaller numerically) sequence of slowly damped oscillations. The separate question was treated in II of the behavior of finite systems, that is, those for which the usual "thermodynamic limit" of statistical mechanics, with its assumption of unlimited

extent and an infinite number of degrees of freedom, has not been taken. This treatment, using the same method of the master equation and of a weak-coupling scheme, produced a solution which manifested a complicated time dependence and, in particular, the existence of Poincaré recurrences. What the results of both I and II had in common was the appearance of contributions to the solutions of the master equation which were not analytic in the coupling constant when that quantity tended to zero. This effect, which appeared to be essentially involved in the weak-coupling approximation, gave rise to an unusual and disturbing feature, namely that the diagonal elements of the density matrix, quantities which, as probabilities, should always lie between zero and unity, could become negative. In this paper, an alternative approach to the same problem is presented which obviates this difficulty.

The model to be discussed has already been specified in I and II. The Hamiltonian [Eq. (I-6)] is

$$H = \epsilon_1 \alpha\alpha^* + \epsilon_2 \alpha^*\alpha + \sum_\lambda [\tfrac{1}{2}\hbar\omega_\lambda (a_\lambda^* a_\lambda + 1)] + \sum_\lambda (h_\lambda^* \alpha^* a_\lambda + h_\lambda \alpha a_\lambda^*), \quad (1)$$

in which ϵ_1, ϵ_2 are respectively the energies of the ground state $|1\rangle$ and the excited state $|2\rangle$ of the two-level system. The operators a_λ and α are defined by the following equations:

$$\alpha = |1\rangle\langle 2|,$$
$$\alpha^* = |2\rangle\langle 1|,$$
$$\langle n_\lambda | a_\lambda | m_\lambda \rangle = [2(n_\lambda + 1)]^{\frac{1}{2}} \delta^{\mathrm{Kr}}(m_\lambda - n_\lambda - 1) = \langle m_\lambda | a_\lambda^* | n_\lambda \rangle.$$

The state $|n_\lambda\rangle$ is one with n_λ $(= 0, 1, 2, \cdots)$ photons in the λth mode, and $\delta^{\mathrm{Kr}}(\cdots)$ denotes a Kronecker delta. Following the arguments of I and II, we shall choose the coupling h_λ such that

$$|h_\lambda|^2 = \hbar^2 \alpha c E / L,$$

where $\hbar E = \epsilon_2 - \epsilon_1$, c is the velocity of light, and L is the length of the system. The normal modes λ of the boson field can be characterized by the wave numbers k_λ which satisfy the following relations:

$$\omega_\lambda = c\,|k_\lambda|,$$
$$k_\lambda = \frac{2\pi n}{L},$$

where the integers n replace λ as the label. The dimensionless coupling constant α replaces, for a one-dimensional system, the fine-structure constant of quantum electrodynamics. The states of the system, between which matrix elements of the Hamiltonian [Eq. (1)] are to be taken, are given by

$$|i; \{n_\lambda\}\rangle = |i\rangle \prod_{\lambda\in\Lambda} |n_\lambda\rangle,$$

with $i = 1, 2$ and $n_\lambda = 0, 1, 2, \cdots$ for each mode λ. It should be noted that, for the problem of spontaneous emission, the initial state of the system is chosen as

$$|2; \{0\}\rangle. \quad (2)$$

Further, the Hamiltonian admits no terms proportional to either αa_λ or $\alpha^* a_\lambda^*$ which would induce nonenergy-conserving transitions.

With the model that we have now specified, an expression was obtained in II [Eq. (II-40)] for that diagonal element of the density matrix which expresses the probability at any moment that the 2-level system be in its excited state, if the initial state (2) is taken at time zero. [This expression is reproduced in Eq. (43).] It consists of a constant term plus a sum of terms each of which depends on time through a cosine factor. The frequencies appearing in these factors are all linearly independent. Because of this fact, this sum of time-dependent terms assumes negative values as frequently as positive ones, and its greatest lower bound has the same absolute value as its least upper bound. Thus, certainly when the constant term in the expression is less than $\frac{1}{2}$—which it is for most choices of the parameters of the system—the probability will be negative for some times. If, however, the spontaneous emission problem is treated, not by the weak-coupling master equation, which is an approximation to the Liouville–von Neumann equation for the *density matrix*, but by the methods of quantum mechanics starting from the Schrödinger equation for the *state vector* of the system, then the probability of finding the system in a given quantum state will be the square of the modulus of some complex number lying on or within the unit circle, and so will lie between zero and unity, whether or not a weak-coupling approximation is made. This remark motivates the study of the present paper. The problem of the spontaneous emission of the Wigner–Weisskopf atom is dealt with by quantum mechanics, both for a finite system and the limiting case of an infinite one; within the framework of weak coupling, a detailed comparison is made between the different expressions resulting from this treatment and those of I and II. It should be emphasized that we may expect answers of an essentially different structure here, since it is reasonable to suppose that the frequencies which will appear in the quantum-mechanical treatment will be linearly independent, just like those previously

obtained, and so the time-dependence of the probability cannot be of the same form as that of Eq. (II-40). It will be seen that this is indeed so.

The next section contains the formulation of our problem in terms of the Schrödinger equation, which is then solved by a resolvent technique very similar to those used in the statistical-mechanical theory based on the master equation. With this solution to hand, a weak-coupling scheme is introduced and, in Sec. III, the dynamics presented in the limit of a large system. Some substantial differences from the results of I are pointed out. In Sec. IV, the finite system is treated. A solution analogous to that of II is obtained and it is seen to differ from it in many respects. These are highlighted by a detailed discussion of the asymptotic form of the result for large but not infinite systems. The actual numerical estimation of the solution is also presented in this section. This makes possible the investigation of Sec. V, where there is made a characterization of the various expressions previously obtained which facilitates a comparison of their properties with those of the results of II, as well as a quantitative discussion of the phenomenon of Poincaré recurrences. The conclusions to be drawn from this work are discussed in the final section, in which an attempt is made to elucidate the nature of the weak-coupling approximations used. Further work is suggested which might lead to a better understanding of the problems—important ones, we feel, in the study of irreversible processes—which arise in approximate or incomplete descriptions of the dynamics of large systems.

II. THE QUANTUM-MECHANICAL FORMULATION OF THE PROBLEM

The Schrödinger equation for the time evolution of the state vector $|\Psi(t)\rangle$ is

$$i\hbar \frac{\partial}{\partial t}|\Psi(t)\rangle = H|\Psi(t)\rangle, \tag{3}$$

for which the solution is

$$|\Psi(t)\rangle = e^{-iHt/\hbar}|\Psi(0)\rangle. \tag{4}$$

To make this solution explicit, we must specify the Hilbert space on which it is defined and provide this space with a suitable orthonormal basis. It is clear from the Hamiltonian, Eq. (1), that if the initial condition (2) appropriate to the problem of spontaneous emission is chosen, the only states accessible to the system are those for which the two-level atom is de-excited in the presence of one excitation in some mode of the radiation field, and that for which the atom is excited in the absence of any excitation of the field. Let us write these states respectively as $|\lambda\rangle$, where the label λ ranges over the modes of the field, and $|\mathcal{N}\rangle$. The Hilbert space spanned by the $|\lambda\rangle$ and $|\mathcal{N}\rangle$ is then taken as that on which Eq. (4) is defined. The orthonormality relations given by

$$\langle\lambda|\lambda'\rangle = \delta_{\lambda\lambda'},$$
$$\langle\lambda|\mathcal{N}\rangle = 0,$$
$$\langle\mathcal{N}|\mathcal{N}\rangle = 1$$

follow from the fact that these states are nondegenerate eigenstates of the unperturbed Hamiltonian H_0 such that

$$H_0|\mathcal{N}\rangle = \hbar E,$$
$$H_0|\lambda\rangle = \hbar\omega_\lambda,$$

where the zero of energy is chosen as that of the de-excited atom plus the zero-point energy of the field. Here

$$\hbar E = \epsilon_2 - \epsilon_1.$$

On the Hilbert space of the problem, the completeness relation

$$|\mathcal{N}\rangle\langle\mathcal{N}| + \sum_\lambda |\lambda\rangle\langle\lambda| = 1 \tag{5}$$

will also hold.

The Eq. (4) may be expressed in terms of the resolvent operator

$$(\mathfrak{z} - H/\hbar)^{-1}$$

as follows:

$$|\Psi(t)\rangle = -\frac{1}{2\pi i}\int_c d\mathfrak{z}\, e^{-i\mathfrak{z}t}\left(\mathfrak{z} - \frac{H}{\hbar}\right)^{-1}|\Psi(0)\rangle,$$

where the Bromwich contour C is taken parallel to the positive direction of the real axis of $\mathfrak{z}$ and above all singularities of the integrand. For the case of spontaneous emission we take $|\Psi(0)\rangle = |\mathcal{N}\rangle$, and so the problem will be solved if the expression

$$(\mathfrak{z} - H/\hbar)^{-1}|\mathcal{N}\rangle$$

can be found. To this end, let us write

$$(\mathfrak{z} - H/\hbar)^{-1}|\mathcal{N}\rangle = |S(\mathfrak{z}, \mathcal{N})\rangle, \tag{6}$$

and then

$$(\mathfrak{z} - H/\hbar)|S(\mathfrak{z}, \mathcal{N})\rangle = |\mathcal{N}\rangle. \tag{7}$$

The ket $|S(\mathfrak{z}, \mathcal{N})\rangle$ can be expanded in the following way:

$$|S(\mathfrak{z}, \mathcal{N})\rangle = S_{\mathcal{N}}(\mathfrak{z}, \mathcal{N})|\mathcal{N}\rangle + \sum_\lambda S_\lambda(\mathfrak{z}, \mathcal{N})|\lambda\rangle. \tag{8}$$

When this is done, there results the following set of linear equations for the coefficients $S_{\mathcal{N}}(\mathfrak{z}, \mathcal{N})$ and $S_\lambda(\mathfrak{z}, \mathcal{N})$:

$$(\mathfrak{z} - E)S_{\mathcal{N}}(\mathfrak{z}, \mathcal{N}) + \sum_\lambda \left(-\frac{h_\lambda\sqrt{2}}{\hbar}\right)S_\lambda(\mathfrak{z}, \mathcal{N}) = 1, \tag{9}$$

$$\left(-\frac{h_\lambda\sqrt{2}}{\hbar}\right)S_{\mathcal{N}}(\mathfrak{z}, \mathcal{N}) + (\mathfrak{z} - \omega_\lambda)S_\lambda(\mathfrak{z}, \mathcal{N}) = 0. \tag{10}$$

We solve these equations and find

$$S_\lambda(\mathcal{Z}, \mathcal{N}) = \frac{h_\lambda \sqrt{2}}{\hbar(\mathcal{Z} - \omega_\lambda)} S_{\mathcal{N}}(\mathcal{Z}, \mathcal{N}),$$

$$S_{\mathcal{N}}(\mathcal{Z}, \mathcal{N}) = \left[\mathcal{Z} - E - \sum_\lambda \frac{2|h_\lambda|^2}{\hbar^2(\mathcal{Z} - \omega_\lambda)}\right]^{-1}. \quad (11)$$

The main object of study will be $\langle \mathcal{N} | \Psi(t) \rangle$, which is given by

$$\langle \mathcal{N} | \Psi(t) \rangle = -\frac{1}{2\pi i} \int_c d\mathcal{Z} e^{-i\mathcal{Z}t} S_{\mathcal{N}}(\mathcal{Z}, \mathcal{N})$$

$$= -\frac{1}{2\pi i} \int_c d\mathcal{Z} e^{-i\mathcal{Z}t} \left[\mathcal{Z} - E - \sum_\lambda \frac{2|h_\lambda|^2}{\hbar^2(\mathcal{Z} - \omega_\lambda)}\right]^{-1}. \quad (12)$$

It can be seen that the choice of $|h_\lambda|^2$ used in I and II, namely

$$|h_\lambda|^2 = \hbar^2 \alpha c E / L, \quad (13)$$

will cause the summation in Eq. (12) to diverge with the usual frequency spectrum

$$\omega_\lambda = |2\pi c n_\lambda / L|, \quad n_\lambda \text{ an integer},$$

and so the following artifice will be used: The $|h_\lambda|^2$ will be made proportional to a negative power $-p$ of ω_λ in such a way that, for $\omega_\lambda = E$, it takes the value given by Eq. (13). Hence, we put

$$|h_\lambda|^2 = \hbar^2 \alpha c^{1-p} E^{1+p} / L \, |k_\lambda|^p. \quad (14)$$

This device is similar to that used in I to avoid the ultraviolet divergence, and here, for this reason and also to avoid the infrared divergence, we must take $0 < p < 1$.

It may seem surprising that an *exact* solution, Eq. (12), has been obtained from a Hamiltonian involving a nontrivial interaction term between a two-level system and a field, and it may not be out of place to discuss this point here. The result has been obtained because it was possible, with the specific choice of the initial state vector $|\Psi(0)\rangle$, to restrict attention to the Hilbert space spanned by $|\mathcal{N}\rangle$ and the $|\lambda\rangle$. Further, the Hamiltonian excludes all processes which lead to unlimited creation of photons of the field, and in this way any consideration of a Fock space for these is obviated. Finally, the absence of any direct photon–photon interaction means that the infinite set of equations (9, 10) can be solved explicitly. There is another way to obtain Eq. (12), which also casts light on the matter. By an analysis exactly analogous to that of the third section of I, a "master equation" can be constructed from the Schrödinger equation (3) for the quantity $\langle \mathcal{N} | \Psi(t) \rangle$:

$$\frac{\partial}{\partial t} \langle \mathcal{N} | \Psi(t) \rangle = -iE \langle \mathcal{N} | \Psi(t) \rangle + \int_0^t d\tau \mathcal{C}(\tau) \langle \mathcal{N} | \Psi(t - \tau) \rangle. \quad (15)$$

This equation may be solved by the method of Laplace transforms to yield

$$S_{\mathcal{N}}(\mathcal{Z}, \mathcal{N}) = -[\psi(\mathcal{Z}) + i(\mathcal{Z} - E)]^{-1} \langle \mathcal{N} | \Psi(0) \rangle, \quad (16)$$

where the "collision operator" $\psi(z)$ is the Laplace transform of $\mathcal{C}(\tau)$ in Eq. (15) and is given by a perturbation expansion

$$\psi(z) = \frac{1}{i\hbar} \sum_{n=1}^{\infty} \langle \mathcal{N}| H_1 \left[\frac{1}{(\hbar\mathcal{Z} - H_0)} H_1\right]^n |\mathcal{N}\rangle_{\text{irr}}. \quad (17)$$

Here the suffix "irr" means that when the matrix element is expanded using the relation (5), the term $|\mathcal{N}\rangle\langle\mathcal{N}|$ is to be omitted. But then one can see that, since

$$\langle \lambda | H_1 | \lambda' \rangle = 0 \quad \text{for} \quad \lambda \neq \lambda',$$

only the first term of Eq. (17) will be nonvanishing. Substitution of this term into Eq. (16) yields the result Eq. (11).

Finally, in this section, the summation

$$\sum_\lambda \frac{2|h_\lambda|^2}{\hbar^2(\mathcal{Z} - \omega_\lambda)}$$

will be performed in the limit of an infinite system, i.e., where $L \to \infty$. With the expression (14) for $|h_\lambda|^2$ the sum becomes

$$4 \sum_{n=1}^{\infty} \left(\frac{\alpha c^{1-p} E^{1+p}}{L \left(\frac{2\pi n}{L}\right)^p \left(z - \frac{2\pi c n}{L}\right)} \right), \quad (18)$$

which in the limit $L \to \infty$ goes over to an integral

$$-\frac{2}{\pi} \alpha E^{1+p} \int_0^\infty dk \frac{1}{k^p(k - \mathcal{Z})} = -2\alpha E^{1+p} \mathcal{Z}^{-p} e^{ip\pi} \csc(p\pi) \quad (19)$$

for $\operatorname{Im} \mathcal{Z} > 0$. In Eq. (19), $\arg \mathcal{Z}$ is chosen to lie between zero and π. In this limit, then, the solution (12) is

$$\langle \mathcal{N} | \Psi(t) \rangle = -\frac{1}{2\pi i} \int_c d\mathcal{Z} e^{-i\mathcal{Z}t} [\mathcal{Z} - E + 2\alpha E^{2+p} \mathcal{Z}^{-p} e^{ip\pi} \csc(p\pi)].$$

III. THE LIMIT OF AN INFINITE SYSTEM

The calculations performed from this point on will all be, in some sense, in the limit of weak coupling.

This limit has been discussed at length in I and II for the statistical-mechanical formalism which springs from the Liouville–von Neumann equation. It has been seen there that the introduction of a scaled dimensionless time

$$\tau = \alpha E t$$

is necessary for the proper taking of the limit. In terms of τ, then, Eq. (12) is, by the use of the result (19),

$$\langle \mathcal{N} \mid \Psi(\tau)\rangle = -\frac{1}{2\pi i}\int_c d\xi e^{-i\xi\tau}\frac{\alpha}{\alpha\xi - 1 + 2\alpha^{1-p}e^{ip\pi}\csc(p\pi\xi^{-p})}, \tag{20}$$

where

$$\xi = \frac{\mathcal{Z}}{\alpha E}.$$

The usual procedure is now to let α tend to zero and to call the result the weak-coupling solution. But Eq. (20) vanishes if this is done. The remedy is to extract a phase factor of

$$e^{-iEt},$$

which is the solution of the problem without interaction and which, when written as

$$e^{-i\tau/\alpha},$$

has no weak-coupling limit. When this is done, the result is

$$\begin{aligned} e^{iEt}\langle \mathcal{N} \mid \Psi(t)\rangle &= -\frac{1}{2\pi i}\int_c d\mathcal{Z} e^{-i\mathcal{Z}t}[\mathcal{Z} + 2\alpha E^{1+p}e^{ip\pi}\csc p\pi(\mathcal{Z} + E)^{-p}]^{-1} \\ &= -\frac{1}{2\pi i}\int_c d\xi e^{-i\xi\tau}[\xi + 2e^{ip\pi}\csc p\pi(\alpha\xi + 1)^{-p}]^{-1} \end{aligned} \tag{21}$$

in terms of dimensionless variables. The limit of $\alpha \to 0$ can now be taken, yielding the weak-coupling solution

$$\begin{aligned} e^{iEt}\langle \mathcal{N} \mid \Psi(\tau)\rangle &= -\frac{1}{2\pi i}\int_c d\xi e^{-i\xi\tau}[\xi + 2e^{ip\pi}\csc(p\pi)]^{-1} \\ &= \exp[2i\tau\csc(p\pi)e^{ip\pi}]. \end{aligned} \tag{22}$$

The probability that the state $|\mathcal{N}\rangle$ be occupied at the time denoted by τ is given by

$$|\langle \mathcal{N} \mid \Psi(\tau)\rangle|^2 = |e^{iEt}\langle \mathcal{N} \mid \Psi(\tau)\rangle|^2 = \rho(\tau)$$

say, in conformity with the notation of the density-matrix formalism. Then,

$$\begin{aligned} \rho(\tau) &= |\exp[2i\tau\csc(p\pi)e^{ip\pi}]|^2 \\ &= \exp[2\,\mathrm{Re}\,(2i\tau\csc(p\pi)e^{ip\pi})] \\ &= e^{-4\tau}. \end{aligned} \tag{23}$$

This is the usual result.

It is interesting to see if we can go beyond the result given by Eq. (23) and obtain expressions akin to those of I which manifest a nonexponential contribution to $\rho(\tau)$. That contribution arose as a result of a perturbative solution of the generalized master equation in which only the leading term of the perturbation series was taken, but in which no $\alpha \to 0$ limit was imposed. It has been seen in the preceding section that this kind of approach to the Schrödinger equation leads to the *exact* result. Another kind of perturbation theory would thus be necessary to achieve results similar to those of I. Such a theory is that based on the identity

$$\left(\mathcal{Z} - \frac{H}{\hbar}\right)^{-1} = \left(\mathcal{Z} - \frac{H_0}{\hbar}\right)^{-1} + \left(\mathcal{Z} - \frac{H_0}{\hbar}\right)^{-1}\frac{H_1}{\hbar}\left(\mathcal{Z} - \frac{H}{\hbar}\right)^{-1}$$

of which the right-hand side can be iterated

$$\left(\mathcal{Z} - \frac{H}{\hbar}\right)^{-1} = \sum_{n=0}^{\infty}\left(\mathcal{Z} - \frac{H_0}{\hbar}\right)^{-1}\left[\frac{H_1}{\hbar}\left(\mathcal{Z} - \frac{H_0}{\hbar}\right)^{-1}\right]^n. \tag{24}$$

With this expression, Eq. (6) gives

$$S_{\mathcal{N}}(\mathcal{Z}, \mathcal{N}) = \sum_{n=0}^{\infty}(\mathcal{Z} - E)^{-1}\langle \mathcal{N}|\left[\frac{H_1}{\hbar}\left(\mathcal{Z} - \frac{H_0}{\hbar}\right)^{-1}\right]^n|\mathcal{N}\rangle. \tag{25}$$

The $n = 1$ term vanishes, and so the first correction to the unperturbed result $(\mathcal{Z} - E)^{-1}$ comes from the next term. Accordingly we shall take as the lowest-order perturbation solution

$$\begin{aligned} S_{\mathcal{N}}(\mathcal{Z}, \mathcal{N}) &= (\mathcal{Z} - E)^{-1}\left\{1 + \langle \mathcal{N}|\left[\frac{H_1}{\hbar}\left(\mathcal{Z} - \frac{H_0}{\hbar}\right)^{-1}\right]^2|\mathcal{N}\rangle\right\} \\ &= (\mathcal{Z} - E)^{-1}\left(1 + (\mathcal{Z} - E)^{-1}\sum_{\lambda}\frac{2|h_\lambda|^2}{\hbar^2(\mathcal{Z} - \omega_\lambda)}\right) \\ &= (\mathcal{Z} - E)^{-1}[1 - 2\alpha E^{1+p}(\mathcal{Z} - E)^{-1}\mathcal{Z}^{-p}e^{ip\pi}\csc(p\pi)] \end{aligned}$$

by use of Eq. (19) for the infinite system. From this expression is obtained easily

$$e^{iEt}\langle \mathcal{N} \mid \Psi(\tau)\rangle = 1 + \frac{e^{ip\pi}\csc(p\pi)}{\pi i}\int_c d\xi\frac{e^{-i\xi\tau}}{\xi^2(\alpha\xi + 1)^p}. \tag{26}$$

The integral in the right-hand side can be evaluated by closing the contour by a large semicircle in the lower half-plane of ξ, and then it becomes the sum of two terms, one from the residue of the double pole at $\xi = 0$ and the other from the integral of the discontinuity of the integrand across a cut from $\xi = -1/\alpha$ to $\xi = -\infty$ along the negative real axis. With

this, Eq. (26) yields

$$e^{iEt}\langle \mathcal{N} \mid \Psi(\tau)\rangle = 1 + e^{ip\pi}\Big[2i\tau \csc(p\pi) + 2\alpha p \csc(p\pi) - \frac{2}{\pi}\int_0^\alpha d\xi e^{i\tau/\xi}\Big(\frac{\xi}{\alpha-\xi}\Big)^p\Big]. \quad (27)$$

The cut integral is now indeed seen to furnish a contribution which is nonanalytic in α at $\alpha = 0$ and so might be expected to yield a nonexponential contribution to $\rho(\tau)$ similar to that in Eq. (I-62). However, although the present perturbational scheme has given rise to a nonanalyticity, it is of a different kind from that which arose in I. First, the terms on the right-hand side of the expression (27) other than the cut integral are not exponential. It is apparent from the nature of the expansion (24) that no finite order of that perturbation can yield an exponential term of the sort obtained in Eq. (22); rather only the *sum* of the contributions from the poles at $\xi = 0$ which occur in each successive term of the expansion (25) will give rise to an exponential contribution. This is the price paid for obtaining the cut integral in Eq. (27). On the other hand, from the structure of the remaining terms of that equation it is clear that the final exponential term will have the form

$$[1 + 2\alpha p e^{ip\pi}\csc(p\pi) + O(\alpha^2)] \times \exp[2i\tau \csc(p\pi)e^{ip\pi} + O(\alpha)],$$

which yields, for $p = 0$, the same contribution, namely

$$\Big(1 + \frac{4\alpha}{\pi}\Big)e^{-4\tau},$$

as is obtained in Eq. (I-62) for this order from the statistical-mechanical treatment. Second, if the square of the modulus of the right-hand side of Eq. (27) is taken, the nonanalytic contribution to $\rho(\tau)$ which results from the cut integral is not at all of the same form as the nonexponential term in Eq. (I-62). Thus, although the present analysis has shown considerable similarities to that of I, the results obtained are essentially different. The two perturbational schemes, while each gives rise to nonanalytic terms, yield expressions for $\rho(\tau)$ which cannot be considered as comparable.

IV. THE FINITE SYSTEM

We shall now return to the general solution, Eq. (12), and consider the weak-coupling dynamics when L is kept at a finite value. This entails an evaluation of the infinite series (18). It is shown in the Appendix that the following result holds: For Re $a < 0$, Im < 0,

$$\sum_{n=0}^{\infty}[n^p(n+a)]^{-1} = \pi(-a)^{-p}e^{ip\pi}\csc(p\pi) + 2\pi i(-a)^{-p}/(e^{2\pi i a}-1) + 2\int_0^\infty dy[y^{p-1}(e^{2\pi y}-1)(a^2+y^2)]^{-1} \times \{\cos(p\pi/2) + (y\pi a)^{-1}\sin(p\pi/2) \times [2\pi a^2 - y^{-1}(e^{2\pi y}-1)(a^2+y^2)]\}. \quad (28)$$

From this, the series (18) can be calculated. First, it is convenient at this stage to rewrite Eq. (18) in terms of the dimensionless scaled variables which facilitate the taking of the weak-coupling limit. Not only should one use

$$\xi = \mathcal{Z}/\alpha E$$

[see Eq. (20)], but, as has been seen in II, the length L should appear through the variable

$$\sigma^2 = \alpha EL/c.$$

The series then becomes

$$-\frac{4\alpha^{1-p}E\sigma^{2p}}{(2\pi)^{p+1}}\sum_{n=1}^{\infty}\Big[n^p\Big(n - \frac{\sigma^2\xi}{2\pi}\Big)\Big]^{-1}. \quad (29)$$

The $\alpha \to 0$ limit in Eq. (12) requires, just as in the infinite system calculation, that the factor e^{-iEt} be divided out, and so it is not the expression (29) itself which will be needed, but rather (29) with ξ replaced by $\xi + 1/\alpha$, so as to obtain the result corresponding to Eq. (21). This leads to the result

$$S(\xi) \equiv \alpha^{-p}\sum_{n=1}^{\infty}\Big\{n^p\Big[n - \frac{\sigma^2}{2\pi}(\xi + 1/\alpha)\Big]\Big\}^{-1} \quad (30)$$

$$= \pi[\sigma^2(\alpha\xi+1)/2\pi]^{-p}e^{ip\pi}\csc p\pi + 2\pi i[\sigma^2(\alpha\xi+1)/2\pi]^{-p}(e^{-[\sigma^2(\xi+1/\alpha)]}-1)^{-1} + 2\int_0^\infty dy\alpha^{1-p}\Big\{y^{p-1}(e^{2\pi y}-1) \times \Big[\alpha^2y^2 + \frac{\sigma^4}{4\pi^2}(\alpha\xi+1)^2\Big]\Big\}^{-1} \times \Big\{\alpha\cos\frac{p\pi}{2} - \frac{2}{y\sigma^2(\alpha\xi+1)}\sin\Big(\frac{p\pi}{2}\Big) \times \Big[\frac{\sigma^4}{2\pi}(\alpha\xi+1)^2 - y^{-1}(e^{2\pi y}-1) \times \Big(\alpha^2y^2 + \frac{\sigma^4}{4\pi^2}(\alpha\xi+1)^2\Big)\Big]\Big\} \quad (31)$$

for Im $\xi > 0$, Re $\xi > -1/\alpha$. From Eq. (12), we obtain

$$e^{iEt}\langle \mathcal{N} \mid \Psi(\tau)\rangle = -\frac{1}{2\pi i}\int_c d\xi e^{-i\xi\tau}\Big(\xi + \frac{4\sigma^{2p}}{(2\pi)^{p+1}}S(\xi)\Big)^{-1} \quad (32)$$

This is the exact solution of the Schrödinger equation. At this point, we may let σ^2 become infinite and verify that Eq. (21) is recovered:

$$\lim_{\sigma^2\to\infty}[4\sigma^{2p}/(2\pi)^{p+1}]S(\xi) = 2e^{ip\pi}\csc p\pi(\alpha\xi+1)^{-p}$$

since, for Im $\xi > 0$, $e^{-i\sigma^2(\xi+1/\alpha)} \to \infty$ as $\sigma^2 \to \infty$, and in Eq. (31), the integrand is $O(\sigma^{-2})$. Now we may take the limit $\alpha \to 0$ in Eq. (31). It can be remarked that this limit could not be taken at an earlier stage, since it is evident from Eq. (30) that the limit cannot be defined without explicitly performing the summation. All that is needed is

$$\begin{aligned}\lim_{\alpha\to 0} S(\xi) &= \pi(\sigma^2/2\pi)^{-p}e^{ip\pi}\csc p\pi \\ &\quad + 2\pi i(\sigma^2/2\pi)^{-p}(e^{-i\sigma^2\xi}e^{-i\sigma^2/\alpha}-1)^{-1} \\ &= \pi(\sigma^2/2\pi)^{-p}\left(\cot p\pi + i\frac{e^{-i\sigma^2(\xi+1/\alpha)}+1}{e^{-i\sigma^2(\xi+1/\alpha)}-1}\right) \\ &= \pi\left(\frac{\sigma^2}{2\pi}\right)^{-p}\left[\cot p\pi - \cot\frac{\sigma^2}{2}(\xi+1/\alpha)\right].\end{aligned}$$

Again, as in the calculation of II, a nonanalyticity in α has appeared for the finite system even in the weak-coupling limit. The weak-coupling solution, from Eq. (32), is then

$$\begin{aligned}&e^{iEt}\langle\mathcal{N}\,|\,\Psi(\tau)\rangle \\ &\quad = -\frac{1}{2\pi i}\int_c d\xi \\ &\qquad \times e^{-i\xi\tau}\left\{\xi + 2\left[\cot p\pi - \cot\frac{\sigma^2}{2}\left(\xi+\frac{1}{\alpha}\right)\right]\right\}^{-1}.\end{aligned}$$

It is convenient to remove the term $2\cot p\pi$ from the denominator of the integrand by extracting a further exponential factor of $e^{-2i\tau\cot p\pi}$. This yields the result

$$\begin{aligned}&e^{iEt}e^{-2i\tau\cot p\pi}\langle\mathcal{N}\,|\,\Psi(\tau)\rangle \\ &= -\frac{1}{2\pi i}\int_c d\xi \\ &\quad \times e^{-i\xi\tau}\left[\xi - 2\cot\frac{\sigma^2}{2}\left(\xi - 2\cot p\pi + \frac{1}{\alpha}\right)\right]^{-1}. \quad (33)\end{aligned}$$

To evaluate the integral in this expression, the contour C can be closed by a large semicircle in the lower half-plane of ξ. The singularities of the integrand within the closed contour are then the zeros of the expression

$$\xi - 2\cos(\sigma^2/2)(\xi - 2\cot p\pi + 1/\alpha), \quad (34)$$

and these may be seen to be distributed along the real axis so that there is one in each interval of length $2\pi/\sigma^2$. The effect of the nonanalyticity (through the term $1/\alpha$ in the argument of the cotangent) and that of p (through the term $2\cot p\pi$) can be grouped into some phase angle, φ say, to be determined in a numerical calculation by the chosen values of p and α, but for theoretical purposes essentially arbitrary. The expression (34) will now be written simply as

$$\xi - 2\cot(\sigma^2/2)(\xi+\varphi).$$

If a zero of this be denoted by ξ_q, then the residue of the integrand in Eq. (33) at ξ_q is

$$\begin{aligned}&e^{-i\xi_q\tau}\left[\frac{d}{d\xi}\left(\xi - 2\cot\frac{\sigma^2}{2}(\xi+\varphi)\right)\Big|_{\xi=\xi_q}\right]^{-1} \\ &\quad = e^{-i\xi_q\tau}\{1+\sigma^2[1+\cot^2(\sigma^2/2)(\xi_q+\varphi)]\}^{-1} \\ &\quad = e^{-i\xi_q\tau}[1+\sigma^2(1+\tfrac{1}{4}\xi_q^2)]^{-1}\end{aligned}$$

by the condition that ξ_q be a zero. The expression (33) can thus be written

$$\sum_{\xi_q}\frac{e^{-i\xi_q\tau}}{1+\sigma^2(1+\frac{1}{4}\xi_q^2)}, \quad (35)$$

where the summation runs over all the zeros ξ_q of Eq. (34). In the limit $\sigma^2 \to \infty$, this expression can be evaluated by noticing that there is one ξ_q in each interval of length $2\pi/\sigma^2$, and thus replacing Eq. (35) by the Riemann integral

$$\begin{aligned}&\lim_{\sigma^2\to\infty}\frac{\sigma^2}{2\pi}\int_{-\infty}^{+\infty}d\xi e^{-i\xi\tau}[1+\sigma^2(1+\tfrac{1}{4}\xi^2)]^{-1} \\ &\qquad = \frac{2}{\pi}\int_{-\infty}^{+\infty}d\xi\frac{e^{-i\xi\tau}}{\xi^2+4} = e^{-2\tau}.\end{aligned}$$

Thus, in the limit,

$$\begin{aligned}e^{iEt}\langle\mathcal{N}\,|\,\Psi(\tau)\rangle &= e^{-2\tau}e^{2i\tau\cot p\pi} \\ &= \exp(2i\tau\csc p\pi e^{ip\pi})\end{aligned}$$

in accord with the result stated in Eq. (22). We shall be interested, as in the preceding section, in

$$\rho(\tau) = |\langle\mathcal{N}\,|\,\Psi(\tau)\rangle|^2.$$

From Eq. (35) there results

$$\begin{aligned}\rho(\tau) &= \left|\sum_{\xi_q}\frac{e^{-i\xi_q\tau}}{1+\sigma^2(1+\frac{1}{4}\xi_q^2)}\right|^2 \\ &= \sum_{\xi_q}[1+\sigma^2(1+\tfrac{1}{4}\xi_q^2)]^{-1} \\ &\quad + \sum_{\xi_q>\xi_{q'}}\frac{2\cos(\xi_q-\xi_{q'})\tau}{[1+\sigma^2(1+\frac{1}{4}\xi_q^2)][1+\sigma^2(1+\frac{1}{4}\xi_{q'}^2)]} \\ &= \rho_c + \rho_T \quad (36)\end{aligned}$$

say, in notation analogous to that of II, Sec. VII, for the two contributions to $\rho(\tau)$. The former, ρ_c, is independent of time and equal to the time average of $\rho(\tau)$, and the latter, ρ_T, contains all information on the time dependence of the system. It can be seen at once that the result (36) has quite different structure from that of Eq. (II–40), the corresponding result

obtained from the Liouville–von Neumann equation. Firstly, the value of the right-hand side of Eq. (36) can never become negative, by virtue of the method of calculation. Secondly, the frequencies $\xi_q - \xi_{q'}$ which appear in ρ_T are by no means linearly independent, although the ξ_q, taken separately, are so, being the solutions of the transcendental equation

$$\xi - 2\cot\frac{\sigma^2}{2}(\xi + \varphi) = 0. \tag{37}$$

However, both results, Eqs. (36) and (II–40), give quasiperiodic expressions for $\rho(\tau)$, as one would expect.

An analysis similar to that of Sec. VI of II can be performed for Eq. (36). By replacing the ξ_q by $2n\pi/\sigma^2$, with n ranging over the integers, an asymptotic expression for $\rho(\tau)$ can be obtained for large (but *not* infinite) σ^2. First, for ρ_c there results from Eq. (36)

$$\begin{aligned}\rho_c &\sim \sum_{n=-\infty}^{+\infty}(n^2\pi^2/\sigma^2 + \sigma^2)^{-2} \\ &= \frac{\sigma^4}{\pi^4}\sum_{n=-\infty}^{+\infty}[n^2 + (\sigma^2/\pi)^2]^{-1} \\ &= \frac{1}{2\sigma^2}\coth\sigma^2 + \tfrac{1}{2}\operatorname{csch}^2\sigma^2. \end{aligned}\tag{38}$$

This result has to be compared with Eq. (II–35), where it is found that

$$\rho_c \sim (1 + 2\sigma\tanh 2\sigma)^{-1} \tag{39}$$

is obtained from Eq. (II–40) under similar assumptions. It is seen that the two expressions, Eqs. (38) and (39), are intrinsically different, in that Eq. (38) depends only on σ^2, while Eq. (39) depends only on σ. Even in the asymptotic region, then, the two approaches leading respectively to Eqs. (36) and (II–40) are not coincident. Next, for ρ_T, one obtains

$$\rho_T(\tau) \sim \sum_{n=-\infty}^{+\infty}{\sum_{m=-\infty}^{+\infty}}' \frac{\cos(2\pi/\sigma^2)(n-m)\tau}{(\sigma^2 + n^2\pi^2/\sigma^2)(\sigma^2 + m^2\pi^2/\sigma^2)},$$

where the prime indicates that the terms in which $n = m$ are to be omitted from the double sum. On simplification, this sum becomes

$$\begin{aligned}\rho_T(\tau) &\sim \sigma^4\left(\sum_{n=-\infty}^{\infty}\frac{\cos(2\pi n\tau/\sigma^2)}{n^2\pi^2 + \sigma^4}\right)^2 \\ &\quad - \sigma^4\sum_{n=-\infty}^{\infty}(n^2\pi^2 + \sigma^4)^{-2} \qquad (40) \\ &= \operatorname{csch}^2\sigma^2\cosh^2(\sigma^2 - 2\tau) - \tfrac{1}{2}\operatorname{csch}^2\sigma^2 \\ &\quad - (1/2\sigma^2)\coth\sigma^2 \quad \text{for} \quad \tau < \sigma^2. \qquad (41)\end{aligned}$$

This closed expression has no analog in the treatment of II, where the summations in Eq. (II–41) cannot be performed explicitly. When Eqs. (38) and (41) are grouped together, the full result is obtained,

$$\rho_{as}(\tau) = \operatorname{csch}^2\sigma^2\cosh^2(\sigma^2 - 2\tau) \quad \text{for} \quad \tau < \sigma^2, \tag{42}$$

and where $\tau > \sigma^2$ the answer comes from observing that the summation in the first term of the right-hand side of Eq. (40) is periodic in τ with period σ^2. The asymptotic expression (42) thus has none of the more intricate structure of the solution, Eq. (36), and is merely a periodic function which is graphed in Fig. 1.

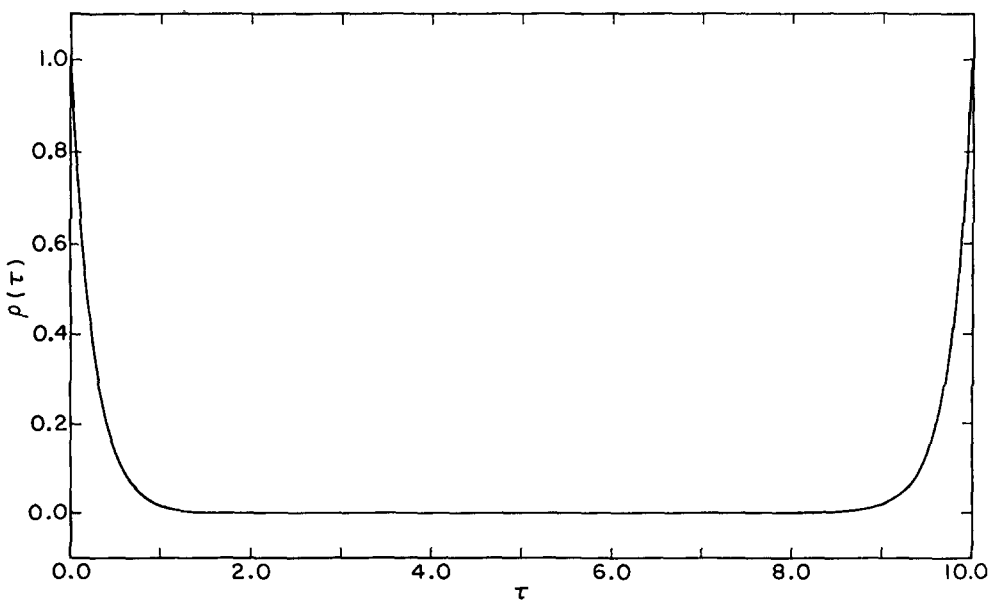

FIG. 1. A plot of the asymptotic expression, Eq. (42), for $\sigma^2 = 10$. The intercept at $\tau = 0$ is given by $\coth^2\sigma^2 = 1$, and the value of $\rho_{as}(\tau)$ at $\tau = \sigma^2/2$ is $\operatorname{csch}^2\sigma^2 \simeq 0$.

It is quickly seen that when $\sigma^2 \to \infty$

$$\rho_{as}(\tau) \to e^{-4\tau}.$$

We shall now return our attention to Eq. (36) itself. A numerical study of $\rho(\tau)$ has been made, based on this equation. For the purposes of comparison with the results presented in II, two values have been chosen for σ^2, namely 1 and 10, and the function $\rho(\tau)$ calculated for these values over a considerable range of τ. It was necessary first of all to determine the locations of the solutions ξ_q of Eq. (37). This equation can be rewritten

$$\xi \sin (\sigma^2/2)(\xi + \varphi) = 2 \cos (\sigma^2/2)(\xi + \varphi),$$

so that by calculating the left- and right-hand sides of this separately and noting where they coincide, numerical estimates of the ξ_q can be found. For the case $\sigma^2 = 1$, the first 82 solutions on either side of the origin $\xi = 0$ were obtained, and for $\sigma^2 = 10$ the first 404 solutions were obtained. An upper bound for the truncation error committed by restricting oneself to these numbers of terms in the summations of Eq. (36) can be obtained by noting the defect in $\rho(0)$ from unity. When $\tau = 0$ all the cosines in τ_T have argument zero, and so the truncated higher terms in the sums have their maximum value. It is found that for $\rho^2 = 1$, with 82 poles on either side of the origin taken into account, that $\rho(0) = 0.9900$ which corresponds to an error of 1%. For $\sigma^2 = 10$, with 404 poles taken into account, $\rho(0) = 0.9800$ which corresponds to an error of 2%. Since for positive values of τ the arguments of the cosines will never be exactly in phase, the truncation errors can be expected to be much smaller after the initial moment. A further check on these errors is found by estimating the higher terms in Eq. (36) by their approximating integral. This means that, for $\sigma^2 = 1$, the contribution to $\rho(0)$ from higher terms above the first 82 on either side of $\xi = 0$ is approximately

$$\frac{4}{2\pi} \int_{251}^{\infty} \frac{d\xi}{2 + \frac{1}{4}\xi^2} \sim 0.010.$$

For $\sigma^2 = 10$, the contributions from the terms other than the first 404 on either side are

$$\frac{40}{2\pi} \int_{122}^{\infty} \frac{d\xi}{11 + \frac{5}{2}\xi^2} \sim 0.020.$$

This check thus yields the same error estimates as above. Finally, the phase angle φ was chosen as zero for most of the work, but checks were made with $\varphi = \pi/2$ for both $\sigma^2 = 1$ and $\rho^2 = 10$ to see if the choice of φ made an appreciable difference to the results when ρ_e and $\rho_T(\tau)$ were calculated from Eq. (36).

In Fig. 2 we have plotted the dependence of $\rho_T(\tau)$ on τ for $\sigma^2 = 10$ and for a phase $\varphi = 0$. A calculation for $\sigma^2 = 10$ with a phase $\varphi = \pi/2$ yielded results which were indistinguishable from those plotted in Fig. 2. On the other hand, it was found that for $\sigma^2 = 1$, a change in phase from $\varphi = 0$ to $\varphi = \pi/2$ did produce a change in the profile of the $\rho_T(\tau)$ vs τ curve. In Fig. 3 $\rho_T(\tau)$ vs τ is plotted for $\sigma^2 = 1$ and $\varphi = 0$, and in Fig. 4 for $\sigma^2 = 1$ and $\varphi = \pi/2$. These results will be discussed in Sec. VI.

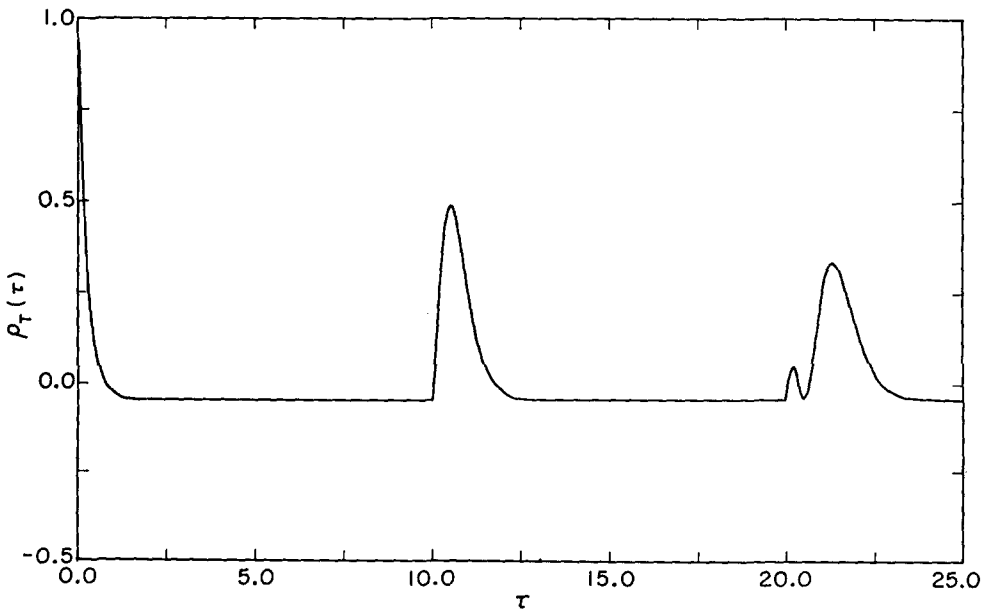

FIG. 2. A plot of $\rho_T(\tau)$ vs τ computed for $\sigma^2 = 10$ and $\varphi = 0$ using 404 pole locations.

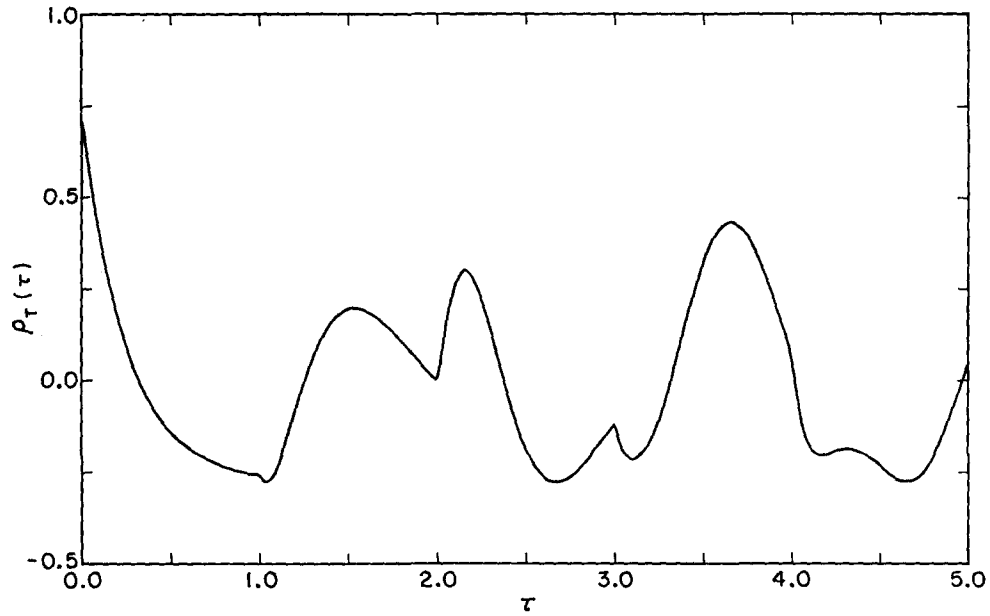

FIG. 3. A plot of $\rho_T(\tau)$ vs τ computed for $\sigma^2 = 1$ and $\varphi = 0$ using 82 pole locations.

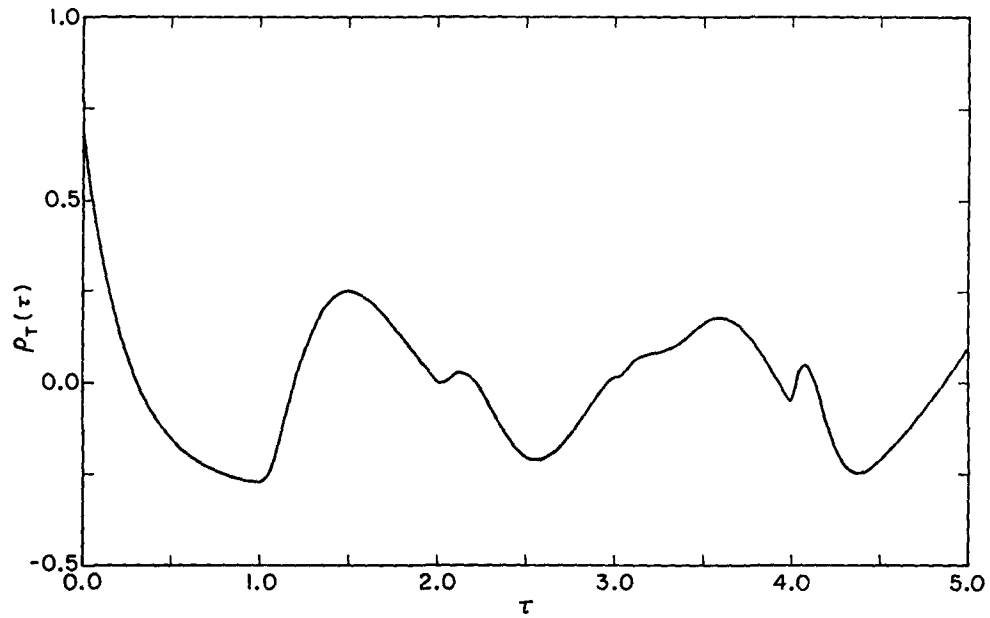

FIG. 4. A plot of $\rho_T(\tau)$ vs τ computed for $\sigma^2 = 1$ and $\varphi = \pi/2$ using 45 pole locations.

V. PROPERTIES OF THE SOLUTIONS FOR FINITE SYSTEMS

Two expressions have now been obtained for $\rho(\tau)$ in the region of weak coupling—the statistical-mechanical one, $\rho_L(\tau)$ say, given by Eq. (II-40), and the quantum-mechanical one, $\rho_S(\tau)$ say, given by Eq. (36). For convenience, Eq. (II-40) will be rewritten here:

$$\rho_L(\tau) = \left(1 + \frac{2\sigma \sinh 2\sigma}{\cosh 2\sigma - \epsilon}\right)^{-1} + \sum_{n=1}^{\infty} \frac{2\theta_n^2}{\theta_n^2 + \frac{4}{\sigma^2}} \cos\left[\tau\left(\theta_n^2 + \frac{4}{\sigma^2}\right)^{\frac{1}{2}}\right]$$

$$\times \left[1 + (-1)^n \frac{\sigma^2}{(1-\epsilon^2)^{\frac{1}{2}}}\left(\theta_n^2 + \frac{16}{1-\epsilon^2}\right)^{\frac{1}{2}} + \frac{\sigma^2}{4}\left(\theta_n^2 + \frac{16}{1-\epsilon^2}\right)\right]^{-1}, \tag{43}$$

where ϵ is a nonanalyticity parameter [equal to $\cos(\sigma^2/\alpha)$] and the θ_n are the successive positive roots of the transcendental equation (II-36),

$$4 \sin(\sigma^2\theta) = \theta[\epsilon - \cos(\sigma^2\theta)].$$

It is of importance in the study of these two solutions to have a means of characterizing them so that the differences and similarities can be seen. The means chosen here will be to obtain the average frequency $L(q)$ with which the solution $\rho(\tau)$ achieves the value q. For several reasons this is a convenient characterization. The occurrence of negative values of $\rho_L(\tau)$ will be evidenced by $L(q)$ remaining nonzero for negative q, whereas, for $\rho_S(\tau)$, only values of q between 0 and 1 need be considered. Since, particularly for large τ, the details of the time dependence of both $\rho_L(\tau)$ and $\rho_S(\tau)$ are at once intricate and largely uninteresting, the function $L(q)$ will contain the relevant information for a meaningful comparison of the two solutions. The matter of Poincaré recurrences has already been brought up in II, and it is by means of $L(q)$ that a quantitative discussion of these can be made. Clearly, expressions like Eq. (36) and Eq. (43), involving as they do linearly independent sets of frequencies, will never *exactly* regain their initial value of 1, for which all the phases would need to be integral multiples of 2π, but they may approach this value to within any arbitrarily small number. Lastly, although in the case of $\rho_S(\tau)$, $L(q)$ can be determined only by plotting the function $\rho_S(\tau)$ and counting the number of times its graph intersects a given ordinate over a fixed range of τ, in the study of $\rho_L(\tau)$ there exists another independent (and less tedious) procedure for computing $L(q)$. Use may be made of a theorem of Kac[5] which deals with the average frequency with which a sum of cosines of the form

$$f(t) = \sum_{n=1}^{N} a_n \cos(\omega_n t) \tag{44}$$

achieves a value q. The coefficients a_n are required to be real, and the frequencies ω_n are required to be linearly independent. [It is for the latter reason that the Kac theorem is not applicable to Eq. (36).] The use of the theorem was suggested by the work of Mazur and Montroll[5] who discussed Poincaré recurrences in assemblies of coupled harmonic oscillators; these authors also provided a simplified proof of the theorem. The result is that the mean frequency with which the expression (44) has the value q is given by

$$L(q) = \frac{1}{2\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y^{-2} \cos(qx) \left\{ \prod_{n=1}^{N} J_0(|a_n| x) - \prod_{n=1}^{N} J_0(|a_n| [x^2 + y^2\omega_n^2]^{\frac{1}{2}}) \right\} dx\, dy. \tag{45}$$

Here J_0 is the Bessel function of order zero. The sum of cosines in Eq. (44) is of course an infinite one, and so strictly the theorem is not applicable to it. But since as $\theta_n \to \infty$ the coefficients a_n of the cosines decrease like θ_n^{-2}, it is possible to truncate the sum after some properly chosen finite number of terms, N say. The maximum error which is committed by this procedure can be made as small as one likes by increasing N, and in the numerical estimation of Eq. (43) exactly the same approximation is made.

In Fig. 5 we have plotted $L(q)$ vs q for $\sigma^2 = 10$ for the time-dependent parts ρ_T of $\rho_L(\tau)$ (solid line) and $\rho_S(\tau)$ (dotted line). (The constant parts ρ_c are given in the caption.) In order to obtain an accurate estimate of the dependence of $L(q)$ on q, both $\rho_S(\tau)$ and $\rho_L(\tau)$ were computed at intervals of 0.1 of τ for values in the range $0 \leq \tau \leq 400$. It was found that $\rho(\tau)$ became completely "dephased" for values of $\tau > 100$, and since it is the long-time behavior of the function $\rho(\tau)$ that we wish to study, values of $L(q)$ were determined by a direct counting procedure for the range $100 \leq \tau \leq 400$. Although, in principle, values of $L(q)$ vs q computed using the counting procedure should be determined from a study of $\rho(\tau)$ vs τ for all τ, such a program is obviously impossible. In this and in the following cases where the counting procedure is used, a range of τ is chosen so that the smaller, more sensitive values of q occur about 300 times: Such a range in the present case $\sigma^2 = 10$ is $100 \leq \tau \leq 400$.

As mentioned in Sec. IV, although the choice of phase was immaterial for $\sigma^2 = 10$ in that exactly the same curve was generated for $\varphi = 0$ and $\varphi = \pi/2$, for the case $\sigma^2 = 1$ this is not so, as may be seen from Figs. 3 and 4. It is of some interest then to examine whether this difference in phases is significant in terms of the long-time behavior of $\rho_S(\tau)$. A study of $L(q)$ is well suited to an examination of this point, and in Fig. 6 we have plotted it for $\sigma^2 = 1$ and for the two phases $\varphi = 0$ and $\varphi = \pi/2$. To construct these plots $\rho_S(\tau)$ was calculated at intervals of 0.02 of τ for the range $0 \leq \tau \leq 150$. Since, as can be seen from an examination of Figs. 3 and 4, $\rho(\tau)$ dephases almost at once, the entire range of τ was considered in counting up the number of times a given q occurred. The results given in Fig. 6 reveal some difference in the long-time behavior of $\rho_S(\tau)$ for $\varphi = 0$ and $\varphi = \pi/2$. A more interesting feature here, however, is the appearance in both these plots of two rather well-defined maxima

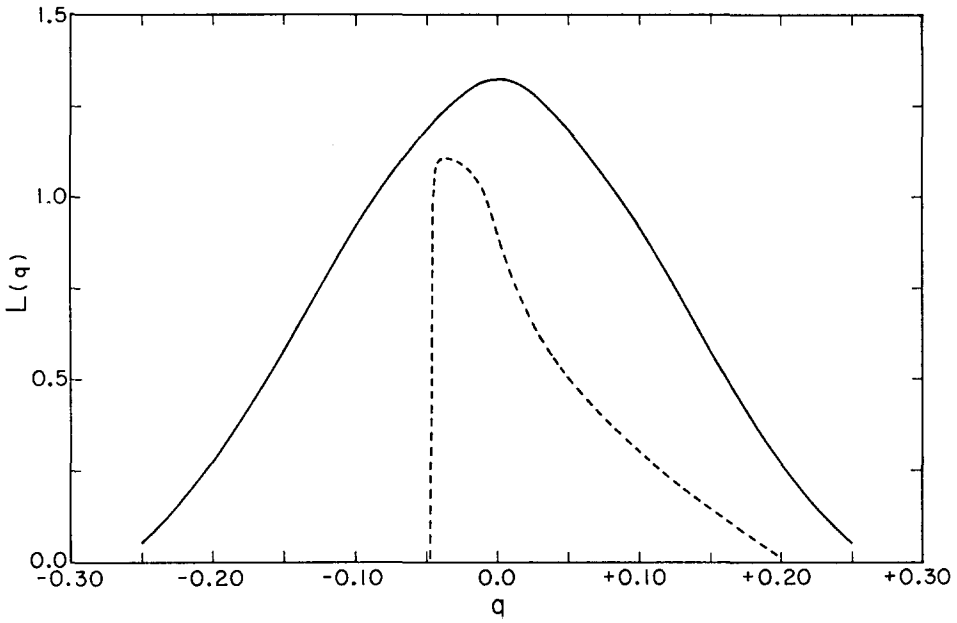

FIG. 5. A plot of $L(q)$ vs q for $\sigma^2 = 10$. The solid line was constructed from an analysis of the time-dependent part ρ_T of $\rho_L(\tau)$ using the counting procedure discussed in the text. For completeness, we note that for $\rho_L(\tau)$, $\rho_C = 0.1362$, and $\rho_T(0) = 0.8504$. The dotted line was constructed from a similar study of the time-dependent part ρ_T of $\rho_S(\tau)$. Here, $\rho_C = 0.0465$ and $\rho_T(0) = 0.9335$.

disposed on either side of the origin $q = 0$. These are not of equal amplitude, the one for positive q being larger. This behavior is so different from that observed in an examination of the long-time behavior of either $\rho_S(\tau)$ or $\rho_L(\tau)$ for $\sigma^2 = 10$ that a detailed study of $L(q)$ corresponding to $\rho_L(\tau)$ for $\sigma^2 = 1$ was thought necessary. A study of this case is particularly instructive since, in addition to the direct-counting method employed in determining the results presented above, it is practical for this case to make use of the Kac formula, Eq. (45). In Fig. 7 we plotted $L(q)$ for $\rho_L(\tau)$ for $\sigma^2 = 1$. The full curve corresponds to the dependence of $L(q)$ on q determined by the straightforward method of counting. The circles correspond to values of $L(q)$ computed using the Kac formula. It is necessary at this point to discuss briefly the details of the numerical calculation. First of all, in constructing Fig. 7, only the first 20 terms in Eq. (43) were computed, as opposed to the 190 that were used

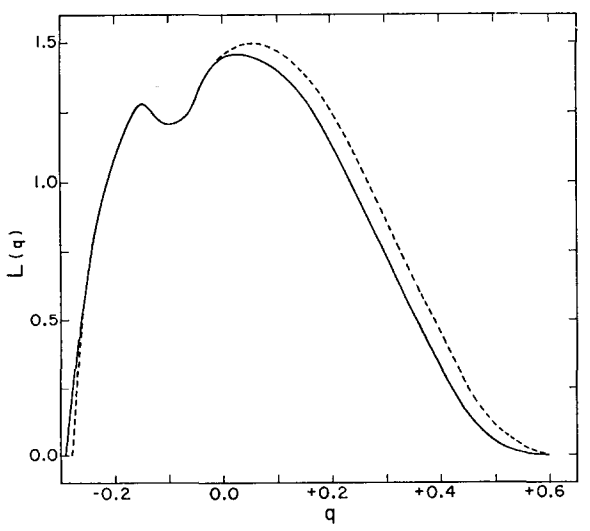

FIG. 6. A plot of $L(q)$ versus q for $\sigma^2 = 1$, constructed from an analysis of the time-dependent part ρ_T of $\rho_S(\tau)$, using the counting procedure. The dotted line refers to a phase $\varphi = 0$; for this choice of phase $\rho_C = 0.2784$ and $\rho_T(0) = 0.7116$. The solid line refers to a phase $\varphi = \pi/2$; for this phase, $\rho_C = 0.2930$ and $\rho_T(0) = 0.6886$.

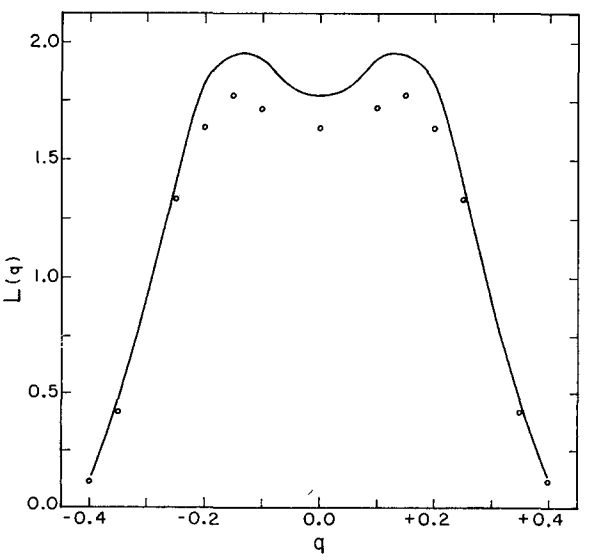

FIG. 7. A plot of $L(q)$ vs q for $\sigma^2 = 1$. Here, the solid line describes the dependence of $L(q)$ on q as determined from an analysis of the time-dependent part ρ_T of $\rho_L(\tau)$ using the counting procedure. The circles refer to the values of $L(q)$ computed using the Kac formula. We note that for 20 pole locations, $\rho_C = 0.3881$ and $\rho_T(0) = 0.5715$.

in constructing Fig. 7 of II. The reason for this limitation has to do with the amount of computer time required to calculate values of $L(q)$ accurately using Eq. (45). It is to be emphasized here, however, that the use of a smaller number of terms in this case is not a serious limitation, since the values of $\rho_L(\tau)$ generated were not significantly different from those computed using 190 terms: The most serious discrepancy appears to be at $\tau = 0$, where for 190 terms we found

$$\rho_L(0) = 0.6076,$$

whereas in the present case for 20 terms we have

$$\rho_T(0) = 0.5715,$$

an error of less than 6%. The values of $L(q)$ predicted by the Kac formula were accordingly computed and compared with those obtained by direct counting for exactly 20 terms. But since it takes somewhat over one hour of computer time per point to determine a value of $L(q)$ that does not change by more than 1% when the area of integration is increased fivefold, it is thus out of the question to compute, using Eq. (45), values of $L(q)$ for $\rho_L(\tau)$ with $\sigma^2 = 10$. There one requires far more than 20 terms to approximate the behavior of $\rho_L(\tau)$ calculated using the full complement of 600 terms as was done in constructing Fig. 8 of II. Lastly, we note that although the limits of integration on the Kac formula are $(-\infty, +\infty)$, a careful error analysis of the values calculated from Eq. (45) revealed that these limits, for our purposes, could be replaced by $(-80, +80)$ without committing an error of more than 1%, and this was done for a grid spacing of 0.1. As regards the counting procedure, we have, as in our study of $L_S(q)$ for $\sigma^2 = 1$, determined values of $L_L(q)$ from the time evolution of $\rho_L(\tau)$ in the range $0 \leq \tau \leq 150$. Here again, one must expect some error since, as mentioned above, only 300 occurrences of the sensitive values of q were used to determine each value of $L_L(q)$. Given the possible errors that can arise in calculations based both on Eq. (45) and on the counting procedure, it is worth noting that the maximum discrepancy between the values of $L(q)$ determined using the two procedures is 11%. Furthermore, despite the slight disagreement in the two results, both calculations show $L_L(q)$ to have *two* maxima disposed on either side of the origin $q = 0$. These maxima, unlike those of $L_S(q)$, are of equal amplitude and are symmetrically disposed about $q = 0$. Finally, examination of Figs. 6 and 7 reveals that, for $\sigma^2 = 1$ as for $\sigma^2 = 10$, $L_S(q) < L_L(q)$ for all q.

VI. DISCUSSION AND CONCLUSIONS

In this paper, a detailed comparison has been made between two essentially different "weak-coupling approximations" to the spontaneous emission of the Wigner–Weisskopf atom, both in an infinite system and in a finite one. The difficulties which had arisen in II in the treatment based on the Liouville equation, principally the occurrence of negative values of $\rho(\tau)$, can indeed be removed by proceeding instead from the Schrödinger equation, but the nature of the weak-coupling limit of the new solution is also far from clear. Both results, $\rho_L(\tau)$ and $\rho_S(\tau)$, have one important feature in common however: The limit $\alpha \to 0$ gives rise to a nonanalyticity—that is, except under certain circumstances, the limit is not well defined. The exceptional cases are of course that where the thermodynamic limit is taken, thereby suppressing any dependence on a nonanalyticity parameter, and that of the asymptotic solution for large but not infinite systems, Eq. (42). Otherwise, although numerical results are readily obtainable, different solutions to the same problem appear, depending on the approach chosen for the calculation. Especially as σ^2 becomes small, near unity, the influence of nonanalyticity parameters, the φ of this paper or the ϵ of II [see Eq. (43)], becomes significant, and so the meaningfulness of the approximation scheme is very dubious in this régime. This matter has been discussed in II, where other reasons are given to doubt the possibility of a weak-coupling approach for small σ^2. Perhaps, then, this point should be passed over for present purposes. But it is by no means the whole problem.

The solutions $\rho_L(\tau)$ and $\rho_S(\tau)$ obtained here and in II, despite the numerical similarities remarked earlier, are essentially different in almost *all* of their aspects. Which of them provides a better approximation to the exact dynamics prescribed by the Hamiltonian, Eq. (1), cannot be seen at once. The statistical-mechanical result, $\rho_L(\tau)$, certainly has more structure than $\rho_S(\tau)$, and may thus seem to provide more detailed information about the system's evolution. But it has features which cannot accord with the exact solution. The negative values not only defy the probabilistic interpretation of the density matrix, but they indicate that this object, ρ say, does not evolve under a unitary time-development operator. Further, an initial configuration $\rho(0)$ which corresponds to a pure quantum state (and not an ensemble of different states averaged) does not remain so; and the Gibbs entropy

$$\mathrm{Tr}\,[\rho(t) \log \rho(t)]$$

is not a constant of the motion. These properties are

of course incompatible with the exact Liouville–von Neumann equation, and by necessity with *any* description, however otherwise inadequate, based on a state function rather than a density matrix.

It seems to us that the only clear way to resolve the difficulties brought up here is to compare both $\rho_S(\tau)$ and $\rho_L(\tau)$ with the *exact* solution of the quantum-mechanical problem. This solution has been obtained in this paper in Eq. (32), but its properties have not yet been examined either analytically or numerically. It will be the task of the next paper in this series to perform such an investigation, and one also of the simpler infinite-system solution, Eq. (21). In this way, we hope not only to cast light on the question raised here regarding weak-coupling approximations and their relative merits, but also to gain a better understanding of the thermodynamic limit in the context of an irreversible model where an exact solution, however complicated, is known, and thereby of the formal processes of the statistical-mechanical theory based on the generalized master equation [Eq. (I-24)]. In principle, the way is open also to an examination of our model in the light of ergodic theory.

ACKNOWLEDGMENTS

One of us (R. D.) acknowledges financial support from the Robert C. Welch Foundation. The other (J. J. K.) has benefited from funds provided to the University of Notre Dame by the National Science Foundation.

APPENDIX

We have to prove the result given by Eq. (28) in Sec. IV. This is done in a way similar to that outlined in Whittaker and Watson[6] for a formula due to Plana. We shall consider the function

$$[\mathcal{z}^p(\mathcal{z}+a)(e^{-2\pi i\mathcal{z}}-1)]^{-1} \qquad \text{(A1)}$$

of a complex variable $\mathcal{z}$ for $\operatorname{Re} a < 0$, $\operatorname{Im} a < 0$. This function is to be integrated around the indented rectangular contour shown in the upper half of Fig. 8, with corners at $\mathcal{z} = 0$, N (an integer), $N + i\infty$, and $i\infty$. The cut from $\mathcal{z} = 0$ to $-\infty$ needed to make the function (A1) regular is outside the contour and of no importance in the calculation which follows. The only singularity of Eq. (A1) lying inside the contour then is the simple pole at $\mathcal{z} = -a$ where there is a residue of

$$[(-a)^p(e^{2\pi ia}-1)]^{-1}.$$

At the integral sites $\mathcal{z} = 1, 2, \cdots, N$ on the real axis, the function (A1) has simple poles—at $\mathcal{z} = n$ the residue is

$$-[2\pi i n^p(n+a)]^{-1}.$$

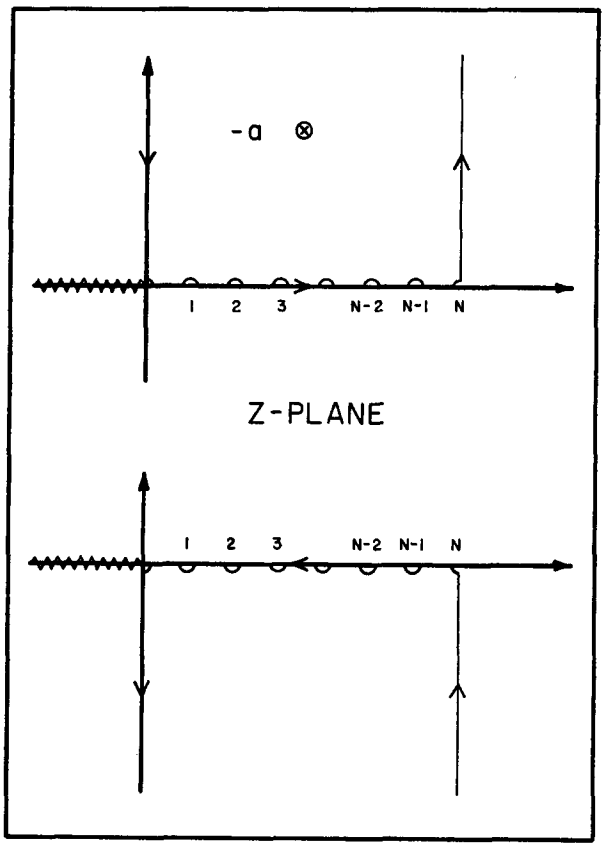

FIG. 8. The two contours of integration used in establishing the result Eq. (A4).

Thus the contribution to the contour integral from the indentations at these points is

$$\frac{1}{2}\sum_{n=1}^{N-1}\frac{1}{n^p(n+a)}+\frac{1}{4N^p(N+a)}+\mathcal{O}(r),$$

where r is the radius of the indentations. The contributions from that part of the contour lying on the real axis is

$$\mathcal{P}\int_0^N d\mathcal{z}\,\frac{1}{\mathcal{z}^p(\mathcal{z}+a)(e^{-2\pi i\mathcal{z}}-1)}+\mathcal{O}(r),$$

where $\mathcal{P}$ indicates that the principal value of the integral is to be taken at the singularities of the integrand. From the sides of the rectangle parallel to the imaginary axis comes the term

$$+\frac{1}{i}\int_r^\infty dy(e^{2\pi y}-1)^{-1}\{[(iy)^p(a+iy)]^{-1} - [(N+iy)^p(N+a+iy)]^{-1}\}.$$

The quarter-circle at $\mathcal{z} = 0$ gives rise to the following contribution:

$$-i\int_{\pi/2}^0 d\theta r^{1-p}e^{i(1-p)\theta}(2\pi iare^{i\theta})^{-1}+\mathcal{O}(r) = \frac{ir^{-p}}{2\pi ap}(e^{-ip\pi/2}-1)+\mathcal{O}(r).$$

Clearly the top of the rectangle, when it is removed to infinity, does not contribute to the integral. The application of Cauchy's theorem now yields the result

$$\tfrac{1}{2}\sum_{n=1}^{N-1}[n^p(n+a)]^{-1}+[4N^p(N+a)]^{-1}+\Im\int_0^N[\zeta^p(z+a)(e^{-2\pi i\zeta}-1)]^{-1}\,d\zeta+\frac{1}{i}\int_r^\infty dy[e^{2\pi y}-1]^{-1}[(iy)^p(a+iy)]^{-1}-[(N+iy)^p(N+a+iy)]^{-1}+\frac{ir^{-p}}{2\pi ap}(e^{-ip\pi/2}-1)=\frac{2\pi i}{(-a)^p(e^{2\pi ia}-1)}+\mathcal{O}(r). \tag{A2}$$

A similar consideration of the function

$$[z^p(z+a)(e^{2\pi i\zeta}-1)]^{-1}$$

around the contour described in the lower half of Fig. 8 gives the further result

$$\tfrac{1}{2}\sum_{n=1}^{N-1}[n^p(n+a)]^{-1}+[4N^p(N+a)]^{-1}+\Im\int_0^N[z^p(z+a)(e^{2\pi i\zeta}-1)]^{-1}\,d\zeta-\frac{1}{i}\int_r^\infty dy(e^{2\pi y}-1)^{-1}\{[(-iy)^p(a-iy)]^{-1}-[(N-iy)^p(N+a-iy)]^{-1}\}+\frac{ir^{-p}}{2\pi ap}(e^{ip\pi/2}-1)=\mathcal{O}(r). \tag{A3}$$

When Eqs. (A2) and (A3) are summed, there results

$$\sum_{n=1}^{N-1}[n^p(n+a)]^{-1}+[2N^p(N+a)]^{-1}=\int_0^N\frac{d\zeta}{\zeta^p(\zeta+a)}-\frac{r^{-p}}{\pi ap}\sin\frac{p\pi}{2}+I(N,r)+\frac{2\pi i}{(-a)^p(e^{2\pi ia}-1)}+\frac{1}{i}\int_r^\infty 2\,dy[y^p(e^{2\pi y}-1)(a^2+y^2)]^{-1}\times\left(a\sin\frac{p\pi}{2}+y\cos\frac{p\pi}{2}\right)+\mathcal{O}(r), \tag{A4}$$

where $I(N, r)$ represents that part of the integral term which comes from the second term in square brackets of the integrands in Eqs. (A2) and (A3).

Our result is obtained by letting $N\to\infty$, $r\to 0$ in Eq. (A4). To do this, the following observations are made. First, we have

$$\lim_{r\to 0}\left(\frac{-r^{-p}}{\pi ap}\sin\frac{p\pi}{2}+\int_r^\infty dy\,\frac{2a\sin(p\pi/2)}{y^p(e^{2\pi y}-1)(a^2+y^2)}\right)=\lim_{r\to 0}\int_r^\infty dy\left(\frac{2a\sin(p\pi/2)}{y^p(e^{2\pi y}-1)(a^2+y^2)}-\frac{\sin(p\pi/2)}{\pi ay^{p+1}}\right)=\int_0^\infty dy\,\frac{\sin(p\pi/2)}{\pi ay^{p+1}(e^{2\pi y}-1)(a^2+y^2)}\times[2\pi a^2y-(e^{2\pi y}-1)(a^2+y^2)],$$

which is a well-defined integral, since the integrand behaves like y^{-p} as $y\to 0$. Secondly, we have

$$I(N,r)=\frac{1}{i}\int_r^\infty dy(e^{2\pi y}-1)^{-1}\times\{[(N+iy)^p(N+a+iy)]^{-1}-[(N-iy)^p(N+a-iy)]^{-1}\},$$

and because the expression in square brackets behaves for small y like

$$\frac{-2iy-2ipy[N/(N+a)]^{-1}+\mathcal{O}(y)}{(N+a)^2N^p},$$

then the integrand of $I(N, r)$ is finite as $y\to 0$ and so the integral tends to zero as $N\to\infty$. When all these results are grouped together, Eq. (A4) yields

$$\sum_{n=1}^{\infty}[n^p(n+a)]^{-1}=\int_0^\infty\frac{d\zeta}{\zeta^p(\zeta+a)}+\frac{2\pi i}{(-a)^p(e^{2\pi ia}-1)}+2\int_0^\infty dy[y^{p-1}(e^{2\pi y}-1)(a^2+y^2)]^{-1}\times\left[\cos\frac{p\pi}{2}+\frac{1}{\pi a}\sin\frac{p\pi}{2}\times\left(\frac{2\pi a^2-y^{-1}(e^{2\pi y}-1)(a^2+y^2)}{y}\right)\right],$$

which, along with the relation

$$\int_0^\infty\frac{d\zeta}{\zeta^p(\zeta+a)}=\pi(-a)^{-p}e^{ip\pi}\csc p\pi$$

for $\operatorname{Re} a<0$, $\operatorname{Im} a<0$, with $0<\arg(-a)<\pi/2$, gives the desired result, Eq. (28).

[1] R. Davidson and J. J. Kozak, J. Math. Phys. **11**, 189 (1970).
[2] R. Davidson and J. J. Kozak, J. Math. Phys. **11**, 1420 (1970).
[3] R. W. Zwanzig, "Statistical Mechanics of Irreversibility," in *Lectures in Theoretical Physics*, W. E. Britten, B. W. Downs, and J. Downs, Eds. (Interscience, New York, 1961), Vol. III.
[4] I. Prigogine, *Non-Equilibrium Statistical Mechanics* (Interscience, New York, 1962); P. Resibois, "A Perturbative Approach to Irreversible Statistical Mechanics," in *Physics of Many-Particle Systems* (Gordon & Breach, New York, 1966), Vol. I.
[5] M. Kac, Am. J. Math. **65**, 609 (1943); P. Mazur and E. Montroll, J. Math. Phys. **1**, 70 (1960).
[6] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge U.P., New York, 1962).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

A Method for Generating Solutions of Einstein's Equations

ROBERT GEROCH
Department of Physics, University of Texas, Austin, Texas 78712

(Received 16 September 1970)

A method is described for constructing, from any source-free solution of Einstein's equations which possesses a Killing vector, a one-parameter family of new solutions. The group properties of this transformation are discussed. A new formalism is given for treating space-times having a Killing vector.

INTRODUCTION

One of the obstacles to a better understanding of the physical implications of general relativity is the relative scarcity of exact solutions of Einstein's equations. Although approximate methods are available, we have as yet no prescription, given a physical situation, for writing down a class of exact solutions which might represent that situation. (In electrodynamics, for example, Green's functions provide such a prescription.) Although it is unlikely that a similar prescription will be available for the gravitational field in the near future, a substantial increase in the number of known solutions would be a useful first step.

We shall here introduce a technique for generating explicit, exact, source-free solutions of Einstein's equations. More precisely, there is associated with any exact solution which has a Killing vector a one-parameter family of solutions, each with a Killing vector. The family forms a circle. Repeating the transformation merely results in a further rotation within the original circle of solutions; no additional solutions are obtained in this way. Even applied to Minkowski space, the transformation described here generates nonflat exact solutions of Einstein's equations. Our method generalizes work of Buchdahl,[1] Ehlers,[2] and Harrison.[3]

The transformations have an interesting group structure. Given a solution with a Killing vector, the group of transformations on that solution emerges initially as the three-dimensional Lie group $SL(2, R)$. There is a (two-dimensional) subgroup N of $SL(2, R)$ which corresponds to "pure gauge transformations," i.e., whose elements change the parameters describing the solution without substantially altering the solution itself. The collection of new solutions is given by the quotient set, $SL(2, R)/N$. Although it is possible to introduce a "gauge-free" transformation, the situation is then complicated by the fact that N is not a normal subgroup of $SL(2, R)$ and, hence, that $SL(2, R)/N$ has no natural group structure.

With each solution there is associated three curl-free bivectors. One is identically a curl, one is the vector discovered by Komar, and the third is new. Under a transformation of the solution, the three forms (or, alternatively, the numbers obtained by integrating them over a compact 2-surface) behave as the components of a symmetric, second-rank tensor under $SL(2, R)$. This behavior is closely related to the "wire singularities" present, e.g., in the original formulation of NUT space.[4]

Unfortunately, many of the solutions resulting from the method have no obvious physical interpretation.

The Appendix consists of a new, coordinate-independent treatment of the Einstein equations in the presence of a Killing vector. This formulation considerably simplifies many of our calculations.

1. THE TRANSFORMATION

The transformation with which we shall be concerned can be described as follows. Consider a source-free solution of Einstein's equations, i.e., a four-dimensional manifold M with metric g_{ab} for which the Ricci tensor R_{ab} vanishes. Suppose, in addition, that there is a Killing field ξ^a. We define the norm and twist of ξ^a, respectively, by

$$\lambda = \xi^a \xi_a, \tag{1}$$

$$\omega_a = \epsilon_{abcd} \xi^b \nabla^c \xi^d. \tag{2}$$

Taking the curl of (2) and using $R_{ab} = 0$, we see that, at least locally, ω_a is the gradient of a scalar field ω. Einstein's and Killing's equations imply, furthermore, that the right-hand sides of

$$\nabla_{[a}\alpha_{b]} = \tfrac{1}{2}\epsilon_{abcd}\nabla^c \xi^d$$

and

$$\nabla_{[a}\beta_{b]} = 2\lambda \nabla_a \xi_b + \omega \epsilon_{abcd} \nabla^c \xi^d \tag{3}$$

are curl-free; hence, there exist, locally, solutions α_a and β_a of (3). Normalize α_a and β_a by[5]

$$\xi^a \alpha_a = \omega,$$
$$\xi^a \beta_a = \omega^2 + \lambda^2 - 1. \tag{4}$$

Then, for each real number θ, the metric

$$\begin{aligned} g'_{ab} &= [(\cos\theta - \omega\sin\theta)^2 + \lambda^2\sin^2\theta]g_{ab} \\ &\quad + 2\sin\theta\xi_{(a}[2\alpha_{b)}\cos\theta - \beta_{b)}\sin\theta] \\ &\quad + \lambda[(\cos\theta - \omega\sin\theta)^2 + \lambda^2\sin^2\theta]^{-1} \\ &\quad \times \sin^2\theta[2\alpha_a\cos\theta - \beta_a\sin\theta] \\ &\quad \times [2\alpha_b\cos\theta - \beta_b\sin\theta] \end{aligned} \tag{5}$$

is an exact, source-free solution of Einstein's equations. Each metric (5) admits ξ^a as a Killing vector.[6] The solutions corresponding to θ's differing by a multiple of π are, of course, identical. Otherwise, the solutions are in general distinct.

We shall first describe how this transformation arises, and show that the Ricci tensor of (5) does indeed vanish.

Since the metric g_{ab} admits a Killing vector, one would expect to be able to "divide out" by the action of the symmetry to obtain a three-dimensional formulation of the solution. A technique for doing this is discussed in the Appendix. The solution is characterized by a three-dimensional manifold S with a nonsingular metric h_{ab} and the two scalar fields λ and ω, subject to [cf. (A18)]

$$\begin{aligned} \mathcal{R}_{ab} &= \tfrac{1}{2}\lambda^{-2}[(D_a\omega)(D_b\omega) - h_{ab}(D^m\omega)(D_m\omega)] \\ &\quad + \tfrac{1}{2}\lambda^{-1}D_aD_b\lambda - \tfrac{1}{4}\lambda^{-2}(D_a\lambda)(D_b\lambda), \\ D^2\lambda &= \tfrac{1}{2}\lambda^{-1}(D^m\lambda)(D_m\lambda) - \lambda^{-1}(D^m\omega)(D_m\omega), \\ D^2\omega &= \tfrac{3}{2}\lambda^{-1}(D^m\lambda)(D_m\omega), \end{aligned} \tag{6}$$

where D_a is the (covariant) derivative and $\mathcal{R}_{ab}$ the Ricci tensor[7] with respect to h_{ab}. [Strictly speaking, (6) is applicable only when $\lambda \neq 0$. If $\lambda < 0$, h_{ab} is positive definite, while if $\lambda > 0$, h_{ab} has signature $(-, +, +)$.] Equations (6) can be simplified considerably by the substitution $\tilde{h}_{ab} = \lambda h_{ab}$, $\tau = \omega + i\lambda$:

$$\begin{aligned} \tilde{\mathcal{R}}_{ab} &= -2(\tau - \bar{\tau})^{-2}(\tilde{D}_{(a}\tau)(\tilde{D}_{b)}\bar{\tau}), \\ \tilde{D}^2\tau &= 2(\tau - \bar{\tau})^{-1}(\tilde{D}_m\tau)(\tilde{D}_n\tau)\tilde{h}^{mn}, \end{aligned} \tag{7}$$

where $\tilde{D}_a$ and $\tilde{\mathcal{R}}_{ab}$ are the derivative and Ricci tensor with respect to $\tilde{h}_{ab}$ ($\tilde{D}^2 = \tilde{h}^{ab}\tilde{D}_a\tilde{D}_b$). We are given a solution, $\tilde{h}_{ab}$, τ, of (7): We wish to write down a new solution, $\tilde{h}'_{ab}$, τ'. The form of Eqs. (7) suggests that we consider $\tilde{h}'_{ab} = \tilde{h}_{ab}$, $\tau' = \tau'(\tau)$. Substituting into (7), we find that the only solution of this form is

$$\tau' = (a\tau + b)/(c\tau + d), \tag{8}$$

where a, b, c, and d are real numbers which can, without loss of generality, be normalized by $ac - bd = 1$. Expressed in terms of our original quantities—h_{ab}, λ, and ω—the transformation (8) takes the form

$$\begin{aligned} h'_{ab} &= \lambda(\lambda')^{-1}h_{ab}, \\ \lambda' &= \lambda[(c\omega + d)^2 + c^2\lambda^2]^{-1}, \\ \omega' &= [(a\omega + b)(c\omega + d) + ac\lambda^2] \\ &\quad \times [(c\omega + d)^2 + c^2\lambda^2]^{-1}. \end{aligned} \tag{9}$$

In fact, the *only* substitution of this type—$h'_{ab} = f(\lambda, \omega)h_{ab}$, $\lambda' = \lambda'(\lambda, \omega)$, $\omega' = \omega'(\lambda, \omega)$—which preserves (6) is (9).[8] Thus, the general transformation is defined by a real 2×2 matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with unit determinant, i.e., by an element of $SL(2, R)$. Successive application of two such transformations [cf. (8)] yields the transformation associated with the product of the matrices.

Finally, we wish to recover the expression (5). Set

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{10}$$

so that (9) becomes

$$\begin{aligned} h'_{ab} &= \lambda(\lambda')^{-1}h_{ab}, \\ \lambda' &= \lambda[(\cos\theta - \omega\sin\theta)^2 + \lambda^2\sin^2\theta]^{-1}, \\ \omega' &= [(\omega\cos\theta + \sin\theta)(-\omega\sin\theta + \cos\theta) \\ &\quad - \sin\theta\cos\theta\lambda^2] \\ &\quad \times [(\cos\theta - \omega\sin\theta)^2 + \lambda^2\sin^2\theta]^{-1}. \end{aligned} \tag{11}$$

According to the Appendix, the transition from three-dimensional quantities (11) to four-dimensional quantities is accomplished as follows. Solve

$$\nabla_{[a}\eta_{b]} = \tfrac{1}{2}(\lambda')^{-2}\epsilon_{abcd}\xi^c h'^{dm}\nabla_m\omega',$$

and

$$\xi^a\eta_a = 1 \tag{12}$$

for η_a, and set

$$g'_{ab} = \lambda(\lambda')^{-1}(g_{ab} - \lambda^{-1}\xi_a\xi_b) + \lambda'\eta_a\eta_b. \tag{13}$$

Substituting (11) into (12), we see that the solution is

$$\eta_a = (\lambda')^{-1}\xi_a + 2\alpha_a\cos\theta\sin\theta - \sin^2\theta\beta_a, \tag{14}$$

where α_a and β_a are given by Eqs. (3) and (4). Equation (5) now follows from (13).

It is easily checked that the metric g'_{ab} defined by (5) again has ξ^a as a Killing vector. The norm and twist of ξ^a, with respect to g'_{ab}, are given by (11).

Thus, the fundamental group of transformations (8) is the three-dimensional Lie group $SL(2, R)$. Unfortunately, two of the three dimensions correspond to pure gauge. Consider the subgroup N of $SL(2, R)$ consisting of matrices of the form $\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$ (so $d = a^{-1}$).

For such a matrix, (9) becomes

$$\begin{aligned} h'_{ab} &= d^2 h_{ab}, \\ \lambda' &= d^{-2}\lambda, \\ \omega' &= d^{-1}(a\omega + b). \end{aligned} \tag{15}$$

We see that (15) represents merely a constant conformal transformation, a rescaling of the Killing vector, and the addition of a constant to ω. Such transformations do not in any essential way alter the original solution of Einstein's equations. The subgroup N thus represents gauge transformations. The group $SL(2, R)$ is two-to-one isomorphic to $O(2, 1)$: the three-dimensional Lorentz group. In terms of $O(2, 1)$, N is a subgroup of null rotations [a subgroup of $O(2, 1)$ which leaves invariant a fixed null direction]. The collection of effective transformations is $SL(2, R)/N$, a circle. However, since N is not a normal subgroup of $SL(2, R)$, there is no natural group structure on $SL(2, R)/N$. The fact that the collection of "effective transformations" does not form a group stems from the following circumstance. The original transformation, as described at the beginning of this section, did not, strictly speaking, act on a solution of Einstein's equations. Before applying the transformation it was necessary to make a decision: which twist scalar ω to use. [ω is determined by (2) only up to an additive constant.] For a given θ, the resulting metric (5) depends on the choice of this constant. However, the collection of metrics obtained, as θ varies from 0 to π, is the same no matter which ω was chosen originally. In other words, the only effect of adding a constant to ω is to alter the way in which the metrics (5) are parameterized by θ. (The situation is similar, but simpler, with regard to multiplying the original metric, or the original Killing vector, by a constant factor.)

Our restriction of the general element of $SL(2, R)$ to the form (10) may now be described as follows. Since $SL(2, R)/N$ has no natural group structure, we (somewhat artificially) impose such a structure on it. This is done by choosing a subgroup of $SL(2, R)$ which intersects each coset of N exactly once. The elements of the form (10) constitute such a subgroup. [For some purposes the subgroup consisting of elements of the form $\left(\begin{smallmatrix} 1 & 0 \\ c & 1 \end{smallmatrix}\right)$, which intersects each coset except one, is more convenient.]

To summarize, the most general transformation is expressed in four-dimensional form, without redundancy, by (5). It is often more convenient, however, to consider instead the three-dimensional quantities—h_{ab}, λ, and ω—and retain the full group $SL(2, R)$.

Finally, we discuss the transformation properties of certain exact forms associated with the solutions. It is easily verified from (7) that the three real vector fields

$$\begin{aligned} V_1^a &= (\tau - \bar\tau)^{-2}\tilde h^{am}(\tilde D_m\tau + \tilde D_m\bar\tau), \\ V_2^a &= (\tau - \bar\tau)^{-2}\tilde h^{am}(\bar\tau\tilde D_m\tau + \tau\tilde D_m\bar\tau), \\ V_3^a &= (\tau - \bar\tau)^{-2}\tilde h^{am}(\bar\tau^2\tilde D_m\tau + \tau^2\tilde D_m\bar\tau), \end{aligned} \tag{16}$$

on S are divergence-free (using $\tilde D_a$). [Remarkably enough, $(\tau - \bar\tau)^{-2}\tilde h^{am}(\bar\tau^s\tilde D_m\tau + \tau^s\tilde D_m\bar\tau)$ is divergence-free only for $s = 0$, 1, or 2.] Under (8), the three fields (16) are taken into linear combinations of themselves:

$$\begin{aligned} V_1^a &\to d^2V_1^a + 2cdV_2^a + c^2V_3^a, \\ V_2^a &\to bdV_1^a + (ad + bc)V_2^a + acV_3^a, \\ V_3^a &\to b^2V_1^a + 2abV_2^a + a^2V_3^a. \end{aligned} \tag{17}$$

But (17) is the transformation law for the components of a symmetric, second-rank tensor under $SL(2, R)$ [or, alternatively, of a vector under $O(2, 1)$]. To express (16) in terms of four-dimensional quantities, we contract each vector with $\tilde\epsilon_{abc}$ to obtain three skew, second-rank, covariant, curl-free, tensor fields on S. The mapping $\psi: M \to S$ (see Appendix) then induces corresponding curl-free bivectors on M. These are

$$\begin{aligned} &\nabla_{[a}(\lambda^{-1}\xi_{b]}), \\ &\nabla_{[a}(\omega\lambda^{-1}\xi_{b]}) - \tfrac{1}{2}\epsilon_{abcd}\nabla^c\xi^d, \\ &\nabla_{[a}(\lambda^{-1}(\omega^2 + \lambda^2)\xi_{b]}) - 2\lambda\nabla_a\xi_b - \omega\epsilon_{abcd}\nabla^c\xi^d, \end{aligned} \tag{18}$$

[compare (3)]. Now suppose that in our original space-time M there is a compact two-dimensional submanifold K which is not the boundary of any compact three-dimensional submanifold of M (i.e., K represents a nonzero element of the second homology group of M). We would expect to have such a surface, for example, in the exterior field of a star (e.g., the Schwarzschild or Kerr solutions): K would be a 2-sphere surrounding the star at one instant of time. Integrating the forms (18), we obtain three numbers[9] associated with K:

$$\begin{aligned} I_1 &= \int_K \nabla_{[a}(\lambda^{-1}\xi_{b]})\, dS^{ab}, \\ I_2 &= \int_K [\nabla_{[a}(\omega\lambda^{-1}\xi_{b]}) - \tfrac{1}{2}\epsilon_{abcd}\nabla^c\xi^d]\, dS^{ab}, \\ I_3 &= \int_K [\nabla_{[a}(\lambda^{-1}(\omega^2 + \lambda^2)\xi_{b]}) \\ &\qquad - 2\lambda\nabla_a\xi_b - \omega\epsilon_{abcd}\nabla^c\xi^d]\, dS^{ab}. \end{aligned} \tag{19}$$

These numbers, of course, obey the same transformation law (17) as the corresponding forms. In particular, there is an "invariant" associated with a symmetric, second-rank tensor over $SL(2, R)$:

$$I = I_1I_3 - (I_2)^2. \tag{20}$$

Note that the first integral (19) vanishes identically (the integrand is a curl). How can this observation be consistent with the fact that the integrals (19) transform as the components of a tensor over $SL(2, R)$? After (9), the quantity I_1 will, in general, no longer be zero. The answer, of course, is that the nonvanishing of either of the last two integrals (19) implies that α_a or β_a of (3) cannot be regular everywhere on M. Singularities in α_a or β_a will result in singularities in the transformed metric g'_{ab} of (5). Thus, the transformation (5) will generate singularities on certain 2-surfaces in M, these surfaces having the property that they intersect every compact 2-surface K on which the integrals (19) fail to vanish.

Consider, as an example, the Schwarzschild solution, where ξ^a is the timelike Killing vector. In this case, I_2 is essentially the Schwarzschild mass, while $I_3 = 0$. Thus, the transformed metric (5) will have a singularity on a 2-surface. In fact, g'_{ab} turns out to be the Taub–NUT metric, and the singularity the "wire singularities" found in the original formulation of that metric.[4,2] It is clear from this formulation that the singularity is not of the "curvature type." It is perhaps not surprising that it can be eliminated.[10]

2. CONCLUSION

Although the technique described here produces large new classes of solutions of Einstein's equations, the method offers very little insight into the physical implications of the metrics. Presumably, such insight must come from a detailed study of examples. It turns out, however, that, even if one begins with a moderately simple metric g_{ab}, the resulting g'_{ab} can be quite complicated. For example, starting from the general Killing vector in the Schwarzschild solution (linear combination of time–translation and rotation), the resulting metrics, although they can easily be written out explicitly, are too involved to admit any simple interpretation.

More encouraging is the situation in the case of two Killing vectors. If a space–time has two commuting Killing vectors (e.g., a Weyl solution), then the transformation can be applied with respect to any linear combination of the Killing fields. The resulting g'_{ab} certainly has one Killing field: But does it necessarily have two? The answer is yes. Furthermore, the two Killing vectors in the new metric also commute, and so the transformations can be iterated. If one applies two transformations, first with respect to one Killing vector and then with respect to another, the result depends on the order in which the transformations are applied. Thus, by iterating transformations, one expects to obtain families of solutions which involve many parameters—perhaps even arbitrary functions—starting, for example, from just one Weyl metric. Can one, by means of such iterations, obtain simple metrics? cause the wire singularities to cancel each other out? These questions will be dealt with in a subsequent paper.

APPENDIX: A THREE-DIMENSIONAL FORMALISM FOR SPACE-TIMES WITH ONE KILLING VECTOR

In this Appendix we shall introduce a formalism, based in part on one developed by Ehlers,[2] for dealing with a space–time on which there is given a preferred Killing vector.

Let M, g_{ab} be a space–time with Killing vector field ξ^a, and suppose further that ξ^a is either everywhere timelike or everywhere spacelike. (Otherwise, it is necessary to consider the regions $\xi^a\xi_a < 0$ and $\xi^a\xi_a > 0$ separately.[11]) Let S denote the collection of all trajectories of ξ^a. That is, an element of S is an (inextendible) curve in M which is everywhere tangent to ξ^a. We define a mapping ψ from M onto S as follows: For each point p of M, $\psi(p)$ is the trajectory of ξ^a passing through p. Our final assumption (which always holds locally) is that S may be given the structure of a differentiable 3-manifold such that ψ is a smooth mapping. This assumption serves to eliminate certain global situations in which a trajectory of ξ^a "passes arbitrarily near to itself" in M.

If the Killing field ξ^a were hypersurface orthogonal, then it would be possible to represent S as one of the hypersurfaces in M which is everywhere orthogonal to ξ^a. Each trajectory of ξ^a would intersect this hypersurface in exactly one point. In the nonhypersurface orthogonal case, however, there is no natural way of introducing such surfaces in M. That is to say, it is most natural in the general case to regard S as a quotient space of M (i.e., $\psi: M \to S$) rather than a subspace ($S \to M$). As we shall see, there is a one-to-one correspondence between tensor fields and tensor operations on S and certain tensor fields and tensor operations on M. The differential geometry of S will, in this sense, be mirrored in M.

We first consider the representation of tensor fields on S as certain tensor fields on M. In fact, we shall show that *there is a one-to-one correspondence between tensor fields* $T'^{b\cdots d}_{a\cdots c}$ *on* S *and tensor fields* $T^{b\cdots d}_{a\cdots c}$ *on* M *which satisfy*[12]

$$\xi^a T^{b\cdots d}_{a\cdots c} = 0, \quad \cdots, \quad \xi_d T^{b\cdots d}_{a\cdots c} = 0,$$
$$\pounds_\xi T^{b\cdots d}_{a\cdots c} = 0. \tag{A1}$$

Let μ' be a (smooth) scalar field on S, so μ' assigns a

real number to each ξ-trajectory in M. Set $\mu = \mu' \circ \psi$: That is, μ is the scalar field on M which assigns to the point $p \in M$ the value assigned by μ' to the trajectory through p. This μ is evidently constant along each trajectory of ξ^a, i.e., we have $\pounds_\xi \mu = 0$. Conversely, any smooth scalar field μ on M, subject to $\pounds_\xi \mu = 0$, is constant along each ξ-trajectory, and so defines a unique scalar field μ' on S. We next consider covariant vector fields. Any covariant vector field A'_a on S may be written in the form

$$A'_a = \alpha' D_a \beta' + \cdots + \mu' D_a \nu', \tag{A2}$$

where D_a is the gradient operator on S and $\alpha', \beta', \cdots, \nu'$ are scalar fields on S. We define the corresponding covariant vector field on M by

$$A_a = \alpha \nabla_a \beta + \cdots + \mu \nabla_a \nu, \tag{A3}$$

where ∇_a is the gradient and $\alpha, \beta, \cdots, \nu$ are the corresponding scalar fields on M. Evidently, A_a satisfies (A1) and is uniquely determined by A'_a. (This fact, of course, is well known: A covariant tensor field can be "pulled back" through a differentiable mapping.) Conversely, any covariant vector field on M, subject to (A1), can be written in the form (A3) with scalars which are constant along the ξ-trajectories and, hence, via (A2), defines a covariant vector field on S. Interpret contravariant vector fields as derivations on the scalars.[13] Then there is clearly a one-to-one correspondence between contravariant vector fields on S and derivations on the scalar fields on M which are constant along the ξ-trajectories. But such derivations are, in turn, in one-to-one correspondence with vector fields on M satisfying (A1). Our result now follows from the fact that tensors of higher valence can be constructed from sums of outer products of (contravariant and covariant) vectors.

Note furthermore that the operations of addition, outer product, and contraction commute with the passage from M to S. That is to say, if one of these operations is first carried out in M and the result taken over to S, the result is the same as that obtained if the fields are first taken to S and the corresponding operation carried out there. Thus, the entire tensor algebra on S is completely and uniquely mirrored by tensors on M subject to (A1). While it is useful conceptually to have the three-dimensional manifold S, it plays no further logical role in the formalism. We shall hereafter drop the primes: We shall continue to speak of tensor fields being *on* S merely as a shorthand way of saying that the field (formally, on M) satisfies (A1).

In particular, the following are tensor fields on S:

$$h_{ab} = g_{ab} - (\xi^m \xi_m)^{-1} \xi_a \xi_b, \tag{A4}$$

$$h^{ab} = g^{ab} - (\xi^m \xi_m)^{-1} \xi^a \xi^b, \tag{A5}$$

$$h^b_a = \delta^b_a - (\xi^m \xi_m)^{-1} \xi_a \xi^b, \tag{A6}$$

$$\epsilon_{abc} = (\pm \xi^m \xi_m)^{-\frac{1}{2}} \epsilon_{abcd} \xi^d. \tag{A7}$$

Equations (A4) and (A5) define the metric and inverse metric on S. Note that the indices of any tensor on S can be raised or lowered with either h or g with the same result. Equation (A6) defines the Kronecker delta on S. Alternatively, (A6) can be interpreted as the projection operator onto S. That is to say, if a tensor satisfies the last equation (A1), then its projection, by h^b_a, satisfies (A1). Finally, (A7) is the alternating tensor on S. (The sign within the parenthesis is to be so chosen that ϵ_{abc} will be real.) Note that $\epsilon_{abc}\epsilon^{abc} = 6$.

We next introduce the (covariant) derivative on S. If $T^{b\cdots d}_{a\cdots c}$ is any tensor field on S, define

$$D_e T^{b\cdots d}_{a\cdots c} = h^p_e h^m_a \cdots h^n_c h^b_r \cdots h^d_s \nabla_p T^{r\cdots s}_{m\cdots n}. \tag{A8}$$

Clearly, (A8) is again a tensor field on S. Furthermore, D_a satisfies the following conditions:

1. The derivative of the outer product of two tensor fields on S is equal to the first times the derivative of the second plus the second times the derivative of the first (Leibnitz rule).
2. The contraction of the derivative of any tensor field on S equals the derivative of its contraction.
3. If μ is any scalar field on S, then $D_a \mu$ is the gradient of μ, and $D_{[a} D_{b]} \mu = 0$ (torsion-free).
4. The derivative of the sum of two tensors on S is the sum of their derivatives.
5. The derivative of the metric is zero.

But these are the axioms for the unique covariant derivative operator on a manifold with metric.[13] We conclude that D_a is that operator.

To illustrate these remarks (and to obtain an expression we shall need later), we evaluate the Riemann tensor of S, the analog of the first Gauss–Codazzi equation. Let k_c be an arbitrary vector field on S. Then

$$\begin{aligned} D_a D_b k_c &= h^p_a h^q_b h^r_c \nabla_p (h^s_q h^t_r \nabla_s k_t) \\ &= h^p_a h^s_b h^t_c \nabla_p \nabla_s k_t \\ &\quad - (\xi^m \xi_m)^{-1} h^p_a h^q_b h^r_c (\nabla_p \xi_q) \xi^s \nabla_s k_r \\ &\quad - (\xi^m \xi_m)^{-1} h^p_a h^q_b h^r_c (\nabla_p \xi_r) \xi^t \nabla_q k_t. \end{aligned} \tag{A9}$$

Now antisymmetrize over a and b. We eliminate the derivatives of k_c on the right, using, for the second term, the fact that $\pounds_\xi k_r = 0$ and, for the third term,

the fact that $\xi^t k_t = 0$. Thus,

$$D_{[a}D_{b]}k_c = h_a^p h_b^q h_c^r \nabla_{[p}\nabla_{q]}k_r + (\xi^m\xi_m)^{-1}h_a^p h_b^q h_c^r(\nabla_p\xi_q)(\nabla_r\xi_s)k^s + (\xi^m\xi_m)^{-1}h_{[a}^p h_{b]}^q h_c^r(\nabla_p\xi_r)(\nabla_q\xi_t)k^t. \quad \text{(A10)}$$

But k_c is arbitrary. Therefore, the Riemann tensor $\mathcal{R}_{abcd}$ of S is related to the Riemann tensor R_{abcd} of M by

$$\mathcal{R}_{abcd} = h_{[a}^p h_{b]}^q h_{[c}^r h_{d]}^s [R_{pqrs} + 2(\xi^m\xi_m)^{-1}(\nabla_p\xi_q)(\nabla_r\xi_s) + 2(\xi^m\xi_m)^{-1}(\nabla_p\xi_r)(\nabla_q\xi_s)]. \quad \text{(A11)}$$

The basic equations for a space-time with a Killing vector are a set of differential equations on three quantities, the metric h_{ab}, the norm (1), and twist (2) of the Killing vector. Note that λ and ω_a are fields on S. It follows immediately from (1), (2), and Killing's equation that the derivative of the Killing vector may be expressed directly in terms of λ and ω_a:

$$\nabla_a\xi_b = \tfrac{1}{2}\lambda^{-1}\epsilon_{abcd}\xi^c\omega^d + \lambda^{-1}\xi_{[b}D_{a]}\lambda. \quad \text{(A12)}$$

We shall also require the formula for the second derivative of the Killing vector:

$$\nabla_a\nabla_b\xi_c = R_{dabc}\xi^d. \quad \text{(A13)}$$

[To prove (A13), begin with $\nabla_a\nabla_b\xi_c$ and interchange first the first two indices, then the last two, then the first two, etc., until the indices have been restored to their original positions. Interchanging the first two indices gives a term involving the Riemann tensor; interchanging the last two contributes a minus sign, by Killing's equation.] Taking the curl and divergence of (2), using (A12) and (A13), we obtain, respectively,

$$D_{[a}\omega_{b]} = -\epsilon_{abmn}\xi^m R_p^n\xi^p, \quad \text{(A14)}$$

$$D^a\omega_a = \tfrac{3}{2}\lambda^{-1}\omega_m D^m\lambda. \quad \text{(A15)}$$

Applying $D^2 = D^aD_a$ to (1), using (A12) and (A13), we obtain

$$D^2\lambda = \tfrac{1}{2}\lambda^{-1}(D^m\lambda)(D_m\lambda) - \lambda^{-1}\omega^m\omega_m - 2R_{mn}\xi^m\xi^n. \quad \text{(A16)}$$

Finally, contracting (A11) once, and again using (A12) and (A13),

$$\mathcal{R}_{ab} = \tfrac{1}{2}\lambda^{-2}[\omega_a\omega_b - h_{ab}\omega^m\omega_m] + \tfrac{1}{2}\lambda^{-1}D_aD_b\lambda - \tfrac{1}{4}\lambda^{-2}(D_a\lambda)(D_b\lambda) + h_a^m h_b^n R_{mn}. \quad \text{(A17)}$$

The basic equations for a space-time with a Killing vector are (A14), (A15), (A16), and (A17).

In the source-free case ($R_{ab} = 0$), (A14) implies that ω_a is a gradient, $\omega_a = D_a\omega$. The equations (A14)–(A17) then take the form

$$\begin{aligned}\mathcal{R}_{ab} &= \tfrac{1}{2}\lambda^{-2}[(D_a\omega)(D_b\omega) - h_{ab}(D^m\omega)(D_m\omega)] + \tfrac{1}{2}\lambda^{-1}D_aD_b\lambda - \tfrac{1}{4}\lambda^{-2}(D_a\lambda)(D_b\lambda),\\ D^2\lambda &= \tfrac{1}{2}\lambda^{-1}(D^m\lambda)(D_m\lambda) - \lambda^{-1}(D^m\omega)(D_m\omega),\\ D^2\omega &= \tfrac{3}{2}\lambda^{-1}(D^m\lambda)(D_m\omega).\end{aligned} \quad \text{(A18)}$$

Finally, we show how it is possible to recover the original 4-dimensional space-time from its 3-dimensional formulation. Suppose we are given a 3-manifold S with a metric h_{ab}, a scalar field λ, and a vector field ω_a [subject, of course, to the following condition: Either $\lambda < 0$ and h_{ab} is positive-definite, or else $\lambda > 0$ and h_{ab} has signature $(-, +, +)$]. Suppose in addition that these fields satisfy (A15). We shall show that there is an essentially unique four-dimensional space-time M, g_{ab} with a Killing vector ξ^a and a mapping $\psi\colon M \to S$ which reproduces the given fields on S. It will then follow, in particular, that the full content of Einstein's equations in the presence of a Killing vector is expressed by (A18), for all the components of the Ricci tensor are involved in (A14)–(A17). The construction proceeds as follows. Choose an arbitrary four-dimensional manifold M along with a nowhere-vanishing vector field ξ^a on M. Let ψ be a smooth mapping from M to S which induces a diffeomorphism between S and the manifold of ξ-trajectories in M: In particular, ψ maps each trajectory of ξ^a into a single point of S. The idea is to use ψ and the various fields on S to obtain the metric g_{ab} of M. By (A15), the skew field

$$F'_{ab} = -\tfrac{1}{2}\lambda^{-\frac{3}{2}}\epsilon_{abc}D^c\omega \quad \text{(A19)}$$

on S is curl-free. The "pull-back" of F'_{ab} is therefore a curl-free skew field F_{ab} on M. Let η_a be a vector field on M satisfying

$$\nabla_{[a}\eta_{b]} = F_{ab}, \qquad \xi^a\eta_a = 1. \quad \text{(A20)}$$

[Note, from (A12), that η_a will eventually be $\lambda^{-1}\xi_a$.] Finally, let H_{ab} denote the symmetric tensor field on M obtained, via ψ, from h_{ab}. (So, in particular, $\xi^a H_{ab} = 0$.) Then the required metric on M is

$$g_{ab} = H_{ab} + \lambda\eta_a\eta_b. \quad \text{(A21)}$$

It is easily verified, from (A12) and (A20), that ξ^a is a Killing vector of g_{ab} and that the norm and twist of ξ^a, with respect to g_{ab}, are just λ and ω, respectively.

[1] H. Buchdahl, Quart. J. Math. **5** (1954).
[2] J. Ehlers, in *Les théories relativistes de la gravitation* (CNRS, Paris, 1959).
[3] B. K. Harrison, J. Math. Phys. **9**, 1744 (1968).
[4] E. T. Newman, L. Tamburino, and T. Unti, J. Math. Phys. **4**, 915 (1963).

[5] Since we are permitted by (3) to add a gradient to α_a and β_a, these fields can always be chosen to satisfy (4). Note, however, that even (3) and (4) together do not determine the fields uniquely: There remains the freedom to add the gradient of any scalar field which is constant along the trajectories of ξ^a. While the addition of such a gradient does formally change the resulting metric (5), the change is of a trivial sort: It can be effected by a diffeomorphism on M.

[6] That $\mathfrak{L}_\xi\alpha_a = \mathfrak{L}_\xi\beta_a = 0$ is a consequence of (3), (4), and the following identity. If ξ^a is any vector field and $F_{a_1\cdots a_s}$ any skew tensor field, then

$$s\nabla_{[a_1}(\xi^m F_{a_2\cdots a_s]m}) - (s+1)\xi^m\nabla_{[a_1}F_{a_2\cdots a_s m]} + (-1)^s\mathfrak{L}_\xi F_{a_1\cdots a_s} = 0.$$

[7] Our conventions for the Riemann and Ricci tensors are

$$D_{[a}D_{b]}k_c = \tfrac{1}{2}\mathcal{R}_{abcd}k^d, \quad \mathcal{R}_{ab} = \mathcal{R}^m{}_{amb}.$$

[8] We are ignoring a trivial transformation: $h'_{ab} = (\text{const})h_{ab}$, $\lambda' = \lambda, \omega' = \omega$.

[9] Since K is compact, the curls do not, of course, contribute to the integrals. The second integral (19) is well known. The third appears to be new.

[10] C. W. Misner, J. Math. Phys. 4, 924 (1963).

[11] Note, however, that the transformation as originally given by (5) does not require any assumptions whatever on $\xi^m\xi_m$.

[12] Note that we use Latin indices for both tensors on M and tensors on S. The reason for this will emerge shortly.

[13] See, for example, S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1963).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Wavefunction Expansion in Terms of Spherical Bessel Functions*

A. GERSTEN†
Department of Physics, Texas A & M University, College Station, Texas

(Received 5 January 1970)

A formalism is derived for expanding the solutions of the Schrödinger equation in terms of spherical Bessel functions. The regular and the irregular solutions are treated. A relation between the expansion coefficients and the phase shifts is derived. As an application, the expansion coefficients of both the irregular and regular Coulomb wavefunctions are given in a form of a simple recurrence relation. The expansions have been checked numerically and found to be very suitable for calculating the regular Coulomb wavefunction in a very large region of the coordinate and the Coulomb parameter.

1. INTRODUCTION

The spherical Bessel functions are the most common functions used in scattering calculations. In the absence of a potential, they are the solutions of the Schrödinger equation. They can be easily and accurately calculated with the aid of the backward recurrence relations.[1] For these reasons it seemed appealing to look at the general problem of expanding the solutions of the Schrödinger equation in terms of spherical Bessel functions.

The expansions in terms of Bessel functions have properties similar to power series expansions, and the theorems concerning expansions of analytic functions in a Laurent-type expansion are almost the same with the exception that powers are replaced by Bessel functions. For a detailed discussion, the reader is referred to Watson's book.[2] In the case of spherical Bessel functions, this analogy can be made by replacing the nth power by a spherical Bessel function of order n. Accordingly, if a solution of the Schrödinger equation can be expanded in a power series, it can also be expanded in spherical Bessel functions.

In this paper a formalism is given for the expansion of the solution of the Schrödinger equation in terms of spherical Bessel functions (Secs. 2, 4, and 5). It appears that the relation between the expansion coefficients and the phase shifts is very simple (see Sec. 3). In this paper both the regular and the irregular solutions are treated. As a practical application of this method, the expansion coefficients of the Coulomb wavefunctions are derived in the form of a recurrence relation. The expansion of the regular Coulomb wavefunctions was known before,[3] but its derivation was not published. We derive here the expansions of the irregular as well as the regular Coulomb wavefunctions (Sec. 6).

The difficulties of computing the Coulomb wavefunctions are well known.[4] We found by numerical checking that the expansion of the regular Coulomb wavefunction in terms of spherical Bessel functions is very satisfactory in a wide region of the coordinate and the Coulomb parameter (see Sec. 7), for which no other single method[4] can be used effectively.

2. THE EXPANSION IN SPHERICAL BESSEL FUNCTIONS

Let us write the Schrödinger equation for the l-partial wave for a particle with mass μ and energy E:

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} + k^2 - \frac{l(l+1)}{r^2}\right]\psi_\lambda(r) = \frac{2\mu}{\hbar^2}V(r)\psi_\lambda(r), \tag{2.1}$$

[5] Since we are permitted by (3) to add a gradient to α_a and β_a, these fields can always be chosen to satisfy (4). Note, however, that even (3) and (4) together do not determine the fields uniquely: There remains the freedom to add the gradient of any scalar field which is constant along the trajectories of ξ^a. While the addition of such a gradient does formally change the resulting metric (5), the change is of a trivial sort: It can be effected by a diffeomorphism on M.

[6] That $\mathfrak{L}_\xi\alpha_a = \mathfrak{L}_\xi\beta_a = 0$ is a consequence of (3), (4), and the following identity. If ξ^a is any vector field and $F_{a_1\cdots a_s}$ any skew tensor field, then

$$s\nabla_{[a_1}(\xi^m F_{a_2\cdots a_s]m}) - (s+1)\xi^m\nabla_{[a_1}F_{a_2\cdots a_s m]} + (-1)^s\mathfrak{L}_\xi F_{a_1\cdots a_s} = 0.$$

[7] Our conventions for the Riemann and Ricci tensors are

$$D_{[a}D_{b]}k_c = \tfrac{1}{2}\mathcal{R}_{abcd}k^d, \quad \mathcal{R}_{ab} = \mathcal{R}^m{}_{amb}.$$

[8] We are ignoring a trivial transformation: $h'_{ab} = (\text{const})h_{ab}$, $\lambda' = \lambda, \omega' = \omega$.

[9] Since K is compact, the curls do not, of course, contribute to the integrals. The second integral (19) is well known. The third appears to be new.

[10] C. W. Misner, J. Math. Phys. 4, 924 (1963).

[11] Note, however, that the transformation as originally given by (5) does not require any assumptions whatever on $\xi^m\xi_m$.

[12] Note that we use Latin indices for both tensors on M and tensors on S. The reason for this will emerge shortly.

[13] See, for example, S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1963).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Wavefunction Expansion in Terms of Spherical Bessel Functions*

A. GERSTEN†
Department of Physics, Texas A & M University, College Station, Texas

(Received 5 January 1970)

A formalism is derived for expanding the solutions of the Schrödinger equation in terms of spherical Bessel functions. The regular and the irregular solutions are treated. A relation between the expansion coefficients and the phase shifts is derived. As an application, the expansion coefficients of both the irregular and regular Coulomb wavefunctions are given in a form of a simple recurrence relation. The expansions have been checked numerically and found to be very suitable for calculating the regular Coulomb wavefunction in a very large region of the coordinate and the Coulomb parameter.

1. INTRODUCTION

The spherical Bessel functions are the most common functions used in scattering calculations. In the absence of a potential, they are the solutions of the Schrödinger equation. They can be easily and accurately calculated with the aid of the backward recurrence relations.[1] For these reasons it seemed appealing to look at the general problem of expanding the solutions of the Schrödinger equation in terms of spherical Bessel functions.

The expansions in terms of Bessel functions have properties similar to power series expansions, and the theorems concerning expansions of analytic functions in a Laurent-type expansion are almost the same with the exception that powers are replaced by Bessel functions. For a detailed discussion, the reader is referred to Watson's book.[2] In the case of spherical Bessel functions, this analogy can be made by replacing the nth power by a spherical Bessel function of order n. Accordingly, if a solution of the Schrödinger equation can be expanded in a power series, it can also be expanded in spherical Bessel functions.

In this paper a formalism is given for the expansion of the solution of the Schrödinger equation in terms of spherical Bessel functions (Secs. 2, 4, and 5). It appears that the relation between the expansion coefficients and the phase shifts is very simple (see Sec. 3). In this paper both the regular and the irregular solutions are treated. As a practical application of this method, the expansion coefficients of the Coulomb wavefunctions are derived in the form of a recurrence relation. The expansion of the regular Coulomb wavefunctions was known before,[3] but its derivation was not published. We derive here the expansions of the irregular as well as the regular Coulomb wavefunctions (Sec. 6).

The difficulties of computing the Coulomb wavefunctions are well known.[4] We found by numerical checking that the expansion of the regular Coulomb wavefunction in terms of spherical Bessel functions is very satisfactory in a wide region of the coordinate and the Coulomb parameter (see Sec. 7), for which no other single method[4] can be used effectively.

2. THE EXPANSION IN SPHERICAL BESSEL FUNCTIONS

Let us write the Schrödinger equation for the l-partial wave for a particle with mass μ and energy E:

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} + k^2 - \frac{l(l+1)}{r^2}\right]\psi_\lambda(r) = \frac{2\mu}{\hbar^2}V(r)\psi_\lambda(r), \tag{2.1}$$

where $k^2 = 2\mu E\hbar^{-2}$ and $\hbar$ is Planck's constant divided by 2π. We assume that $V(r)$ can have at most singularity of order $1/r$ near the origin. The solutions of Eq. (2.1) will be specified according to

$$\begin{aligned} \lambda &= l \qquad \text{for the regular solution,} \\ &= -l-1 \quad \text{for the irregular solution.} \end{aligned} \tag{2.2}$$

Let us now write the solution $\psi_\lambda(r)$ as a series of spherical Bessel functions:

$$\psi_\lambda(r) = \sum_{m=0}^{\infty} f_m^\lambda(k) j_{\lambda+m}(kr), \tag{2.3}$$

where the j_m are the spherical Bessel functions of order m. For negative orders, the j_m coincide with the spherical Neumann functions n_m according to

$$n_m(kr) = (-1)^{m+1} j_{-m-1}(kr). \tag{2.4}$$

When $V(r) = 0$, the solution to Eq. (2.1) is the spherical Bessel function of order λ,

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} + k^2 - \frac{l(l+1)}{r^2}\right] j_\lambda(kr) = 0, \tag{2.5}$$

where λ is given by Eq. (2.2). This equation will be used in the following calculations. Let us make the expansion

$$2\mu\hbar^{-2} r^2 V(r) j_m(kr) = \sum_{n=0}^{\infty} b_n^m(k) j_{m+1+n}(kr). \tag{2.6}$$

In Sec. 4 are given formulas for calculating the coefficients $b_n^m(k)$ in terms of the power series coefficients of the potential $V(r)$. In order to calculate the coefficients $f_m^\lambda(k)$ of the expansion (2.3), we insert this expansion in Eq. (2.1). Using Eqs. (2.5) and (2.6) and comparing the coefficients of the same spherical Bessel functions, we get

$$\begin{aligned} f_m^\lambda[(\lambda+m)(\lambda+m+1) - l(l+1)] &= \sum_{n=0}^{m-1} f_n^\lambda b_{m-n-1}^{\lambda+n} \\ &= \sum_{n=0}^{m-1} f_{m-1-n}^\lambda b_n^{\lambda+m-1-n}, \quad m = 1, 2, \cdots. \end{aligned} \tag{2.7}$$

Equation (2.7) enables us to calculate successively the coefficients $f_m^l(k)$ of the regular solution ($\lambda = l$). They are given in terms of the b_m^n and the arbitrary coefficient $f_0^l(k)$. In Sec. 5 we will treat the irregular solution.

3. THE PHASE SHIFTS

The spherical Bessel functions have the following asymptotic behavior:

$$j_l(kr) \sim \sin(kr - \tfrac{1}{2}\pi l)/(kr) \quad \text{for} \quad kr \gg l. \tag{3.1}$$

On the other hand, the asymptotic behavior of the wavefunction $\psi_l(r)$ is given (up to a multiplicative factor) by

$$\psi_l(r) \sim [\sin(kr - \tfrac{1}{2}\pi l) + \tan\delta_l \cos(kr - \tfrac{1}{2}\pi l)]/(kr), \tag{3.2}$$

where δ_l is the phase shift. Comparing Eqs. (2.3), (3.1), and (3.2), we find a simple formula for the phase shifts:

$$\tan\delta_l = -(f_1^l - f_3^l + f_5^l - \cdots)/(f_0^l - f_2^l + f_4^l - \cdots). \tag{3.3}$$

This expression is derived under the assumption of finite range forces.

4. CALCULATION OF THE COEFFICIENTS $b_n^m(k)$

In order to determine the coefficients $b_n^m(k)$ of Eq. (2.6), we shall use the Cauchy identity for an analytic function $f(z)$:

$$f(z_0) = \frac{1}{2\pi i}\oint_C \frac{f(z)}{z - z_0}\, dz, \tag{4.1}$$

where C is a circle of radius R with center at the origin, and z_0 is any point inside it. Let us now expand

$$\frac{krj_\lambda(kr)}{kx - kr} = \sum_{n=0}^{\infty} A_n^\lambda(kx) j_{\lambda+1+n}(kr), \tag{4.2}$$

where the A_n^λ are functions which will be determined later on, λ is an integer, and the j_λ are the spherical Bessel functions. From Eqs. (4.1) and (4.2) we have

$$\begin{aligned} &2\mu\hbar^{-2} r^2 V(r) j_\lambda(kr) \\ &\quad = \left[\frac{1}{2\pi i}\oint_C \frac{2\mu\hbar^{-2} zV(z)\, dz}{z - r}\right] rj_\lambda(kr) \\ &\quad = \frac{\mu}{\pi i\hbar^2}\sum_{n=0}^{\infty}\left[\oint zV(z) A_n^\lambda(kz)\right] j_{\lambda+1+n}(kr), \end{aligned} \tag{4.3}$$

where the path of integration encloses the point $z = r$ ($r \neq 0$ for λ negative). Comparing Eq. (2.6) with Eq. (4.3),we get

$$b_n^\lambda(k) = \frac{\mu\hbar^{-2}}{\pi i}\oint zV(z) A_n^\lambda(kz)\, dz. \tag{4.4}$$

Let us now determine the functions $A_n^\lambda(kz)$. Let us write Eq. (4.2) in a more concise form:

$$zj_\lambda(z)/(t - z) = \sum_{n=0}^{\infty} A_n^\lambda(t) j_{\lambda+1+n}(z). \tag{4.5}$$

Multiplying both sides by $(t - z)/z$ and substituting

$$j_\lambda(z)/z = [j_{\lambda-1}(z) + j_{\lambda+1}(z)]/(2\lambda + 1),$$

we get

$$\begin{aligned} j_\lambda(z) = &\sum_{n=0}^{\infty} \frac{tA_n^\lambda(t)}{2\lambda + 2n + 3} j_{\lambda+n}(z) - \sum_{n=1}^{\infty} A_{n-1}^\lambda(t) j_{\lambda+n}(z) \\ &+ \sum_{n=2}^{\infty} \frac{tA_{n-2}^\lambda(t)}{2\lambda + 2n - 1} j_{\lambda+n}(z). \end{aligned} \tag{4.6}$$

Equating the coefficients of the spherical Bessel functions, we get

$$A_0^\lambda(t) = (2\lambda + 3)t^{-1}, \quad A_1^\lambda(t) = (2\lambda + 3)(2\lambda + 5)t^{-2}, \tag{4.7a}$$

and the following recurrence relation:

$$A_n^\lambda(t) = (2\lambda + 2n + 3)A_{n-1}^\lambda(t)/t - [(2\lambda + 3 + 2n)/(2\lambda - 1 + 2n)]A_{n-2}^\lambda(t), \quad n > 1. \tag{4.7b}$$

From Eqs. (4.7) we can calculate all functions $A_n^\lambda(t)$, for example,

$$A_2^\lambda(t) = (2\lambda + 3)(2\lambda + 5)(2\lambda + 7)t^{-3} - (2\lambda + 7)t^{-1},$$
$$A_3^\lambda(t) = (2\lambda + 3)(2\lambda + 5)(2\lambda + 7)(2\lambda + 9)t^{-4} - 2(2\lambda + 5)(2\lambda + 9)t^{-2}.$$

We see that $A_n^\lambda(t)$ are polynomials in t^{-1} of order $n + 1$ and satisfy

$$A_n^\lambda(-t) = (-1)^{n+1}A_n^\lambda(t). \tag{4.8}$$

Let us write

$$A_n^\lambda(t) = \sum_{i=1}^{n+1} a_i^{\lambda,n} t^{-i}, \tag{4.9}$$

and expand the potential $V(r)$ in a power series:

$$2m\hbar^{-2}V(r) = v_{-1}r^{-1} + v_0 + v_1 r + v_2 r^2 + \cdots. \tag{4.10}$$

Substituting Eqs. (4.9) and (4.10) in Eq. (4.4), we get

$$b_n^\lambda(k) = \sum_{i=1}^{n+1} a_i^{\lambda,n} v_{i-2} k^{-i}, \tag{4.11}$$

which is the desired formula for the $b_n^\lambda(k)$.

5. THE IRREGULAR SOLUTIONS

We can write the irregular solution $I_l(r) \equiv \psi_{-l-1}(r)$ of Eq. (2.1) in the form

$$I_l(r) = A\psi_l(r) \ln (2r) + R_l(r), \tag{5.1}$$

where $\psi_l(r)$ is the regular solution, A is an arbitrary constant, and $R_l(r)$ has a pole of order $l + 1$ at $r = 0$. One important feature of the irregular solution (5.1) is that we can add to it a regular solution and have another irregular solution.

Substituting Eq. (5.1) in Eq. (2.1), we get an inhomogeneous equation for R_l:

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} + k^2 - \frac{l(l+1)}{r^2}\right]R_l(r) = 2\mu\hbar^{-2}V(r)R_l(r) - A[r^{-2}\psi_l(r) + 2r^{-1}\psi_l'(r)]. \tag{5.2}$$

Now we expand

$$R_l(r) = \sum_{m=0}^{\infty} g_m^l(k) j_{-l-1+m}(kr). \tag{5.3}$$

In order to determine the coefficients g_m^l, it is necessary to calculate the coefficients d_m^l of the following expansion:

$$A[\psi_l(r) + 2r\psi_l'(r)] = \sum_{m=0}^{\infty} d_m^l j_{l+m}(kr). \tag{5.4}$$

Using the expansion (2.3) and the relation

$$zj_l'(z) = lj_l(z) - zj_{l+1}(z),$$

we get

$$\psi_l(r) + 2r\psi_l'(r) = \sum_{m=0}^{\infty} (2l + 2m + 1)f_m^l j_{l+m}(kr) - 2kr\sum_{m=0}^{\infty} f_m^l j_{l+m+1}(kr). \tag{5.5}$$

The second sum on the rhs of Eq. (5.5) should be treated separately:

$$\sigma(r) = 2kr\sum_{m=0}^{\infty} f_m^l j_{l+m+1}(kr). \tag{5.6}$$

Using the method of Sec. 4, we find

$$krj_{l+m+1}(kr) = \sum_{n=0}^{\infty}(-1)^n(2l + 2m + 4n + 5)j_{l+m+2n+2}(kr). \tag{5.7}$$

Substituting Eq. (5.7) in (5.6), we get

$$\begin{aligned}\sigma(r) &= 2\sum_{m=0}^{\infty} f_m^l \sum_{n=0}^{\infty}(-1)^n \times (2l + 2m + 4n + 5)j_{l+m+2n+2}(kr) \\ &= 2\sum_{s=0}^{\infty}\left[\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} f_m^l(-1)^n \times (2l + 2m + 4n + 5)\delta_{m+2n,s}\right]j_{l+s+2}(kr) \\ &= 2\sum_{s=0}^{\infty}(2l + 2s + 5)F_s^l j_{l+s+2}(kr),\end{aligned} \tag{5.8}$$

where

$$F_s^l = \sum_{n=0}^{[s/2]}(-1)^n f_{s-2n}^l. \tag{5.9}$$

From Eqs. (5.4), (5.5), (5.8), and (5.9) we get

$$\begin{aligned}d_m^l &= (2l + 2m + 1)A(f_m^l - 2F_{m-2}^l) \\ &= (2l + 2m + 1)A(2F_m^l - f_m^l).\end{aligned} \tag{5.10}$$

Substituting Eqs. (5.3), (5.4), and (2.6) in Eq. (5.2), we get ($d_m^l = 0$ for m negative)

$$g_m^l[(-l - 1 + m)(-l + m) - l(l + 1)] = -d_{m-2l-1}^l + \sum_{n=0}^{m-1} b_n^{-l-1+m-1-n} g_{m-1-n}^l. \tag{5.11}$$

From Eq. (5.11) we see that the coefficient g_{2l+1}^l is undetermined. This results from the fact that any regular solution can be added to the irregular one and the resulting combination is another irregular solution.

The g^l_{2l+1} is the coefficient which multiplies the contribution of the regular solution. For $m = 2l + 1$, Eq. (5.11) reduces to

$$\sum_{n=0}^{2l} g^l_{2l-n} b_n^{l-1-n} = d^l_0. \tag{5.12}$$

Equation (5.12) together with Eq. (5.11) for $0 \leq m \leq 2l$ constitutes a system of linear equations from which the coefficients $g^l_0, g^l_1, \cdots, g^l_{2l}$ are uniquely determined.

6. THE COULOMB WAVEFUNCTIONS

In order to find the expansion of the Coulomb wavefunctions in spherical Bessel functions, we have to find the expansion coefficients $b^m_n(k)$ of Eq. (2.6) in terms of the expansion coefficients of Eqs. (4.9) and (4.10). From Eqs. (4.7), (4.8), and (4.9) we have

$$\begin{aligned} a^{\lambda,n}_1 &= 0 && \text{for odd } n, \\ &= (-1)^{n/2}(2\lambda + 2n + 3) && \text{for even } n. \end{aligned} \tag{6.1}$$

For the Coulomb potential $V(r)$, Eq. (4.10) has the form

$$2\mu\hbar^{-2}V(r) = 2k\eta/r, \tag{6.2}$$

where η is the Coulomb parameter. The regular and irregular Coulomb wavefunctions are usually denoted by $F_l(\eta, kr)$ and $G_l(\eta, kr)$, respectively. Those functions divided by r are the solutions of the Schrödinger equation (2.1) with the potential given by (6.2). Substituting the results of Eqs. (6.1) in Eq. (4.11) we get

$$\begin{aligned} b^\lambda_n(k) &= 0 && \text{for odd } n, \\ &= (-1)^{n/2} 2\eta(2\lambda + 2n + 3) && \text{for even } n. \end{aligned} \tag{6.3}$$

From Eqs. (6.3) and (2.7), we get the following recurrence relation for the expansion coefficients f^l_m:

$$\begin{aligned} s_l(m) f^l_m/(2l + m + 1) \\ = 2\eta(f^l_{m-1} - f^l_{m-3} - f^l_{m-5} - \cdots), \end{aligned} \tag{6.4}$$

where we denote

$$s_\lambda(m) = (\lambda + m)(\lambda + m + 1) - l(l + 1), \tag{6.5}$$

and the lower indices of f on the rhs of Eq. (6.4) are not negative. Adding Eq. (6.4) to the same equation but with $m - 2$ instead of m, we get

$$\frac{s_l(m)}{2l + 2m + 1} f^l_m - 2\eta f^l_{m-1} + \frac{s_l(m-2)}{2l + 2m - 3} f^l_{m-2} = 0, \quad m \geq 2, \tag{6.6a}$$

and

$$S_l(1) f^l_1/(2l + 3) = 2\eta f^l_0. \tag{6.6b}$$

The value of f^l_0 is still undetermined. It can be determined from the known power series expansion (5.4) of $F_l(\eta, kr)$ near $r = 0$, and is equal to

$$f^l_0 = (2l + 1)!!\, C_l(\eta), \tag{6.6c}$$

where[4]

$$C_l(\eta) = \frac{2^l}{(2l + 1)(2l)!} \times \left(\frac{1(1 + \eta^2)(2^2 + \eta^2) \cdots (l^2 + \eta^2) 2\pi\eta}{\exp(2\pi\eta) - 1}\right)^{\frac{1}{2}}. \tag{6.7}$$

Thus, the regular Coulomb wavefunction is given by

$$F_l(\eta, kr) = r \sum_{n=0}^{\infty} f^l_n j_{l+n}(kr), \tag{6.8}$$

where f^l_n are given in Eq. (6.6). In order to derive the expansion of the irregular Coulomb wavefunction, we proceed along the method described in Sec. 5. The coefficients are given in Eqs. (5.9), (5.10), and (5.11). From Eqs. (5.9) and (6.4) we have

$$F^l_m = s_l(m + 1) f^l_{m+1}/[2\eta(2l + 2m + 3)], \tag{6.9}$$

where $s_l(m)$ is defined by Eq. (6.5) and f^l_m are the coefficients of the expansion of the regular Coulomb wavefunction given in Eq. (6.6).

On the basis of Eqs. (5.1), (5.3), (5.10), (5.11), (6.3), (6.5), and (6.9), the irregular Coulomb wavefunction $G_l(\eta, kr)$ satisfies

$$\begin{aligned} r^{-1}G_l(\eta, kr) = Ar^{-1}F_l(\eta, kr) \ln(2r) \\ + \sum_{n=0}^{\infty} g^l_n j_{-l-1+n}(kr), \end{aligned} \tag{6.10}$$

where the coefficients g^l_n are related through

$$\begin{aligned} g^l_m S_{-l-1}(m) = 2\eta(2m - 2l - 1) \\ \times (g^l_{m-1} - g^l_{m-3} + g^l_{m-5} - \cdots) - d^l_{m-2l-1}, \end{aligned} \tag{6.11}$$

where

$$d^l_{m-2l-1} = (2m - 2l - 1)A[2F^l_{m-2l-1} - f_{m-2l-1}]. \tag{6.12}$$

Similar to the derivation of Eq. (6.6a), we can get from Eq. (6.11) the following recurrence relation for the coefficients g^l_m:

$$\begin{aligned} \frac{g^l_m S_{-l-1}(m) + d^l_{m-2l-1}}{2m - 2l - 1} - 2\eta g^l_{m-1} \\ + \frac{g^l_{m-2} S_{-l-1}(m-2) + d^l_{m-2l-3}}{2m - 2l - 5} = 0. \end{aligned} \tag{6.13}$$

In Eqs. (6.12) and (6.13), the coefficients A and g^l_{2l+1} are still undetermined; those are the two arbitrary

constants of an irregular solution. We can determine them from the known properties of the Coulomb wavefunctions. The formulas for the power series expansions were given by Yost *et al.*,[5] and are collected in a compact form in Refs. 1 and 4. The irregular Coulomb wavefunction is given by

$$G_L(\eta, \rho) = [2\eta/C_0^2(\eta)]F_L(\eta, \rho) \times [\ln 2\rho + q_L(\eta)p/_L(\eta)] + \theta_L(\eta, p), \tag{6.14}$$

where $F_L(\eta, \rho)$ is the regular Coulomb wavefunction and $C_0(\eta)$ is given by Eq. (6.7). The ratios $q_L(\eta)/p_L(\eta)$ are too elaborate to be given here. They are given in Refs. 1, 3, and 5, and some of them are tabulated in Ref. 4. $\theta_L(\eta, \rho)$ is a power series in ρ, and near $\rho = 0$,

$$\theta_L(\eta, \rho) \approx \rho^{-L}/[(2L + 1)C_L(\eta)]. \tag{6.15}$$

Comparing Eq. (6.14) with Eq. (6.10), we find

$$A = 2\eta/C_0^2(\eta). \tag{6.16}$$

As we mentioned before, the coefficient g_{2l+1}^l multiplies a regular solution; therefore, from Eqs. (6.14) and (6.16),

$$g_{2l+1}^l = Af_0^l q_l(\eta)/p_l(\eta). \tag{6.17}$$

From Eq. (6.15) we can also determine the value of the coefficient g_0^l. Near $\rho = 0$

$$j_{-l-1}(\rho) \approx (-1)^l/[(2l + 1)!!\,\rho^{l+1}]. \tag{6.18}$$

On the basis of Eqs. (6.10), (6.14), (6.15), and (6.18) we have

$$g_0^l = (-1)^l(2l + 1)!!/[(2l + 1)C_L(\eta)]. \tag{6.19}$$

Now, with the aid of Eq. (6.10) and the recurrence relation of Eq. (6.13) and Eqs. (6.12), (6.16), (6.17), and (6.19), the irregular Coulomb wavefunction can be determined.

7. COMPUTATION OF THE COULOMB WAVEFUNCTION

We checked numerically the expansion of $F_0(\eta, \rho)$ for $1 \leq \rho \leq 20$ and $1 \leq \eta \leq 20$. In Fig. 1 are depicted the number of terms of the expansion in spherical Bessel functions, for which five-digit accuracy is achieved. The spherical Bessel functions were computed with the aid of a backward recurrence relation.[1] One can see that the number of terms depends mainly on ρ and less on η. It was previously assumed[3] that the spherical Bessel functions expansion is good for large ρ and smaller η. From our analysis it appears

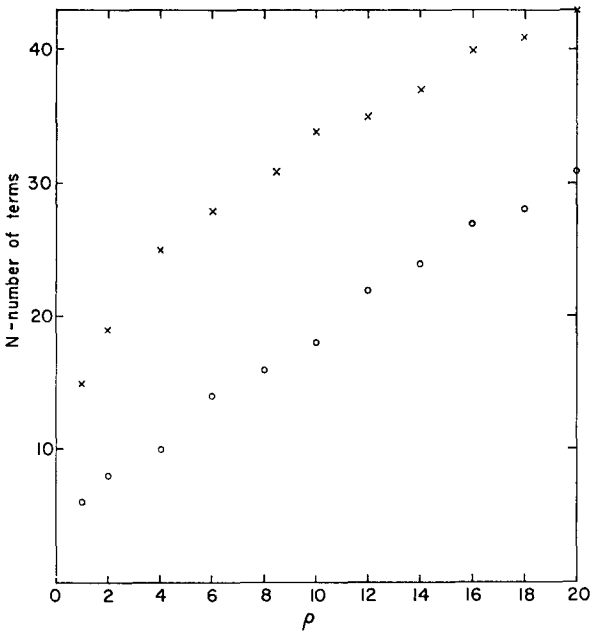

FIG. 1. The number of terms of the expansion of $F_0(\eta, \rho)$ in spherical Bessel functions, for which five-digit accuracy is achieved for two values of η. ○ depicts $\eta = 1$; ×, $\eta = 20$.

that for smaller ρ the convergence is faster and η can exceed ρ.

The irregular Coulomb wavefunction $G_0(\eta, \rho)$ can be calculated according to Eq. (6.10). The sum of the rhs of Eq. (6.10) can be calculated with almost the same accuracy as the regular Coulomb wavefunction, but for increasing values of $\rho\eta$ the two terms on the rhs of Eq. (6.10) almost cancel each other, and here the accuracy depends on the number of digits with which the calculations are performed. For example, in performing the calculations with seven digits of accuracy, one will get inaccurate or wrong answers for approximately $\rho\eta > 5$.

8. SUMMARY AND DISCUSSION

In the preceding sections all necessary formulas for finding the expansion coefficients of the expansion of the Schrödinger equation in terms of spherical Bessel functions are derived. In Sec. 4, we dealt with the expansion of a given function multiplied by a spherical Bessel function in terms of spherical Bessel functions. This we achieve by introducing the polynomials $A_n^\lambda(t)$ given by Eq. (4.2) and by finding the recursion relation (4.7) which determines the values of these polynomials. Next, the expansion coefficients are obtained using Eqs. (4.9)–(4.11).

Having these results, the solutions of the Schrödinger equation can be found using the recurrence relations (2.7) for the regular solution and (5.11) for the irregular solution. The expansion coefficients

of the regular solution are simply related to the phase shifts through Eq. (3.3).

This method is applied to find the expansion coefficients of the regular and irregular Coulomb wavefunctions in Sec. 6. In Sec. 7 we find that this expansion is also suitable for computing the regular Coulomb wavefunction. This example also indicates that the expansion in terms of spherical Bessel functions can be a useful representation of wavefunctions.

ACKNOWLEDGMENTS

The author would like to thank Professor R. A. Bryan for encouragement, and both him and Dr. J. Binstock for reading the manuscript and giving helpful suggestions.

* Work supported by the Air Force Office of Scientific Research, Office of Aerospace Research, U.S. Air Force, under Grant No. 69-1817.

† Present address: Institute of Theoretical Physics, University of Nijmegen, Nijmegen, Netherlands.

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (U.S. Department of Commerce, Natl. Bur. Stds., 1964), p. 452.

[2] G. N. Watson, *Theory of Bessel Functions* (University Press, Cambridge, 1944), Ch. 16.

[3] *Tables of Coulomb Wave Functions* (Natl. Bur. Stds., U.S. GPO, Washington, D.C., 1952), Vol. I. The expansion is due to P. M. Morse

[4] C. E. Fröberg, Rev. Mod. Phys. **27**, 399 (1954).

[5] F. L. Yost *et al.*, Phys. Rev. **49**, 174 (1936).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

An Approximate Solution of the Vacuum Static Case of Spherical Symmetry in Brans-Dicke Theory

M. N. MAHANTA AND D. R. K. REDDY
Department of Mathematics, Indian Institute of Technology, New Delhi 29, India

(Received 16 September 1970)

Starting from the usual variational principle, the gravitational field equations for the vacuum static case of spherical symmetry are obtained in Brans–Dicke [Phys. Rev. **124**, 925 (1961)] scalar–tensor theory following a technique due to Weyl [H. Weyl, *Space-Time-Matter* (Dover, New York, 1922)]. The field equations being highly nonlinear, an approximate solution to the second order is presented. It is observed that the results of gravitational redshift, deflection of light, and the rotation of the perihelion of Mercury are in agreement with the earlier results obtained by Brans, Dicke, and Heckmann. But the method, being simpler, can be used in solving other problems in the theory.

1. INTRODUCTION

Several attempts have been made in recent years to incorporate Mach's principle into the general theory of relativity. One of them is the theory by Brans and Dicke.[1] A scalar–tensor theory of gravitation has been developed in their paper.[1] A scalar ϕ is introduced by them into the usual variational principle of general relativity, viz.,

$$\delta\int[R + (16\pi G)L](-g)^{\frac{1}{2}}\,d^4x = 0, \qquad (1)$$

and a generalization of (1) is obtained in the form

$$\delta\int[\phi R + (16\pi)L - \omega\phi_{,i}\phi^{,i}/\phi](-g)^{\frac{1}{2}}d^4x = 0, \quad (2)$$

where R is the scalar curvature, L the Lagrangian density of matter, ω the dimensionless constant, and the velocity of light being considered as unity. Here ϕ plays the role of G^{-1}, G being the constant of gravitation. From (2) they have obtained field equations expressing the line element in the isotropic form.[2] Heckmann[3] also considered the spherically symmetric case of the above and presented an exact solution. In the above two cases the equations are very complicated and so a simpler solution is well worth discussion.

In this paper we have started from the same variational principle (2) and obtained field equations for the vacuum static case following a technique used first by Weyl[4] and then by Pauli.[5] The field equations being highly nonlinear, an approximate solution of the vacuum static case of spherical symmetry is obtained. It is found that the results obtained using the above solution for the three tests of general relativity are in agreement with the results of Brans–Dicke and others. It is realized that even though the solution is an approximate one, it is simpler and may be used in solving other problems in this theory. Also, in the existing earlier exact solutions, an approximation is what is required for practical calculations.

of the regular solution are simply related to the phase shifts through Eq. (3.3).

This method is applied to find the expansion coefficients of the regular and irregular Coulomb wavefunctions in Sec. 6. In Sec. 7 we find that this expansion is also suitable for computing the regular Coulomb wavefunction. This example also indicates that the expansion in terms of spherical Bessel functions can be a useful representation of wavefunctions.

ACKNOWLEDGMENTS

The author would like to thank Professor R. A. Bryan for encouragement, and both him and Dr. J. Binstock for reading the manuscript and giving helpful suggestions.

* Work supported by the Air Force Office of Scientific Research, Office of Aerospace Research, U.S. Air Force, under Grant No. 69-1817.

† Present address: Institute of Theoretical Physics, University of Nijmegen, Nijmegen, Netherlands.

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (U.S. Department of Commerce, Natl. Bur. Stds., 1964), p. 452.

[2] G. N. Watson, *Theory of Bessel Functions* (University Press, Cambridge, 1944), Ch. 16.

[3] *Tables of Coulomb Wave Functions* (Natl. Bur. Stds., U.S. GPO, Washington, D.C., 1952), Vol. I. The expansion is due to P. M. Morse

[4] C. E. Fröberg, Rev. Mod. Phys. **27**, 399 (1954).

[5] F. L. Yost *et al.*, Phys. Rev. **49**, 174 (1936).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

An Approximate Solution of the Vacuum Static Case of Spherical Symmetry in Brans-Dicke Theory

M. N. MAHANTA AND D. R. K. REDDY
Department of Mathematics, Indian Institute of Technology, New Delhi 29, India

(Received 16 September 1970)

Starting from the usual variational principle, the gravitational field equations for the vacuum static case of spherical symmetry are obtained in Brans–Dicke [Phys. Rev. **124**, 925 (1961)] scalar–tensor theory following a technique due to Weyl [H. Weyl, *Space-Time-Matter* (Dover, New York, 1922)]. The field equations being highly nonlinear, an approximate solution to the second order is presented. It is observed that the results of gravitational redshift, deflection of light, and the rotation of the perihelion of Mercury are in agreement with the earlier results obtained by Brans, Dicke, and Heckmann. But the method, being simpler, can be used in solving other problems in the theory.

1. INTRODUCTION

Several attempts have been made in recent years to incorporate Mach's principle into the general theory of relativity. One of them is the theory by Brans and Dicke.[1] A scalar–tensor theory of gravitation has been developed in their paper.[1] A scalar ϕ is introduced by them into the usual variational principle of general relativity, viz.,

$$\delta\int[R + (16\pi G)L](-g)^{\frac{1}{2}}\,d^4x = 0, \qquad (1)$$

and a generalization of (1) is obtained in the form

$$\delta\int[\phi R + (16\pi)L - \omega\phi_{,i}\phi^{,i}/\phi](-g)^{\frac{1}{2}}d^4x = 0, \quad (2)$$

where R is the scalar curvature, L the Lagrangian density of matter, ω the dimensionless constant, and the velocity of light being considered as unity. Here ϕ plays the role of G^{-1}, G being the constant of gravitation. From (2) they have obtained field equations expressing the line element in the isotropic form.[2] Heckmann[3] also considered the spherically symmetric case of the above and presented an exact solution. In the above two cases the equations are very complicated and so a simpler solution is well worth discussion.

In this paper we have started from the same variational principle (2) and obtained field equations for the vacuum static case following a technique used first by Weyl[4] and then by Pauli.[5] The field equations being highly nonlinear, an approximate solution of the vacuum static case of spherical symmetry is obtained. It is found that the results obtained using the above solution for the three tests of general relativity are in agreement with the results of Brans–Dicke and others. It is realized that even though the solution is an approximate one, it is simpler and may be used in solving other problems in this theory. Also, in the existing earlier exact solutions, an approximation is what is required for practical calculations.

2. FIELD EQUATIONS

Following Pauli,[5] we consider the line element for the static spherically symmetric case in the form

$$(ds)^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 + l(x^1\,dx^1 + x^2\,dx^2 + x^3\,dx^3) + g_{44}(dx^4)^2. \tag{3}$$

The scalar curvature R can be calculated from Ref. 4, $R_{ik} = [R_{22}]\delta_i^k + ([R_{11}] - [R_{22}])x^i x^k/r^2$, $i, k = 1, 2, 3$, as

$$[R_{11}] = \Delta r^{-2} g_{44}^{-1}\tfrac{1}{2}(d/dr)(r^2 g'_{44}\Delta^{-1}) - 2\Delta' r^{-1}\Delta^{-1},$$
$$[R_{22}] = -r^{-2}\Delta^{-1}(d/dr)(r g_{44}\Delta^{-1}) - r^{-2}, \tag{4}$$
$$R_{44} = -g_{44} r^{-2}\Delta^{-1}\tfrac{1}{2}(d/dr)(r^2 g'_{44}\Delta^{-1}),$$

where

$$h^2 = 1 + l/r^2, \quad \Delta = h(-g_{44})^{\frac{1}{2}} = (-g)^{\frac{1}{2}}. \tag{5}$$

(Here dashes denote differentiation w.r.t. r.) and $[R_{ii}]$ is the value of R_{ii} at the point $x^1 = r$, $x^2 = x^3 = 0$. Now R is given by

$$R = g^{ik}R_{ik} = g^{11}R_{11} + g^{22}R_{22} + g^{33}R_{33} + g^{44}R_{44},$$

which is equal to ($x^1 = r$, $x^2 = x^3 = 0$)

$$-r^{-2}\Delta^{-1}(d/dr)(r^2 g'_{44}\Delta^{-1}) + 2\Delta' r^{-1}\Delta^{-3} g_{44} - 2r^{-2}\Delta^{-1}(d/dr)(r g_{44}\Delta^{-1}) - 2r^{-2}. \tag{6}$$

Now returning to the variational principle, we can write (2) as

$$\delta\int\left\{\phi\left[-\frac{1}{r^2\Delta}\frac{d}{dr}\left(\frac{r^2 g'_{44}}{\Delta}\right) + \frac{2}{r}\frac{\Delta'}{\Delta^3}g_{44} - \frac{2}{r^2\Delta}\frac{d}{dr}\left(\frac{r g_{44}}{\Delta}\right) - \frac{2}{r^2}\right] + \frac{\omega g_{44}}{\Delta^2}\frac{\phi'^2}{\phi}\right\} r^2\Delta\,dr = 0, \tag{7}$$

where

$$\phi_{,i}\phi^{,i} = g^{11}\phi'^2\phi^{-1} = h^{-2}\phi'^2\phi^{-1} = -g_{44}\phi'^2\Delta^{-2}\phi^{-1}$$

and the space is empty $L = 0$. Also

$$(-g)^{\frac{1}{2}} = \Delta$$

and

$$d^4x = d^4x^4 \cdot d\Omega \cdot r^2\,dr$$

(where $d\Omega$ is the elementary solid angle at the origin).[4] Now variation with respect to ϕ, Δ, and g_{44} in (7), respectively, leads to the following field equations:

$$-(d/dr)(r^{-2}g'_{44}\Delta^{-1}) + 2r g_{44}\Delta'\Delta^{-2} - 2(d/dr)(r g_{44}\Delta^{-1}) - 2\Delta = \omega r^2 g_{44}\Delta^{-1}\phi'^2\phi^{-2} + 2\omega(d/dr)(r^2 g_{44}\Delta^{-1}\phi'\phi^{-1}), \tag{8}$$

$$-2r g'_{44}\Delta^{-2} - 2g_{44}\Delta^{-2} - 2 = \omega r^2 g_{44}\phi'^2\Delta^{-2}\phi^{-2}, \tag{9}$$

$$2r\Delta'\Delta^{-2} = -\omega r^2\phi'^2\Delta^{-1}\phi^{-2}. \tag{10}$$

3. SOLUTIONS OF FIELD EQUATIONS

It can be easily verified that, by putting $\phi = \text{const}$, the above equations reduce to the well-known case of the Schwarzschild solution as obtained by Weyl[4] and Pauli.[5]

We try the solution

$$(d/dr)(r^2 g_{44}\phi'\Delta^{-1}\phi^{-1}) = 0. \tag{11}$$

And so from (8)

$$(d/dr)(r^2 g'_{44}\Delta^{-1}) + 2r g_{44}\Delta'\Delta^{-2} - 2(d/dr)(r g_{44}\Delta^{-1}) - 2\Delta = \omega r^2 g_{44}\phi'^2\Delta^{-1}\phi^2. \tag{8'}$$

Also, we have

$$-2r g'_{44}\Delta^{-2} - 2g_{44}\Delta^{-2} - 2 = \omega r^2 g_{44}\Delta^{-2}\phi'^2\phi^{-2} \tag{9'}$$

and

$$2r\Delta'\Delta^{-2} g_{44} = -\omega r^2 g_{44}\Delta^{-1}\phi'^2\phi^{-2}. \tag{10'}$$

Equations (8'), (9'), and (10') constitute only two independent equations which can be taken as

$$(d/dr)(r g_{44}\Delta^{-1}) = -\Delta,$$
$$(d/dr)(r^2 g'_{44}\Delta^{-1}) = 4r g_{44}\Delta'\Delta^{-2}. \tag{12}$$

From these equations we may obtain the following equation (after some reduction):

$$y^2 y'' = Ar^3 y' \tag{13}$$

where

$$y = r g_{44}\Delta^{-1}, \quad y' = dy/dr, \quad y'' = d^2y/dr^2,$$

and A is a constant.

In view of the difficulty of finding an exact solution of (13), we fell back on (12) for an approximate solution correct up to the second order in $1/r$, by using the method of successive approximation.

Let us consider the field equations (8), (9), and (10). When $\phi = \text{const}$, we get

$$\Delta = 1, \quad g_{44} = -1 + G/r, \quad G = \text{const}.$$

Now substituting these values in rhs's of Eqs. (12) and integrating with respect to r, we get

$$g_{44}/\Delta = -1 + 2\lambda/r, \quad \lambda = \text{const}, \tag{14}$$

$$g'_{44}/\Delta = C/r^2, \quad C = \text{const}. \tag{15}$$

Eliminating Δ between (14) and (15) and integrating with respect to r, we get

$$g_{44} = -(1 - 2\lambda/r)^{-C/2\lambda}. \tag{16}$$

Again from (14) and (16) we get

$$\Delta = (1 - 2\lambda/r)^{-(C/2\lambda+1)}.$$

Using these values of g_{44} and Δ in (5), we get

$$h^2 = g_{11} = (1 - 2\lambda/r)^{-(C/2\lambda+2)}.$$

Again from (11) ϕ can be determined as

$$\phi = p(1 - 2\lambda/r)^{-B/2\lambda}, \quad B = \text{const}, \quad p = \text{const}.$$

Thus the solution to the second approximation can be written as

$$\begin{aligned} g_{44} &= -(1 - 2\lambda/r)^{-C/2\lambda}, \\ g_{11} &= (1 - 2\lambda/r)^{-(C/2\lambda+2)}, \\ \phi &= p(1 - 2\lambda/r)^{-B/2\lambda}, \end{aligned} \tag{17}$$

where λ, C, B, and ϕ are constants which are to be determined.

Now in (17) expanding g_{44} to the second order and ϕ and g_{11} to the first order in $1/r$, we get

$$\begin{aligned} g_{44} &\approx -1 + (-C/r) + (-C/2)(-C - 2\lambda)(1/r^2), \\ g_{11} &\approx 1 + (C + 2\lambda)(1/r), \\ \phi &\approx p(1 + B/r). \end{aligned}$$

Now equating the coefficients of $1/r$ to Brans–Dicke [1] weak-field approximations, we get (since to the first order these values should be the same)

$$\begin{aligned} -C &= 2M\phi_0^{-1}[1 + 1/(3 + 2\omega)], \\ C + 2\lambda &= 2M\phi_0^{-1}[1 - 1/(3 + 2\omega)], \end{aligned}$$

so that $2\lambda = 4M\phi_0^{-1}$, $\lambda = 2M\phi_0^{-1}$,

$$p = \phi_0, \quad \text{and} \quad B = 2M\phi_0^{-1}(3 + 2\omega)^{-1}.$$

Here M stands for the finite mass of the visible universe and ϕ_0 is a constant and is to be computed to first order in mass densities.[1]

Hence (17) now becomes

$$\begin{aligned} g_{44} &\approx -1 + 2M\phi_0^{-1}r^{-1}[1 + 1/(3 + 2\omega)] \\ &\quad + 4M^2\phi_0^{-2}r^{-2}(4 + 2\omega)(1 + \omega)(3 + 2\omega)^{-2}, \\ g_{11} &\approx 1 + 2M\phi_0^{-1}r^{-1}[1 - 1/(3 + 2\omega)], \\ \phi &\approx \phi_0[1 + (2M\phi_0^{-1})(3 + 3\omega)^{-1}r^{-1}]. \end{aligned} \tag{17'}$$

Now putting[1]

$$G_0 = \phi_0^{-1}(4 + 2\omega)(3 + 2\omega)^{-1}, \tag{18}$$

we get

$$\begin{aligned} g_{44} &\approx -1 + (2MG_0/r) \\ &\quad + (2M^2G_0^2/r^2)((1 + \omega)/(2 + \omega)), \\ g_{11} &\approx 1 + 2MG_0r^{-1}(1 + \omega)(2 + \omega)^{-1}, \\ \phi &\approx \phi_0[1 + (2MG_0)(4 + 2\omega)^{-1}r^{-1}]. \end{aligned} \tag{17''}$$

Thus an approximate solution, correct up to $1/r^2$ to the field equations (8), (9), and (10) for the metric (3), can be written in the form (17″).

4. THE THREE TESTS OF GENERAL RELATIVITY

The above solution, to the first order, is sufficient to discuss the gravitational redshift and the deflection of light. But to discuss the rotation of the perihelion of Mercury's orbit requires a solution, to the second order, for g_{44}.

The gravitational redshift is given by[6]

$$1 - ds'/ds = 1 - (g'_{44}/g_{44})^{\frac{1}{2}} = MG_0g^{-1} \tag{19}$$

(here g'_{44} does not mean differentation of g_{44},)

where g is the sun's radius and G_0 is given by (18).

The deflection of light is seen to be[1]

$$g_{11}/g_{44} = \delta\theta = 4G_0MR^{-1}(3 + 2\omega)(4 + 2\omega)^{-1}, \tag{20}$$

where R denotes the closest approach distance of the light ray to the sun's mass. Thus it differs from the general relativity value by the factor in the brackets.

The derivation of the precision of the perihelion of Mercury is a crucial test for general relativity theory which requires the knowledge of g_{44} in the line element to second order in MG_0/r. (In other words nonlinear terms in the field equations are required.)

The motion of a planet (assumed to be infinitesimal in comparison with the sun's mass) is represented by a geodesic world line of its four equations,[4,5]

$$\frac{d^2x^i}{ds^2} + \begin{Bmatrix} \alpha\beta \\ i \end{Bmatrix} \frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0.$$

The equation corresponding to $i = 4$ gives, for the static gravitational field, the energy integral

$$g_{44}\frac{dx^4}{ds} = \text{const}.$$

Now by using the law of areas

$$x^1(dx^2/ds) - x^2(dx^1/ds) = \text{const}$$

and introducing

$$x^1 = r\cos\psi, \quad x^2 = r\sin\psi,$$

the integral of area is

$$r^2(d\psi/ds) = \text{const} = b. \tag{21}$$

Now the energy integral becomes

$$g_{44}[1 - g_{11}(dr/ds)^2 - r^2(d\psi/ds)^2] = \text{const} = E,$$

where E is a constant of motion which can be taken to be unity to the lowest order in $1/r$.[7–9]

Now using (21) and substituting for g_{44}, we arrive

at the orbit of the planet

$$(d\rho/d\psi)^2 = 2MG_0\rho b^{-2}(E - 1 - \omega)(2 + \omega)^{-1} + (E + 1)b^{-2} - \rho^2[1 - 2M^2G_0^2b^{-2}(2 + \omega)^{-1}] + 2MG_0(1 + \omega)(2 + \omega)^{-1}\rho^3, \quad (22)$$

where $\rho = 1/r$ and b^2 is also constant of motion such that $b^2 = MG_0a(1 - e^2)$ (a being the semimajor axis for the orbit and e the eccentricity of the orbit). Now ψ can be expressed in terms of ρ by an elliptic integral of the first kind and, hence, conversely, ρ is an elliptic function of ψ. So we have

$$\psi = \int d\theta\{2MG_0(1 + \omega)/(2 + \omega) \times [\rho_0 - (\rho_1 + \rho_2)/2 - (\rho_1 - \rho_2)/2 \cos\theta]\}^{-\frac{1}{2}}, \quad (23)$$

where ρ_0, ρ_1, ρ_2 are the positive roots of (23) such that

$$1/\rho_1 = a(1 - e), \quad 1/\rho_2 = a(1 + e),$$

so that

$$\rho_1 + \rho_2 = 2a^{-1}(1 - e^2)^{-1} = 2MG_0b^{-2} \quad (24)$$

and

$$\rho_0 + \rho_1 + \rho_2 = [1 - (2M^2G_0^2)(2 + \omega)^{-1}b^{-2} \times (2 + \omega)(2MG_0)^{-1}(1 + \omega)^{-1}].$$

The perihelion is characterized by the values $\theta = 0, 2\pi$, etc. The increase of the azimuth ψ after a full revolution from perihelion to perihelion is furnished by the above integral taken between the limits 0 and 2π; with sufficient accuracy this increase may be set

$$2\pi\{2MG_0(1 + \omega)(2 + \omega)^{-1}[\rho_0 - (\rho_1 + \rho_2)/2]\}^{-\frac{1}{2}} = 2\pi\{1 - 6M^2G_0^2b^{-2} \times [(1 + \omega)(2 + \omega)^{-1} + (6 + 3\omega)^{-1}]\}^{-\frac{1}{2}} = 2\pi[1 + 3M^2G_0^2b^{-2}(3\omega + 4)(3\omega + 6)^{-1}].$$

The advance of the perihelion per revolution is

$$\delta\psi = 6\pi M^2G_0^2b^{-2}(3\omega + 4)(3\omega + 6)^{-1}, \quad (25)$$

which is simply the general relativity value multiplied by the factor $(3\omega + 4)(3\omega + 6)^{-1}$.

5. CONCLUSION

Thus, even though this is an approximate solution we see that the results to the second order in $1/r$ are in agreement with the results of Brans–Dicke[1] and Heckmann.[2] But this method of solution of the field equations is simpler than earlier methods.

ACKNOWLEDGMENTS

The authors express their gratitude to Professor C. Brans for his valuable communication to the junior author, and also to Dr. P. K. Bhattacharyya and Dr. S. R. K. Iyengar of the Indian Institute of Technology, Delhi, for helpful discussions. Financial support by the Council of Scientific and Industrial Research, India, is gratefully acknowledged.

[1] C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1961).
[2] C. Brans, Ph.D. thesis, Princeton University, 1961.
[3] P. Jordan, Schwerkraft und Weltall (Friedrich Vieweg und Sohn, Braunschweig, 1955).
[4] H. Weyl, *Space-Time-Matter* (Dover, New York, 1922).
[5] W. Pauli, *Theory of Relativity* (Pergamon, New York, 1958).
[6] J. L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960).
[7] R. Adler, M. Bazin, and M. Schiffer, *Introduction to General Relativity* (McGraw-Hill, New York, 1965).
[8] L. D. Landau and E. M. Lifshitz, *Mechanics* (Pergamon, New York, 1960).
[9] R. F. O. Connel, Am. J. of Phys. **36**, 757 (1968).

Asymptotic Equivalence of Equilibrium Ensembles of Classical Statistical Mechanics*

LUIGI GALGANI, LUCIANO MANZONI,† AND ANTONIO SCOTTI†
Istituto di Fisica dell'Università, Milano, Italia

(Received 19 June 1970)

We start from the solution of the problem of the thermodynamic limit for microcanonical entropy as given by Ruelle and extend his method to solve the analogous problem of consistency, within the same ensemble, for the Legendre transforms of entropy. We solve then in a unified way the problem of the asymptotic equivalence of ensembles.

I. INTRODUCTION

The motivation of the present paper is the following. On the one hand there is now available the solution of the problem of the thermodynamic limit for microcanonical entropy as given by Ruelle,[1] which in particular provides a rigorous treatment at the boundary of the region where the limit entropy is defined. On the other hand the same author does not prove the asymptotic equivalence of ensembles in a unified way.

A previous attempt[2] at giving such a unified treatment was not completely satisfactory because of lack of rigor in dealing with the behavior at the above mentioned boundary. Moreover, a great simplification obtains if, following Griffiths,[3] one first demonstrates the possibility of exchanging the thermodynamic limit with the taking of the Legendre transforms of microcanonical entropy, and then passes to the equivalence of ensembles.[4]

In Sec. II we consider the microcanonical ensemble: The results by Ruelle on the limit microcanonical entropy are recalled (Theorem 1) and corresponding results on the Legendre transforms are proven (corollary). The asymptotic equivalence of ensembles is then established in Sec. III (Theorem 2). Details of the proof of the corollary are given in the Appendix.

II. THERMODYNAMIC LIMIT FOR THE MICROCANONICAL ENSEMBLE

We consider a classical continuous system composed of one species of particles of mass $m > 0$ having only translational degrees of freedom in a ν dimensional space. For each integer $n > 0$ and $x_1, \cdots, x_n$, $p_1, \cdots, p_n \in R^\nu$ the Hamiltonian

$$H(x_1, \cdots, x_n, p_1, \cdots, p_n) = \sum_{i=1}^{n} \frac{p_i^2}{2m} + U(x_1, \cdots, x_n)$$

is defined. U is the potential energy, which is assumed to be Lebesgue measurable with values in $R \cup \{\infty\}$, invariant under permutations of its arguments and under translations. In addition, the following well-known properties of stability and weak tempering will be considered:

(i) There exist $B \geq 0$ such that $U(x_1, \cdots, x_n) \geq -nB$ for all $n > 0$ and $x_1, \cdots, x_n \in R^\nu$ (stability);

(ii) there exist $\lambda > \nu$ and $R_0 > 0$, $A \geq 0$, such that

$$U(x_1', \cdots, x_{n_1}', x_1'', \cdots, x_{n_2}'') - U(x_1', \cdots, x_{n_1}') - U(x_1'', \cdots, x_{n_2}'') \leq A n_1 n_2 r^{-\lambda}$$

whenever $|x_j'' - x_i'| \geq r \geq R_0$ for all $i = 1, \cdots, n_1$, $j = 1, \cdots, n_2$ (weak tempering).

Let Λ be a bounded Lebesgue measurable subset of R^ν with volume $V(\Lambda)$; for $E \in R$ we introduce the microcanonical partition function

$$\Omega(\Lambda, n, E) = \frac{1}{n!} \int_{R^{n\nu}} dp_1 \cdots dp_n \int_{\Lambda^n} dx_1 \cdots dx_n \times \vartheta[E - H(x_1, \cdots, x_n, p_1, \cdots, p_n)], \tag{1}$$

where ϑ is the characteristic function of the interval $(0, \infty)$. We further define the microcanonical entropy

$$S(\Lambda, n, E) = \log \Omega(\Lambda, n, E), \tag{2}$$

where S may take the value $-\infty$ (if $\Omega = 0$).

The following theorem has been established by Ruelle.

Theorem 1 (thermodynamic limit for microcanonical entropy):

Let $S(\Lambda, n, E)$ be defined by (1), (2) for a stable tempered interaction. There exist

(a) $\rho_{cp} > 0$ or $\rho_{cp} = +\infty$,

(b) a convex continuous function ϵ_0 on the interval $[0, \rho_{cp})$ such that $\epsilon_0(0) = 0$ and $\epsilon_0(\rho) \geq -\rho B$,

(c) a concave continuous function s on the region $\Theta = \{(\rho, \epsilon): 0 \leq \rho < \rho_{cp}, \epsilon > \epsilon_0(\rho)\}$, increasing in ϵ for fixed ρ and such that $s(0, \epsilon) = 0$ for $\epsilon > 0$.

Let $\Lambda \to \infty$ in the sense of Fisher and

$$\lim_{\Lambda \to \infty} V(\Lambda)^{-1} n = \rho, \quad \lim_{\Lambda \to \infty} V(\Lambda)^{-1} E = \epsilon;$$

(α) if $(\rho, \epsilon) \in \Theta$. then

$$\lim_{\Lambda \to \infty} V(\Lambda)^{-1} S(\Lambda, n, E) = s(\rho, \epsilon),$$

(β) if (ρ, ϵ) belongs to the boundary of Θ, then

$$\overline{\lim_{\Lambda \to \infty}} V(\Lambda)^{-1} S(\Lambda, n, E) \leq s^*(\rho, \epsilon),$$

where $s^*(\rho, \epsilon) = \overline{\lim}\, s(\rho^*, \epsilon^*)$ when $(\rho^*, \epsilon^*) \in \Theta$,

$$(\rho^*, \epsilon^*) \to (\rho, \epsilon),$$

(γ) if (ρ, ϵ) belongs to the complement of the closure of Θ, then

$$\lim_{\Lambda \to \infty} V(\Lambda)^{-1} S(\Lambda, n, E) = -\infty.$$

We come now to the problem of the thermodynamic limit for the Legendre transforms of entropy within the microcanonical ensemble. Let $\beta > 0$, $\mu \in R$; we introduce the functions

$$\Phi^+(\Lambda, n, \beta) = \sup_E e^{-\beta E} \Omega(\Lambda, n, E), \tag{3}$$

$$\Psi^+(\Lambda, \mu, E) = \sup_n e^{-\mu n} \Omega(\Lambda, n, E), \tag{4}$$

$$\Xi^+(\Lambda, \mu, \beta) = \sup_n e^{-\mu n} \Phi^+(\Lambda, n, \beta), \tag{5}$$

where the suprema are over $E \in R$ or integers $n > 0$; the logarithms of such functions are the generalized Legendre transforms of $S(\Lambda, n, E)$.

Define $\epsilon_{\inf} = \inf_{0 \leq \rho < \rho_{cp}} \epsilon_0(\rho)$; it is $\epsilon_{\inf} \leq 0$ or $\epsilon_{\inf} = -\infty$. For $\epsilon > \epsilon_{\inf}$ define further $\rho_1(\epsilon)$, $\rho_2(\epsilon)$ as the abscissas of the points of the boundary of Θ having ordinate ϵ; it is $0 \leq \rho_1(\epsilon) < \rho_2(\epsilon) < \rho_{cp}$. Let $0 \leq \rho < \rho_{cp}$, $\epsilon > \epsilon_{\inf}$, $\beta > 0$; $\mu \in R$; then the Legendre transforms of $s(\rho, \epsilon)$,

$$f(\beta, \rho) = \sup_{\epsilon > \epsilon_0(\rho)} [s(\rho, \epsilon) - \beta\epsilon], \tag{6}$$

$$g(\mu, \epsilon) = \sup_{\rho_1(\epsilon) < \rho < \rho_2(\epsilon)} [s(\rho, \epsilon) - \mu\rho], \tag{7}$$

$$p(\mu, \beta) = \sup_{0 \leq \rho < \rho_{cp}} [f(\rho, \beta) - \mu\rho], \tag{8}$$

are defined. The following functions,

$$f^*(\beta) = \overline{\lim_{\rho \to \rho_{cp}, \rho < \rho_{cp}}} f(\rho, \beta), \tag{9}$$

$$g^*(\mu) = \overline{\lim_{\epsilon \to \epsilon_{\inf}, \epsilon > \epsilon_{\inf}}} g(\mu, \epsilon), \tag{10}$$

will also be considered. From the properties of s stated in Theorem 1, the following properties of f, g, p, f^*, g^* are easily proved[5]: f is a continuous concave function of ρ and a continuous convex function of β, g is a continuous convex function of μ and a continuous concave function of ϵ, p is a continuous convex function of μ and β, f^* is a continuous convex function of β, g^* is a continuous convex function of μ, and in addition

$$f(\rho, \beta) \geq s^*[\rho, \epsilon_0(\rho)] - \beta\epsilon_0(\rho), \tag{11}$$

$$g(\mu, \epsilon) \geq s^*[\rho_i(\epsilon), \epsilon] - \mu\rho_i(\epsilon), \quad i = 1, 2. \tag{12}$$

The following corollary of Theorem 1 is proved by a modified but closely related version of the method used by Ruelle in another context.[6]

Corollary (thermodynamic limit for the Legendre transforms of microcanonical entropy):

Let $\Phi^+(\Lambda, n, \beta)$, $\Psi^+(\Lambda, \mu, \epsilon)$, $\Xi^+(\Lambda, \mu, \beta)$, $f(\rho, \beta)$, $g(\mu, \epsilon)$, $p(\mu, \beta)$, $f^*(\beta)$, $g^*(\mu)$ be defined by (3), (4), (5), (6), (7), (8), (9), (10). Let $\Lambda \to \infty$ in the sense of Fisher.

(i) Let $V(\Lambda)^{-1} n \to \rho$:

(α) If $0 \leq \rho < \rho_{cp}$, then

$$\lim_{\Lambda \to \infty} V(\Lambda)^{-1} \log \Phi^+(\Lambda, n, \beta) = f(\rho, \beta);$$

(β) if $\rho = \rho_{cp}$, then

$$\overline{\lim_{\Lambda \to \infty}} V(\Lambda)^{-1} \log \Phi^+(\Lambda, n, \beta) \leq f^*(\beta);$$

(γ) if $\rho > \rho_{cp}$, then

$$\lim_{\Lambda \to \infty} V(\Lambda)^{-1} \log \Phi^+(\Lambda, n, \beta) = -\infty.$$

(ii) Let $V(\Lambda)^{-1} E \to \epsilon$:

(α) If $\epsilon > \epsilon_{\inf}$, then

$$\lim_{\Lambda \to \infty} V(\Lambda)^{-1} \log \Psi^+(\Lambda, \mu, E) = g(\mu, \epsilon);$$

(β) if $\epsilon = \epsilon_{\inf}$, then

$$\overline{\lim_{\Lambda \to \infty}} V(\Lambda)^{-1} \log \Psi^+(\Lambda, \mu, E) \leq g^*(\mu);$$

(γ) if $\epsilon < \epsilon_{\inf}$, then

$$\lim_{\Lambda \to \infty} V(\Lambda)^{-1} \log \Psi^+(\Lambda, \mu, E) = -\infty.$$

(iii) $\lim_{\Lambda \to \infty} V(\Lambda)^{-1} \log \Xi^+(\Lambda, \mu, \beta) = p(\mu, \beta).$

Details of the proof are given in the Appendix.

III. THERMODYNAMIC LIMIT FOR THE OTHER ENSEMBLES

Let $\beta > 0$, $\mu \in R$. The partition functions of the ensembles of interest are

$$\Phi(\Lambda, n, \beta) = \beta \int_{-\infty}^{+\infty} dE e^{-\beta E} \Omega(\Lambda, n, E), \tag{13}$$

$$\Psi(\Lambda, \mu, E) = \sum_{n=1}^{\infty} e^{-\mu n} \Omega(\Lambda, n, E), \tag{14}$$

$$\Xi(\Lambda, \mu, \beta) = \sum_{n=1}^{\infty} e^{-\mu n} \Phi(\Lambda, n, \beta). \tag{15}$$

The following inequalities,[7]

$$\begin{aligned}\Phi(\Lambda, n, \beta) &\geq e^{-\beta E}\Omega(\Lambda, n, E),\\ \Psi(\Lambda, \mu, E) &\geq e^{-\mu n}\Omega(\Lambda, n, E),\\ \Xi(\Lambda, \mu, \beta) &\geq e^{-\mu n}\Phi(\Lambda, n, \beta),\end{aligned} \tag{16}$$

will be used in the proof of the theorem below.

Theorem 2 (asymptotic equivalence of ensembles): Let $\Phi(\Lambda, n, \beta)$, $\Psi(\Lambda, \mu, E)$, $\Xi(\Lambda, \mu, \beta)$, $f(\rho, \beta)$, $g(\mu, \epsilon)$, $p(\mu, \beta)$, $f^*(\beta)$, $g^*(\mu)$ be defined by (13), (14), (15), (6), (7), (8), (9), (10). Then, that which in the statement of the corollary of Sec. II is predicated of the daggered quantities is true for the corresponding undaggered ones.

Proof: Consider first case (i). It is, for all n and all $\beta > 0$,

$$e^{-\beta E}\Omega(\Lambda, n, E) \leq \Phi^+(\Lambda, n, \beta) \leq \Phi(\Lambda, n, \beta);$$

using the first inequality in the form $\Omega(\Lambda, n, E) \leq \Phi^+(\Lambda, n, \beta)e^{\beta E}$, we get from (9), for any $\beta' < \beta$,

$$\begin{aligned}\Phi(\Lambda, n, \beta) &\leq \beta\Phi^+(\Lambda, n, \beta)\int_{-nB}^{+\infty} e^{(\beta'-\beta)E}\,dE\\ &= \Phi^+(\Lambda, n, \beta')\frac{\beta}{\beta - \beta'}e^{(\beta-\beta')nB},\end{aligned}$$

so that

$$\begin{aligned}\Phi^+(\Lambda, n, \beta) &\leq \Phi(\Lambda, n, \beta)\\ &\leq \Phi^+(\Lambda, n, \beta')\frac{\beta}{\beta - \beta'}e^{(\beta-\beta')nB},\end{aligned}$$

which altogether concludes the proof by referring to part (i) of the corollary and to the continuity of f and f^* as functions of β.

The other two cases are treated in an analogous way.

IV. CONCLUSION

The results obtained in this paper can be summarized as follows:

(i) Within microcanonical theory a corollary to the fundamental theorem of Ruelle on the existence of the limit entropy (Theorem 1) is given. It guarantees the possibility of exchanging the limit with the taking of the Legendre transforms of entropy.

(ii) The asymptotic equivalence of the ensembles is established in a straightforward and unified way (Theorem 2).

We have explicitly considered four ensembles. The application of our method to the others is straightforward.

APPENDIX: PROOF OF THE COROLLARY OF SECTION II

Consider for example case i(α). Given $\delta > 0$, from Theorem 1(α) it follows that for sufficiently large Λ one can find $n(\Lambda)$ such that

$$e^{-\mu n(\Lambda)}\Omega(\Lambda, n(\Lambda), E) \geq e^{V(\Lambda)[g(\mu,\epsilon)-\delta]}, \tag{A1}$$

so that, *a fortiori*, it is

$$\Psi^+(\Lambda, \mu, E) \geq e^{V(\Lambda)[g(\mu,\epsilon)-\delta]}. \tag{A2}$$

On the other hand, for sufficiently large Λ and all n it is also

$$e^{-\mu n}\Omega(\Lambda, n, E) \leq e^{V(\Lambda)[g(\mu,\epsilon)+\delta]}, \tag{A3}$$

so that

$$\Psi^+(\Lambda, \mu, E) \leq e^{V(\Lambda)[g(\mu,\epsilon)+\delta]}. \tag{A4}$$

Indeed, if (A3) were not true, there should exist a sequence

$$(\Lambda_i, n_i, E_i), \quad V(\Lambda_i) \to \infty, \quad V(\Lambda_i)^{-1}E_i \to \epsilon > \epsilon_{\inf}$$

such that

$$e^{-\mu n_i}\Omega(\Lambda_i, n_i, E_i) > e^{V(\Lambda_i)[g(\mu,\epsilon)+\delta]},$$

considering then all possible limit points of $V(\Lambda_i)^{-1}n_i$, we would contradict Theorem 1 (α), (β), or (γ); in particular the contradiction to (β) is a consequence of (11).

From (A2) and (A4), part i(α) of the corollary immediately follows. All other cases are treated in an analogous way.

* Work supported in part by C.N.R.

† At C.C.R. Euratom, Ispra, Italy.

[1] D. Ruelle, *Statistical Mechanics, Rigorous Results* (Benjamin, New York, 1969), Chap. 3.

[2] L. Galgani, A. Scotti, and F. Valz Gris, Physica, **47**, 601 (1970).

[3] R. B. Griffiths, J. Math. Phys. **10**, 1447 (1965).

[4] L. Galgani, Proceedings of the Crete Summer School (1969).

[5] Consider for example the cases concerning f. The stated properties of f are well-known properties of Legendre transforms: see L. Galgani and A. Scotti, Physica **42**, 242 (1969)—the proofs given there may easily be modified in order to insure continuity. The stated property of f^* is an immediate consequence of the convexity of f as function of β. Finally, (11) is proven by remarking that, by the continuity of f as function of ρ, given $\rho \in [0, \rho_{cp})$ and $\delta > 0$ there exists a neighborhood I of ρ such that $f(\rho, \beta) \geq \sup_{\rho' \in I} f(\rho', \beta) - \delta$, while

$$\sup_{\rho' \in I} f(\rho', \beta) = \sup_{\rho' \in I} \sup_{\epsilon > \epsilon_0(\rho')} [s(\rho', \epsilon) - \beta\epsilon] \geq s^*[\bar{\rho}, \epsilon_0(\bar{\rho})] - \beta\epsilon_0(\bar{\rho})$$

for any $\bar{\rho} \in I$.

[6] D. Ruelle, Ref. 1, Sec. 3.4.1.

[7] The first of these inequalities may be proven by integration by parts, as in Ref. 2.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Scattering by a Nonspherical Potential*

MING CHIANG LI
Department of Physics, Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

(Received 26 October 1970)

The scattering of a simple nonspherical potential with spheroidal symmetry is discussed. The spheroidal partial wave analysis method is used. It is found that the spheroidal phase shifts have the similar physical meaning as the conventional spherical phase shifts. The integral equation for the spheroidal phase shift is given. The scattering by a discontinuous spheroidal potential is also presented. The WKB approximation is used to evaluate the spheroidal phase shift at high energy. For small angle scattering our formulation reduces to a simple form as in the eikonal description.

1. INTRODUCTION

A study of a nonspherical potential scattering problem is presented here. We shall only discuss the potential with oblate spheroidal symmetry,[1] not just for its mathematical simplicity, but also for some deep-rooted physical reasons.

Up till now, the most often used potential in the scattering theory is a spherical one, or a modification, for example, through spin–orbit coupling. The choice is suitable for systems where the relative distance between the objects is large compared to their particular sizes, since the potential between them will be essentially spherical. Such a potential will not violently disturb the internal states of the system in the scattering process; thus the important fundamental conservation laws can be directly and explicitly implemented without having to introduce an excessive number of variables. In the special case for pointlike objects with some intrinsic properties, the spherical potential is in fact the only possible choice, since the relative distance is large compared to their zero spatial extent and no other internal excitation is available.

However, most physical systems are more complex. In a high energy scattering process, even the elementary particles are no longer treated as pointlike objects, and the relative potential between them depends not only on their distance but also on their shapes, mass distribution, internal structures, and motions. The effect of the above dependences becomes more pronounced at small distances relative to the size of the objects. This leads us to the belief that an investigation of a nonspherical scattering problem will be worthwhile and should provide us with some new insights into the problems of high energy scattering processes.

For practical purposes and to avoid overcomplexity we have disregarded certain physical phenomena associated with a nonspherical potential, namely the change in the shape and mass distribution of the objects, etc. Consequently, the nonspherical potential we choose to study is that of an oblate spheroidal symmetry. Such a potential is closely related to a spherical one, as the latter may be considered as a special type of spheroidal potential. The choice was made on the basis of a new approach to the high energy scattering process of elementary particle physics.

It has been suggested[2] that one should treat elementary particles as extended objects in the high energy scattering process, rather than a mathematical point with some intrinsic properties. This idea is quite closely related to the quark model approach in which the elementary particles are no longer considered as elementary, but as composite. In the quark model one often stresses the internal composition of the constituents. In the extended object model, which sometimes is called the droplet model, the spatial distribution and character of the objects are emphasized. In order to take into account the above properties, Byers and Yang[3] pictured the high energy scattering process as a wave passing through the Lorentz-contracted optical medium. The medium has a particular shape which is in the form of a disk or ellipsoid. These shapes are all spheroidal in nature. The scattering by a spheroidal shaped optical medium can be viewed as the scattering by a spheroidal potential. The same picture is also used in the reflection model[4] which treats the high energy backward scattering as a process at the discontinuous boundary of the disk. The successes of the Byers and Yang model and the reflection model stimulated the present investigation.

The spheroidal potential is not totally nonspherical, and is characterized by an interfocal distance. At a large distance compared with the interfocal distance the potential can be treated as spherical. At a closer distance the nonspherical nature appears. Hence the proposed potential satisfies the intuitive physical requirement on the relative potential of two extended objects. In other words, a spherical potential is expected at large distance compared to the size of the objects, and the distorted nonspherical potential is

confined to their immediate vicinity, so physically it is a very meaningful nonspherical potential. The importance of the spheroidal potential in modern physics was mentioned by Rainwater, Granger, and Spence[5] nearly twenty years ago. They pointed out the necessity for introducing such a potential in the determination of nuclear energy levels.

For a spherical potential one uses the conventional partial wave analysis method. For a spheroidal potential one may employ the spheroidal partial wave analysis method,[6] which is adopted in the present investigation. These two analytical methods bear a very close resemblance. Furthermore, the spherical potential can be viewed as a special kind of spheroidal potential for which the interfocal distance is small.

The spheroidal partial wave analysis has been used in the classical scattering theory of sound and electromagnetic waves on a spheroidal object. The problems treated there have the character of boundary value problems. The basic method of the spheroidal analysis is well known and will not be repeated here. In Sec. 2 the problem of the present study is formulated. In Sec. 3 an integral equation for the spheroidal phase shift is given. The equation is very similar to the one in the spherical phase shift analysis. In Sec. 4 we discuss the scattering by discontinuous oblate spheroidal potentials. It is found that the low order spheroidal phase factor $2\delta_n$ is due to the difference of the optical path caused by the existence of the oblate spheroidal square well. This is the same result as for the spherical phase factor $2\delta_l$ in the scattering by a discontinuous spherical potential. In Sec. 5 the WKB approximation is discussed in the high energy scattering process. It is pointed out that the low order spheroidal partial wave can be treated as a hyperbolic trajectory. In Sec. 6 we deal with the very high energy scattering amplitude at small angles. It is shown that the spheroidal scattering amplitude at small angles approaches a limiting form. The limiting form sometimes is called the eikonal description, which is widely used in high energy scattering processes.

2. OBLATE SPHEROIDAL COORDINATES

The spheroidal partial wave analysis is conveniently described by spheroidal coordinate systems. There are two such systems. The one that is related to our problem is the oblate spheroidal coordinate system. The oblate spheroidal coordinates for the relative distance $\mathbf{r}$ are defined by following equations:

$$\begin{aligned} x &= (d/2)[(1-\eta^2)(\xi^2+1)]^{\frac{1}{2}}\cos\phi, \\ y &= (d/2)[(1-\eta^2)(\xi^2+1)]^{\frac{1}{2}}\sin\phi, \\ z &= (d/2)\eta\xi, \end{aligned} \tag{2.1}$$

with $0 \leq \xi < \infty$, $0 \leq \eta \leq 1$, $0 \leq \phi \leq 2\pi$, and d the interfocal distance. The surface $\xi = \text{const}$,

$$\frac{x^2+y^2}{\xi^2+1} + \frac{z^2}{\xi^2} = \left(\frac{d}{2}\right)^2, \tag{2.2}$$

is a flattened ellipsoid, which we will refer to as a spheroid, of revolution with major axis of length $d(\xi^2+1)^{\frac{1}{2}}$, minor axis of length $d\xi$, and eccentricity $e = 1/\xi$. The surface $|\eta| = \text{const} < 1$,

$$\frac{x^2+y^2}{1-\eta^2} - \frac{z^2}{\eta^2} = \left(\frac{d}{2}\right)^2, \tag{2.3}$$

is a hyperboloid of revolution with an asymptotic cone inclined at the angle $\theta = \cos^{-1}\eta$ to the z axis. In the limit when the interfocal distance d becomes zero, the oblate spheroidal coordinate system reduces to a spherical coordinate system; the limiting relations are

$$d \to 0, \quad \tfrac{1}{2}\,d\xi \to r, \quad \text{and} \quad \eta \to \cos\theta, \tag{2.4}$$

where r and θ are the spherical coordinates. Relation (2.4) indicates that any formulation based on a spherical coordinate system is a limiting form of a corresponding one in the spheroidal system.

The spheroidal analysis of the scattering amplitude has been discussed in the classical scattering theory of sound and electromagnetic waves.[7] The problems treated are the classical boundary value problems of the scattering of sound waves by a rigid spheroid and of the scattering of electromagnetic waves by a perfectly conducting spheroid. (A thin rod and a circular disk may be considered as two special spheroids. These two cases have attracted most attention.) The problem proposed here is a potential problem in quantum mechanics. The Schrödinger equation

$$\frac{-\hbar^2}{2\mu}\nabla^2\psi + V(\xi,\eta,\phi)\psi = \frac{\hbar^2k^2}{2\mu}\psi \tag{2.5}$$

has the form

$$\begin{aligned} \Big[\frac{\partial}{\partial\eta}(1-\eta^2)\frac{\partial}{\partial\eta} + \frac{\partial}{\partial\xi}(\xi^2+1)\frac{\partial}{\partial\xi} \\ + \left(\frac{1}{1-\eta^2} - \frac{1}{1+\xi^2}\right)\frac{\partial^2}{\partial\phi^2} + (\tfrac{1}{2}kd)^2(\xi^2+\eta^2) \\ - \frac{\mu d^2}{2\hbar^2}(\xi^2+\eta^2)V(\xi,\eta,\phi)\Big]\psi = 0, \end{aligned} \tag{2.6}$$

where μ is the reduced mass and k the incident momentum in the center of mass frame. The potential $V(\xi,\eta,\phi)$ has the property

$$\frac{\mu d^2}{2\hbar^2}(\xi^2+\eta^2)V(\xi,\eta,\phi) \equiv U(\xi)$$

and

$$U(\xi) \xrightarrow{\xi\to\infty} 0, \tag{2.7}$$

or

$$V(\mathbf{r}) = V'(\mathbf{r}), \quad \xi < \xi_0,$$
$$= 0, \quad \xi > \xi_0, \tag{2.8}$$

where $V'(\mathbf{r})$ is a continuous function. Potential (2.8) might be called a spheroidal optical medium in the case of a complex potential $V'(\mathbf{r})$, and $\frac{1}{2}\,d\xi_0$ denotes the size of the medium. The scattering picture given by this potential resembles the eikonal description of Byers and Yang[3] in the droplet model.

The conventional phase shift analysis is one of the variable separation methods in mathematical physics for solving partial differential equations. The choice of a particular separation method depends on the problem. It is known from classical scattering theory that for these potentials in Eqs. (2.7) and (2.8) the conventional phase shift analysis is inadequate and should be replaced by the spheroidal phase shift analysis. Then the scattering amplitude $f_k(\theta)$ should be expressed as

$$f_k(\theta) = \frac{1}{ik}\sum_n \frac{1}{N_{0n}(-ic)} \times S_{0n}(-ic, 1)S_{0n}(-ic, \eta)(e^{2i\delta_n} - 1), \tag{2.9}$$

where $S_{0n}(-ic, \eta)$ are the oblate spheroidal angle functions $S_{mn}(-ic, \eta)$ with $m = 0$. The normalization constant is determined by

$$\int_{-1}^{1} S_{mn}(-ic, \eta)S_{mn'}(-ic, \eta)\,d\eta = \delta_{nn'}N_{mn}(-ic), \tag{2.10}$$

where

$$c = \tfrac{1}{2}kd. \tag{2.11}$$

The phase shift δ_n defined here has the same physical meaning as the phase shift δ_l in the spherical scattering case. Namely, δ_n is the phase difference between the asymptotic solution and the potential free asymptotic solution of the spheroidal radial equations. The conventional spherical analysis can be viewed as a special case of Eq. (2.9) in the limit given in Eq. (2.4). A comment should be made about Eq. (2.9). It is known that the scattering amplitude $f_k(\theta)$ is defined in the asymptotic region of the scattered wave. In this region there is no difference between the values η and $\cos\theta$.

3. INTEGRAL EQUATION FOR PHASE SHIFTS

In this section we would like to derive an explicit expression for the spheroidal phase shifts δ_n in terms of the potential (2.7). It is well known from scattering theory that the eigenfunction $\psi_k^{(+)}(\mathbf{r})$ satisfies the integral equation

$$\psi_k^{(+)}(\mathbf{r}) = e^{ikr\cos\theta} - \frac{\mu}{2\pi\hbar^2} \times \int \frac{\exp(ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|} V(\mathbf{r}')\psi_k^{(+)}(\mathbf{r}')d^3\mathbf{r}'. \tag{3.1}$$

The expansion of the eigenfunction $\psi_k^{(+)}(\mathbf{r})$ in terms of oblate spheroidal functions can be written as

$$\psi_k^{(+)}(\mathbf{r}) = 2\sum_n i^n \frac{1}{N_{0n}(-ic)} \times S_{0n}(-ic, 1)S_{0n}(-ic, \eta)T_{0n}(-ic, i\xi). \tag{3.2}$$

The radial eigenfunction $T_{0n}(-ic, i\xi)$ has the asymptotic expression

$$T_{0n}(-ic, i\xi) \xrightarrow[\xi\to\infty]{} e^{i\delta_n}[R_{0n}^{(1)}(-ic, i\xi)\cos\delta_n - R_{0n}^{(2)}(-ic, i\xi)\sin\delta_n]. \tag{3.3}$$

The spheroidal expansion of the free space Green's function is written as

$$\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r} - \mathbf{r}'|} = \frac{ik}{2\pi}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{2 - \delta_{0m}}{N_{mn}(-ic)}S_{mn}(-ic, \eta)S_{mn}(-ic, \eta') \times \cos m(\phi - \phi')\begin{cases} R_{mn}^{(1)}(-ic, i\xi)R_{mn}^{(3)}(-ic, i\xi'), & \xi < \xi', \\ R_{mn}^{(1)}(-ic, i\xi')R_{mn}^{(3)}(-ic, i\xi), & \xi > \xi', \end{cases} \tag{3.4}$$

where $R_{mn}^{(1)}(-ic, i\xi)$, $R_{mn}^{(2)}(-ic, i\xi)$, and $R_{mn}^{(3)}(-ic, i\xi)$ are the spheroidal radial functions; the vector $\mathbf{r}'$ has the spheroidal coordinates ξ', η', and ϕ'. For the potential given in Eq. (2.7), one can express Eq. (3.1) as follows:

$$\psi_k^{(+)}(\mathbf{r}) = e^{ikr\cos\theta} - \frac{ikd}{4\pi}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{2 - \delta_{0m}}{N_{mn}(-ic)}S_{mn}(-ic, \eta) \times \sum_{n'} i^{n'}\frac{2}{N_{0n'}(-ic)}S_{0n'}(-ic, 1) \times \int d\phi' \cos m(\phi - \phi') \times \int d\eta' S_{mn}(-ic, \eta')S_{0n}(-ic, \eta') \times \left(R_{mn}^{(3)}(-ic, i\xi)\int_0^{\xi} d\xi' R_{mn}^{(1)}(-ic, i\xi') \times U(\xi')T_{0n'}(-ic, i\xi') + R_{mn}^{(1)}(-ic, i\xi) \times \int_{\xi}^{\infty} d\xi' R_{mn}^{(3)}(-ic, i\xi')U(\xi')T_{0n'}(-ic, i\xi')\right). \tag{3.5}$$

Using the orthogonality relation (2.10) and integrating over the range of the variables η' and ϕ', we get

$$\psi_k^{(+)}(r) = e^{ikr\cos\theta} - ikd \times \sum_{n=0}^{\infty} \frac{i^n}{N_{0n}(-ic)} S_{0n}(-ic,\eta)S_{0n}(-ic,1) \times \left(R_{0n}^{(3)}(-ic,i\xi)\int_0^{\xi} d\xi' R_{0n}^{(1)}(-ic,i\xi') \times U(\xi')T_{0n}(-ic,i\xi') + R_{0n}^{(1)}(-ic,i\xi) \times \int_{\xi}^{\infty} d\xi' R_{0n}^{(3)}(-ic,i\xi')U(\xi')T_{0n}(-ic,i\xi')\right). \tag{3.6}$$

From the plane wave expansion

$$e^{ikr\cos\theta} = 2\sum_n i^n \frac{1}{N_{0n}(-ic)} \times S_{0n}(-ic,1)S_{0n}(-ic,\eta)R_{0n}^{(1)}(-ic,i\xi), \tag{3.7}$$

one obtains

$$T_{0n}(-ic,i\xi) = R_{0n}^{(1)}(-ic,i\xi) - \frac{ikd}{2} R_{0n}^{(3)}(-ic,i\xi) \times \int_0^{\xi} d\xi' R_{0n}^{(1)}(-ic,i\xi')U(\xi')T_{0n}(-ic,i\xi') - \frac{ikd}{2} R_{0n}^{(1)}(-ic,i\xi) \times \int_{\xi}^{\infty} d\xi' R_{0n}^{(3)}(-ic,i\xi')U(\xi')T_{0n}(-ic,i\xi'). \tag{3.8}$$

The integral equation for the phase shifts follows from the asymptotic behavior of Eq. (3.8). By utilizing the asymptotic forms[1]

$$\begin{aligned} R_{mn}^{(1)}(-ic,i\xi) &\xrightarrow[c\xi\to\infty]{} j_n(kr), \\ R_{mn}^{(2)}(-ic,i\xi) &\xrightarrow[c\xi\to\infty]{} n_n(kr), \\ R_{mn}^{(3)}(-ic,i\xi) &\xrightarrow[c\xi\to\infty]{} h_n^{(1)}(kr) \end{aligned} \tag{3.9}$$

and asymptotic expansions for the spherical Bessel function j_n, Neumann function n_n, and Hankel function h_n, the following equation can be obtained from Eq. (3.7) in the limit $c\xi \to \infty$:

$$e^{i\delta_n}\cos[c\xi - \tfrac{1}{2}(n+1)\pi + \delta_n] = \cos[c\xi - \tfrac{1}{2}(n+1)\pi] - \frac{ikd}{2}\exp i[c\xi - \tfrac{1}{2}(n+1)] \times \int_0^{\infty} d\xi' R_{0n}^{(1)}(-ic,i\xi')U(\xi')T_{0n}(-ic,i\xi'). \tag{3.10}$$

Let us introduce a new radial function $U_{0n}(-ic,i\xi')$ which is defined by

$$U_{0n}(-ic,i\xi) \equiv T_{0n}(-ic,i\xi)e^{-i\delta_n}. \tag{3.11}$$

Finally we arrive at the simple formula from Eq. (3.9)

$$\sin\delta_n = -\tfrac{1}{2}kd\int_0^{\infty} d\xi' R_{0n}^{(1)}(-ic,i\xi')U(\xi')U_{0n}(-ic,i\xi'). \tag{3.12}$$

It is only a matter of a few steps to check that, in the limit as $c \to 0$, Eq. (3.12) reduces to the integral equation satisfied by the conventional phase shifts.

4. DISCONTINUOUS OBLATE SPHEROIDAL POTENTIAL

The potential given in Eq. (2.8) is discontinuous on the boundary of a spheroid $\xi = \xi_0$. Outside the spheroid the potential is zero, and the wavefunction $\psi_k^{(+)}(\mathbf{r})$ can be written as

$$\psi_k^{(+)}(\mathbf{r}) = 2\sum_n i^n \frac{1}{N_{0n}(-ic)} S_{0n}(-ic,1)S_{0n}(-ic,\eta) \times e^{i\delta_n}[R_{0n}^{(1)}(-ic,i\xi)\cos\delta_n - R_{0n}^{(2)}(-ic,i\xi)\sin\delta_n], \quad \xi > \xi_0. \tag{4.1}$$

The value of the phase shift δ_n is determined by solving the Schrödinger equation inside the spheroid ($\xi < \xi_0$), with the regularity requirement at the origin ($\xi = 0$, $\eta = 0$), and by joining the interior solution smoothly onto the exterior solution (4.1) at $\xi = \xi_0$. The final expression for the phase shift δ_n depends on the explicit form of potential V' in Eq. (2.8). In this section we shall present a discussion for two simple potentials V'.

(1) The potential discussed has the form

$$V'(\mathbf{r}) \equiv V'(\xi,\eta,\phi) = \frac{\hbar^2}{2\mu}(k^2 - k'^2)\frac{\xi^2}{\xi^2+\eta^2}, \quad \xi < \xi_0, \tag{4.2}$$

where k' will be called the interior wavenumber. The regularity at the origin requires $R_{0n}^{(1)}(-ic',i\xi)$ as the interior radial function, where $c' = \frac{1}{2}k'd$. By joining smoothly the interior and exterior radial functions at the boundary $\xi = \xi_0$, we obtain the relation satisfied by the spheroidal phase shift δ_n:

$$e^{i\delta_n}\sin\delta_n = e^{2i\chi_n}\frac{S_n}{\beta_n - \Delta_n - iS_n} + e^{-i\chi_n}\sin\chi_n. \tag{4.3}$$

The parameters χ_n, S_n, β_n, Δ_n are determined by

$$\begin{aligned} \Delta_n + iS_n &= -\frac{R_{0n}^{(3)\prime}(-ic,i\xi_0)}{R_{0n}^{(3)}(-ic,i\xi_0)}, \\ \beta_n &= \frac{R_{0n}^{(1)\prime}(-ic',i\xi_0)}{R_{0n}^{(1)}(-ic',i\xi_0)}, \\ e^{2i\chi_n} &= -\frac{R_{0n}^{(4)}(-ic,i\xi_0)}{R_{0n}^{(3)}(-ic,i\xi_0)}, \end{aligned} \tag{4.4}$$

where $R_{0n}^{(4)}(-ic, \xi)$ is the spheroidal radial function $R_{mn}^{(4)}(-ic, i\xi)$ with $m = 0$ and the prime denotes the derivative with respect to the variable ξ.

(2) The potential discussed has the form

$$V'(\mathbf{r}) \equiv V'(\xi, \eta, \phi) = (\hbar^2/2\mu)(k^2 - k'^2), \quad \xi < \xi_0. \tag{4.5}$$

Potential (4.5) might be called a spheroidal square well with range $\frac{1}{2}d\xi_0$ and interior wavenumber k'. From the regularity requirement, the interior wavefunction has the form

$$\psi_k^{(+)}(\mathbf{r}) = \sum_n \frac{A_n}{N_{0n}(-ic')} R_{0n}^{(1)}(-ic\,, i\xi) S_{0n}(-ic', \eta), \quad \xi < \xi_0, \tag{4.6}$$

where A_n are arbitrary parameters and $c' = \frac{1}{2}k'd$. The smoothness on the boundary yields the following relation:

$$\begin{aligned}
&[R_{0n'}^{(1)}(-ic', i\xi_0)]^{-1}\Big(2\sum_n \frac{1}{N_{0n}} S_{0n}(-ic, 1)A_{n'n}e^{i\delta_n} \\
&\quad \times [R_{0n}^{(1)}(-ic, i\xi_0)\cos\delta_n - R_{0n}^{(2)}(-ic, i\xi_0)\sin\delta_n]\Big) \\
&= [R_{0n'}^{(1)'}(-ic', i\xi_0)]^{-1}\Big(2\sum_n \frac{1}{N_{0n}} S_{0n}(-ic, 1)A_{n'n}e^{i\delta_n} \\
&\quad \times [R_{0n}^{(1)'}(-ic, i\xi_0)\cos\delta_n - R_{0n}^{(2)'}(-ic, i\xi_0)\sin\delta_n]\Big),
\end{aligned} \tag{4.7}$$

where

$$A_{n'n} = \int_{-1}^{1} S_{0n'}(-ic', \eta) S_{0n}(-ic, \eta)\, d\eta. \tag{4.8}$$

In order to exploit the physical meaning of the spheroidal phase shift δ_n implied by Eq. (4.7), we shall restrict ourselves to a special case with

$$|c - c'| \ll c \quad \text{and} \quad c \to \infty. \tag{4.9}$$

In the limit as $c \to \infty$, the spheroidal angle functions have the form

$$S_{0n}(-ic, \eta) \xrightarrow[c\to\infty]{} \text{const} \begin{cases} e^{-c(1-\eta)}L_\nu[2c(1-\eta)], & 0 < \eta \le 1, \\ (-)^n e^{-c(1+\eta)}L_\nu[2c(1+\eta)], & -1 \le \eta < 0, \\ e^{-c}\frac{1}{2}[L_\nu(2c) + (-)^n L_\nu(2c)], & \eta = 0, \end{cases} \tag{4.10}$$

where

$$\begin{aligned} \nu &= n/2, && n \text{ even}, \\ &= \tfrac{1}{2}(n-1), && n \text{ odd}, \end{aligned} \tag{4.11}$$

and $L_\nu(x)$ is the Laguerre polynomial. The arbitrary constant in Eq. (4.10) depends on the normalization for the spheroidal angle function $S_{0n}(-ic, \eta)$, and should not appear in the expression for the scattering amplitude. For simplicity, this constant is omitted in the discussion. In the limit as $c \to \infty$, the coefficient $A_{nn'}$ can be expressed as

$$\begin{aligned}
A_{n'n} \xrightarrow[c\to\infty]{} &\int_0^1 e^{-(c+c')(1-\eta)} \\
&\times L_{\nu'}[2c'(1-\eta)]L_\nu[2c(1-\eta)]\, d\eta \\
&+ \int_{-1}^{0} (-)^{n+n'} e^{-(c+c')(1+\eta)} \\
&\times L_{\nu'}[2c'(1+\eta)]L_\nu[2c(1+\eta)]\, d\eta \\
= &\frac{1}{2c}\int_0^\infty [1 + (-)^{n+n'}] \\
&\times e^{-(1+\epsilon)x} L_{\nu'}[(1+2\epsilon)x]L_\nu[x]\, dx,
\end{aligned} \tag{4.12}$$

where

$$\begin{aligned} \nu' &= n'/2, && n' \text{ even}, \\ &= \tfrac{1}{2}(n'-1), && n' \text{ odd}, \end{aligned} \tag{4.13}$$

and

$$\epsilon = (c' - c)/2c. \tag{4.14}$$

After evaluating the above integration, one arrives at

$$\begin{aligned}
A_{n'n} = (1/2c)[1 + (-)^{n+n'}](-)^{\nu_M+\nu'} \\
\times \frac{\Gamma(\nu_M + 1)}{\Gamma(\nu_m + 1)\Gamma(\nu_M - \nu_m + 1)} \epsilon^{\nu_M - \nu_m},
\end{aligned} \tag{4.15}$$

where

$$\nu_m = \min(\nu, \nu')$$

and

$$\nu_M = \max(\nu, \nu'). \tag{4.16}$$

The renormalization constant can be found in the same way:

$$N_{0n} = 1/c. \tag{4.17}$$

By using the above approximations, Eq. (4.7) can be expressed as

$$\begin{aligned}
&[R_{0n}^{(1)}(-ic', i\xi_0)]^{-1}\{e^{i\delta_n}[R_{0n}^{(1)}(-ic, i\xi_0)\cos\delta_n \\
&\quad - R_{0n}^{(2)}(-ic, i\xi_0)\sin\delta_n] \\
&\quad - (\nu+1)\epsilon e^{i\delta_{n+2}}[R_{0n+2}^{(1)}(-ic, i\xi_0)\cos\delta_{n+2} \\
&\quad - R_{0n+2}^{(2)}(-ic, i\xi_0)\sin\delta_{n+2}] \\
&\quad - \nu\epsilon e^{i\delta_{n-2}}[R_{0n-2}^{(1)}(-ic, i\xi_0)\cos\delta_{n-2} \\
&\quad - R_{0n-2}^{(2)}(-ic, i\xi_0)\sin\delta_{n-2}]\} \\
&= [R_{0n}^{(1)'}(-ic', i\xi_0)]^{-1}\{e^{i\delta_n}[R_{0n}^{(1)'}(-ic, i\xi_0)\cos\delta_n \\
&\quad - R_{0n}^{(2)'}(-ic, i\xi_0)\sin\delta_n] \\
&\quad - (\nu+1)\epsilon e^{i\delta_{n+2}}[R_{0n+2}^{(1)'}(-ic, i\xi_0)\cos\delta_{n+2} \\
&\quad - R_{0n+2}^{(2)'}(-ic, i\xi_0)\sin\delta_{n+2}] \\
&\quad - \nu\epsilon e^{i\delta_{n-2}}[R_{0n-2}^{(1)'}(-ic, i\xi_0)\cos\delta_{n-2} \\
&\quad - R_{0n-2}^{(2)'}(-ic, i\xi_0)\sin\delta_{n-2}]\}.
\end{aligned} \tag{4.18}$$

A simpler form can be obtained by neglecting the higher order contribution with respect to parameter ϵ in Eq. (4.18):

$$\begin{aligned}&[R_{0n}^{(1)}(-ic', i\xi_0)]^{-1}\\&\quad\times\{(1+i\delta_n)[R_{0n}^{(1)}(-ic, i\xi_0)-\delta_n R_{0n}^{(2)}(-ic, i\xi_0)]\\&\quad-(\nu+1)\epsilon R_{0n+2}^{(1)}(-ic, i\xi_0)-\nu\epsilon R_{0n-2}^{(1)}(-ic, i\xi_0)\}\\&=[R_{0n}^{(1)'}(-ic', i\xi_0)]^{-1}\\&\quad\times\{(1+i\delta_n)[R_{0n}^{(1)'}(-ic, i\xi_0)-\delta_n R_{0n}^{(2)'}(-ic, i\xi_0)]\\&\quad-(\nu+1)\epsilon R_{0n+2}^{(1)'}(-ic, i\xi_0)-\nu\epsilon R_{0n-2}^{(1)'}(-ic, i\xi_0)\}.\end{aligned}\tag{4.19}$$

The spheroidal radial functions have the following simple asymptotic forms at large c and small n:

$$\begin{aligned}&R_{0n}^{(1)}(-ic, i\xi_0)\\&=[c(1+\xi_0^2)]^{-1}[\cos(c\xi_0-\tfrac{1}{2}n\pi)+\xi_0\sin(c\xi_0-\tfrac{1}{2}n\pi)]\\&=(1/c)(1+\xi_0^2)^{-\frac{3}{2}}\cos(c\xi_0-\tfrac{1}{2}n\pi-\alpha),\\&R_{0n}^{(2)}(-ic, i\xi_0)\\&=[c(1+\xi_0^2)]^{-1}[\sin(c\xi_0-\tfrac{1}{2}n\pi)+\xi_0\cos(c\xi_0-\tfrac{1}{2}n\pi)]\\&=(1/c)(1+\xi_0^2)^{-\frac{3}{2}}\sin(c\xi_0-\tfrac{1}{2}n\pi+\alpha),\end{aligned}\tag{4.20}$$

$$\begin{aligned}R_{0n}^{(1)'}(-ic, i\xi_0)&=-(1+\xi_0^2)^{-\frac{3}{2}}\sin(c\xi_0-\tfrac{1}{2}n\pi-\alpha),\\R_{0n}^{(2)'}(-ic, i\xi_0)&=(1+\xi_0^2)^{-\frac{3}{2}}\cos(c\xi_0-\tfrac{1}{2}n\pi+\alpha),\end{aligned}$$

where

$$\tan\alpha=\xi_0.\tag{4.21}$$

In terms of the above asymptotic forms, Eq. (4.19) can be rewritten as

$$\tan(c'\xi_0-\tfrac{1}{2}n\pi-\alpha)=\tan(c\xi_0-\tfrac{1}{2}n\pi-\alpha+\beta),\tag{4.22}$$

where

$$\tan\beta=\frac{\delta_n\cos 2\alpha}{1+i\delta_n+(2\nu+1)\epsilon-\delta_n\sin 2\alpha}\approx\delta_n.\tag{4.23}$$

Finally, the phase shift δ_n can be approximately expressed by

$$2\delta_n=2(c'-c)\xi_0=(k'-k)\,d\xi_0.\tag{4.24}$$

The length $d\xi_0$ is the minor axis of the spheroid, in which the spheroidal square well is confined. For a spherical square well the spherical phase shift δ_l can be also expressed in the same form as in Eq. (4.24). The physical interpretation of Eq. (4.24) is simple. It states that the spherical phase factor $2\delta_n$ is due to the difference of the optical path caused by the existence of the oblate spheroidal square well.

In the spherical square well scattering process, the Legendre functions are used for expanding the scattering amplitude, and the partial wave amplitudes are expressed in the forms of the spherical phase shifts. In the case of the spheroidal square well, the spheroidal angle functions are used, and the spheroidal partial wave amplitudes are expressed in terms of the spheroidal phase shifts. Since the spherical and spheroidal phase shifts describe the same physical consequence, we may conclude that, at least in the square well cases, the shape of the potential determines the form of the expansion basis functions, and not the partial wave amplitudes.

5. WKB APPROXIMATION

The semiclassical approximate WKB method[7] for solving the radial part of the wave equation (2.6) with potential (2.7) is presented in this section. This method is only applicable to high energy scattering processes. Let

$$T_{0n}(-ic, i\xi)=(1+\xi^2)^{-\frac{1}{2}}G_{0n}(-ic, i\xi),\tag{5.1}$$

then the radial wave equation has the form

$$\begin{aligned}&\frac{d^2}{d\xi^2}G_{0n}(-ic, i\xi)\\&\quad-\frac{1}{1+\xi^2}\left[\frac{1}{1+\xi^2}+\lambda_{0n}+U(\xi)-c^2\xi^2\right]\\&\quad\times G_{0n}(-ic, i\xi)=0,\end{aligned}\tag{5.2}$$

where λ_{0n} is the eigenvalue associated with the spheroidal angle function $S_{0n}(-ic, \eta)$. On making the transformation

$$c\xi=e^{\chi}\tag{5.3}$$

and on setting

$$G_{0n}(-ic, i\xi)=e^{\chi/2}U_{0n}(c\chi),\tag{5.4}$$

Eq. (5.2) has the form

$$\left[\frac{d^2}{d\chi^2}+Q^2(\chi)\right]U_{0n}(c, \chi)=0,\tag{5.5}$$

where

$$\begin{aligned}-Q^2(\chi)=\tfrac{1}{4}+\Bigg[&\frac{c^2}{c^2+e^{2\chi}}+\lambda_{0n}\\&+U(\chi)-e^{2\chi}\Bigg]\frac{e^{2\chi}}{c^2+e^{2\chi}}.\end{aligned}\tag{5.6}$$

Using the standard WKB procedure, the solution for the differential equation (5.5) can be given as

$$U_{0n}(c, \chi)=F[Q(\chi)]^{-\frac{1}{2}}\sin\left[\frac{\pi}{4}+\int_{\chi_0}^{\chi}Q(S)\,dS\right],\quad\chi>\chi_0,\tag{5.7}$$

where χ_0 is the turning point

$$Q(\chi_0)=0\tag{5.8}$$

and F is a normalization constant. It is apparent that the chosen potential (2.7) should give one and only one turning point χ_0 for Eq. (5.7) to hold. After setting

$$Q^2(\chi) = \xi^2 p^2(\xi), \tag{5.9}$$

where

$$p^2(\xi) = [c^2\xi^2 - 1/(1+\xi^2) - \lambda_{0n} - U(\xi)](1+\xi^2)^{-1} - 1/4\xi^2, \tag{5.10}$$

the radial wavefunction $T_{0n}(-ic, i\xi)$ may be expressed by

$$T_{0n}(-ic, i\xi) = F\sqrt{c}\,(1+\xi^2)^{-\frac{1}{2}}p^{-\frac{1}{2}}(\xi)\sin\left[\frac{\pi}{4} + \int_{\xi_n}^{\xi} p(\xi)\,d\xi\right], \quad \xi > \xi_n, \tag{5.11}$$

where

$$\xi_n \equiv c^{-1}e^{\chi_0}. \tag{5.12}$$

From Eq. (3.3) and the asymptotic form of the radial spheroidal function, we have

$$T_{0n}(-ic, i\xi) \xrightarrow[\xi\to\infty]{} (1/c\xi)\cos[\delta_n + c\xi - \tfrac{1}{2}(n+1)\pi]. \tag{5.13}$$

However, from Eq. (5.11) the asymptotic form should be

$$T_{0n}(-ic, i\xi) \xrightarrow[\xi\to\infty]{} \frac{F\sqrt{c}}{c\xi}\sin\left(\frac{\pi}{4} + \int_{\xi_n}^{\infty}[p(\xi) - c]\,d\xi + c\xi - c\xi_n\right). \tag{5.14}$$

Comparing Eqs. (5.13) with (5.14), the phase shift δ_n is determined by

$$\delta_n = \frac{\pi}{2}(n + \tfrac{1}{2}) - c\xi_n + \int_{\xi_n}^{\infty}[p(\xi) - c]\,d\xi + 2n'\pi, \tag{5.15}$$

where n' is an arbitrary integer. In the case of zero potential, there should be no scattering and the phase shift δ_n is zero. Thus the phase shift δ_n in Eq. (5.15) is obviously the difference

$$\delta_n = \int_{\xi_n}^{\infty}[p(\xi) - p_0(\xi)]\,d\xi, \tag{5.16}$$

where

$$p_0^2(\xi) = [c^2\xi^2 - (1+\xi^2)^{-1} - \lambda_{0n}](1+\xi^2)^{-1} - 1/4\xi^2. \tag{5.17}$$

In the eikonal approximation, the Born term is used to evaluate the phase shift in Eq. (5.16)

$$\delta_n^{(1)} = -\frac{1}{2}\int_{\xi_n}^{\infty}\frac{U(\xi)\,d\xi}{(1+\xi^2)p_0(\xi)}, \tag{5.18}$$

and in the approximation the parameter ξ_n is determined by the equation

$$p_0^2(\xi_n) \approx 0, \tag{5.19}$$

and can be expressed as

$$\xi_n^2 \approx (2c\xi^2)^{-1}[(\lambda_{0n}^2 + c^2)^{\frac{1}{2}} + \lambda_{0n}]. \tag{5.20}$$

The eigenvalue $\lambda_{0n}(-ic)$ has a simple form for large value c:

$$\lambda_{0n}(-ic) \xrightarrow[c\to\infty]{} -c^2 + 2c(2\nu + 1), \tag{5.21}$$

where ν is given by Eq. (4.11). In the applicable region of the eikonal approximation, the eigenvalue $\lambda_{0n}(-ic)$ determined by Eq. (5.21) would be large for most spheroidal order n. The phase shift $\delta_n^{(1)}$ in Eq. (5.18) can be written in terms of the parameter ξ:

$$\delta_n^{(1)} = -\frac{1}{2c}\int_{\xi_n}^{\infty}\frac{\xi U(\xi)\,d\xi}{\left[(1+\xi^2)(\xi^2 - \xi_n^2)\left(\xi^2 + \frac{1}{4c^2\xi_n^2}\right)\right]^{\frac{1}{2}}}. \tag{5.22}$$

The above integration is only carried out over $\xi > \xi_n$. It means that the phase shift $\delta_n^{(1)}$ of the spheroidal partial wave only depends on the potential in the region

$$R \equiv \tfrac{1}{2}\,d\xi \geq \tfrac{1}{2}\,d\xi_n, \tag{5.23}$$

where R is the distance from the scattered center. The parameter ξ_n is a monotonically increasing function of the spheroidal partial order n. These facts lead us to the picture that the high order spheroidal partial waves are the waves at large distance from the scattering center. The picture is the same as that observed in the spherical scattering.

The physical meaning of the phase shift $\delta_n^{(1)}$ can be more explicitly understood in the low partial orders. From Eqs. (5.20) and (5.21) the quantity ξ_n at small order n can be approximated by

$$\xi_0 = \tfrac{1}{2}[1 + (2\nu + 1)/c], \quad n \ll c. \tag{5.24}$$

By using the above approximation, Eq. (5.22) has the form

$$\delta_n^{(1)} = -\frac{1}{2c}\int_{\xi_n}^{\infty}\left[(1+\xi^2)\left(1 + \xi^2 - 2\frac{2\nu+1}{c}\right)\times(\xi^2 - \xi_0^2)\right]^{-\frac{1}{2}}U(\xi)\xi\,d\xi \approx -\frac{1}{2c}\int_0^{\infty}\frac{U[(\xi^2+\xi_n^2)^{\frac{1}{2}}]\,d\xi}{\xi^2 + \xi_n^2 + 1 - \frac{2\nu+1}{c}}.$$

The above expression can be rewritten in terms of

potential (2.7),

$$2\delta_n^{(1)} = -\left(\frac{ud}{\hbar^2}\right)\frac{1}{k}\int_0^\infty V((\xi^2+\xi_n^2)^{\frac{1}{2}}, [1-(2\nu+1)/c]^{\frac{1}{2}}, 0)\,d\xi. \quad (5.25)$$

Equation (5.25) implies that the spheroidal phase shift $\delta_n^{(1)}$ with small order n only depends on a spatial integration along a hyperbolic trajectory. Since potential (2.7) is cylindrically symmetric with respect to the z axis, the trajectory can be viewed as a trajectory on the xz plane,

$$x^2[(2\nu+1)/c]^{-1} - z^2[1-(2\nu+1)/c]^{-1} = (d/2)^2. \quad (5.26)$$

The closest distance x_n between the origin and the trajectory is equal to

$$x_n = (d/2)[(2\nu+1)/c]^{\frac{1}{2}} = (d/2)(n/c)^{\frac{1}{2}}. \quad (5.27)$$

From the above discussion we can draw a physical picture. In the eikonal approximation of the high energy scattering process, each low order spheroidal partial wave can be considered as a classical trajectory, and the trajectory passes the scattered center at a distance x_n. The picture is the same as in the eikonal approximation of the spherical potential scattering process. In such a process the spherical partial wave with order n_s is pictured as a linear trajectory with an impact parameter $b = (n_s + \frac{1}{2})/k$.

6. HIGH ENERGY SCATTERING AMPLITUDE AT SMALL ANGLE

The spheroidal expansion of the scattering amplitude $f_k(\theta)$ has a complicated form. If one is interested only in the small angle scattering process at high energy, then a simpler form can be obtained.

For small arguments x there exists a relation between the Laguerre polynomial $L_\nu(x)$ and the Bessel function[8]

$$L_\nu(x) \simeq J_0[2(\nu x)^{\frac{1}{2}}]e^{x/2}, \quad x \text{ small}. \quad (6.1)$$

By utilizing the above relation and Eq. (4.10) the scattering amplitude in Eq. (2.10) at small angle can be written as

$$f_k(\theta) \approx \frac{c}{ik}\sum_{\nu=0}^{\infty} J_0\{2[2c\nu(1-\eta)]^{\frac{1}{2}}\} \times [(e^{2i\delta_{2\nu}}-1) + (e^{2i\delta_{2\nu+1}}-1)]. \quad (6.2)$$

For the high energy scattering process, the terms involved in the above series are very many. It is a good approximation to replace the above summation by an integration, and neglect the difference of the adjacent phase shifts

$$f_k(\theta) \approx \frac{2c}{ik}\sum_{\nu=0}^{\infty} J_0\{2[2c\nu(1-\eta)]^{\frac{1}{2}}\}[e^{2i\delta_{2\nu}}-1]$$
$$\approx ik\int_0^\infty (1-e^{2i\delta(B)})J_0(B\sqrt{-t})B\,dB, \quad (6.3)$$

where

$$B = (2/k)(c\nu)^{\frac{1}{2}}$$

and

$$\delta(B) = \delta_{2\nu}. \quad (6.4)$$

The momentum $-t$ is defined in the usual way:

$$-t = 2k^2(1-\cos\theta) = 2k^2(1-\eta). \quad (6.5)$$

The small angle scattering amplitude is expressed in terms of the Fourier–Bessel transform of the spheroidal partial wave amplitude. The expression is often seen in discussions of high energy scattering processes under the name of the eikonal description. In the conventional method, the Fourier–Bessel transform comes from the spherical partial wave analysis of the scattering amplitude. However, Schiff and Wu[9] have shown that, even for a nonspherical continuous potential, the Fourier–Bessel transform can also be used to express the scattering amplitude. They wrote

$$f_k(\theta) \approx ik\int_0^\infty \left\{1 - \exp\left[-i\frac{\mu}{\hbar^2 k}\right]\int_{-\infty}^{\infty} V_c(b,z)\,dz\right\} \times J_0(b\sqrt{-t})b\,db, \quad (6.6)$$

where b is the impact parameter; $V_c(b,z)$ is any spherical or nonspherical continuous potential. The spheroidal potential Eq. (2.7) used in the present paper may have two singularities at focal points. Equation (6.3) shows that, even in the existence of such a singularity, the eikonal description may be used. It is also interesting to see whether Eq. (6.3) is reducible to the form of Schiff and Wu, Eq. (6.6). We shall discuss such a problem. By comparing Eq. (6.3) and Eq. (6.6), the spheroidal phase shift in the sense of Schiff and Wu is expressed as

$$\delta_s(B) \equiv -\frac{\mu}{2\hbar^2 k}\int_{-\infty}^{\infty} V(B,z)\,dz, \quad (6.7)$$

where the potential V is given in Eq. (2.8). In the oblate spheroidal coordinates system the above integration has the form

$$\delta_s(B) = -\frac{1}{2c}\int_{\xi_n'}^{\infty} \frac{U(\xi)\,d\xi}{[(\xi^2+1)(\xi^2+1-4B^2d^{-2})]^{\frac{1}{2}}}, \quad (6.8)$$

where

$$\xi_n' = \begin{cases} 0, & B < d/2, \\ (2/d)\left(B^2 - \frac{d^2}{4}\right)^{\frac{1}{2}}, & B > d/2. \end{cases} \quad (6.9)$$

For the high energy scattering process, the spheroidal phase shift $\delta_n^{(1)}$ in Eq. (5.22) evaluated by the WKB approximation may be used for evaluating the spheroidal phase shift $\delta(B)$ in Eq. (6.3):

$$\delta(B) = -\frac{1}{2c}\int_{\xi_n}^{\infty} \frac{\xi U(\xi)\, d\xi}{\left[(1+\xi^2)(\xi^2-\xi_n^2)\left(\xi^2+\frac{1}{4c^2\xi_n^2}\right)\right]^{\frac{1}{2}}}. \tag{6.10}$$

If the function $U(\xi)$ in Eqs. (6.8) and (6.10) has the property

$$\frac{1}{\xi}U(\xi) \xrightarrow{\xi\to 0} \text{finite}, \tag{6.11}$$

then one can prove that

$$\frac{|\delta_s(B)-\delta(B)|}{|\delta(B)|} \approx O(1/\sqrt{c}), \quad \text{at} \quad c\to\infty. \tag{6.12}$$

The potential $V(\xi, \eta, \phi)$ in Eq. (2.7), with the condition Eq. (6.11), might have two singularities at focal points of the spheroidal coordinate system. It shows that even for some discontinuous potentials the formulation given by Schiff and Wu may still hold. Although our scattering amplitude at small angle can be reduced to the general Schiff and Wu form in the high energy limit, at finite energies a difference between them does exist.

The thin-disk potential is a special kind of oblate spheroidal square well. The oblate spheroid within which the potential is confined has a small minor axis of length $d\xi_0$. We have already discussed the spheroidal phase shifts from the oblate spheroidal potential in Sec. 4. The exact determination of the scattering amplitude through these spheroidal phase shifts, even for a disk type potential, is tedious. At the end of Sec. 5 it was pointed out that, for a chosen spheroidal partial wave, the phase shift only depends on the potential at region $R \geq \frac{1}{2}\, d\xi_0$. Since, for the disk well, the potential is zero outside the spheroid ($\xi \geq \xi_0$), the spheroidal partial waves with $\xi_n > \xi_0$ are not disturbed in the scattering process, and their corresponding partial wave amplitudes should vanish. For the thin-disk well the spheroid partial waves with small ξ_n will contribute to the scattering amplitude. From Eqs. (5.20) and (5.21) the contributed partial wave has

$$\lambda_{0n} \leq 0 \tag{6.13}$$

for the high energy scattering process. From Eqs. (4.24), (6.3), and (6.13), the small angle scattering amplitude from the thin-disk well has the approximate form

$$f_k(\theta) \approx ik(1-e^{i(k'-k)d\xi_0})\int_0^{d/2} J_0(B\sqrt{-t})B\, dB$$

$$= ic(1-e^{i(k'-k)d\xi_0})\frac{J_1[(d/2)\sqrt{-t}]}{\sqrt{-t}}. \tag{6.14}$$

The above expression is a familiar one as it is often encountered in the problem of scattering by a disk.[10]

ACKNOWLEDGMENTS

In conclusion, the author wishes to express his appreciation to Prof. C. N. Yang for several discussions. He also wishes to thank Prof. S. P. Almeida for interesting conversations on several matters presented here.

* Work supported in part by the National Science Foundation.

[1] C. Flammer, *Spheroidal Wave Functions* (Stanford U.P., Stanford, Cal., 1957); J. A. Stratton, P. M. Morse, L. J. Chu, J. D. C. Little, and F. J. Corbato, *Spheroidal Wave Fanctions* (Wiley, New York, 1956); P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).

[2] T. T. Wu and C. N. Yang, Phys. Rev. **B 137**, 708 (1965).

[3] N. Byers and C. N. Yang, Phys. Rev. **B 142**, 976 (1966).

[4] M. C. Li, Phys. Rev. **D 1**, 326 (1970); S. P. Almeida and M. C. Li, Particles and Nuclei **1**, 24 (1970).

[5] J. Rainwater, Phys. Rev. **79**, 432 (1950); S. Granger and R. D. Spence, Phys. Rev. **83**, 460 (1951).

[6] R. D. Spence and S. Granger, J. Acoust. Soc. Am. **23**, 701 (1951); J. Meixner and W. Andrejewski, Ann. Phys. **7**, 15 (1950); C. Flammer, J. Appl. Phys. **26**, 1218 (1955); **24**, 1224 (1953); C. Yeh, J. Acoust. Soc. Am. **42**, 518 (1967); **46**, 797 (1969).

[7] H. Jeffreys, Proc. London Math. Soc. **23**, 428 (1928); G. Wentzel, Z. Physik **38**, 518 (1926); H. A. Kramers, *ibid.*, **39**, 828 (1926); M. Brillouin, J. Phys. **3**, 65 (1922).

[8] A. Erdélyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill, New York, 1955); G. N. Watson, *Theory of Bessel Functions* (Cambridge U.P., London, 1966).

[9] L. I. Schiff, Phys. Rev. **103**, 443 (1956); T. T. Wu, *ibid.*, **108**, 466 (1957).

[10] S. Fernbach, R. Serber, and T. B. Taylor, Phys. Rev. **75**, 1352 (1949).

Stability of Time-Dependent Particlelike Solutions in Nonlinear Field Theories. II

DAVID L. T. ANDERSON
*Department of Theoretical Physics, University of St. Andrews, Fife, Scotland**

(Received 6 August 1970)

It is shown that particlelike solutions of the form $\psi = \varphi(r)e^{-i\omega t}$ to the nonlinear field

$$\nabla^2\psi - c^{-2}\partial^2\psi/\partial t^2 = \kappa^2\psi - \mu^2\psi\psi^*\psi + \lambda\psi\psi^*\psi\psi^*\psi$$

can exist for a certain range of the field parameters. The stability of the lowest-order solution of the above form is examined by first-order perturbation theory and also by direct integration of the perturbed field. Both methods indicate the existence of stable particlelike states, some of which have the further asset of a positive-definite energy density.

1. INTRODUCTION

In a previous paper[1] (hereafter referred to as I), the stability of the lowest-order (nodeless) solution of the form

$$\psi = \varphi(r)e^{-i\omega t}, \quad \omega > 0, \tag{1}$$

of the nonlinear field equation

$$\nabla^2\psi - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = \kappa^2\psi - \mu^2\psi\psi^*\psi \tag{2}$$

was examined. It was found that the solution was highly unstable for all choice of parameters κ, μ, and ω, that the lifetime/size ratio was essentially a constant independent of the field parameters, that the energy density was not positive definite,and that, as a direct consequence, a physically unacceptable (singular) decay mode was possible. In the present work, a generalization of (2) is considered which overcomes many of the above difficulties. For example, it is possible, by judicious choice of the field parameters, to have stable time-dependent particlelike solutions for which the energy density is positive definite. These particlelike solutions cannot possess absolute stability if sufficiently severe disturbances are allowed, but, when a decay is induced, it cannot take place by the singular mode.

The particular generalization of (2) considered is

$$\nabla^2\psi - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = \kappa^2\psi - \mu^2\psi\psi^*\psi + \lambda\psi\psi^*\psi\psi^*\psi, \tag{3}$$

where λ is a real, usually positive parameter. One can derive (2) from the Lagrangian density. We have

$$\mathfrak{L} = \frac{1}{c^2}\left|\frac{\partial\psi}{\partial t}\right|^2 - |\nabla\psi|^2 - \kappa^2|\psi|^2 + \tfrac{1}{2}\mu^2|\psi|^4 - \tfrac{1}{3}\lambda|\psi|^6, \tag{4}$$

the corresponding energy density $\mathcal{E}$ being given by

$$\mathcal{E} = \frac{1}{c^2}\left|\frac{\partial\psi}{\partial t}\right|^2 + |\nabla\psi|^2 + \kappa^2|\psi|^2 - \tfrac{1}{2}\mu^2|\psi|^4 + \tfrac{1}{3}\lambda|\psi|^6. \tag{5}$$

As in I, we shall restrict the analysis to spherically symmetric solutions of the form (1), when (3) can be reduced to

$$\frac{d^2\varphi'}{dr'^2} + \frac{2}{r'}\frac{d\varphi'}{dr'} = \varphi' - \varphi'^3 + B\varphi'^5, \tag{6}$$

where φ', r', and B are given by

$$\varphi' = \mu\varphi(1 - \omega'^2)^{-\frac{1}{2}}, \quad r' = \kappa r(1 - \omega'^2)^{+\frac{1}{2}}, \tag{7}$$

$$B = \lambda\kappa^2(1 - \omega'^2)/\mu^4. \tag{8}$$

2. PARTICLELIKE SOLUTIONS OF (6)

A. Phase Space Analysis

In order to exploit the analogy with forces on macroscopic point particles and their equations of motion, we will write (6) in the form

$$\ddot{y} = \frac{d(\frac{1}{2}\dot{y}^2)}{dy} = -\frac{2}{t}\dot{y} + y - y^3 + By^5, \qquad \varphi' \to y, \quad r' \to t, \tag{9}$$

where the dot denotes differentiation with respect to time. Let us define a new equation from which the $\dot{y}$ term is absent, viz.,

$$\ddot{y} = y - y^3 + By^5. \tag{10}$$

This can be integrated to give

$$\tfrac{1}{2}\dot{y}^2 + (-\tfrac{1}{2}y^2 + \tfrac{1}{4}y^4 - \tfrac{1}{6}By^6) = E, \tag{11}$$

where E is a constant of integration. For a particle of unit mass, (10) represents motion under the conservative force $(y - y^3 + By^5)$. The energy of this motion can be expressed as $E = T + V$, where T, the

kinetic energy, is $\frac{1}{2}\dot{y}^2$, and V, the potential energy, is given by

$$V = -\tfrac{1}{2}y^2 + \tfrac{1}{4}y^4 - \tfrac{1}{6}By^6. \tag{12}$$

The equilibrium points of the motion are given by

$$\left(\frac{\partial v}{\partial y}\right) = 0 = -y + y^3 - By^5. \tag{13}$$

Equation (13) has solutions

$$y = 0, \quad y = \pm A_1, \quad y = \pm A_2,$$

where A_1 and A_2 are given by

$$A_1 = \left(\frac{1 - (1 - 4B)^{\frac{1}{2}}}{2B}\right)^{\frac{1}{2}}, \quad A_2 = \left(\frac{1 + (1 - 4B)^{\frac{1}{2}}}{2B}\right)^{\frac{1}{2}}. \tag{14}$$

The trajectories in phase space $(\dot{y}, y)$ of the motion defined by (10) are given by $E = \text{const}$. The motion we want to analyze is (9), for which

$$\frac{dE}{dt} = \frac{-2(\dot{y})^2}{t}. \tag{15}$$

Since this is always negative, it follows that the trajectories of particles obeying (9) always move to decreasing E.

B. The Phase Trajectories

The notation used in the phase analysis is as follows. The solutions of (9) sought in this work are those for which $\dot{y} = 0$ at $t = 0$ and $y \to D(k)e^{-t}/t$ as $t \to \infty$, where $D(k)$ is a constant for any given solution and the index k denotes the number of nodes possessed by the solution. Let us denote the value of y at $t = 0$ by a_0. Particlelike solutions will occur only for certain discrete values of a_0 denoted $A(k)$. For convenience, we consider starting from a trial point $\dot{y} = 0$, $y = a_0$ at $t = 0$ and integrating forwards in time. Only for certain, if any, values of a_0 [i.e., $A(k)$] will the solution have an asymptotic exponential

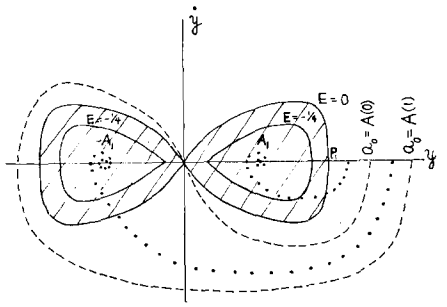

FIG. 1. Phase space for $B = 0$. The nodeless and one-node solutions (dashed lines) are shown. The solid curves are the phase trajectories of (10) $E = \text{const}$. The hatched area indicates the region of negative E while the dotted curves show the phase trajectories of particles obeying the equation of motion (9), but for which $a_0 \neq A(k)$. These trajectories end up spiraling around ± 1.

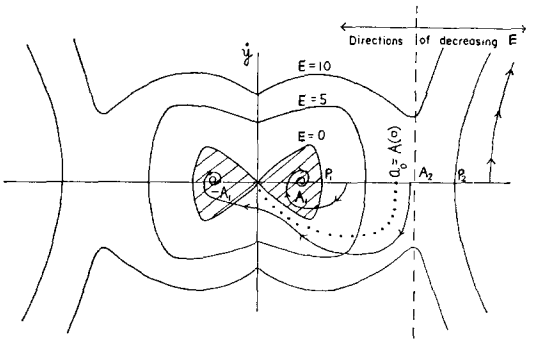

FIG. 2. Typical phase space for B in the range $0 < B < \frac{3}{16}$. The dotted curve indicates the nodeless solution to (9). As in Fig. 1, if $a_0 \neq A(k)$, then the trajectory spirals around $\pm A_1$. If $a_0 > A_2$, the direction of decreasing E is away from the origin.

form. Because different values of B give rise to different types of phase trajectories, it is convenient to split the analysis into four separate regions.

i. $B = 0$

In Fig. 1 we show the appropriate phase space diagram. The solid lines are the curves $E = \text{const}$ and are therefore the trajectories of a particle obeying (10). We want the phase trajectory of a particle obeying (9), and these are given by (14), i.e., the trajectory always moves to decreasing E. The figure of eight lying round the origin is the curve $E = 0$. Once a trajectory enters such a region, it cannot escape and must end up on the lowest point in the lobe which it enters, i.e., ± 1. For certain values of y (when $\dot{y} = 0$), there are trajectories which go to the origin. These values of y are $A(k)$, $k = 0, 1, 2, \cdots$. The trajectories for the two lowest-order solutions to (9) for $B = 0$ are marked on Fig. 1. The trajectories of the higher-order solutions can be obtained by a simple extension of the above process.

ii. $B < 0$

The phase space for B negative is very similar to that for the $B = 0$ case except that $A_1 < 1$. Otherwise, the analysis goes through as for the $B = 0$ case. (The above process would break down if the point P_1 at which the $E = 0$ trajectory cuts the line $\dot{y} = 0$ were to lie nearer the origin than A_1. However, it is easy to show that this cannot occur for any B.)

iii. $0 < B < \frac{3}{16}$

When B is positive, A_2 is real and finite, and thus one expects the phase space for this case to be involved. In Fig. 2 a typical phase space for $0 < B < \frac{3}{16}$ is drawn. Let us split up the analysis into three categories.

(*a*) $a_0 < P_1$: Then the trajectory lies wholly within the hatched region ($E < 0$) since the trajectory starts within the $E < 0$ region and can therefore not escape.

It will spiral around $+A_1$ if a_0 is chosen positive.

(*b*) $a_0 > A_2$: The direction of decreasing E is away from the origin, and such a trajectory can never cross the line $y = A_2$. Thus there are no particlelike solutions for any $A(k) > A_2$. This result is also proved in Appendix A.

(*c*) $P_1 < a_0 < A_2$: All $A(k)$ must lie between P_1 and A_2. Some trajectories for values of a_0 in this range are shown in Fig. 2. The behavior is essentially the same as for the $B < 0$ and $B = 0$ cases, i.e., except for certain values of a_0 and $A(k)$ the trajectories enter the hatched regions, from which they cannot escape, and end up spiraling around $\pm A_1$.

iv. $B > \frac{3}{16}$

The diagram of Fig. 2 is typical only of B in the range $0 < B < \frac{3}{16}$. Let us define P_1 and P_2 as the points at which the $E = 0$ trajectory cuts the line $\dot{y} = 0$. This gives

$$P_1 = \left(\frac{\frac{1}{2} - (\frac{1}{4} - \frac{4}{3}B)^{\frac{1}{2}}}{\frac{2}{3}B}\right)^{\frac{1}{2}}, \quad P_2 = \left(\frac{\frac{1}{2} + (\frac{1}{4} - \frac{4}{3}B)^{\frac{1}{2}}}{\frac{2}{3}B}\right)^{\frac{1}{2}}. \tag{16}$$

In Fig. 3 we draw A_1, P_1, A_2, P_2 as a function of B. From this graph it is clear that the order of points O, A_1, P_1, A_2, P_2 is preserved up to $B = \frac{3}{16}$. For $B > \frac{3}{16}$, no real P_1 and P_2 exist; a typical phase space for the case $\frac{3}{16} < B < \frac{1}{4}$ is shown in Fig. 4. It is clear that for any $a_0 < A_2$, we are starting within a region for which E is negative, and, because the trajectory always moves to decreasing E, the origin (at which E is zero) can never be reached. For $a_0 > A_2$, we move in a direction away from the origin to a region of negative E.

Thus there are no values of a_0 for which particlelike solutions can exist.

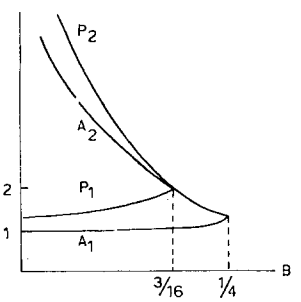

FIG. 3. Plot of A_1, P_1, A_2, and P_2 against B. For $B > \frac{3}{16}$ no real P_1 or P_2 exist. At $B = \frac{3}{16}$, $P_1 = A_2 = P_2$.

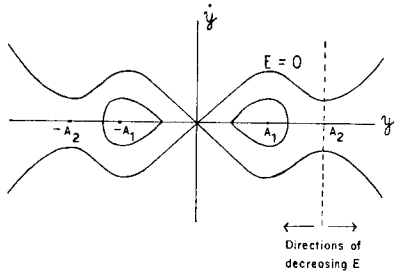

FIG. 4. Typical phase space for $\frac{3}{16} < B < \frac{1}{4}$. The whole axis $\dot{y} = 0$, except the origin, lies in a region of negative E.

Conclusion

From an analysis of (9) in the phase plane, it has been shown that particlelike solutions can exist for B in the range $-\infty < B < \frac{3}{16}$, but that such solutions cannot exist for $B > \frac{3}{16}$. When solutions do exist, the values of $A(k)$ can be bounded:

$$P_1 < A(k), \qquad \text{for} \quad -\infty < B < 0,$$

$$P_1 < A(k) < A_2, \quad \text{for} \quad 0 < B < \tfrac{3}{16}.$$

When $a_0 \neq A(k)$, the trajectory oscillates about one of the special solutions $\pm A_1$.

C. Numerical Solutions

Solutions to (9) were obtained numerically. A brief discussion of the methods used will be given elsewhere. In the following analysis the stability of the nodeless solution is discussed, and for this reason we plot only this solution in Fig. 5 for various B values. Note how strongly B controls the "size" of the solution. In Table I values of $A(1)$ and $D(1)$ are tabulated for some values of B.

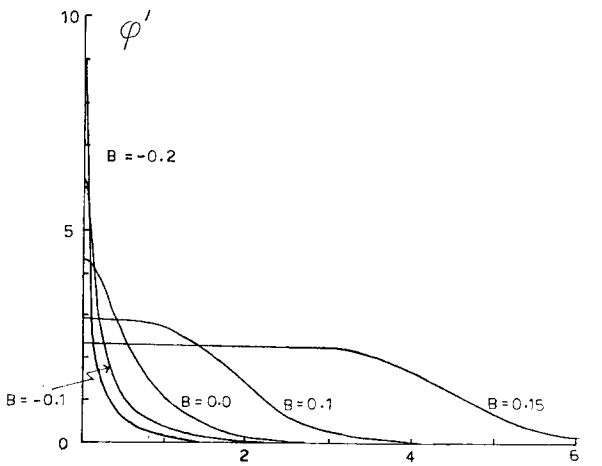

FIG. 5. Graph of the nodeless solution of (6) for various values of B. The larger the value of B, the more distended the solution. As $B \to \frac{3}{16}$, $\varphi'(r') \to 2$.

TABLE I. Table of numerically obtained values of $A(0)$ and $D(0)$ together with the values of A_2 for various B. Note how, as $B \to \frac{3}{16}$, $A(0) \to A_2 \to 2$. For $B > 0.1$, A_2 gives a good upper bound to $A(0)$.

B	$A(0)$	A_2	$D(0)$
−0.1	6.232 ± 0.002	—	1.09 ± 0.01
−0.04	5.023 ± 0.001	—	1.72 ± 0.01
0.0	4.337 ± 0.001	∞	2.71 ± 0.01
0.05	3.578 ± 0.001	4.35	5.94 ± 0.04
0.1	2.906 ± 0.001	2.98	22.7 ± 0.2
0.11	2.781 ± 0.001	2.82	34.6 ± 0.3
0.12	2.661 ± 0.001	2.68	58.2 ± 0.5
0.13	2.5456 ± 0.0005	2.55	113.0 ± 1.0
0.14	2.4358 ± 0.0003	2.437	279.0 ± 3.0
0.15	2.3325 ± 0.0001	2.3327	1040.0 ± 10.0
0.16	2.23606 ± 0.00005	2.236068	8960.0 ± 20.0

3. STABILITY BY FIRST-ORDER PERTURBATION THEORY

The stability of the nodeless solution of (13) to small perturbations is considered for $0 < B < \frac{3}{16}$. Let us denote by ψ_0 the undisturbed state, by ψ the perturbed state, and by $\psi_1 = \psi - \psi_0$ the perturbation, considered small at least initially.

By keeping up to first order in ψ_1 and ψ_1^* in (13) and using the transformations (7) and (8), one can derive the eigenvalue problem (see I)

$$\left(\nabla'^2 + \frac{(\Omega' + \omega')^2 - 1}{(1 - \omega'^2)} + 2\varphi_0'^2 - 3B\varphi_0'^4\right)\eta = (-\varphi_0'^2 + 2B\varphi_0'^4)\chi,$$

$$\left(\nabla'^2 + \frac{(\Omega' - \omega')^2 - 1}{(1 - \omega'^2)} + 2\varphi_0'^2 - 3B\varphi_0'^4\right)\chi = (-\varphi_0'^2 + 2B\varphi_0'^4)\eta, \quad (17)$$

where ψ_1 has been expressed in the form

$$\psi_1 = (\eta e^{-i\Omega t} + \chi^* e^{+i\Omega^* t})e^{-i\omega t}$$

and $\Omega' = \Omega/KC = \Omega_r' + i\Omega_i'$. The problem of determining the stability of ψ_0 has been reduced to finding the eigenvalues Ω_r' and Ω_i' of (17) as a function of ω' for given B.

In the $B = 0$ case, nontrivial solutions to (17) were found only for η and χ spherically symmetric, when Ω' was purely imaginary. In the present case a more complicated situation exists. Provided that $0 < B < \frac{3}{16}$, there exists a critical value of ω', denoted $\omega_c'(B)$, at which the nature of the eigenvalue Ω' changes. The eigenvalues of (17) are then of the form

$$\text{(i)}\ \Omega_r' = 0, \quad \Omega_i' \neq 0, \quad 0 < \omega' < \omega_c',$$
$$\text{(ii)}\ \Omega_r' \neq 0, \quad \Omega_i' = 0, \quad \omega_c' < \omega' < 1. \quad (18)$$

No solutions for which Ω_r' and Ω_i' were simultaneously nonzero were found. These results, and the results obtained in I, suggest that Ω'^2 should be real for this type of eigenvalue problem, but the author has been unable to prove this.

In Fig. 6(a), Ω_i' is plotted against Ω' for the values of $B = 0.05$, 0.075, 0.1, and 0.15375 while, in Fig. 6(b), Ω_r' is plotted against ω' for the same B values.

4. STABILITY BY DIRECT PERTURBATION METHODS

A. Introduction

The approach resembles that used in I. By defining $\rho = k\mathbf{r}$ and $\tau = kct$, one can reduce (3) to

$$\nabla_\rho^2 \Psi - \frac{\partial^2 \Psi}{\partial \tau^2} = \Psi - \Psi\Psi^*\Psi + \tilde{B}(\Psi\Psi^*)^2\Psi, \quad (19)$$

where

$$B = \lambda\kappa^2/\mu^4 = B/(1 - \omega'^2),$$
$$\Psi = \mu\psi/\kappa = (1 - \omega'^2)^{\frac{1}{2}}\varphi_0' e^{-i\omega'\tau}. \quad (20)$$

As before, φ_0' is the nodeless solution of (6). It is now obvious that $\tilde{B}$ is the important parameter with

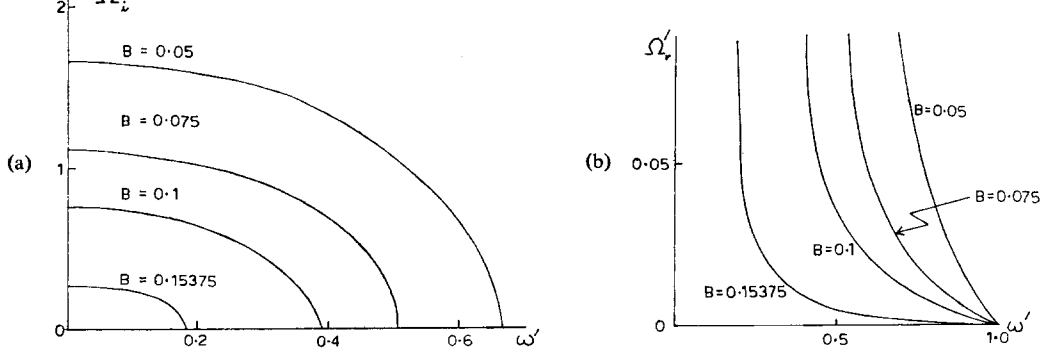

FIG. 6. Plot of eigenvalues Ω_r' and Ω_i' against ω' for fixed B as obtained by numerical solution of (17). Note the existence of a critical value of ω' for each positive B and that the nature of the solution changes from the form ($\Omega_r' = 0$, $\Omega_i' \neq 0$) to ($\Omega_r' \neq 0$, $\Omega_i' = 0$) at this point.

regard to stability. To solve (19), one must choose a value of ω' and $\tilde{B}$. If $\tilde{B} > \frac{3}{16}$, there is no particlelike solution to (19) of the form (1) unless ω' is chosen greater than $\omega'_L = (1 - \frac{3}{16}\tilde{B})^{\frac{1}{2}}$ since particlelike solutions exist only for $B < \frac{3}{16}$.

Before considering the stability of solutions by direct perturbation, it is worthwhile examining the energy density. Equation (5) can be written in transformed notation as

$$\mathcal{E} = \frac{\kappa^4}{\mu^2}\mathcal{E}' = \frac{\kappa^4}{\mu^2}\left(\left|\frac{\partial\Psi}{\partial\tau}\right|^2 + |\nabla_\rho\Psi|^2 + |\Psi|^2 - \tfrac{1}{2}|\Psi|^4 + \tfrac{1}{3}\tilde{B}|\Psi|^6\right). \tag{21}$$

Then $\mathcal{E}$ will be positive definite provided $\tilde{B} > \frac{3}{16}$, which implies that a state with positive-definite energy density can occur only if the particlelike solution is time dependent. In Fig. 7, $\tilde{B}$ is plotted against ω' along curves of constant B and the regions of (i) no particlelike solutions, (ii) particlelike solutions with positive-definite energy density, and (iii) particlelike solutions for which the energy density need not be positive definite are marked. One can also incorporate the results of Sec. 3 by marking on the curve of critical ω'_c as a function of B and ω'. This shows that the results of first-order perturbation theory imply the existence of (a) stable solutions for which $\mathcal{E}$ is positive definite, (b) stable solutions for which $\mathcal{E}$ need not be positive definite, and (c) unstable solutions. The curve of ω'_c is not plotted in the region $\omega' \to 1$ because of the numerical inaccuracy in obtaining the value of $\tilde{B}$ corresponding to a pair of values (ω'_c, B) in this region. [The error in $\tilde{B}$ denoted $\Delta\tilde{B} = 2B\omega'_c\Delta\omega'_c/(1 - \omega'^2_c)$ can be large for even a small error $(\Delta\omega'_c)$ in ω'_c when $\omega' \to 1$.]

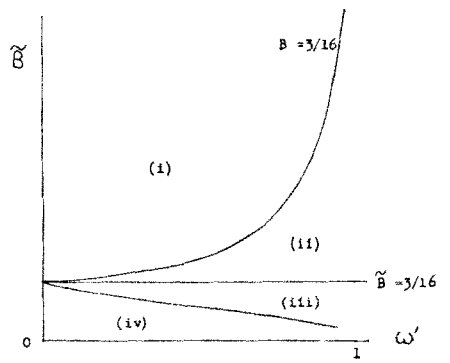

FIG. 7. Graph of $\tilde{B}$ vs ω' showing how the space splits into four regions: (i) region of no particlelike solutions $(B > \frac{3}{16})$, (ii) region where the energy density is positive definite, (iii) region of stable solutions with $\mathcal{E}$ not positive definite, and (iv) region of unstable particlelike solutions.

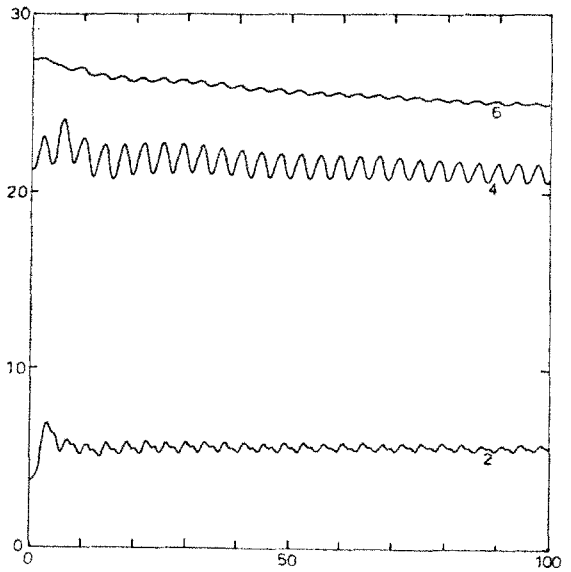

FIG. 8. Time development of the state Ψ for the point S after it has been forced to oscillate at an unnatural frequency. The state oscillates with an angular frequency of ~ 1.6 radians/τ-unit and radiates some "excess energy."

We will test the conclusions (a), (b), and (c) of first-order perturbation theory by examining the stability of typical points S, R, and U in the three regions (see Fig. 7). It will be shown that S and R are very stable, but that U is highly unstable.

B. Direct Perturbation Results

For an unstable particle, a convenient method of illustrating graphically the time development of the particle state, when disturbed, is to plot the energy density against the reduced radial distance ρ for a series of reduced times τ (see I). For a stable particle this is unsatisfactory, and one must adopt an alternative method. Let us define $\bar{E}(a)$ by

$$\bar{E}(a) = \int_0^a \mathcal{E}\, dv. \tag{22}$$

Let us consider the consequences of applying various disturbances to S $(\tilde{B} = \frac{5}{32}, \omega' = 0.8)$. Figure 8 shows the results of trying to force the particle to oscillate at an unnatural frequency. The initial disturbed state is defined by

$$\Psi]_{\tau=0} = 0.6\varphi'_0, \quad \frac{\partial\Psi}{\partial\tau}\Big] = -0.3i\varphi'_0, \tag{23}$$

where φ'_0 is the nodeless solution of (6) corresponding to S (i.e., $B = 0.1$, $\omega' = 0.8$). Plotted in Fig. 8 is $\bar{E}(a)$ for the values of $a = 2$, 4, 6, and on all three levels oscillations with an angular frequency of ~ 1.6 radians τ-unit can be seen. In Table II the values of $\bar{E}(a)$ for various a for the undisturbed state S are given for comparison. Although $\bar{E}(\infty)$ is not plotted

TABLE II. Values of $\bar{E}(a)$ for $a = 2, 4, 6, 8, 10$ for the undisturbed state S.

a	2	4	6	8	10
$\bar{E}(a)$	6	29.9	37.4	38.1	38.2

in Fig. 8, it is only marginally larger than $\bar{E}(6)$ since all the energy is effectively contained within a sphere of radius $\rho = 6$. Comparison of $\bar{E}(\infty)$ for the disturbed and undisturbed states shows that, although the energy of the perturbed state is less than that of the undisturbed state, there is no decay. Some energy is radiated as can be seen from the fall of $\bar{E}(6)$, but this curve quickly levels out. That stable perturbed states of lower energy than the initial state can exist is an undesirable feature of this field, resulting from the fact that there exists a continuous range of solutions for any $\tilde{B}$.

Figures 9–12 show how the particle responds to disturbances initiating outside the particle radius. For the case shown in Fig. 9, the energy of the disturbance is as large as that of the unperturbed state. Such a disturbance cannot be considered small in any sense; yet a decay is not induced. Instead, the particle breaks into oscillation and radiates some excess energy. The disturbance applied in this case can be given by

$$\Psi_1|_{\tau=0} = \zeta_1(4), \quad \left.\frac{\partial \Psi_1}{\partial \tau}\right|_{\tau=0} = -3\zeta_1(4), \tag{24}$$

where $\zeta_1(4)$ is as defined in Appendix B.

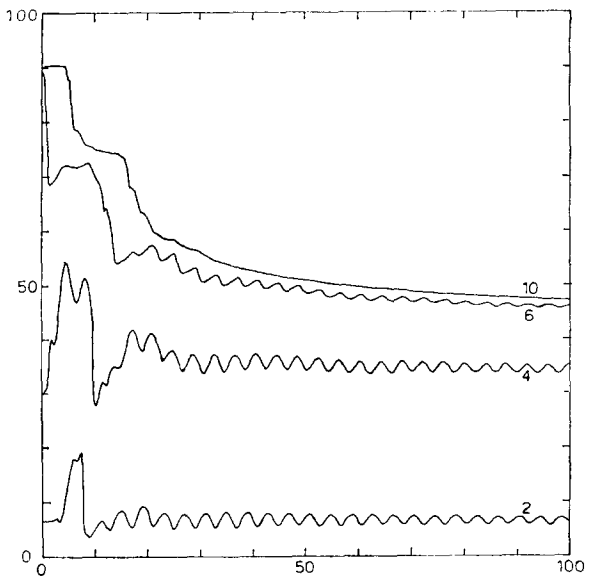

FIG. 9

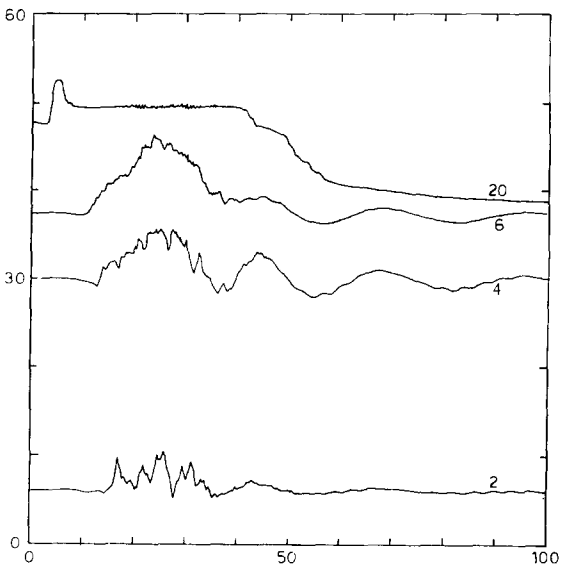

FIG. 10

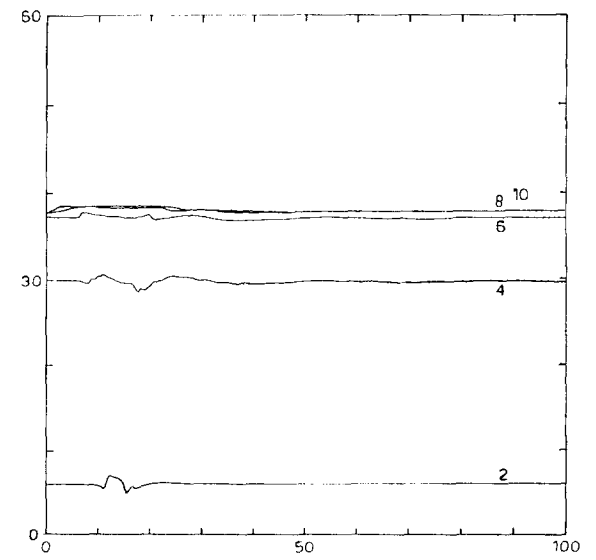

FIG. 11

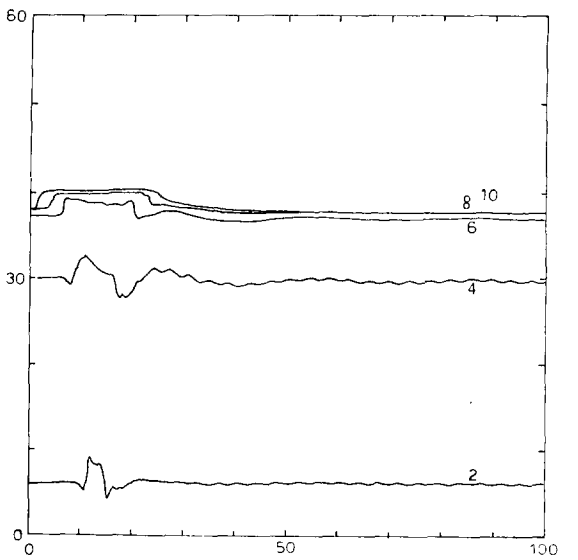

FIG. 12

FIGS. 9–12. Plots of $\bar{E}(a)$ vs τ for S for various applied disturbances.

In Figs. 11 and 12 the applied disturbance can be seen moving in on the particle center and also, though less clearly, can be seen propagating outwards after exciting the particle.

The graphs in Figs. 8–12 inclusive are all for rather large disturbances. One might ask what happens when small random disturbances are applied (e.g., similar to those applied in I). In this case the particle does not decay, but the perturbed levels $\bar{E}(a)$ remain so close to the unperturbed levels (which are simply straight lines parallel to the τ axis) that they do not illustrate any new interesting behavior. In order to produce a development noticeably different from the unperturbed state, one must apply a moderate to severe disturbance.

Examination of the stability of the point R ($\tilde{B} = \frac{5}{18}$, $\omega' = 0.8$) leads one to the conclusion that it is also a very stable state, exhibiting behavior similar to S, but that this time the energy density need not be positive definite. The fact that during the time evolution the energy density might become negative in some region of space for some disturbances does not seem to render R less stable than a state such as S for which $\mathcal{E}$ can never be negative anywhere.

Although S and R are, as suggested by first-order perturbation theory, found by direct perturbation methods to be extremely stable, no state can be unconditionally stable. Figure 13 shows $\bar{E}(a)$ vs τ, $a = 2, 4, 6, 8, 10$, for a disturbance which induced a dissipative decay in S.

The point U ($\tilde{B} = 0.1$, $\omega' = 0.2$) is, as expected, extremely unstable, decaying at the slightest provocation. In Fig. 14 we illustrate a new type of decay

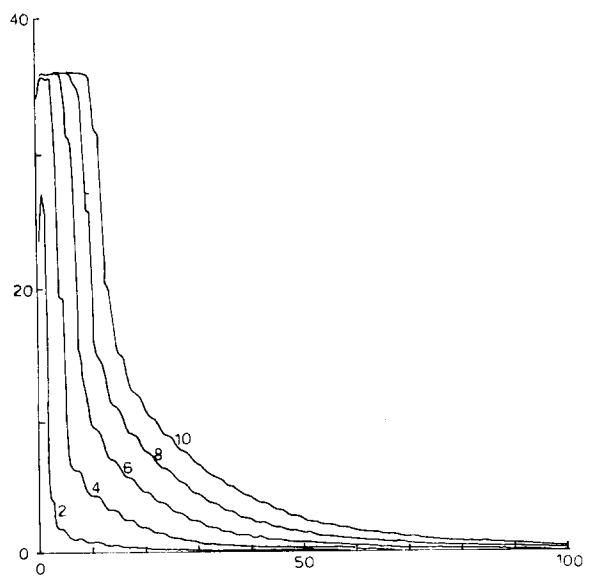

FIG. 13. Example of a dissipative mode of decay for S. $\bar{E}(a)$ is plotted against τ for the values of $a = 2, 4, 6, 8, 10$.

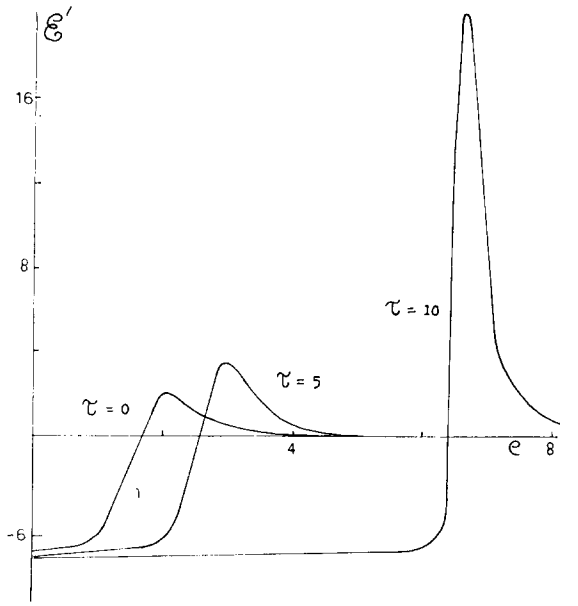

FIG. 14. Graph of the reduced energy density $\mathcal{E}'$ against the reduced radial distance ρ at a series of reduced times τ for the point U. Notice how $\mathcal{E}'$ tends to assume the constant value (≈ -7.2) over an increasingly large volume of space.

exhibited by U. In this case, $\mathcal{E}$ tends to a constant value over an increasingly large area of space. Energy conservation requires that the positive bump increase in magnitude to compensate for the increase in negative energy as for the singular decay of I, but it must travel outwards in this case.

C. Summary

The results of first-order perturbation theory are confirmed by direct perturbation techniques. The stability of points, for which $\mathcal{E}$ must be positive definite (S) and for which it need not be (R, U), is considered. From Sec. 3, it is suggested that S and R are stable, but that U is unstable. Direct perturbation methods confirm this. It is found that S and R exhibit very much the same behavior: that both are extremely stable and that both tend to break into oscillation and radiate some excess energy when excited, and that both can be destroyed.

5. CONCLUSION

Stable time-dependent particlelike solutions are found to exist to the field equation (3), for suitable choice of the field parameters κ, μ, λ, and ω. One can have the further refinement that the energy density of the particle is positive definite if $\lambda\kappa^2/\mu^4 > \frac{3}{16}$ while $\lambda(\kappa^2 - \omega^2/c^2)/\mu^4 < \frac{3}{16}$. We would like to be able to consider an elementary particle as being represented by a local concentration of the energy density and thus intuitively feel that the energy density ought to be positive definite. This has the advantage of eliminating

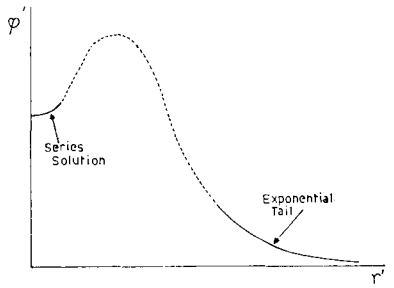

FIG. 15. Shape of a nodeless solution to (6) for $a_0 > A_2$ showing the series solution for $r' \sim 0$, the exponential decay for $r' \to \infty$, and the necessity for a maximum at some finite value of r'.

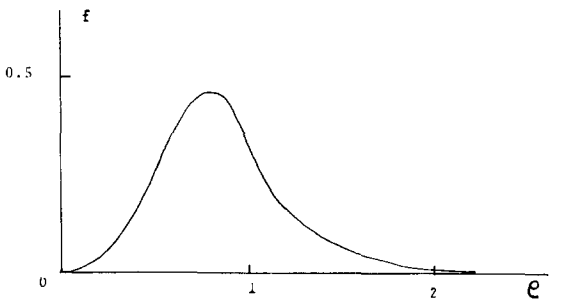

FIG. 16. Plot of the function $f(\rho)$ against ρ.

the unphysical decay modes which can occur if δ can be negative. However, although the field here considered is a considerable improvement over that discussed in I, it cannot be considered as a sensible one for an elementary particle. For example, the results imply the existence of stable spinless massive particles, and no such particles have been observed to date. Further, the problem of assigning values to the field parameters is no nearer solution here than it was in I. Nevertheless, the results encourage further investigation of such fields. The onset of stability as a function of ω has not been considered in this paper, but there is an observed correlation between the critical value of ω' and the extrema of the total integrated energy. It is hoped that this may be the subject of a later work.

ACKNOWLEDGMENTS

The author would like to express his gratitude to G. H. Derrick for many helpful discussions and his indebtedness to the Carnegie Trust for the Universities of Scotland for a research scholarship.

APPENDIX A: PROOF THAT THERE CAN BE NO PARTICLELIKE SOLUTIONS TO (6) IF $a_0 > A_2$

For small r', Eq. (6) has the series solution

$$\varphi' = a_0 + a_2 r'^2 + a_4 r'^4 + \cdots, \qquad \text{(A1)}$$

where $a_2 = \frac{1}{6}a_0(1 - a_0^2 + Ba_0^4)$ will be positive if $a_0 > A_2$.

Thus, for $a_0 > A_2$, φ' is an increasing function in the neighborhood of the origin. The asymptotic form of a particlelike solution to (9) for large r' is $\varphi' \to De^{-r'}/r'$. Any particlelike solution must be a combination of these two forms. For the sake of example, let us consider a nodeless solution. Then it must be of the form shown in Fig. 15. Since φ' increases in the $n'h'd$ of the origin, φ' must have a maximum. The condition for a maximum is $d\varphi'/dr' = 0$, $d^2\varphi'/dr'^2 < 0$. When $d\varphi'/dr' = 0$, Eq. (6) has the form

$$\frac{d^2\varphi'}{dr'^2} = \varphi' - \varphi'^3 + B\varphi'^5. \qquad \text{(A2)}$$

But because φ' must be greater than A_2 at this point, the right-hand side of (A2) is positive, implying that (6) cannot have a maximum if $\varphi' > A_2$, which in turn implies that (6) can have no nodeless particlelike solutions if $a_0 > A_2$. The nonexistence of the higher-node solutions follows as a simple extension of this argument. This argument holds for any B in the range $0 < B < \frac{1}{4}$.

Corollary: There can be no particlelike solution to (6) for any a_0 if $B > \frac{1}{4}$. If $B > \frac{1}{4}$, the coefficient a_2 is positive for any positive a_0, and so is the rhs of (A2). Thus it is not possible to have particlelike solutions for any a_0 in this case.

APPENDIX B: DEFINITION OF $\zeta_1(b)$

When we were applying disturbances to Ψ_0, it was convenient to have in the computer a series of basic disturbances from which other disturbances could be constructed. Such a disturbance is ζ_1.

By $\zeta_1(b)$ is meant

$$\zeta_1 = 0, \qquad \rho < b,$$
$$\zeta_1 = f(\rho - b), \qquad \rho \geq b,$$

where $f(\rho)$ is shown graphically in Fig. 16.

* Present address: Commonwealth Meteorology Research Centre, P.O. Box 5089AA, Melbourne, Australia.

[1] D. L. T. Anderson and G. H. Derrick, J. Math. Phys. **11**, 1336 (1970).

$O(2, 1)$ and the Harmonic Oscillator Radial Function*

LLOYD ARMSTRONG, JR.
The Johns Hopkins University, Baltimore, Maryland 21218

(Received 1 October 1970)

The radial functions of the three-dimensional isotropic harmonic oscillator are shown to form bases for unitary representations of the noncompact group $O(2, 1)$. The functions r^s are shown to transform simply under the action of the generators of this group. As a result, matrix elements of r^s can easily be evaluated. Selection rules on these matrix elements are obtained by studying Kronecker products of representations of $O(2, 1)$.

I. INTRODUCTION

In a previous paper,[1] we analyzed the radial wavefunctions of hydrogen using the noncompact group $O(2, 1)$. In this paper, we wish to carry out a like analysis of the radial wavefunctions of the harmonic oscillator.

The eigenvalue equation for a particle moving in a harmonic oscillator potential can, of course, be solved exactly. For that reason, this potential is used in many problems in physics and chemistry to approximate potentials whose exact form is not known, or is too complicated to treat conveniently.[2] As a result, the harmonic oscillator has been studied using a variety of techniques.[2–5]

One of the techniques which has been used for these studies is, of course, group theory. That the radial functions of the harmonic oscillator form bases for representation of $O(2, 1)$ has been known for several years.[4] It appears, however, that this symmetry has not been exploited in the evaluation of matrix elements.[5] In this paper, we utilize the group properties of these functions in order to evaluate matrix elements of r^s (s a positive or negative integer). In the process, we obtain a number of interesting selection rules for matrix elements on r^s. These rules are analogous to those obtained by Pasternack and Sternheimer[6] for the hydrogen atom.

II. THE ALGEBRA AND ITS REPRESENTATIONS

The complexification of the algebra of $O(2, 1)$ is composed of three operators, J_+, J_-, and J_3, which satisfy the commutation relations

$$[J_3, J_\pm] = \pm J_\pm,$$
$$[J_+, J_-] = 2J_3. \tag{1}$$

A realization of these generators in a two-dimensional (z, t) space is given by[7]

$$J_\pm = e^{it}\left(z\frac{\partial}{\partial z} \mp i\frac{\partial}{\partial t} \mp \frac{z}{2}\right),$$
$$J_3 = i\frac{\partial}{\partial t}. \tag{2}$$

A basis for an irreducible representation of this algebra is formed by the states f_{ab}:

$$f_{ab} = \left[\frac{(l+\frac{1}{2})}{4\pi\beta^{\frac{1}{2}}}\right]^{\frac{1}{2}} e^{i(n+\frac{1}{2})t/2} z^{\frac{1}{4}} R_{nl}(z), \tag{3}$$

where

$$a = \tfrac{1}{2}(l - \tfrac{1}{2}),$$
$$b = \tfrac{1}{2}(n + \tfrac{1}{2}),$$

and $R_{nl}(\beta r^2)/r$ is the radial wavefunction for the three-dimensional isotropic harmonic oscillator:

$$R_{nl}(\beta r^2) = \left[\frac{(\frac{1}{2}n - \frac{1}{2}l - \frac{1}{2})!\, 2\beta^{\frac{1}{2}}}{\Gamma(\frac{1}{2}n + \frac{1}{2}l + 1)^3}\right]^{\frac{1}{2}} (\beta r^2)^{\frac{1}{2}(l+1)} e^{-\beta r^2/2} \times L^{l+\frac{1}{2}}_{\frac{1}{2}(n-l-1)}(\beta r^2). \tag{4}$$

In Eqs. (3) and (4), $\beta = m\omega/\hbar$, where $\omega/2\pi$ is the classical frequency of oscillation and L^c_a is the Laguerre polynomial of Morse and Feshbach.[8]

The action of the operators given by Eqs. (2) on the states f_{ab} is described by the equations

$$J_3 f_{ab} = b f_{ab},$$
$$J_\pm f_{ab} = \pm[(b \mp a)(b \pm a \pm 1)]^{\frac{1}{2}} f_{ab\pm1}. \tag{5}$$

The Casimir operator for this algebra, denoted by J^2, can also be easily investigated. One finds

$$J^2 f_{ab} = (J_+J_- + J_3^2 - J_3) f_{ab} = a(a+1) f_{ab}.$$

By using Eqs. (2), (4), and setting $z = \beta r^2$, this equation can be written in the familiar form

$$\left[\frac{d^2}{dr^2} - \frac{m^2\omega^2}{\hbar^2} r^2 + \frac{2m\omega}{\hbar}(n + \tfrac{1}{2})\right] R_{nl} = \frac{l(l+1)}{r^2} R_{nl}.$$

The second of Eqs. (5) indicates that

$$J_- f_{aa+1} = 0,$$

thus providing a lower bound to the basis. Clearly, however, $J_+ f_{ab} \neq 0$ for $b > a$; there is, therefore, no upper bound to the representation, i.e., the representation is infinite dimensional. It is also clearly irreducible. Further, the eigenvalues of the operators J_+J_- and J_-J_+ are real and negative definite. This condition, in conjunction with the reality of a and b, implies that

the representation formed by the states f_{ab} $(b > a)$ is unitary.[9] This representation is, of course, just the positive discreet representation[10] $\mathfrak{D}_a^+$.

The states f_{ab} can be used to form the basis for a Hilbert space[9,10]; we define the scalar product in this space by

$$(f_{a'b'} \mid f_{ab}) = \int f^*_{a'b'} f_{ab}\, d\Omega, \tag{6}$$

where $d\Omega = dt\, dz/z^2$. Using this definition, we find

$$(f_{a'b'} \mid f_{ab}) = \delta(a, a')\delta(b, b'), \tag{7}$$

and

$$(f_{a'b'} \mid J_3 \mid f_{ab}) = (J_3 f_{a'b'} \mid f_{ab}),$$
$$(f_{a'b'} \mid J_\pm \mid f_{ab}) = -(J_\mp f_{a'b'} \mid f_{ab}).$$

III. OPERATORS AND MATRIX ELEMENTS

The purpose of the present section is to utilize the algebra of the previous section in the calculation of matrix elements of z^k. To facilitate the calculation, we introduce the operators

$$T^{(k)}{}_q = z^{-k} e^{iqt}, \quad 2k \geq 0, \tag{8}$$
$$P^{(k)}{}_q = z^{k+1} e^{iqt}, \quad 2k \geq -1, \tag{9}$$

where $2k$ is an integer and $2q$ is a positive or negative integer. The transformation properties of these operators with respect to the operations of the group is determined by their commutators with respect to the algebra:

$$\begin{aligned} [J_\pm, T^{(k)}{}_q] &= -(k \mp q) T^{(k)}{}_{q\pm 1}, \\ [J_3, T^{(k)}{}_q] &= q T^{(k)}{}_q, \\ [J_\pm, P^{(k)}{}_q] &= (k \pm q + 1) P^{(k)}{}_{q\pm 1}, \\ [J_3, P^{(k)}{}_q] &= q P^{(k)}{}_q. \end{aligned} \tag{10}$$

In order to determine the dependence of the matrix element $(f_{a'b'} \mid T^{(k)}{}_q \mid f_{ab})$ on the "magnetic" quantum numbers b, b', and q, we evaluate matrix elements of the first of the Eqs. (10), and use Eq. (5):

$$\begin{aligned} &[(b' \pm a')(b' \mp a' \mp 1)]^{\frac{1}{2}} (f_{a'b'\mp 1} \mid T^{(k)}{}_q \mid f_{ab}) \\ &\quad - [(b \mp a)(b \pm a \pm 1)]^{\frac{1}{2}} (f_{a'b'} \mid T^{(k)}{}_q \mid f_{ab\pm 1}) \\ &\quad = -(k \mp q)(f_{a'b'} \mid T^{(k)}{}_{q\pm 1} \mid f_{ab}), \end{aligned}$$

where

$$a' = \tfrac{1}{2}(l' - \tfrac{1}{2}),$$
$$b' = \tfrac{1}{2}(n' + \tfrac{1}{2}).$$

The top signs in the above equation result from taking the commutator of $T^{(k)}{}_q$ with J_+, the lower, from the commutator with J_-. Repeatedly using the relationship resulting from the commutator with J_+, we obtain

$$(f_{a'b'} \mid T^{(k)}{}_q \mid f_{ab}) = \left(\frac{\Gamma(2a+2)\Gamma(a'+b'+1)(b'-a'-1)!\,(b-a-1)!}{\Gamma(a+b+1)}\right)^{\frac{1}{2}}$$
$$\times \sum_t (b-a-1-t)!^{-1} [\Gamma(q+a+a'+t+2)(a-a'+q+t)!]^{-\frac{1}{2}} \binom{k-q}{t} (f_{a'q+t+a+1} \mid T^{(k)}{}_{q+t} \mid f_{aa+1}).$$

In a like manner, repeated use of the relationship resulting from the commutator of $T^{(k)}{}_q$ with J_- leads to the result

$$(f_{a'b'} \mid T^{(k)}{}_q \mid f_{aa+1}) = \left(\frac{\Gamma(2a'+2)(b'-a'-1)!}{\Gamma(a'+b'+1)}\right)^{\frac{1}{2}} \binom{k+q}{b'-a'-1} (f_{a'a'+1} \mid T^{(k)}{}_{a'-a} \mid f_{aa+1}).$$

Combining these two results, we find that we can write

$$(f_{a'b'} \mid T^{(k)}{}_q \mid f_{qb}) = A(kq, ab \mid a'b')(f_{a'a'+1} \mid T^{(k)}{}_{a'-a} \mid f_{aa+1}), \tag{11}$$

where the coefficient $A(kq, ab \mid a'b')$ is defined as

$$A(kq, ab \mid a'b') = \left(\frac{\Gamma(2a+2)\Gamma(2a'+2)\Gamma(a'+b'+1)(b'-a'-1)!\,(b-a-1)!}{\Gamma(a+b+1)}\right)^{\frac{1}{2}}$$
$$\times \sum_t (-1)^{q+a-a'+t} [(b-a-1-t)!\,\Gamma(a+a'+q+t+2)]^{-1} \binom{a-a'-k-1}{q+a-a'+t} \binom{k-q}{t}. \tag{12}$$

In the above equations, the binomial coefficients are defined as

$$\binom{n}{r} = \frac{\Gamma(n+1)}{\Gamma(n-r+1)r!} \quad \text{if } 2n \text{ is an odd integer,}$$

$$\binom{n}{r} = \frac{n!}{(n-r)!\,r!} \quad \text{if } 2n \text{ is an even integer} \geq 0,$$

$$\binom{n}{r} = \frac{(-1)^r(r-n-1)!}{(-n-1)!\,r!} \quad \text{if } 2n \text{ is an even integer} \leq 0. \tag{13}$$

Further, the factorial of a negative number is not defined, and any term in Eq. (12) involving such a negative factorial in either the numerator or denominator should be set equal to zero.

Matrix elements of the operators $P^{(k)}{}_q$ can be studied using the same techniques. In this case, one finds

$$(f_{a'b'}|\, P^{(k)}{}_q\, |f_{ab}) = A(-k-1\ q, ab \mid a'b') \times (f_{a'a'+1}|\, P^{(k)}{}_{a'-a}\, |f_{aa+1}), \tag{14}$$

where $A(-k-1\ q,\ ab \mid a'b')$ is obtained by substituting $-k-1$ for k in Eq. (12).

Finally, to complete the calculation, one must obtain the matrix elements on the right-hand sides of Eqs. (11) and (14). This is easily done by recalling that[8] $L_0^s = \Gamma(s+1)$ and using Eqs. (3), (4), (6), (8) and (9):

$$(f_{a'a'+1}|\, T^{(k)}{}_{a'-a}\, |f_{aa+1}) = \frac{\Gamma(a+a'-k+1)}{[\Gamma(2a+1)\Gamma(2a'+1)]^{\frac{1}{2}}},$$

$$(f_{a'a'+1}|\, P^{(k)}{}_{a'-a}\, |f_{aa+1}) = \frac{\Gamma(a+a'+k+2)}{[\Gamma(2a+1)\Gamma(2a'+1)]^{\frac{1}{2}}}. \tag{15}$$

IV. SELECTION RULES

In the previous section, we used properties of the algebra to obtain matrix elements of the operators $P^{(k)}$ and $I^{(k)}$. In this section, we shall obtain selection rules for these matrix elements by considering Kronecker products of representations of $O(2, 1)$.

In order to carry out such a study, it is convenient to define states which satisfy the relations

$$\begin{aligned} J_\pm t_{kq} &= -(k \mp q)t_{kq\pm1}, \\ J_3 t_{kq} &= q t_{kq}, \\ J_\pm p_{kq} &= (k \pm q + 1)p_{kq\pm1}, \\ J_3 p_{kq} &= q p_{kq}. \end{aligned} \tag{16}$$

The transformation properties of the states produced by the action of the operator $T^{(k)}{}_q$ on the state f_{ab} are the same as the transformation properties of the states formed by the product $t_{kq} f_{ab}$. We can now utilize the well-known result that the matrix element

$$(f_{a'b'}|\, T^{(k)}{}_q\, |f_{ab})$$

vanishes identically unless $\mathfrak{D}_{a'}^+$ is contained at least once in the decomposition of the Kronecker product $\mathfrak{D}(k) \times \mathfrak{D}_a^+$, where $\mathfrak{D}(k)$ is a representation which has as basis the states t_{kq}. Corresponding relationships hold, of course, between the properties of the operators $P^{(k)}{}_q$ and the states p_{kq}.

We must now determine the representations whose bases are formed by the states of Eqs. (16). We consider first the case in which either both $2k$ and $2q$ are even integers, or they are both odd integers. Clearly, in this case, the states t_{kq} form a representation which is not fully reducible. However, an irreducible representation can be formed in the subspace with $|q| \leq k$. This finite-dimensional irreducible representation must be nonunitary because $O(2, 1)$ has no finite-dimensional unitary representation (except the trivial one-dimensional identity representation).[9] We denote this irreducible representation as $\mathfrak{D}(k)$. The Kronecker product $\mathfrak{D}(k) \times \mathfrak{D}_a^+$ has previously been studied[1]; the results can be stated by the decomposition

$$\mathfrak{D}(k) \times \mathfrak{D}_a^+ = \sum_{(a'=|k-a|)}^{k+a} \mathfrak{D}_{a'}^+ + \text{other}. \tag{17}$$

Here, "other" refers to representations for which $a' < |k-a|$; such representations are of no real interest here since matrix elements $(f_{a'b'}|\, T^{(k)}{}_q\, |f_{ab})$, where $a' < |k-a|$, are always equal to infinity. For our purposes, then, if $|q| = |b-b'| \leq k$, matrix elements of $T^{(k)}$ must vanish when $|a-a| > k$.

Considering still only the case in which both $2k$ and $2q$ are either even or odd integers, one finds that the states p_{kq} also form a representation which is not fully reducible. In this case, two irreducible representations can be formed by considering specific subspaces only. In particular, the states p_{kq} with $q \geq k+1$ form a basis for the representation $\mathfrak{D}_k^+$, and the states with $-q \geq k+1$ form a basis for the representation $\mathfrak{D}_k^-$, the negative discrete representation.[10] Both of these representations are unitary and irreducible. The Clebsch–Gordan series for $\mathfrak{D}_k^+ \times \mathfrak{D}_a^+$ and $\mathfrak{D}_k^- \times \mathfrak{D}_a^+$ have been studied by several authors.[11,12] Their results can be summarized as follows:

$$\begin{aligned} \mathfrak{D}_k^+ \times \mathfrak{D}_a^+ &= \sum_{a'=a+k+1}^{\infty} \mathfrak{D}_{a'}^+, \\ \mathfrak{D}_k^- \times \mathfrak{D}_a^+ &= \sum_{a'=0}^{a-k-1} \mathfrak{D}_{a'}^+ + \mathfrak{D}, \quad a > k, \\ &= \sum_{a'=0}^{k-a-1} \mathfrak{D}_{a'}^- + \mathfrak{D}, \quad a < k, \end{aligned} \tag{18}$$

where $\mathfrak{D}$ signifies continuous representations.[9,10] The exact specification of the continuous representations of no interest to us, since we are interested only in representations of $\mathfrak{D}^{+}_{a'}$, which appear on the right-hand side of Eq. (18). From this equation, then, we obtain the result that when $q = b' - b \geq k + 1$, matrix elements of $P^{(k)}$ vanish if $a' < a + k + 1$; when $-q = b - b' \geq k + 1$, matrix elements of $P^{(k)}$ vanish if $a' > a - k - 1$.

Finally, if we consider the case in which $2k$ is an even (odd) integer, and $2q$ is an odd (even) integer, we find that no irreducible unitary representations are formed by the states t_{kq} or p_{kq}. Correspondingly, there are no selection rules present in this case.

V. DISCUSSION

We can collect the results of the previous sections in order to write out explicitly values of the desired matrix elements:

$$\int R_{n'l'}R_{nl}r^{s}\,dr = \beta^{-s/2}(-1)^{\frac{1}{2}(n-n'+l'-l)}\Gamma[\tfrac{1}{2}(l+l'+s+3)]\left(\frac{\Gamma[\frac{1}{2}(n'+l')+1][\frac{1}{2}(n-l-1)]!\,[\frac{1}{2}(n'-l'-1)]!}{\Gamma[\frac{1}{2}(n+l)+1]}\right)^{\frac{1}{2}} \times \sum_{t}\frac{(-1)^{t}}{[\frac{1}{2}(n-l-1)-t]!\,\Gamma[\frac{1}{2}(l+l'+1+n'-n)+t+1]}\binom{\frac{s+l-l'}{2}}{\frac{n'-n+l-l'}{2}+t}\binom{\frac{n-n'-s-2}{2}}{t}, \tag{19}$$

valid for all integers. Selection rules obtained from the use of group theory indicate that this integral must be set equal to zero when $|n - n'|$ and s are either both even integers or both odd integers and

$$\text{if } s+2>0,\quad n'-n\geq s+2>l'-l \quad \text{or} \quad n-n'\geq s+2>l-l'$$

$$\text{or, if } s+2\leq 0,\quad |n-n'|\leq -s-2<|l-l'|.$$

The reason that these selection rules are not explicitly contained in Eq. (19) is easily understood. When algebraic techniques such as those of Sec. III are used to determine the coefficient $A(kq, ab \mid a'b')$, selection rules appear only when some normalization condition is applied to $A(kq, ab \mid a'b')$.[1,11] We have, of course, made no attempt to carry out such a normalization.

In a previous paper,[1] we noted that matrix elements of r^s evaluated with hydrogenic radial functions are proportional to the Clebsch–Gordan coefficients for $O(3)$. The same type of proportionality exists for the matrix elements discussed in this paper. Consider a Clebsch–Gordan coefficient of $O(3)$ having one of the angular momenta and its corresponding magnetic quantum number fixed at some explicit numerical value, with all of the other parameters left unspecified. Such a coefficient can generally be written as a simple algebraic function of the unspecified parameters. For examples of such functions, see Edmonds.[13] We call these "algebraic" Clebsch–Gordan coefficients. One can, of course, carry out the same procedure for a Clebsch–Gordan coefficient of $O(2, 1)$. One finds, after writing out the algebraic coefficients for both groups, that they are identical to within a phase.[1] The function $A(kq, ab \mid a'b')$ is clearly equal to a Clebsch–Gordan coefficient of $O(2, 1)$ to within a phase and a factor depending only on k, a, and a'; thus $A(kq, ab \mid a'b')$ must have the same dependence in q, b, and b' as does $(kq, ab \mid a'b')_{\text{alg}}$, the algebraic form of the Clebsch–Gordan coefficient of $O(3)$. Then, for example, using Eqs. (5), (8), (9), and (14), we find

$$\int R_{n'l'}R_{nl}r^{s}\,dr \sim A(-\tfrac{1}{2}s-1\ b'-b, ab \mid a'b') \sim (s/2\ b'-b, ab \mid a'b')_{\text{alg}} \tag{20}$$

for $s > 0$. In arriving at the second line of (20), we have used the well-known invariance of the Clebsch–Gordan coefficient of $O(3)$ under the interchange $k \to -k - 1$. That this proportionality is correct is easily seen by using the tables of Shaffer[3] for the radial integral above and the tables of Edmonds[13] for the algebraic Clebsch–Gordan coefficient. For example, from Shaffer[3] we have

$$\int R_{n-1\,l+1}R_{nl}r^{3}\,dr \sim (\tfrac{3}{2}n+\tfrac{1}{2}l+\tfrac{1}{2})[(n-l-1)/2]^{\frac{1}{2}},$$

while from Edmonds we find

$$(\tfrac{3}{2}\ \tfrac{1}{2}, ab \mid a+\tfrac{1}{2}\ b-\tfrac{1}{2})_{\text{alg}} \sim (a+3b)(b-a-1)^{\frac{1}{2}} = (\tfrac{3}{2}n+\tfrac{1}{2}l+\tfrac{1}{2})[(n-l-1)/2]^{\frac{1}{2}}.$$

Proportionalities such as (20) also exist, of course, for $s < 0$.

ACKNOWLEDGMENTS

Much of this work was done during the summer of 1970 while I was visiting the Laboratoire Aime Cotton, Orsay. I would like to thank Prof. Jacquinot and Dr. S. Feneuille for their hospitality during this period.

* Supported in part by the U.S. Atomic Energy Commission.

[1] L. Armstrong, Jr., J. Phys. **31**, C4 (1970).
[2] M. Moshinsky, *Harmonic Oscillator in Modern Physics: From Atoms to Quark* (Gordon & Breach, New York, 1969).
[3] W. H. Shaffer, Rev. Mod. Phys. **16**, 245 (1944); S. Bell, J. Phys. **B3**, 735, 745 (1970).
[4] S. Goshen and H. L. Lipkin, Ann. Phys. (N.Y.) **6**, 301 (1959); H. Bacry and J. L. Richard, J. Math. Phys. **8**, 2231 (1967).
[5] J. D. Louck and W. H. Shaffer, J. Mol. Spectry. **4**, 285 (1960); J. D. Louck, *ibid.* 298, 334 (1960). These authors have also used an algebraic technique similar to that of Sec. III to study the harmonic oscillator. However, their operators are unrelated to those discussed in this paper.
[6] S. Pasternack and R. M. Sternheimer, J. Math. Phys. **3**, 1280 (1962).
[7] W. Miller, *Lie Theory and Special Functions* (Academic, New York, 1968), has discussed an algebra closely related to the one discussed in this paper.
[8] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vols. I and II.
[9] A. O. Barut and C. Fronsdal, Proc. Roy. Soc. **A287**, 532 (1965).
[10] V. Bargmann, Ann. Math. **48**, 568 (1947).
[11] W. J. Holman, III and L. C. Biedenharn, Jr., Ann. Phys. (N.Y.) **39**, 1 (1960); **47**, 205 (1968).
[12] K-H Wang, J. Math. Phys. **11**, 2077 (1970).
[13] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U.P., Princeton, N.J., 1957).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Sum Rule Functions*

MICHAEL BARNSLEY
Theoretical Chemistry Institute, University of Wisconsin, Madison, Wisconsin

(Received 18 August 1970)

In many cases the quantum mechanical sum rules $S(k)$, as k ranges over all real values such that $S(k)$ is convergent, form a particular example from a class of functions called sum rule functions. If a set of up to five values taken by a sum rule function is given, then, by the use of another class of functions called N-sum rule functions, it is shown how one can impose the very best possible bounds on the sum rule function for arbitrary k, on the basis of the given information. In particular, these results apply to quantum mechanical sum rules.

INTRODUCTION

Many physical properties of an atom or molecule can be expressed in terms of the summations, known as *sum rules*:

$$\tilde{S}^q(k) = \sum_j{}' \epsilon_{qj}^{-k} \tilde{v}_{qj},$$

where $\tilde{v}_{qj}$ is an oscillator strength corresponding to a transition from state q to state j, ϵ_{qj} being the corresponding energy shift. If all the $\tilde{v}_{qj}$ are positive for a particular q, for example $q = 0$, then such a family of summations [as k ranges over all real values such that $\tilde{S}(k)$ is convergent] is a particular example of a *sum rule function*. More generally, when the first N of the $\tilde{v}_{qj}$ for a fixed q are negative, as would occur, for example, in the case of dipole oscillator strengths with $q > 0$, then $\tilde{S}(k)$ takes the form of a sum rule function minus an *N-sum rule function*. Thus, there is an immediate interest in the *nature* of sum rule functions and of N-sum rule functions. This is discussed in this paper.

For atoms, the *dipole* oscillator strengths have been especially studied and it is usually possible to calculate or measure various of the corresponding $\tilde{S}(k)$'s directly.[1] For example, the Reiche–Thomas–Kuhn sum rule gives $\tilde{S}(0)$ = number of electrons of the atom (using *atomic* units). This motivates the current interest in the problem of bounding sum rules in terms of other sum rules and in the problem of interpolation between sum rules. Dalgarno and Kingston[2] have found that for the ground state ($q = 0$) the $\tilde{S}(k)$'s can be approximated by the expression

$$\tilde{S}^0(k) = n[\epsilon_{10} + a(2 \cdot 5 - k)^{-1} + b(2 \cdot 5 - k)^{-2}]^k,$$

providing the $\tilde{v}_{qj}$ are dipole oscillator strengths. Here n is the number of electrons in the atom or molecule, "1" is the first excited state with non-vanishing oscillator strength, and the constants a and b are adjusted to make this equation correct for two selected values of k (usually $k = -1$ and $k = -2$). However, it must be stressed that this is only an approximate expression, with no bounding properties. In this paper we show how the problem of sum rule interpolation can be approached in such a way that the *very best possible* bounds to all (quantum mechanical) sum rules (based on any given *set* of sum rules) are obtained. For a large class of given sets

ACKNOWLEDGMENTS

Much of this work was done during the summer of 1970 while I was visiting the Laboratoire Aime Cotton, Orsay. I would like to thank Prof. Jacquinot and Dr. S. Feneuille for their hospitality during this period.

* Supported in part by the U.S. Atomic Energy Commission.

[1] L. Armstrong, Jr., J. Phys. **31**, C4 (1970).
[2] M. Moshinsky, *Harmonic Oscillator in Modern Physics: From Atoms to Quark* (Gordon & Breach, New York, 1969).
[3] W. H. Shaffer, Rev. Mod. Phys. **16**, 245 (1944); S. Bell, J. Phys. **B3**, 735, 745 (1970).
[4] S. Goshen and H. L. Lipkin, Ann. Phys. (N.Y.) **6**, 301 (1959); H. Bacry and J. L. Richard, J. Math. Phys. **8**, 2231 (1967).
[5] J. D. Louck and W. H. Shaffer, J. Mol. Spectry. **4**, 285 (1960); J. D. Louck, *ibid.* 298, 334 (1960). These authors have also used an algebraic technique similar to that of Sec. III to study the harmonic oscillator. However, their operators are unrelated to those discussed in this paper.
[6] S. Pasternack and R. M. Sternheimer, J. Math. Phys. **3**, 1280 (1962).
[7] W. Miller, *Lie Theory and Special Functions* (Academic, New York, 1968), has discussed an algebra closely related to the one discussed in this paper.
[8] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vols. I and II.
[9] A. O. Barut and C. Fronsdal, Proc. Roy. Soc. **A287**, 532 (1965).
[10] V. Bargmann, Ann. Math. **48**, 568 (1947).
[11] W. J. Holman, III and L. C. Biedenharn, Jr., Ann. Phys. (N.Y.) **39**, 1 (1960); **47**, 205 (1968).
[12] K-H Wang, J. Math. Phys. **11**, 2077 (1970).
[13] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U.P., Princeton, N.J., 1957).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Sum Rule Functions*

MICHAEL BARNSLEY
Theoretical Chemistry Institute, University of Wisconsin, Madison, Wisconsin

(Received 18 August 1970)

In many cases the quantum mechanical sum rules $S(k)$, as k ranges over all real values such that $S(k)$ is convergent, form a particular example from a class of functions called sum rule functions. If a set of up to five values taken by a sum rule function is given, then, by the use of another class of functions called N-sum rule functions, it is shown how one can impose the very best possible bounds on the sum rule function for arbitrary k, on the basis of the given information. In particular, these results apply to quantum mechanical sum rules.

INTRODUCTION

Many physical properties of an atom or molecule can be expressed in terms of the summations, known as *sum rules*:

$$\tilde{S}^q(k) = \sum_j{}' \epsilon_{qj}^{-k} \tilde{v}_{qj},$$

where $\tilde{v}_{qj}$ is an oscillator strength corresponding to a transition from state q to state j, ϵ_{qj} being the corresponding energy shift. If all the $\tilde{v}_{qj}$ are positive for a particular q, for example $q = 0$, then such a family of summations [as k ranges over all real values such that $\tilde{S}(k)$ is convergent] is a particular example of a *sum rule function*. More generally, when the first N of the $\tilde{v}_{qj}$ for a fixed q are negative, as would occur, for example, in the case of dipole oscillator strengths with $q > 0$, then $\tilde{S}(k)$ takes the form of a sum rule function minus an *N-sum rule function*. Thus, there is an immediate interest in the *nature* of sum rule functions and of N-sum rule functions. This is discussed in this paper.

For atoms, the *dipole* oscillator strengths have been especially studied and it is usually possible to calculate or measure various of the corresponding $\tilde{S}(k)$'s directly.[1] For example, the Reiche–Thomas–Kuhn sum rule gives $\tilde{S}(0)$ = number of electrons of the atom (using *atomic* units). This motivates the current interest in the problem of bounding sum rules in terms of other sum rules and in the problem of interpolation between sum rules. Dalgarno and Kingston[2] have found that for the ground state ($q = 0$) the $\tilde{S}(k)$'s can be approximated by the expression

$$\tilde{S}^0(k) = n[\epsilon_{10} + a(2 \cdot 5 - k)^{-1} + b(2 \cdot 5 - k)^{-2}]^k,$$

providing the $\tilde{v}_{qj}$ are dipole oscillator strengths. Here n is the number of electrons in the atom or molecule, "1" is the first excited state with non-vanishing oscillator strength, and the constants a and b are adjusted to make this equation correct for two selected values of k (usually $k = -1$ and $k = -2$). However, it must be stressed that this is only an approximate expression, with no bounding properties. In this paper we show how the problem of sum rule interpolation can be approached in such a way that the *very best possible* bounds to all (quantum mechanical) sum rules (based on any given *set* of sum rules) are obtained. For a large class of given sets

of sum rules, we present an explicit construction for the appropriate interpolation functions.

The rigorous upper and lower bounds to sum rules that we can obtain may often be remarkably close; they also have applications. For example, Pack[3] has given a simple formula for an *upper bound* to the van der Waals force constant in the interaction between any two atoms in terms of the noninteger dipole sum rules $\tilde{S}(-1.5)$. Without concern for units, his formula is $\mathrm{Cab} \leq \frac{3}{4}\tilde{S}a(-1.5)\tilde{S}b(-1.5)$. Using rigorous upper bounds to the $\tilde{S}(-1.5)$'s, we can obtain simple rigorous upper bounds to Cab. This formula was also derived independently by Kramer.[4] Again, the Hylleraas variational principle has been applied by Davison[5] to yield lower bounds to the van der Waals force constants in terms of sum rules. Use of alternative trial functions to the ones used by Davison yield lower bounds in terms of various noninteger sum rules, and bounds on these yield bounds on the constants. Similar remarks apply to a variational principle given by Epstein.[6] Finally, Barnsley has obtained excellent simple approximations to the van der Waals force constants by using the interpolation functions directly. It is among the purposes of this paper to establish an initial reference to the theory of sum rule functions on which results concerning the above-mentioned applications may be based.

More generally, sum rule functions arise whenever a series of Stieltjes[7] occurs, and hence the applications of their theory must be numerous. More precisely, if

$$F(z) = \int_0^\infty (1 + uz)^{-1}\, d\phi(u) = \sum_{n=0}^{\infty} (-z)^n \int_0^\infty u^n\, d\phi(u)$$

is a series of Stieltjes, then the function

$$S(\beta) = \int_0^\infty u^\beta\, d\phi(u)$$

is a sum rule function.

This paper emphasizes quantum mechanical sum rule functions and frequent reference is made to a more detailed monograph[8] in which more general proofs are given, including a description of a method for constructing the interpolation functions corresponding to arbitrary interpolation points, and in which a fuller discussion of the nature of sum rule functions is presented.

I. UNRESTRAINED N-SUM RULE FUNCTIONS, AND N-SUM RULE FUNCTIONS

Definition 1: If a function of a real variable β can be written in the form $\mathcal{S}_N(\beta) = \sum_{n=1}^{N} V_n E_n^{-\beta}$, where $V_n \neq 0$, V_n is real, and $0 < E_1 < E_2 < \cdots < E_N$, then $\mathcal{S}_N(\beta)$ is an *unrestrained N-sum rule function.*

Definition 2: If a function of a real variable β can be written in the form $S_N(\beta) = \sum_{n=1}^{N} V_n E_n^{-\beta}$, where $V_n > 0$, V_n is real and $0 < E_1 < E_2 < \cdots < E_N$, then $S_N(\beta)$ is an *N-sum rule function.* Unless otherwise stated, $\mathcal{S}_N(\beta)$ will denote an *unrestrained N-sum rule function*; $S_N(\beta)$ will denote an *N-sum rule function.*

Theorem 1: $\mathcal{S}_N(\beta)$ has at most $(N-1)$ zeros, the zeros at ∞ not being counted.

Proof: We prove this by induction. True for the case $N = 1$; suppose true $N = 1, 2, \cdots, K$. Consider the zeros of $\mathcal{S}_{K+1}(\beta)$, where

$$\mathcal{S}_{K+1}(\beta) = \sum_{n=1}^{K+1} V_n E_n^{-\beta} = V_1 E_1^{-\beta}\left[\sum_{n=2}^{K+1} \frac{V_n V_1^{-1}}{(E_n E_1^{-1})^\beta} + 1\right],$$

which has the form $\mathcal{S}_{K+1}(\beta) = V_1 E_1^{-\beta}[\mathcal{S}_K(\beta) + 1]$. Since $V_1 E_1^{-\beta} \neq 0$, it follows that $\mathcal{S}_{K+1}(\beta)$ has as many zeros as $f(\beta)$, where $f(\beta) = (\mathcal{S}_K(\beta) + 1)$. Now notice $\partial \mathcal{S}_K(\beta)/\partial\beta$ is an unrestrained M-sum rule function with $M \leq K$. Hence $\partial f(\beta)/\partial\beta$ has at most $(K-1)$ zeros, by the inductive hypothesis. Hence $f(\beta)$ has at most K zeros. Hence $\mathcal{S}_{K+1}(\beta)$ has at most K zeros. This completes the induction.

Theorem 2: Any $\mathcal{S}_N(\beta)$ is uniquely defined by the values $\mathcal{S}_N(\beta_i)$, $i = 1, \cdots, 2N$, where $-\infty < \beta_1 < \cdots < \beta_{2N} < \infty$.

Proof: Suppose $\tilde{\mathcal{S}}_N(\beta)$ is an unrestrained N-sum rule function which agrees with $\mathcal{S}_N(\beta)$ for $\beta = \beta_i$, $i = 1, \cdots, 2N$. Then $\mathcal{S}_M(\beta) = \mathcal{S}_N(\beta) - \tilde{\mathcal{S}}_N(\beta)$ is either an unrestrained M-sum rule function with $1 \leq M \leq 2N$ having $2N$ zeros or else $\mathcal{S}_M(\beta) \equiv 0$. The first alternative is not possible by Theorem 1. Hence $\mathcal{S}_N(\beta)$ is unique.

Theorem 3: If $S_{N+1}(\beta)$ is an arbitrary $(N+1)$-sum rule function and $S_N(\beta)$ is an arbitrary N-sum rule function, then $[S_{N+1}(\beta) - S_N(\beta)]$ has at most $2N$ zeros.

Proof:

$$[S_{N+1}(\beta) - S_N(\beta)] = \mathcal{S}_M(\beta),$$

with $M \leq 2N + 1$. Hence by Theorem 1,

$$[S_{N+1}(\beta) - S_N(\beta)]$$

has at most $2N$ zeros.

II. SUM RULE FUNCTIONS

Definition 3: $S(\beta)$ is a *sum rule function* on the closed interval $I = [a, b]$, where $-\infty < a < b < \infty$, if there exists a sequence of N-sum rule functions $\{S_N(\beta)\}_{N=1}^{\infty}$ which is *uniformly convergent* to $S(\beta)$ for $\beta \in [a, b]$.

In particular, if $\{S_N(\beta)\}_{N=1}^{\infty}$ is uniformly convergent for $\beta \in [a, b]$ for any b such that $a < b < \infty$, then we say that $S(\beta)$ is a sum rule function on $[a, \infty)$. If $S(\beta)$ is a sum rule function on $[a, \infty)$ for all a such that $-\infty < a' < a < \infty$, then we say that $S(\beta)$ is a sum rule function on (a', ∞).

If $\alpha > 0$ and γ are real numbers, then the *positive linear transformation* L operates on the real variable β according to $L\beta = \alpha\beta + \gamma$ and operates on $S(\beta)$ according to $LS(\beta) = S(L\beta)$. If $S(\beta)$ is a sum rule function on $I = [a, b]$, then $LS(\beta)$ is a sum rule function on $[L^{-1}a, L^{-1}b]$. If $[a, b]$ and $[c, d]$ are arbitrary given real intervals, then by defining an L by $La = c$, and $Lb = d$, we see that this L sets up a one-to-one correspondence between all sum rule functions on $[a, b]$ and all sum rule functions on $[c, d]$ because L has a unique inverse. We can also apply L to any N-sum rule function. This leads us to the following important observation.

Theorem 4: Suppose we are given any sum rule function $S(\beta)$ on, say, $[a, b]$ and any positive linear transformation L. Suppose further that we establish some bounding relationship between $S(\beta)$ and an $S_N(\beta)$ at some point $c \in [a, b]$, for example, $S_N(c) > S(c)$. Then this same bounding relationship holds between the sum rule function $\tilde{S}(\beta) = LS(\beta)$ and the N-sum rule function $\tilde{S}_N(\beta) = LS_N(\beta)$ at the point $L^{-1}c$. Similar remarks apply to bounding relationships between M- and N-sum rule functions.

III. INTERPOLATION OF A 2-SUM RULE FUNCTION USING A 1-SUM RULE FUNCTION AND OF A 3-SUM RULE FUNCTION USING A 2-SUM RULE FUNCTION

Theorem 5: If $S_2(\beta)$ is an arbitrary 2-sum rule function and $\beta_0 < \beta_1$ are finite real numbers, then there exists an unique 1-sum rule function $S_1(\beta)$ such that $S_1(\beta_i) = S_2(\beta_i)$ for $i = 0, 1$. Moreover, $S_1(\beta) < S_2(\beta)$ if $\beta \in (-\infty, \beta_0) \cup (\beta_1, \infty)$, and $S_1(\beta) > S_2(\beta)$ if $\beta \in (\beta_0, \beta_1)$.

Proof: We need only prove this for $\beta_0 = 0$, $\beta_1 = 1$, for then by use of a suitable positive linear transformation L (Theorem 4) the proof can at once be generalized to give the theorem.

Take $S_1(\beta) = S_2(0)[S_2(0)/S_2(1)]^{-\beta}$. Then $S_1(\beta)$ *is* a 1-sum rule function such that $S_1(\beta_i) = S_2(\beta_i)$ for $i = 0, 1$. It is unique by Theorem 2.

In fact, the general 1-sum rule function interpolation is

$$S_1(\beta) = S_2(\beta_0)\left[\frac{S_2(\beta_0)}{S_2(\beta_1)}\right]^{(\beta_0-\beta)/(\beta_1-\beta_0)}$$

By Theorem 3, $G(\beta) = [S_2(\beta) - S_1(\beta)]$ has at most two zeros. Hence, in this case, it has exactly two zeros; therefore the bounding relationships given in the theorem must be either true as they stand or else true with the inequality signs reversed. They are true as they stand because, taking $\beta = 2$ in the $\beta_0 = 0$, $\beta_1 = 1$ case,we have

$$S_2(2) - S_1(2) = V_1V_2[1/E_1 - 1/E_2]^2 > 0$$

(see Definition 2). This completes the proof.

Theorem 6: If $S_3(\beta)$ is an arbitrary 3-sum rule function and $\beta_0 < \beta_1 < \beta_2 < \beta_3$ are finite real numbers, then there exists an unique 2-sum rule function $S_2(\beta)$ such that $S_2(\beta_i) = S_3(\beta_i)$ for $i = 0, 1, 2, 3$. Moreover, $S_2(\beta) < S_3(\beta)$ for $\beta \in (-\infty, \beta_0) \cup (\beta_1, \beta_2) \cup (\beta_3, \infty)$ and $S_2(\beta) > S_3(\beta)$ for $\beta \in (\beta_0, \beta_1) \cup (\beta_2, \beta_3)$.

Proof: We will not prove this theorem for general interpolation points $\beta_0 < \beta_1 < \beta_2 < \beta_3$. For this we refer to Ref. 8. Here we will consider only those sets of interpolation points such that $\beta_i = L(i)$ for $i = 0, 1, 2, 3$. Hence (Theorem 4), we need only prove the theorem for $\beta_i = i$, where $i = 0, 1, 2, 3$.

Denote $S_3(i) = S_i$ for $i = 0, 1, 2, 3$. Let ϵ_1, ϵ_2 be the roots of the quadratic equation

$$(S_1S_3 - S_2^2)\epsilon^2 - (S_0S_3 - S_1S_2)\epsilon + (S_0S_2 - S_1^2) = 0,$$

and let

$$x = (S_0S_2 - S_1^2)\epsilon_1^2/(S_0 - 2S_1\epsilon_1 + S_2\epsilon_1^2),$$

$y = (S_0 - x)$. Then, after some algebra,[8] we can verify that $x > 0$, $y > 0$, $\epsilon_1 > 0$, $\epsilon_2 > 0$, and $\epsilon_1 \neq \epsilon_2$. Hence, $S_2(\beta) = x\epsilon_1^{-\beta} + y\epsilon_2^{-\beta}$ defines a 2-sum rule function. Algebraically we can show that $S_2(i) = S_3(i) = S_i$ for $i = 0, 1, 2, 3$. By Theorem 2, $S_2(\beta)$ is unique.

It remains to establish that $S_2(\beta)$ has the stated bounding properties.

Write a general $S_3(\beta)$ in the form

$$S_3(\beta) = S_3(\beta)(V_1, V_2, V_3, \epsilon_1, \epsilon_2, \epsilon_3) = S_3(\beta)(\mathbf{v}, \boldsymbol{\epsilon}) = \sum_{i=1}^{3} V_i\epsilon_i^{-\beta}$$

by Definition 2. Let D denote the connected[9] domain

$$D = \{(\mathbf{v}, \boldsymbol{\epsilon}) = (V_1, V_2, V_3, \epsilon_1, \epsilon_2, \epsilon_3) \mid V_i > 0, \epsilon_i > 0; i = 1, 2, 3; \epsilon_1 < \epsilon_2 < \epsilon_3\}.$$

Then $S_3(\beta)(\mathbf{v}, \boldsymbol{\epsilon})$ is a continuous function for $\beta \in \mathbb{R}$, $(\mathbf{v}, \boldsymbol{\epsilon}) \in D$. Construct $S_2(\beta)(\mathbf{v}, \boldsymbol{\epsilon})$ from $S_3(\beta)(\mathbf{v}, \boldsymbol{\epsilon})$ so that $S_2(i)(\mathbf{v}, \boldsymbol{\epsilon}) = S_3(i)(\mathbf{v}, \boldsymbol{\epsilon})$ for $i = 0, 1, 2, 3$, as we

did above. Then $S_2(\beta)(\mathbf{v}, \boldsymbol{\epsilon})$ is also a continuous function for $\beta \in \mathbb{R}$, $(\mathbf{v}, \boldsymbol{\epsilon}) \in D$, although its functional dependence on $(\mathbf{v}, \boldsymbol{\epsilon})$ is very different from that of $S_3(\beta)(\mathbf{v}, \boldsymbol{\epsilon})$ given above.

Define $G(\beta)(\mathbf{v}, \boldsymbol{\epsilon}) = (S_3(\beta)(\mathbf{v}, \boldsymbol{\epsilon}) - S_2(\beta)(\mathbf{v}, \boldsymbol{\epsilon}))$ on D. By Theorem 3, for any fixed choice of $(\mathbf{v}, \boldsymbol{\epsilon})$, $G(\beta)(\mathbf{v}, \boldsymbol{\epsilon})$ has at most four zeros, and hence it has exactly four.

So, if we can show that $G(-1)(\mathbf{v}, \boldsymbol{\epsilon}) > 0$ for all $(\mathbf{v}, \boldsymbol{\epsilon}) \in D$, where we have arbitrarily taken $\beta = -1$ to establish the point, then we will have proved that the bounding relations given in the theorem are true as they stand. It is easily verified that if $(\mathbf{v}, \boldsymbol{\epsilon}) = (1, 2, 3, 1, 2, 3)$, then $G(-1)(\mathbf{v}, \boldsymbol{\epsilon}) > 0$. Hence, for the corresponding 3-sum rule function the theorem *is* true. Suppose that for some choice of $(\mathbf{v}, \boldsymbol{\epsilon})$, say $(\mathbf{v}, \boldsymbol{\epsilon}) = (\mathbf{v}'', \boldsymbol{\epsilon}'')$, we have $G(-1)(\mathbf{v}'', \boldsymbol{\epsilon}'') \leq 0$. Then, because D is connected, there must exist a $(\hat{\mathbf{v}}, \hat{\boldsymbol{\epsilon}})$ such that $G(-1)(\hat{\mathbf{v}}, \hat{\boldsymbol{\epsilon}}) = 0$. But then $G(\beta)(\hat{\mathbf{v}}, \hat{\boldsymbol{\epsilon}})$ has zeros at $\beta = -1, 0, 1, 2, 3$, in contradiction to Theorem 3. This completes the proof.

IV. INTERPOLATION OF AN N-SUM RULE FUNCTION USING A 1-SUM RULE FUNCTION AND OF AN M-SUM RULE FUNCTION USING A 2-SUM RULE FUNCTION ($N > 1$, $M > 2$)

Theorem 7: If $S_N(\beta)$ is an arbitrary N-sum rule function with $N > 1$ and $\beta_0 < \beta_1$ are finite real numbers, then there exists an unique 1-sum rule function such that $S_1(\beta_i) = S_N(\beta_i)$ for $i = 0, 1$. Moreover, $S_1(\beta) < S_N(\beta)$ if $\beta \in (-\infty, \beta_0) \cup (\beta_1, \infty)$ and $S_1(\beta) > S_N(\beta)$ if $\beta \in (\beta_0, \beta_1)$.

Proof: Take

$$S_1(\beta) = S_N(\beta_0)\left[\frac{S_N(\beta_0)}{S_N(\beta_1)}\right]^{(\beta_0-\beta)/(\beta_1-\beta_0)} \tag{1}$$

Then $S_1(\beta)$ is a 1-sum rule function such that $S_1(\beta_i) = S_N(\beta_i)$ for $i = 0, 1$. It is unique by Theorem 2.

Now notice that if we knew $V_1, \cdots, V_N, E_1, \cdots, E_N$ for $S_N(\beta)$, then we could have constructed $S_1(\beta)$ as follows: Construct an $S_{N-1}(\beta)$ from $S_N(\beta)$ by using the 1-sum rule function $S_1^{(N-1)}(\beta)$ defined by

$$S_1^{(N-1)}(\beta_i) = V_{N-1}E_{N-1}^{-\beta_i} + V_N E_N^{-\beta_i} \quad \text{for} \quad i = 0, 1$$

as in Theorem 5, and then take

$$S_{N-1}(\beta) = \sum_{n=1}^{N-2} V_n E_n^{-\beta} + S_1^{(N-1)}(\beta).$$

Then $S_{N-1}(\beta_i) = S_N(\beta_i)$ for $i = 0, 1$, and, by Theorem 5, $S_{N-1}(\beta) < S_N(\beta)$ if $\beta \in (-\infty, \beta_0) \cup (\beta_1, \infty)$, $S_{N-1}(\beta) > S_N(\beta)$ if $\beta \in (\beta_0, \beta_1)$. Proceed in this manner until a 1-sum rule function is obtained which agrees with $S_N(\beta)$ at β_0 and β_1. This must be the same as $S_1(\beta)$ defined in (1), by Theorem 2. Hence $S_1(\beta)$ must have the bounding properties attributed to it in the theorem. This completes the proof.

Theorem 8: If $S_M(\beta)$ is an arbitrary M-sum rule function with $M > 2$ and $\beta_0 < \beta_1 < \beta_2 < \beta_3$ are finite real numbers, then there exists an unique 2-sum rule function $S_2(\beta)$ such that $S_2(\beta_i) = S_M(\beta_i)$ for $i = 0, 1, 2, 3$. Moreover, $S_2(\beta) < S_M(\beta)$ for $\beta \in (-\infty, \beta_0) \cup (\beta_1, \beta_2) \cup (\beta_3, \infty)$ and $S_2(\beta) > S_M(\beta)$ for $\beta \in (\beta_0, \beta_1) \cup (\beta_2, \beta_3)$.

Proof: As in Theorem 6, we will prove the theorem only for sets of interpolation points such that $\beta_i = L(i)$, where $i = 0, 1, 2, 3$ and L is a positive linear transformation. Hence (Theorem 4) we need only establish the theorem for interpolation points $\beta_i = 0, 1, 2, 3$.

Denote $S_M(i) = S_i$, for $i = 0, 1, 2, 3$. Let ϵ_1, ϵ_2 be the roots of the quadratic

$$(S_1S_3 - S_2^2)\epsilon^2 - (S_0S_3 - S_1S_2)\epsilon + (S_0S_2 - S_1^2) = 0.$$

Let

$$x = (S_0S_2 - S_1^2)\epsilon_1^2/(S_0 - 2S_1\epsilon_1 + S_2\epsilon_1^2),$$

any $y = (S_0 - X)$. Take

$$S_2(\beta) = x/\epsilon_1^\beta + y/\epsilon_2^\beta. \tag{2}$$

If we had known $V_1, \cdots, V_M, E_1, \cdots, E_M$ corresponding to $S_M(\beta)$, then we could have obtained a 2-sum rule function with the bounding properties given in the theorem, in a manner similar to that used in the proof of Theorem 7. This ensures that *an* $S_2(\beta)$ does indeed exist with the cited properties. Theorem 2 tells us that it is unique. Since Eq. (2) defines a function which formally fulfills the requirements of this 2-sum rule function, it must actually define it. This completes the proof in the case of $\beta_i = L(i)$, $i = 0, 1, 2, 3$. For general interpolation points we refer to Ref. 8 .

It should be noticed that Theorem 7 constructs the requisite 1-sum rule function from the values $S_N(\beta_i)$, $i = 0, 1$, and utilizes no further information. A similar remark applies to Theorem 8.

V. INTERPOLATION OF ANY SUM RULE FUNCTION USING A 1-SUM RULE FUNCTION AND USING A 2-SUM RULE FUNCTION

Let $S(\beta)$ be any sum rule function on some closed interval $I = [a, b]$. Then there exists a sequence of N-sum rule functions $\{S_N(\beta)\}_{N=1}^{\infty}$ which is uniformly convergent to $S(\beta)$ on I (Definition 3).

Theorem 9: If $S(\beta)$ is an arbitrary sum rule function on an interval $I = [a, b]$ and $\beta_0 < \beta_1$ are finite real numbers belonging to I, then there exists an unique 1-sum rule function such that $S_1(\beta_i) = S(\beta_i)$ for $i = 0, 1$. Moreover, $S_1(\beta) \leq S(\beta)$ if $\beta \in (-\infty, \beta_0) \cup (\beta_1, \infty)$ and $S_1(\beta) > S(\beta)$, if $\beta \in (\beta_0, \beta_1)$.

Proof: Let $\{S_N(\beta)\}_{N=1}^{\infty}$ be a sequence of N-sum rule functions uniformly convergent to $S(\beta)$ on I.

Let $S_1^{(N)}(\beta)$ be the 1-sum rule interpolation to $S_N(\beta)$ which is exact at β_0, β_1. Then (Theorem 7)

$$\begin{aligned} &S_1^{(N)}(\beta) < S_N(\beta) \quad \text{if} \quad \beta \in (-\infty, \beta_0) \cup (\beta_1, \infty), \\ &S_1^{(N)}(\beta) > S_N(\beta) \quad \text{if} \quad \beta \in (\beta_0, \beta_1). \end{aligned} \tag{3}$$

Let

$$S_1(\beta) = \lim_{N \to \infty} S_1^{(N)}(\beta) = S(\beta_0)[S(\beta_0)/S(\beta_1)]^{(\beta_0 - \beta)/(\beta_1 - \beta_0)}.$$

Then $S_1(\beta)$ *is* a 1-sum rule function which agrees with $S(\beta)$ at β_0, β_1; taking limits as $N \to \infty$ in (3), we see moreover that

$$\begin{aligned} &S_1(\beta) \leq S(\beta) \quad \text{if} \quad \beta \in (-\infty, \beta_0) \cup (\beta_1, \infty), \\ &S_1(\beta) \geq S(\beta) \quad \text{if} \quad \beta \in (\beta_0, \beta_1). \end{aligned}$$

$S_1(\beta)$ is unique by Theorem 2. This completes the proof.

Theorem 10: If $S(\beta)$ is an arbitrary sum rule function on an interval $I = [a, b]$ and $\beta_0 < \beta_1 < \beta_2 < \beta_3$ are finite real numbers belonging to I, then there exists a unique 2-sum rule function $S_2(\beta)$ such that $S_2(\beta_i) = S(\beta_i)$ for $i = 0, 1, 2, 3$. Moreover, $S_2(\beta) \leq S(\beta)$ if $\beta \in (-\infty, \beta_0) \cup (\beta_1, \beta_2) \cup (\beta_3, \infty)$ and $S_2(\beta) \geq S(\beta)$ if $\beta \in (\beta_0, \beta_1) \cup (\beta_2, \beta_3)$.

Proof: The proof is similar to the proof of Theorem 9. We will only give the construction of $S_2(\beta)$. Again we will concern ourselves with sets of interpolation points such that $\beta_i = L(i)$, where $i = 0, 1, 2, 3$, and L is a positive linear transformation, so we will give the construction only for $\beta_i = i$ where $i = 0, 1, 2, 3$.

Let $S(i) \equiv S_i$. Use the construction (2) given in Theorem 8.

VI. QUANTUM MECHANICAL SUM RULE FUNCTIONS

Theorem 11: If a function $S(\beta)$ can be written in the form

$$S(\beta) = \int_0^{\infty} u^{-\beta}\, d\phi(u),$$

the integral being convergent in the Riemann–Stieltjes sense for $\beta = \beta_0$ and $\beta = \beta_1$, where $\beta_0 < \beta_1$ and where $\phi(u)$ is (i) monotone nondecreasing and (ii) taking infinitely many values on $(0, \infty)$, then the integral exists for $\beta \in [\beta_0, \beta_1]$ and $S(\beta)$ is a sum rule function on this interval.

Proof:

$$\begin{aligned} S(\beta) &= \int_1^{\infty} u^{\beta} d[-\phi(1/u)] + \int_{1+}^{\infty} u^{-\beta}\, d\phi(u) \\ &= \int_1^{\infty} u^{\beta}\, d\xi(u) + \int_{1+}^{\infty} u^{-\beta}\, d\phi(u), \end{aligned}$$

where $\xi(u) = -\phi(1/u)$. Both $\phi(u)$ and $\xi(u)$ are monotone nondecreasing in $[1, \infty)$ and at least one of $\phi(u)$ and $\xi(u)$ takes infinitely many values in $[1, \infty)$. For simplicity we will assume that both $\phi(u)$ and $\xi(u)$ take infinitely many values in $[1, \infty)$; then $S(\beta)$ has the form

$$S(\beta) = T'(-\beta) + T''(\beta),$$

where $T'(\beta)$, $T''(\beta)$ are functions of the special form $T(\beta) = \int_1^{\infty} u^{-\beta}\, d\theta(u)$, where $\theta(u)$ is monotone nondecreasing and takes infinitely many values in $[1, \infty)$.

It is easily seen that the convergence of $S(\beta)$ at β_0 and β_1 ensures the convergence of $T'(-\beta)$ for $\beta \in (-\infty, \beta_1]$, and ensures the convergence of $T''(\beta)$ for $\beta \in [\beta_0, \infty)$. Hence $S(\beta)$ certainly converges for $\beta \in [\beta_0, \beta_1]$.

We must show that $T(\beta)$ is a sum rule function on $[\beta', \beta'']$ providing it is convergent at β', where $-\infty < \beta' < \beta'' < \infty$ and β'' is arbitrary.

We will assume that $d\theta(u) = \psi(u)\, du$, where $\psi(u)$ is continuous for $u \in [1, \infty)$. The argument easily extends to treat the most general case.

Let $h > 0$ be prescribed arbitrarily small. Suppose, without loss of generality, that $\beta' > 0$. Since $T(\beta)$ is a monotone decreasing in β and since the integral $T(\beta')$ exists, it follows that there exists a finite real number k such that

$$0 \leq \int_k^{\infty} u^{-\beta} \psi(u)\, du < \tfrac{1}{2}h \quad \text{for} \quad \beta \in [\beta', \beta'']. \tag{4}$$

Define a sequence of equipartitions $\{\Delta_N\}_{N=1}^{\infty}$ on the interval $[1, k]$, Δ_M being the set of $(M + 1)$ equally spaced points

$$1 = a_0^M < a_1^M < \cdots < a_M^M = k.$$

Define two M-sum rule functions on Δ_M:

$$\text{(i)} \quad S_{\Delta M}^U(\beta) = \sum_{n=1}^{M} \Psi_n^M (a_n^M - a_{n-1}^M)/(a_{n-1}^M)^{\beta},$$

$$\text{(ii)} \quad S_{\Delta M}^L(\beta) = \sum_{n=1}^{M} \psi_n^M (a_n^M - a_{n-1}^M)/(a_n^M)^{\beta},$$

where

$$\Psi_n^M = \max \{\psi(u) \mid u \in [a_{n-1}^M, a_n^M]\},$$

and

$$\psi_n^M = \min\{\psi(u) \mid u \in [a_{n-1}^M, a_n^M]\}.$$

Then

$$S_{\Delta N}^U(\beta) \geq \int_1^k u^{-\beta}\psi(u)\,du \geq S_{\Delta N}^L(\beta) \quad \text{for all } N \text{ and for all } \beta \in [\beta', \beta''], \quad (5)$$

and we have

$$\begin{aligned} 0 &\leq S_{\Delta N}^U(\beta) - S_{\Delta N}^L(\beta) \\ &= \sum_{n=1}^{N}\left[\frac{(\Psi_n^N - \psi_n^N)}{(a_{n-1}^N)^\beta} + \psi_n^N\left(\frac{1}{(a_{n-1}^N)^\beta} - \frac{1}{(a_n^N)^\beta}\right)\right](a_n^N - a_{n-1}^N) \\ &\leq \sum_{n=1}^{N}\left[(\Psi_n^N - \psi_n^N) + \psi_n^N\left(\frac{1}{(a_{n-1}^N)^\beta} - \frac{1}{(a_n^N)^\beta}\right)\right](a_n^N - a_{n-1}^N) \\ &\qquad \text{for all } \beta \in [\beta', \beta'']. \end{aligned}$$

Since $\psi(u)$ is continuous on the compact set $u \in [1, k]$, there exists an integer N_1 such that $N > N_1$ implies

$$0 \leq (\Psi_n^N - \psi_n^N) < \tfrac{1}{4}(k-1)^{-1}h.$$

Since ψ_n^N is bounded above for all n and N and since the function $f(a, \beta) = 1/a^\beta$ is continuous on the compact set $a \in [1, k]$, $\beta \in [\beta', \beta'']$, there also exists an integer N_2 such that $N > N_2$ implies

$$0 \leq \psi_n^N\{(a_{n-1}^N)^{-\beta} - (a_n^N)^{-\beta}\} < \tfrac{1}{4}(k-1)^{-1}h \quad \text{for all } \beta \in [\beta', \beta''].$$

Hence, $N > \max\{N_1, N_2\}$ implies

$$0 \leq S_{\Delta N}^U(\beta) - S_{\Delta N}^L(\beta) < \tfrac{1}{2}h \quad \text{for all } \beta \in [\beta', \beta''],$$

whence, from (5), for $N > \max\{N_1, N_2\}$,

$$0 \leq \int_1^k u^{-\beta}\psi(u)\,du - S_{\Delta N}^L(\beta) < \tfrac{1}{2}h \quad \text{for all } \beta \in [\beta', \beta''],$$

and finally, from (4), for $N > \max\{N_1, N_2\}$,

$$0 \leq T(\beta) - S_{\Delta N}^L(\beta) < h \quad \text{for all } \beta \in [\beta', \beta''].$$

We deduce that $T(\beta)$ is a sum rule function as required, and the theorem follows.

Definition 4: Any sum rule function which can be written in the form given in Theorem 11 will be called a *quantum mechanical sum rule function.*

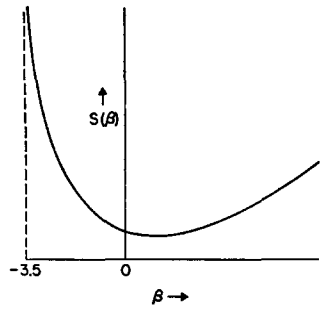

FIG. 1. An example of a quantum mechanical sum rule function. Here $S(\beta)$ corresponds to $q = 0$ and dipole oscillator strengths. $S(\beta)$ diverges to plus infinity at -3.5, is convergent everywhere to the right of this point, and takes its only minimum somewhere in $(-3.5, \infty]$. This figure is only a schematic representation.

In quantum mechanics we are interested in the families of sum rules $\tilde{S}^q(k) = \sum_j' \epsilon_{qj}^{-k}\tilde{V}_{qj}$, where $\tilde{V}_{qj}$ is an oscillator strength corresponding to a transition from state q to state j, ϵ_{qj} being the corresponding transition energy, and where the summation is understood to include an integration over the continuum states of the atom or molecule under consideration. We are concerned only with those cases where all the $\tilde{V}_{qj}$ are positive, for then, by the above theorem $\tilde{S}^q(k)$ is a quantum mechanical sum rule function. (According to our definition, the model harmonic oscillator problem does not yield a quantum mechanical sum rule function.) This occurs, for example, when $q = 0$ and $\tilde{V}_{0j} = f_{0j}$, the f_{qj} being *dipole* oscillator strengths. In such cases we prefer to write

$$\tilde{S}^q(k+1) = S^q(k) = \sum_j' \epsilon_{qj}^{-k-1}\tilde{V}_{qj} = \sum_j' \epsilon_{qj}^{-k}V_{qj},$$

where $\tilde{V}_{qj}\epsilon_{qj}^{-1} = V_{qj}$, so that the value of $S^q(0)$ is independent of the choice of energy scale. Moreover, we will refer to $S^q(\beta)$ rather than $S^q(k)$ to emphasize the functional dependence on the *real variable* β, rather than the *integer variable* k.

What does a quantum mechanical sum rule function look like? Suppose we have an $S^q(\beta)$, as above, where the smallest energy involved, say ϵ_{10}, is positive. Then, if we take ϵ_{10} as the unit of energy, we see that $S^q(\beta)$ becomes a monotone decreasing function of β, and hence that it must be convergent at least on an interval of the form (a', ∞) (see Definition 3). More generally, $S^q(\beta)$ has only one turning point and this is a minimum (see Fig. 1). [Formally, the second derivative of an $S^q(\beta)$ with respect to β is positive whereever it is convergent.]

Given that we are dealing with a (quantum mechanical) sum rule function $S^{(q)}(\beta)$ defined on some interval I, say (a', ∞), and given further the values of $S^{(q)}(\beta_i)$ for various $\beta_i \in I$, where $i = 1, 2, \cdots, j$ and $j = 1, 2, 3, 4$, or 5,[10] then we can apply the results of Sec. V to impose upper and/or lower bounds on $S^{(q)}(\beta)$ for all other $\beta \in I$. In the next section, we show that these bounds are the very best possible that can be obtained on the basis of the given information.

VII. BEST POSSIBLE BOUNDS TO (QUANTUM MECHANICAL) SUM RULE FUNCTIONS ON THE BASIS OF THE GIVEN INFORMATION

Definition 5: If $S(\beta)$ is a sum rule function on $[a, b]$, then the value of $S(\beta')$ for some particular $\beta' \in [a, b]$ is called a *sum rule*.

We will abbreviate the statement "On the basis of the given information" to "OBGI." When we say that a set of bounds on a sum rule function are best possible OBGI we shall mean that they are

(i) *optimal*, i.e., that the bounds are the very closest that can be imposed OBGI,

(ii) *extensive*, i.e., that there is at least one bound corresponding to each value of β,

(iii) *exclusive*, i.e., that OBGI no other nontrivial bounds can be imposed on $S(\beta)$, where zero and plus infinity are considered to be the trivial bounds.

A. Best Possible Bounds on a (Quantum Mechanical) Sum Rule Function When One Sum Rule Is Known, OBGI

Suppose that we know $S(\beta)$ is a (quantum mechanical) sum rule function on some interval, say (a', ∞), and that we know $S(\beta')$ for some $\beta' \in (a', \infty)$; then *it is impossible* to impose nontrivial bounds on $S(\beta)$ for $\beta \neq \beta'$, OBGI.

Proof: If $S(\beta)$ is a (quantum mechanical) sum rule function, then so is $\tilde{S}(\beta) = VE^{-\beta}S(\beta)$ for any $V > 0$, $E > 0$. Let β'' be any point such that $\beta'' \neq \beta'$ and $\beta'' \in (a', \infty)$. Since $S(\beta'')$ must be positive and finite, we can choose V and E such that (i) $VE^{-\beta'} = 1$, (ii) $VE^{-\beta''}S(\beta'')$ is arbitrarily large or arbitrarily small. The existence of $\tilde{S}(\beta)$ thus defined completes the proof.

B. Best Possible Bounds on a (Quantum Mechanical) Sum Rule Function When Two Sum Rules Are Known, OBGI

Suppose that we know $S(\beta)$ is a (quantum mechanical) sum rule function on some interval, say (a', ∞), and that we know $S(\beta_0)$, $S(\beta_1)$ for $\beta_0 < \beta_1$, where $\beta_0, \beta_1 \in (a', \infty)$; then the best possible bounds OBGI are those given by the 1-sum function $S_1(\beta)$ constructed in Theorem 9.

Proof: The bounds given by $S_1(\beta)$ are certainly extensive. We must prove that they are optimal and exclusive.

Let $J \subset (a', \infty)$ be any given closed interval which contains β_0 and β_1. It is easily seen that we can construct a (quantum mechanical) sum rule function $\tilde{S}(\beta)$ which lies arbitrarily close to $S_1(\beta)$ for all $\beta \in J$ and which takes any arbitrarily large finite value at any prescribed point $\beta' \notin J$, $\beta' \in (a', \infty)$. Suppose $\tilde{S}(\beta_0) - S_1(\beta_0) = \varphi_0$ and $\tilde{S}(\beta_1) - S_1(\beta_1) = \varphi_1$. Choose $V > 0$ and $E > 0$ such that

$$VE^{-\beta_0}\tilde{S}(\beta_0) = \tilde{S}(\beta_0) - \varphi_0,$$
$$VE^{-\beta_1}\tilde{S}(\beta_1) = \tilde{S}(\beta_1) - \varphi_1.$$

Then $\tilde{\tilde{S}}(\beta) = VE^{-\beta}\tilde{S}(\beta)$ is a (quantum mechanical) sum rule function. It agrees with $S_1(\beta)$ at β_0 and β_1, lies arbitrarily close to it on J, and is arbitrarily large at β'.

Hence the bounds imposed by $S_1(\beta)$ as constructed in Theorem 9 must be optimal, OBGI.

Moreover, we cannot impose nontrivial upper bounds on $S(\beta)$ outside $[\beta_0, \beta_1]$ OBGI, nor, as can easily be shown by construction of a (quantum mechanical) sum rule function which agrees with $S_1(\beta)$ at β_0, β_1, taking a value arbitrarily close to zero at any prescribed point $\beta'' \in (\beta_0, \beta_1)$, can we impose nontrivial lower bounds on $S(\beta)$ for $\beta \in (\beta_0, \beta_1)$, OBGI. Hence the bounds are exclusive. This completes the proof.

C. Best Possible Bounds on a (Quantum Mechanical) Sum Rule Function When Three Sum Rules Are Known, OBGI

Suppose that we know $S(\beta)$ is a (quantum mechanical) sum rule function on some interval, say (a', ∞), and that we know $S(\beta_0)$, $S(\beta_1)$, $S(\beta_2)$ for $\beta_0 < \beta_1 < \beta_2$, where β_0, β_1, $\beta_2 \in (a', \infty)$, then the best possible bounds OBGI are those given by the two 1-sum rule functions $S_1^i(\beta)$, $i = 0, 1$, where $S_1^i(\beta)$ is the 1-sum rule function which agrees with $S(\beta)$ at β_i and β_{i+1}, as constructed in Theorem 9.

Proof: (See Fig. 2.) Without going into the details (which are given in Ref. 8), we merely point out that it is possible to construct a (quantum mechanical) sum rule function which agrees with $S(\beta)$ at the interpolation points and which would contradict any proposed

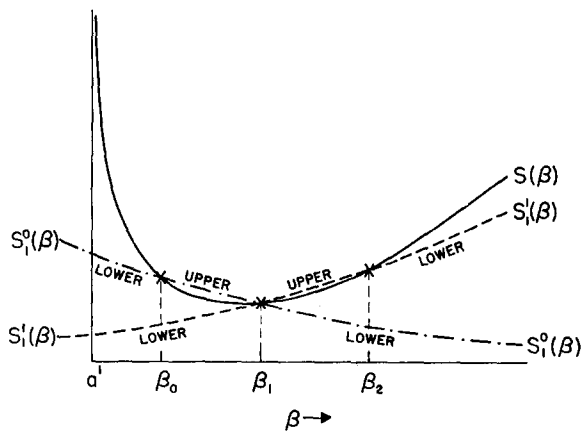

FIG. 2. This figure schematically designates the bounding relationships between the three functions. $S(\beta)$ is a hypothetical sum rule function (quantum mechanical) which we have supposed is divergent to plus infinity at a'.

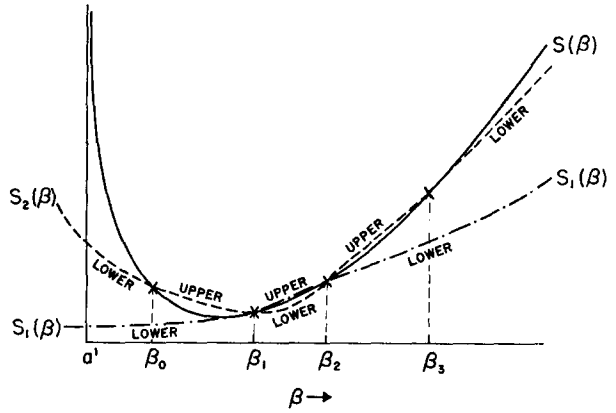

FIG. 3. This figure schematically designates the bounding relationships between the three functions.

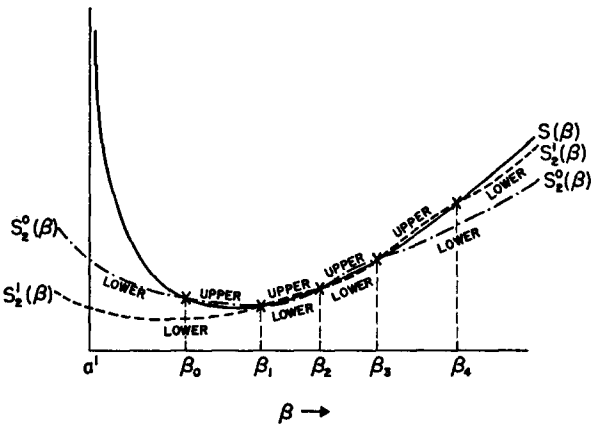

FIG. 4. $S_2^0(\beta)$ and $S_2^1(\beta)$ are the two 2-sum rule functions which agree with $S(\beta)$ at $\beta_0, \beta_1, \beta_2, \beta_3$ and at $\beta_1, \beta_2, \beta_3, \beta_4$, respectively. This figure schematically designates the bounding relationships between the three functions.

improvement (OBGI) on any of the bounds which are supplied by the $S_1^i(\beta)$ functions in the configurations shown in Fig. 2. Moreover, we can again show that it is impossible to prescribe a nontrivial upper bound (OBGI) on $S(\beta)$ outside $[\beta_0, \beta_2]$.

D. Best Possible Bounds on a Quantum Mechanical Sum Rule Function When Four Sum Rules Are Known

Suppose that we know $S(\beta)$ is a (quantum mechanical) sum rule function on some interval, say (a', ∞), and that we know $S(\beta_0)$, $S(\beta_1)$, $S(\beta_2)$, $S(\beta_3)$ for $\beta_0 < \beta_1 < \beta_2 < \beta_3, \beta_i \in (a', \infty)$ for $i = 0, 1, 2, 3$; then the best possible bounds OBGI are those given by the 2-sum rule function which agrees with $S(\beta)$ for $\beta = \beta_i$, $i = 0, 1, 2, 3$, as constructed in Theorem 10, together with the 1-sum rule function $S_1(\beta)$ which agrees with $S(\beta)$ at β_1 and β_2.

Proof: (See Fig. 3.) Without going into the details (which are given in Ref. 8), we merely point out that it is possible to construct a quantum mechanical sum rule function which agrees with $S(\beta)$ at the interpolation points and which would contradict any proposed improvement (OBGI) on any of the bounds which are supplied by $S_1(\beta)$ and $S_2(\beta)$ in the configurations shown in Fig. 3. Moreover, we can show that it is impossible to prescribe a nontrivial upper bound (OBGI) on $S(\beta)$ outside $[\beta_0, \beta_3]$.

E. Best Possible Bounds on a Quantum Mechanical Sum Rule Function When Five Sum Rules Are Known

We simply refer to Fig. 4.

Weinhold[11] presents a method for obtaining *optimal* bounds, having been given some *particular set* of *sum rules*, on a *few other sum rules*. The bounds derived from the given set of sum rules are neither *extensive* nor *exclusive*. Certainly, the results he obtains from sets of two, three, four, or five known sum rules are particular cases of the results obtained here.

Theorem 9 has already been presented by Kramer[4] but the best possible statement is new.

ACKNOWLEDGMENT

This work was motivated by Dr. Saul T. Epstein, who also supplied the uniqueness theorem and its consequences and insisted on the rigorous establishment of the best possible statements. He is thanked for encouraging and assisting me.

* This research was supported by the National Aeronautics and Space Administration (Grant NGL 50-002-001) and by the National Science Foundation (Grant GP 12832).

[1] J. O. Hirschfelder, W. Byers Brown, and S. T. Epstein, "Recent Developments in Perturbation Theory" in *Advances in Quantum Chemistry* (Academic, New York, 1964), Vol. 1.

[2] A. Dalgarno and A. E. Kingston, Proc. Roy. Soc. (London) **A259**, 424 (1960).

[3] R. T. Pack, Chem. Phys. Letters **5**, 257 (1970).

[4] H. L. Kramer, J. Chem. Phys. **53**, 2783 (1970).

[5] W. D. Davison, J. Phys. B., Ser. 2, **1**, 597 (1968).

[6] S. T. Epstein, J. Chem. Phys. **48**, 4716 (1968).

[7] G. A. Baker, Jr., Advan. Theoret. Phys. **1**, 1 (1965).

[8] M. F. Barnsley, Theoretical Chemistry Institute, University of Wisconsin, TCI-394, 1970 (unpublished).

[9] By "connected" we mean here that if (ν', ϵ') and (ν'', ϵ'') belong to D, then there exists a continuous path $(\nu(\alpha), \epsilon(\alpha))$ parametrized by $\alpha \in [0, 1]$, lying entirely within D, such that $(\nu(0), \epsilon(0)) = (\nu', \epsilon')$ and $(\nu(1), \epsilon(1)) = (\nu'', \epsilon'')$.

[10] If an $S_N(\beta)$ *can* be found for interpolation of $S^{(q)}(\beta)$, exact at say $2N$ points, then all the theory will clearly carry through, and upper and lower bounds to $S^q(\beta)$ will be obtained. These will presumably also be "best possible." Walsh has given a method by which such an $S_N(\beta)$ may be obtained in the case of equally spaced interpolation points [Proc. Phys. Soc. (London) **32** (1920)].

[11] F. Weinhold, J. Phys. A, Ser. 2, **1**, 655 (1968).

Analytic Properties of the Phase Matrix*

ALEX J. DRAGT
Center for Theoretical Physics, Department of Physics and Astronomy, University of Maryland College Park, Maryland 20742

(Received 16 October 1970)

We write the S matrix in the form $S = \exp(i\eta)$ and study the analytic properties of the 2-body matrix elements of η as a function of two complex variables. We find that the 2-body matrix elements of η have analytic properties similar to those of the K matrix and the second sheet scattering matrix. In particular, they do not possess Mandelstam analyticity.

1. INTRODUCTION

To implement the requirement of unitarity imposed upon the scattering matrix S, it is tempting to write

$$S = \exp(i\eta) \tag{1.1}$$

and require that η be Hermitian.[1] For example, one might consider a calculational scheme in which η is expanded in powers of a coupling constant.[2] The S matrix would then be unitary in each order of perturbation. We observe also that if η (and hence S) is diagonalized, its eigenvalues are equal to twice the scattering *eigen* phase shifts. In this note we study the two-body matrix elements of η and find that they have analytic properties similar to those of the K-matrix. In particular, the matrix elements of η have singularities in the momentum transfer variable whose position depends upon the energy. Thus, they do not satisfy the Mandelstam representation or the usual energy dispersion relations at fixed momentum transfer.

2. NOTATIONAL PRELIMINARIES

For simplicity, we shall consider a theory in which there are only two kinds of spinless particles. We denote a two particle plane-wave state by $|pk\rangle$ and assign a relativistically invariant scalar product by the rule

$$\langle p'k' \,|\, pk\rangle = E\omega\delta_3(\mathbf{p} - \mathbf{p}')\delta_3(\mathbf{k} - \mathbf{k}'). \tag{2.1}$$

Here E and ω are the "energy" components of p and k respectively. To complete the notation, we set $p^2 = M^2$ and $k^2 = m^2$.

We know that S commutes with the generators of the Poincaré group, and require that η also do so. Thus, we may write

$$\langle p'k'|\,\eta\,|pk\rangle = \delta_4(p + k - p' - k')h(s, t), \tag{2.2}$$

where s and t are the usual Lorentz invariant variables

$$s = (p + k)^2,$$
$$t = (p - p')^2. \tag{2.3}$$

Since η is Hermitian, it follows that h is *real* for s and t physical.

In order to relate the matrix elements of η to phase shifts, it is convenient to introduce two particle states having a definite angular momentum. We define a two-particle angular momentum state at rest in the center-of-momentum frame in terms of linear-momentum states by the equation

$$|Q^0JJ_3\rangle = [4q(2J + 1)]^{\frac{1}{2}}s^{-\frac{1}{4}}\int dR\, D^{J*}_{J_3,0}(R)U(R)\,|p^0k^0\rangle. \tag{2.4}$$

Here dR indicates Haar integration over the rotation group, and p^0, k^0, and Q^0 are the 4-vectors

$$p^0 = (qe_3, E^0),$$
$$k^0 = (-qe_3, \omega^0), \tag{2.5}$$
$$Q^0 = (0, 0, 0, s^{\frac{1}{2}}),$$

with

$$s^{\frac{1}{2}} = E^0 + \omega^0$$

and

$$q = [s - (M + m)^2]^{\frac{1}{2}}[s - (M - m)^2]^{\frac{1}{2}}/(2s^{\frac{1}{2}}).$$

Angular-momentum states with an arbitrary total 4-momentum Q are obtained by writing

$$|QJJ_3\rangle = U(L)\,|Q^0JJ_3\rangle, \tag{2.6}$$

where

$$Q = LQ^0, \tag{2.7}$$

and L is a pure velocity transformation. From the definitions (2.4), (2.6), and the normalization rule (2.1), one obtains the analogous rule

$$\langle Q'J'J_3' \,|\, QJJ_3\rangle = \delta_4(Q' - Q)\delta_{J'J}\delta_{J_3'J_3}. \tag{2.8}$$

Equation (2.4) can also be inverted to express linear-momentum states in terms of angular-momentum states. One finds

$$U(R)\,|p^0k^0\rangle = s^{\frac{1}{4}}(4\pi q)^{-\frac{1}{2}}\sum_{JJ_3} D^J_{J_3,0}(R)(2J + 1)^{\frac{1}{2}}\,|Q^0JJ_3\rangle. \tag{2.9}$$

Now suppose s is physical, but *below the first inelastic threshold*. Then, from Poincaré invariance,

we must have

$$\eta\,|QJJ_3\rangle = h_J(s)\,|QJJ_3\rangle, \tag{2.10}$$

where $h_J(s)$ is real. By the definition of phase shifts in the elastic region, we may write

$$S\,|QJJ_3\rangle = \exp(2i\delta_J)\,|QJJ_3\rangle. \tag{2.11}$$

It follows from Eq. (1.1) that

$$h_J(s) = 2\delta_J(s) \tag{2.12}$$

in the elastic region.

Using Eqs. (2.4) and (2.9), we find that $h_J(s)$ and $h(s, t)$ are related by the formulas

$$h(s, t) = s^{\frac{1}{2}}(4\pi q)^{-1}\sum(2J + 1)h_J(s)P_J(z), \tag{2.13a}$$

$$h_J(s) = 2\pi q s^{-\frac{1}{2}}\int_{-1}^{1} h(s, t)P_J(z)\,dz, \tag{2.13b}$$

where z is the cosine of the scattering angle given by

$$t = -2q^2(1 - z). \tag{2.14}$$

We shall also need a relation between $h_J(s)$ and the partial wave scattering amplitude defined by

$$\langle Q'J'J_3'|\,T\,|QJJ_3\rangle = \delta_{J'J}\delta_{J_3'J_3}\delta_4(Q' - Q)A_J(s), \tag{2.15}$$

with

$$S = 1 + 2iT. \tag{2.16}$$

Comparison of Eqs. (2.11) and (2.15) gives

$$\exp(2i\delta_J) = 1 + 2iA_J(s). \tag{2.17}$$

Thus, in the elastic region, we obtain the relation

$$h_J(s) = -i\log[1 + 2iA_J(s)]. \tag{2.18}$$

Of course, the familiar plane-wave scattering amplitude $A(s, t)$, defined by

$$\langle p'k'|\,T\,|pk\rangle = \delta_4(p + k - p' - k')A(s, t), \tag{2.19}$$

is related to the partial amplitude $A_J(s)$ by formulas similar to Eqs. (2.13a) and (2.13b),

$$A(s, t) = s^{\frac{1}{2}}(4\pi q)^{-1}\sum(2J + 1)A_J(s)P_J(z) \tag{2.20a}$$

and

$$A_J(s) = 2\pi q s^{-\frac{1}{2}}\int_{-1}^{1} A(s, t, u)P_J(z)\,dz. \tag{2.20b}$$

3. ANALYTIC PROPERTIES OF $h(s, t)$

We shall assume that $A(s, t)$ enjoys the analytic properties of the Mandelstam representation. Thus, it is assumed analytic in the topological product of cut s, t, and u planes with u given by

$$s + t + u = 2m^2 + 2M^2. \tag{3.1}$$

To establish a nomenclature, the s plane, for example, is cut from s_1 to $+\infty$ with s_1 real and positive. We take the cuts to include possible poles. The aim of this section is to deduce the analytic properties of $h(s, t)$ from those assumed for $A(s, t)$.

It follows from the Mandelstam analyticity of $A(s, t)$ and Eq. (2.20b) that $A_J(s)$ is analytic in the entire s plane when the latter is cut along various portions of the positive and negative real axes. Further, $A_J(s)$ is known to be analytic along the real axis in the elastic region apart from a branch point at the elastic threshold.[3] It follows from Eq. (2.18) that $h_J(s)$ is also analytic along the real axis within the elastic region. We now extend $h_J(s)$ to the remainder of the s plane by analytic continuation. It is clear that $h_J(s)$ will be analytic everywhere that $A_J(s)$ is, except for possible logarithmic branch points where the argument of Eq. (2.18) vanishes. Such vanishing cannot occur for s in the elastic region since it would violate unitarity. Further, it can at most occur a finite number of times in any compact region of the s plane, since otherwise $A_J(s)$ would have the constant value $\frac{1}{2}i$. Consequently, these possible branch points are isolated.

We should remark that our definition of $h_J(s)$, when analytically extended to inelastic values of s, is at odds with the reality of $h(s, t)$ required by hermiticity. For it is clear that in the process of analytic continuation, Eq. (2.12) should continue to hold and $\delta_J(s)$ has an imaginary part above the first inelastic threshold. Thus, $h(s, t)$, when evaluated in various intervals along the real axis, is *not the boundary value of a single analytic function*. A similar result holds for the K matrix.

In the previous paragraphs, we sketched the analytic properties of $h_J(s)$. The function $h(s, t)$ will have the same analytic properties as a function of s for fixed t provided the series of Eq. (2.13a) converges. The remainder of this section will be devoted to studying the domain of convergence of this series and its analytic continuation outside its convergence domain.

We know from the theory of Legendre series[4] that the sum in Eq. (2.13a) will converge to an analytic function in the z plane within an ellipse having foci at ± 1 and a semimajor axis $\mathcal{M}$ given by

$$\mathcal{M} = \tfrac{1}{2}(R + R^{-1}), \tag{3.2a}$$

with

$$R = \liminf |h_J(s)|^{-1/J} \tag{3.2b}$$

It follows from Eq. (2.18) that

$$\liminf |h_J(s)|^{-1/J} = \liminf |A_J(s)|^{-1/J}, \tag{3.3}$$

since $A_J(s)$ approaches zero exponentially with

increasing J. Consequently the domain of convergence for the expansion (2.13a) is the same as that for Eq. (2.20a).

Again from the theory of Legendre expansions, the series (2.20a) converges within the largest ellipse in the z plane for which $A(s, t)$, when viewed as a function of z for fixed s, is analytic. We denote this ellipse by $E_1(s)$, and the singular points on its boundary by the generic symbol $z_1(s)$. The singularities in z are, of course, the images of singularities in t and u under the mappings given by Eqs. (2.14) and (3.1). Thus, for example, if $A(s, t)$ is singular at t_1, then

$$z_1(s) = 1 + t_1/(2q^2). \tag{3.4}$$

At this point we have learned that $h(s, t)$ is simultaneously analytic in s and t for s not on the previously studied singularities of $h_J(s)$, and t such that $z \in E_1(s)$. Our next task is to continue h analytically in z (and hence in t) beyond $E_1(s)$. We begin by noting that Eq. (2.18) can be rewritten in the form

$$h_J(s) = \int_0^2 d\lambda f_J^\lambda(s), \tag{3.5}$$

where $f_J^\lambda(s)$ is defined by writing

$$f_J^\lambda(s) = A_J(s)[1 + i\lambda A_J(s)]^{-1}. \tag{3.6}$$

Here we take the contour of integration in λ to be such that zeros in the denominator of $f_J^\lambda(s)$ are avoided. Next define $f^\lambda(s, t)$ by the rule

$$f^\lambda(s, t) = s^{\frac{1}{2}}(4\pi q)^{-1} \sum (2J + 1) f_J^\lambda(s) P_J(z). \tag{3.7}$$

By virtue of our definition, we have the relation

$$h(s, t) = \int_0^2 d\lambda f^\lambda(s, t), \tag{3.8}$$

and our problem is reduced to studying the analytic properties of $f^\lambda(s, t)$.

Before continuing further, it is necessary to make a small mathematical digression.

Theorem: Suppose one is given two functions $F(z)$ and $G(z)$, both of which are analytic in the interval $[-1, 1]$ and whose singularity structure is completely known. Expand each in Legendre polynomials

$$F(z) = \sum (2J + 1) F_J P_J(z), \quad \text{etc.} \tag{3.9}$$

Consider the function $H(z)$ defined by the series

$$H(z) = \sum (2J + 1) F_J G_J P_J(z). \tag{3.10}$$

Then $H(z)$, when analytically continued, has singularities only at points z_H given by the relation

$$z_H = z_F z_G + [(z_F^2 - 1)(z_G^2 - 1)]^{\frac{1}{2}}, \tag{3.11}$$

where z_F and z_G are singular points of F and G.[5]

Proof: We note that Legendre polynomials obey the integral relation

$$\int d\Omega_b P_J(\hat{a} \cdot \hat{b}) P_{J'}(\hat{b} \cdot \hat{c}) = 4\pi(2J + 1)^{-1} \delta_{JJ'} P_J(\hat{a} \cdot \hat{c}), \tag{3.12}$$

where $\hat{a}$, $\hat{b}$, and $\hat{c}$ are three unit vectors. Thus, Eq. (3.10) can be rewritten in the form

$$H(\hat{a} \cdot \hat{c}) = (4\pi)^{-1} \int d\Omega_b F(\hat{a} \cdot \hat{b}) G(\hat{b} \cdot \hat{c}). \tag{3.13}$$

Next exploit the known analytic properties of F and G by writing Cauchy representations of the form

$$F(\hat{a} \cdot \hat{b}) = (2\pi i)^{-1} \int_C dw F(w)(w - \hat{a} \cdot \hat{b})^{-1} \quad \text{etc.}, \tag{3.14}$$

where C is some counterclockwise contour encircling the interval $[-1, 1]$. Inserting these representations into Eq. (3.13),we obtain

$$H(z) = \int_{C_1} dw_1 \int_{C_2} dw_2 F(w_1) G(w_2) R(w_1, w_2, z), \tag{3.15}$$

where R is an integral kernel given by

$$R(w_1, w_2, \hat{a} \cdot \hat{c}) = (2\pi i)^{-2}(4\pi)^{-1} \int d\Omega_b (w_1 - \hat{a} \cdot \hat{b})^{-1} (w_2 - \hat{b} \cdot \hat{c})^{-1}. \tag{3.16}$$

The indicated integration for R can be performed directly with the result

$$R(w_1 w_2 z) = (8\pi^2)^{-1} r^{-\frac{1}{2}} \times \log [(w_1 w_2 - z + r^{\frac{1}{2}})/(w_1 w_2 - z - r^{\frac{1}{2}})], \tag{3.17}$$

where

$$r = w_1^2 + w_2^2 + z^2 - 2w_1 w_2 z - 1. \tag{3.18}$$

A careful study of R as a function of z for fixed w_1 and w_2 shows that it is analytic in the z plane cut from z_1 to $+\infty$ with z_1 given by

$$z_1 = w_1 w_2 + [(w_1^2 - 1)(w_2^2 - 1)]^{\frac{1}{2}}. \tag{3.19}$$

In fact, R has the integral representation

$$R(w_1 w_2 z) = -(2\pi)^{-2} \int_{z_1}^{\infty} dw_3 (w_3 - z)^{-1} r^{-\frac{1}{2}}(w_1 w_2 w_3). \tag{3.20}$$

Looking at Eq. (3.15), we see that H is certainly analytic when z differs from the values of z_1 given by Eq. (3.19) as w_1 and w_2 range over the contours C_1 and C_2. Further, the contours C_1 and C_2 can be distorted at will unless one encounters singularities of F and G. Consequently, H will be singular only at points given by Eq. (3.11).

We now return to our earlier discussion. Note that Eq. (3.6) can be rewritten in the form

$$f_J^\lambda(s) = A_J(s) - i\lambda A_J(s) f_J^\lambda(s). \tag{3.21}$$

Multiply both sides of Eq. (3.21) by $s^{\frac{1}{2}}(4\pi q)^{-1}(2J + 1)P_J(z)$ and sum over J. The result is

$$f^\lambda(s, t) = A(s, t) - i\lambda s^{\frac{1}{2}}(4\pi q)^{-1} \times \sum (2J + 1) A_J(s) f_J^\lambda(s) P_J(z). \tag{3.22}$$

Finally, apply the theorem summarized by Eqs. (3.10) and (3.15) to obtain the integral equation

$$f^\lambda(s, t) = A(s, t) - i4\pi q s^{-\frac{1}{2}} \times \int_{C_1} dw_1 \int_{C_2} dw_2 A(s, t_1) f^\lambda(s, t_2) R(w_1 w_2 z). \tag{3.23}$$

Here t_1 and t_2 are given by the relations

$$t_1 = -2q^2(1 - w_1), \quad \text{etc.} \tag{3.24}$$

With the aid of Eq. (3.23), the analytic continuation of $f^\lambda(s, t)$ outside of $E_1(s)$ can be made immediately. We already know that $A(s, t)$ and $f^\lambda(s, t)$ are analytic for $z \in E_1(s)$. It follows that the term in Eq. (3.23) involving the double integral is analytic in a *larger* ellipse $E_2(s)$ whose boundary must lie on (or perhaps beyond) the boundary obtained by inserting into Eq. (3.19) values of w_1 and w_2 lying on the boundary of $E_1(s)$. Thus, the singularities of $f^\lambda(s, t)$ *within* $E_2(s)$ are the *same* as those of $A(s, t)$, $f^\lambda(s, t)$ is also singular at the point or points $z_1(s)$. The analytic continuation of $f^\lambda(s, t)$ from $E_1(s)$ to $E_2(s)$ has been accomplished.

A simple extension of our previous argument permits us to continue f^λ into a still larger ellipse $E_3(s)$. Since we now know that f^λ is analytic within $E_2(s)$ except for singularities at $z_1(s)$, we can conclude that the integral in Eq. (3.23) is analytic in an ellipse $E_3(s)$ [which is larger than $E_2(s)$] except for singularities at points $z_2(s)$ given by

$$z_2(s) = z_1' z_1 + [(z_1'^2 - 1)(z_1^2 - 1)]^{\frac{1}{2}}. \tag{3.25}$$

The general pattern should now be clear. By repeated iteration of our argument, we find that f^λ is analytic in ever larger ellipses except for points z_α given by

$$z_\alpha = z_\beta z_\gamma + [(z_\beta^2 - 1)(z_\gamma^2 - 1)]^{\frac{1}{2}}, \tag{3.26}$$

where z_β is a singular point of A and z_γ is a singular point of f^λ found in previous iterations.[6]

In summary, we have found that $f^\lambda(s, t)$ is simultaneously analytic in s and t except for

(a) the usual singularities of $A_J(s)$ in the s plane,

(b) poles in s due to zeros in the denominator of f^λ for fixed λ. See Eq. (3.6). These poles may also be viewed as simple poles in λ whose position depends upon s.

(c) the usual singularities of $A(s, t)$,

(d) singularities given by Eq. (3.26). Suppose, to give a simple example, that $A(s, t)$ has a singularity in t *only* at $t = t_1$. We then find, by successive iteration, that $f^\lambda(s, t)$ has singularities at the points

$$t_2(s) = 4t_1 + t_1^2/(q^2),$$
$$t_3(s) = 9t_1 + 6t_1^2/(q^2) + t_1^3/(q^4), \quad \text{etc.} \tag{3.27}$$

Finally, let us return to our object of interest, $h(s, t)$. Looking at Eq. (3.8), we see that h will be simultaneously analytic in s and t except at the points enumerated under the headings a, c, and d in the previous paragraph. In addition, h will have logarithmic singularities in s whenever the λ poles [of heading b] coincide with the integration end point $\lambda = 2$. These are just the logarithmic branch points already found in $h_J(s)$.

4. DISCUSSION

It is interesting to compare the analytic properties of the phase matrix η with those of the K matrix and $T^{(2)}$, the second sheet transition matrix. If we denote the usual first sheet transition matrix simply by T as before, we may define η, K, and $T^{(2)}$ by the relations

$$\eta = -i \log [1 + 2iT], \tag{4.1}$$

$$K = 2i(1 - S)(1 + S)^{-1} = 2T[1 + iT]^{-1}, \tag{4.2}$$

$$T^{(2)} = TS^{-1} = T[1 + 2iT]^{-1}. \tag{4.3}$$

Observe that the definition of η can be recast in the form

$$\eta = \int_0^2 d\lambda F(\lambda), \tag{4.4}$$

where

$$F(\lambda) = T(1 + i\lambda T)^{-1}. \tag{4.5}$$

We immediately see the relations

$$K/2 = F(1), \tag{4.6}$$

$$T^{(2)} = F(2). \tag{4.7}$$

We also note that Eq. (4.5) can be rewritten in the form

$$F(\lambda) = T - i\lambda T F(\lambda). \tag{4.8}$$

It is easy to check that Eq. (3.23) is equivalent to Eq. (4.8) if our attention is restricted to two-body matrix elements. We conclude that the analytic properties of K, $T^{(2)}$, and F (and hence η) are essentially the same.

In particular, matrix elements of K, F, and η will possess the Landau singularities [given by Eq. (3.26)] already well known for $T^{(2)}$. Thus, their matrix elements cannot satisfy the Mandelstam representation.

ACKNOWLEDGMENT

The subject of this article arose out of a discussion with C. H. Woo.

* Supported in part by the U.S. Air Force Office of Scientific Research under Grant AFOSR 68-1453A.

[1] H. Lehman and K. Symanzik, Nuovo Cimento **6**, 319, 1957.

[2] I. B. Birula, B. Mielnik, and J. Plebański, Ann. Phys. (N.Y.) **51**, 187, 1969.

[3] W. Zimmerman, Nuovo Cimento **21**, 249, 1961.

[4] G. Szegö, *Orthogonal Polynomials* (Am. Math. Soc., New York, 1959), Vol. XXIII.

[5] This theorem is the Legendre polynomial equivalent of Hadamard's multiplication theorem for Taylor Series. See E. C. Titchmarsh, *Theory of Functions* (Oxford U.P., London, 1939), 2nd ed., p. 157.

[6] For further details, see Ref. 3.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Separable Approximations to Matrices and Functions of Two Variables*

M. BOLSTERLI AND J. L. NORTON
Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87544

(Received 23 November 1970)

Considerable computational time can be saved by the use of separable approximations to matrices and functions of two variables. Such approximations are considered in general.

In solving the Schrödinger equation for a nucleon in a cylindrically symmetric potential, the most efficient method is expansion of the wavefunction in products of oscillator eigenfunctions. The matrix elements of the potential are then of the form

$$V_{ij,kl} = \int \varphi_i^*(z)\chi_j^*(\rho)V(\rho, z)\varphi_k(z)\chi_l(\rho)\rho\, dp\, dz. \quad (1)$$

Typically, there are 60 or so pairs (k, l), so that of order 2,000 integrals of the type (1) must be evaluated for a given function V. If 500 points per integral were used, this would give 10^6 evaluations of the integrand and a major contribution to the total time of the eigenvalue computation.

Stimulated by S. Wahlborn's suggestion that it might be feasible to "factor" the potential V, we have developed the formal theory of such factorization. It turns out to be both possible and extremely useful. Consider first the saving that can be made if V is approximated by

$$V_N(\rho, z) \cong \sum_{\nu=1}^{N} f_\nu(\rho)g_\nu(z). \quad (2)$$

Then the integral in (1) becomes

$$\sum_{\nu=1}^{N} \int \varphi_i^*(z)g_\nu(z)\varphi_k(z)\, dz \int \chi_j^*(\rho)f_\nu(\rho)\chi_l(\rho)\rho\, d\rho \quad (3)$$

so that of the order of 120 one-dimensional integrals, each needing, say, 25 points, must be evaluated. Typically N is less than 10, so that the total number of operations is less than 30,000, a saving of a factor of at least 30.

To justify our consideration of matrices, we note that an integral of the type (1) is generally evaluated by using a Gaussian sum with weights w_ρ and w_z appropriate to the long-range behavior of the integrand

$$V_{ij,kl} \cong \sum_{\alpha=1}^{N_\alpha} \sum_{\beta=1}^{N_\beta} w_\rho(\alpha)w_z(\beta)\varphi_i^*(z_\beta)\chi_j^*(\rho_\alpha) \times V(\rho_\alpha, z_\beta)\varphi_k(z_\beta)\chi_l(\rho_\alpha), \quad (4)$$

so that it is appropriate to approximate not $V(\rho, z)$ but rather the matrix

$$V_{\alpha\beta} = V(\rho_\alpha, z_\beta). \quad (5)$$

For generality, consider the real $N \times M$ matrix A_{ij} with $M \leq N$. We wish to approximate A_{ij} by $S_{ij}^{(R)}$ with

$$S_{ij}^{(R)} = \sum_{\nu=1}^{R} f_{i\nu}g_{j\nu}, \quad R \leq M. \quad (6)$$

As a criterion for the approximation we use the sum of squares

$$J = \sum_{ij} (A_{ij} - S_{ij}^{(R)})^2. \quad (7)$$

Setting the variation of J with respect to $f_{i\nu}$ and $g_j{}^\nu$, respectively, to zero gives the equations for the set of f's and g's that minimizes J:

$$\sum_j A_{ij}g_{j\nu} = \sum_{j\mu} g_{j\mu}g_{j\nu}f_{i\mu},$$

$$\sum_i A_{ij}f_{i\nu} = \sum_{i\mu} f_{i\mu}f_{i\nu}g_{j\nu}. \quad (8)$$

In particular, matrix elements of K, F, and η will possess the Landau singularities [given by Eq. (3.26)] already well known for $T^{(2)}$. Thus, their matrix elements cannot satisfy the Mandelstam representation.

ACKNOWLEDGMENT

The subject of this article arose out of a discussion with C. H. Woo.

* Supported in part by the U.S. Air Force Office of Scientific Research under Grant AFOSR 68-1453A.

[1] H. Lehman and K. Symanzik, Nuovo Cimento **6**, 319, 1957.

[2] I. B. Birula, B. Mielnik, and J. Plebański, Ann. Phys. (N.Y.) **51**, 187, 1969.

[3] W. Zimmerman, Nuovo Cimento **21**, 249, 1961.

[4] G. Szegö, *Orthogonal Polynomials* (Am. Math. Soc., New York, 1959), Vol. XXIII.

[5] This theorem is the Legendre polynomial equivalent of Hadamard's multiplication theorem for Taylor Series. See E. C. Titchmarsh, *Theory of Functions* (Oxford U.P., London, 1939), 2nd ed., p. 157.

[6] For further details, see Ref. 3.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Separable Approximations to Matrices and Functions of Two Variables*

M. BOLSTERLI AND J. L. NORTON
Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87544

(Received 23 November 1970)

Considerable computational time can be saved by the use of separable approximations to matrices and functions of two variables. Such approximations are considered in general.

In solving the Schrödinger equation for a nucleon in a cylindrically symmetric potential, the most efficient method is expansion of the wavefunction in products of oscillator eigenfunctions. The matrix elements of the potential are then of the form

$$V_{ij,kl} = \int \varphi_i^*(z)\chi_j^*(\rho)V(\rho, z)\varphi_k(z)\chi_l(\rho)\rho\, dp\, dz. \quad (1)$$

Typically, there are 60 or so pairs (k, l), so that of order 2,000 integrals of the type (1) must be evaluated for a given function V. If 500 points per integral were used, this would give 10^6 evaluations of the integrand and a major contribution to the total time of the eigenvalue computation.

Stimulated by S. Wahlborn's suggestion that it might be feasible to "factor" the potential V, we have developed the formal theory of such factorization. It turns out to be both possible and extremely useful. Consider first the saving that can be made if V is approximated by

$$V_N(\rho, z) \cong \sum_{\nu=1}^{N} f_\nu(\rho)g_\nu(z). \quad (2)$$

Then the integral in (1) becomes

$$\sum_{\nu=1}^{N} \int \varphi_i^*(z)g_\nu(z)\varphi_k(z)\, dz \int \chi_j^*(\rho)f_\nu(\rho)\chi_l(\rho)\rho\, d\rho \quad (3)$$

so that of the order of 120 one-dimensional integrals, each needing, say, 25 points, must be evaluated. Typically N is less than 10, so that the total number of operations is less than 30,000, a saving of a factor of at least 30.

To justify our consideration of matrices, we note that an integral of the type (1) is generally evaluated by using a Gaussian sum with weights w_ρ and w_z appropriate to the long-range behavior of the integrand

$$V_{ij,kl} \cong \sum_{\alpha=1}^{N_\alpha} \sum_{\beta=1}^{N_\beta} w_\rho(\alpha)w_z(\beta)\varphi_i^*(z_\beta)\chi_j^*(\rho_\alpha) \times V(\rho_\alpha, z_\beta)\varphi_k(z_\beta)\chi_l(\rho_\alpha), \quad (4)$$

so that it is appropriate to approximate not $V(\rho, z)$ but rather the matrix

$$V_{\alpha\beta} = V(\rho_\alpha, z_\beta). \quad (5)$$

For generality, consider the real $N \times M$ matrix A_{ij} with $M \leq N$. We wish to approximate A_{ij} by $S_{ij}^{(R)}$ with

$$S_{ij}^{(R)} = \sum_{\nu=1}^{R} f_{i\nu}g_{j\nu}, \quad R \leq M. \quad (6)$$

As a criterion for the approximation we use the sum of squares

$$J = \sum_{ij} (A_{ij} - S_{ij}^{(R)})^2. \quad (7)$$

Setting the variation of J with respect to $f_{i\nu}$ and $g_j{}^\nu$, respectively, to zero gives the equations for the set of f's and g's that minimizes J:

$$\sum_j A_{ij}g_{j\nu} = \sum_{j\mu} g_{j\mu}g_{j\nu}f_{i\mu},$$

$$\sum_i A_{ij}f_{i\nu} = \sum_{i\mu} f_{i\mu}f_{i\nu}g_{j\nu}. \quad (8)$$

Consider first the case where the f's and g's form orthogonal sets, so that

$$\sum_j g_{j\mu}g_{j\nu} = G_\mu\delta_{\mu,\nu}, \quad \sum_i f_{i\mu}f_{i\nu} = F_\mu\delta_{\mu,\nu}. \tag{9}$$

(It will be shown shortly that the most general case can be reduced to this.) Then Eqs. (8) give

$$\sum_j A_{ij}g_{j\nu} = G_\nu f_{i\nu}, \quad \sum_i A_{ij}f_{i\nu} = F_\nu g_{j\nu}, \tag{10}$$

and there is no coupling between different values of ν. It follows from (10) that

$$\sum_{ij} A_{ik}A_{ij}g_{j\nu} = F_\nu G_\nu g_{k\nu} = \sum_j (\tilde{A}A)_{kj}g_{j\nu},$$

$$\sum (A\tilde{A})_{ki}f_{i\nu} = F_\nu G_\nu f_{k\nu} \tag{11}$$

so that the vector f_ν is an eigenvector of the symmetric positive definite matrix $A\tilde{A}$, with eigenvalue $\lambda_\nu = F_\nu G_\nu \geq 0$, and g_ν is an eigenvector of $\tilde{A}A$ belonging to the same eigenvalue λ_ν. If λ_ν is zero, then either f_ν or g_ν is zero, and the term does not contribute to $S^{(R)}$.

With (10), it follows that

$$J = \sum_{ij} A_{ij}^2 - \sum_{\nu=1}^{R}\lambda_\nu = \mathrm{Tr}\,(\tilde{A}A) - \sum_{\nu=1}^{R}\lambda_\nu. \tag{12}$$

Since $\tilde{A}A$ has at most M eigenvalues λ_ν, it follows that $J = 0$ for $R = M$; for $R < M$, the vectors f_ν and g_ν belonging to the largest eigenvalues λ_ν should be chosen so as to minimize J.

Hence, given $R \leq M$, the procedure for finding $S^{(R)}$ is (a) find the R largest eigenvalues λ_ν of $\tilde{A}A$ and call the corresponding normalized eigenvectors g_ν and (b) let

$$f_\nu = Ag_\nu. \tag{13}$$

Then

$$S_{ij}^{(R)} = \sum_{\nu=1}^{R} f_{i\nu}g_{j\nu}. \tag{14}$$

To show that the f's and g's can be taken orthogonal, suppose the set of f's and g's satisfies (8). Let

$$F_{\mu\nu} = \sum_i f_{i\mu}f_{i\nu}, \quad G_{\mu\nu} = \sum_i g_{i\mu}g_{i\nu}. \tag{15}$$

Then

$$\sum_{ij} A_{ik}A_{ij}g_{j\nu} = \sum_{i\mu} A_{ik}f_{i\mu}G_{\mu\nu} = \sum_{\rho\mu} g_{k\rho}F_{\rho\mu}G_{\mu\nu},$$

$$\sum_{ij} A_{kj}A_{ij}f_{i\nu} = \sum_{\rho\mu} f_{k\rho}G_{\rho\mu}F_{\mu\nu}. \tag{16}$$

Call

$$\sum_\mu F_{\rho\mu}G_{\mu\nu} = N_{\rho\nu}, \tag{17}$$

and let the eigenvectors of $N_{\rho\nu}$ be $v_{\nu\lambda}$ with eigenvalues ξ_λ:

$$\sum_\nu N_{\rho\nu}v_{\nu\lambda} = \xi_\lambda v_{\rho\lambda}. \tag{18}$$

Similarly, the eigenvectors w and eigenvalues η of $\tilde{N}$ satisfy

$$\sum_\rho w_{\rho\lambda}N_{\rho\nu} = \eta_\lambda w_{\rho\lambda}; \tag{19}$$

and, since it easily follows that

$$(\eta_\lambda - \xi_\mu)\sum_\rho w_{\rho\lambda}v_{\rho\mu} = 0, \tag{20}$$

the eigenvalues of N and $\tilde{N}$ are the same, and the w's and v's can be chosen so that

$$\sum_\rho w_{\rho\lambda}v_{\rho\mu} = \delta_{\lambda\mu}, \quad \sum_\mu w_{\rho\mu}v_{\sigma\mu} = \delta_{\rho\sigma}. \tag{21}$$

Now let

$$\sum_\nu f_{i\nu}w_{\nu\lambda} = b_{i\lambda}, \quad \sum g_{j\nu}v_{\nu\lambda} = d_{j\lambda}. \tag{22}$$

Then (16) gives

$$\sum_j (\tilde{A}A)_{kj}d_{j\lambda} = \xi_\lambda d_{k\lambda}, \quad \sum_j (A\tilde{A})_{ki}b_{i\lambda} = \xi_\lambda b_{i\lambda}, \tag{23}$$

so that the d_λ are mutually orthogonal by virtue of being eigenvectors of a symmetric matrix. Similarly, the b_λ are orthogonal. Finally, (21) and (22) give

$$f_{i\nu} = \sum_\lambda b_{i\lambda}v_{\nu\lambda}, \quad g_{j\nu} = \sum_\lambda d_{j\lambda}w_{\nu\lambda} \tag{24}$$

so that

$$S^{(R)} = \sum_{\lambda\mu} b_{i\lambda}d_{j\mu}v_{\nu\lambda}w_{\nu\mu} = \sum_{\lambda=1}^{R} b_{i\lambda}d_{j\lambda} \tag{25}$$

as required.

The result is easily generalized to complex matrices, with transposes replaced by Hermitian adjoints.

Similarly, for functions of two variables, the procedure described in (13) and (14) and the preceding paragraph is changed as follows. Let $A(x, y)$ be the function to be approximated. Let

$$K(x, x') \equiv \int dy A(x, y)A(x', y).$$

Then find the R largest eigenvalues λ_ν of $K(x, x')$, and let the corresponding eigenfunctions be $f_\nu(x)$:

$$\int K(x, x')f_\nu(x')\,dx' = \lambda_\nu f_\nu(x). \tag{26}$$

Let

$$g_\nu(y) = \int A(x, y)f_\nu(x)\,dx. \tag{27}$$

Then

$$S^{(R)}(x, y) = \sum_{\nu=1}^{R} f_\nu(x)g_\nu(y) \tag{28}$$

is the R-term separable approximation to $A(x, y)$ that minimizes

$$J = \int [A(x, y) - S^{(R)}(x, y)]^2\,dx\,dy. \tag{29}$$

The proof of the immediately preceding statements is completely analogous to that sketched for the matrix case.

* Work done under the auspices of the U.S. Atomic Energy Commission.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Density and Current Density as Coordinates for a System of Interacting Bosons at Absolute Zero

B. TSU-SHEN CHANG*
Department of Physics, University of Michigan, Ann Arbor, Michigan 48104

(Received 25 June 1970)

The quantum field Hamiltonian expressed in terms of density and current density variables has been employed together with the equal-time commutation relations among these variables to find the ground state energy and the density fluctuation excitation spectrum of a system of interacting bosons at $T = 0$ °K. The approximation involved consists in assuming that the density fluctuation in space is small compared with the average density. The results easily obtained in the lowest-order approximation agree with those of Bogoliubov. However, in our treatment no condensation of particles in zero-momentum state is assumed or apparent. A connection between the present treatment and the quantum hydrodynamic approach to the irrotational flow of a Bose liquid has been made.

1. INTRODUCTION

For liquid He II near the absolute zero of temperature, it is a well-known fact that there exist density fluctuations associated with the ordinary sound waves.[1,2] To deal with such collective motions, it seems natural to think of the density variable as a proper quantum mechanical coordinate. Previous workers in this connection have used a number of methods, but none of which involved the density variable in a microscopic theory in the way as is presented here.[1-8] Following the recent suggestions of Dashen and Sharp,[9] we have employed the density and the current density as quantum field coordinates together with their equal-time commutation relations to find the ground state energy and the excitation spectrum of an interacting Bose system, a system related to liquid He II.

This method is different from the usual field-theoretic method in that we use the equal-time commutation relations among the density and the current density components, instead of those among the canonical fields ψ and ψ^+. Although the present commutation relations look rather complicated, it turns out to be quite simple to get an approximate energy spectrum for the system on hand by (i) employing a functional representation for the Fourier components of the density and the current density operators and the state vectors of the quantum field for the system and (ii) assuming that the density fluctuation in space is small compared with the average density. The results obtained in the lowest order approximation, which are valid for long wave vectors ($k < 2\pi/r_0$, where r_0 = the average interparticle distance), agree with those of Bogoliubov.[3] However, here no condensation of particles in zero-momentum state is assumed or apparent, and the excitation spectrum refers to the density fluctuations rather than to Bogoliubov's quasiparticles.

The functional representation of the equal-time commutation relations in a representation in which the Fourier components of the density $\rho_{k\neq 0}$ are diagonal is presented in Sec. 2. Section 3 is concerned with finding the energy spectrum of the system in the lowest-order approximation. A connection of the present treatment with the quantum hydrodynamic approach[10,11] to the problem will be made in Sec. 4. Finally in Sec. 5 there will be some relevant discussions about the validity of the approximations used and the results in the low-density limit.

In this paper, we shall adopt units such that the mass of the boson particles $m = 1$ and $\hbar = 1$, unless otherwise stated.

2. FUNCTIONAL REPRESENTATION OF THE COMMUTATION RELATIONS

Consider a system of N spinless bosons interacting through a two-body repulsive central potential and enclosed in a box of volume Ω. The usual quantum field Hamiltonian is

$$H = \tfrac{1}{2}\int d^3x \nabla\psi^+(\mathbf{x})\cdot\nabla\psi(\mathbf{x}) + \tfrac{1}{2}\iint d^3x d^3y \psi^+(\mathbf{x})\psi^+(\mathbf{y})V(|\mathbf{x}-\mathbf{y}|)\psi(\mathbf{y})\psi(\mathbf{x}), \tag{1}$$

where the field operators ψ and ψ^+ obey the usual equal-time canonical relations and $V(|\mathbf{x}-\mathbf{y}|)$ is the two-body interacting potential. This Hamiltonian can be expressed in terms of the density ρ and current density $\mathbf{j}$ by using the identities derived from their definitions,

$$2[\nabla\psi^+(\mathbf{x})]\psi(\mathbf{x}) = \nabla\rho(\mathbf{x}) - 2i\mathbf{j}(\mathbf{x}),$$
$$2\psi^+(\mathbf{x})[\nabla\psi(\mathbf{x})] = \nabla\rho(\mathbf{x}) + 2i\mathbf{j}(\mathbf{x}),$$

and

$$1/\rho(\mathbf{x}) = \psi^{-1}(\mathbf{x})(\psi^+(\mathbf{x}))^{-1},$$

in the kinetic energy part and writing the potential energy in terms of $\rho(\mathbf{x})$ and $\rho(\mathbf{y})$. Thus one has

$$H = \tfrac{1}{8}\int d^3x[\nabla\rho(\mathbf{x}) - 2i\mathbf{j}(\mathbf{x})]\cdot\frac{1}{\rho(\mathbf{x})}[\nabla\rho(\mathbf{x}) + 2i\mathbf{j}(\mathbf{x})] + \tfrac{1}{2}\iint d^3x d^3y\,\rho(\mathbf{x})V(|\mathbf{x}-\mathbf{y}|)\rho(\mathbf{y}) - \tfrac{1}{2}NV(0), \tag{2}$$

where $N = \int d^3x\,\rho(\mathbf{x})$ is the total number operator and $V(0)$ is the interaction potential when $\mathbf{x} = \mathbf{y}$. (We have used N for an operator and for the total number of particles of the system.) The equal-time commutation relations among the densities and the current density components, which can be obtained by aid of the equal-time canonical commutations among ψ and ψ^+, are given as

$$[\rho(\mathbf{x}), \rho(\mathbf{y})] = 0, \tag{3}$$

$$[\rho(\mathbf{x}), j_\alpha(\mathbf{y})] = i\rho(\mathbf{y})\left(\frac{\partial}{\partial y_\alpha}\delta(\mathbf{x}-\mathbf{y})\right) \tag{4}$$

and

$$[j_\alpha(\mathbf{x}), j_\beta(\mathbf{y})] = -ij_\beta(\mathbf{x})\left(\frac{\partial}{\partial x_\alpha}\delta(\mathbf{x}-\mathbf{y})\right) + ij_\alpha(\mathbf{y})\left(\frac{\partial}{\partial y_\beta}\delta(\mathbf{x}-\mathbf{y})\right), \tag{5}$$

with α and β denoting Cartesian components.

We are going to use the above commutation relations, instead of the usual canonical relations for ψ and ψ^+, together with the expression (2) for the Hamiltonian to find the energy spectrum of our system. One way to do this is to resort to a functional representation for the density and the current density operators and for the states of the system. It is found more convenient to deal with the Fourier components of the density and the current density. We present in this section only the functional representation[12] of the commutation relations (3)–(5), a representation in which the Fourier components of the density $\rho_{k\neq 0}$ are diagonal, although they have complex eigenvalues.

Let

$$\rho(\mathbf{x}) = \frac{1}{\Omega^{\frac{1}{2}}}\sum_{\mathbf{k}}\rho_{\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{6}$$

$$\rho_{\mathbf{k}} = \frac{1}{\Omega^{\frac{1}{2}}}\int d^3x\,\rho(\mathbf{x})e^{-i\mathbf{k}\cdot\mathbf{x}}, \tag{7}$$

with the reality condition for $\rho(\mathbf{x})$, $\rho_{\mathbf{k}}^* = \rho_{-\mathbf{k}}$, where the functions $e^{i\mathbf{k}\cdot\mathbf{x}}$ obey periodic boundary conditions. In this $\rho_{\mathbf{k}}$-representation, the functional representation for the $\rho_{\mathbf{k}\neq 0}$ operator is just the c-number function $\rho_{\mathbf{k}}$. Thus the commutation relation (3) is clearly satisfied. It is worthwhile to note that ρ_0, the $k = 0$ Fourier component, has a value by the definition $\rho_0 = N/\Omega^{\frac{1}{2}}$ and is related to the average density by $\rho_0 = \Omega^{\frac{1}{2}}(N/\Omega) = \Omega^{\frac{1}{2}}\rho_{\text{av}}$. Similarly let

$$j_\alpha(\mathbf{x}) = \frac{1}{\Omega^{\frac{1}{2}}}\sum_{\mathbf{k}}j_{\alpha,\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{8}$$

$$j_{\alpha,\mathbf{k}} = \frac{1}{\Omega^{\frac{1}{2}}}\int d^3x j_\alpha(\mathbf{x})e^{-i\mathbf{k}\cdot\mathbf{x}}, \tag{9}$$

and its Hermitian adjoint

$$j_{\alpha,\mathbf{k}}^\dagger = j_{\alpha,-\mathbf{k}}.$$

To find the proper expression for $j_{\alpha,\mathbf{k}}$ operator in $\rho_{\mathbf{k}}$-representation, consider now the commutation relation (4). After putting in it the Fourier components for ρ and j_α, and $\delta(\mathbf{x}) = (1/\Omega)\sum_{\mathbf{q}}e^{i\mathbf{q}\cdot\mathbf{x}}$, one gets

$$\frac{1}{\Omega}\sum_{\mathbf{k},\mathbf{l}}[\rho_{\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{x}}, j_{\alpha,\mathbf{l}}e^{i\mathbf{l}\cdot\mathbf{y}}] = \frac{1}{\Omega^{\frac{3}{2}}}\sum_{\mathbf{p},\mathbf{q}}q_\alpha\rho_{\mathbf{p}}e^{i\mathbf{q}\cdot\mathbf{x}+i(\mathbf{p}-\mathbf{q})\cdot\mathbf{y}}. \tag{10}$$

For fixed values of $\mathbf{k}$ and $\mathbf{l}$, (10) yields, as $\mathbf{q} = \mathbf{k}$, $\mathbf{p} = \mathbf{k} + \mathbf{l}$,

$$[\rho_{\mathbf{k}}, j_{\alpha,\mathbf{l}}] = \frac{1}{\Omega^{\frac{1}{2}}}k_\alpha\rho_{\mathbf{k}+\mathbf{l}}. \tag{11}$$

To be consistent with (11), it is easily seen that a proper functional representation for $j_{\alpha,\mathbf{l}}$ would be

$$j_{\alpha,\mathbf{l}} = -\frac{1}{\Omega^{\frac{1}{2}}}k_\alpha\rho_{\mathbf{k}+\mathbf{l}}\frac{\delta}{\delta\rho_{\mathbf{k}}}, \tag{12}$$

where $\delta/\delta\rho_{\mathbf{k}}$ is a functional derivative with respect to the variable $\rho_{k\neq 0}$. Since the value of $\mathbf{k}$, although fixed, is still arbitrary, we get the following general functional representation for $j_{\alpha,\mathbf{l}}$ by summing over $\mathbf{k}$ on the right-hand side of (12):

$$j_{\alpha,\mathbf{l}} = -\frac{1}{\Omega^{\frac{1}{2}}}\sum_{\mathbf{k}}k_\alpha\rho_{\mathbf{k}+\mathbf{l}}\frac{\delta}{\delta\rho_{\mathbf{k}}}. \tag{13}$$

As pointed out by Grodnik and Sharp,[12] another term should be added to (13) in order to define an inner product on the functionals of $\rho_{\mathbf{k}}$ in such a way that $\rho(\mathbf{x})$ and $\mathbf{j}(\mathbf{x})$ operators are Hermitian. Thus the proper general functional representation is

$$j_{\alpha,\mathbf{l}} = -\frac{1}{\Omega^{\frac{1}{2}}}\sum_{\mathbf{k}}k_\alpha\rho_{\mathbf{k}+\mathbf{l}}\frac{\delta}{\delta\rho_{\mathbf{k}}} - \tfrac{1}{2}l_\alpha\rho_{\mathbf{l}}. \tag{14}$$

It is evident that the expression (14) for $j_{\alpha,\mathbf{l}}$ satisfies the commutation relation (10). The proof for the fact that it is also consistent with the commutation relation (5) is straightforward and hence omitted.

3. APPROXIMATE ENERGY SPECTRUM

With the functional representations obtained above for $\rho_{\mathbf{k}\neq 0}$ and $J_{\alpha,1}$, we are in a position to solve a Schrödinger equation approximately, which contains wavefunctionals, acting just like ordinary wavefunctions for the states of the system. To get the approximate energy spectrum for the system, we shall use an approximate Hamiltonian obtained from (2) and, likewise, an approximate functional representation for the current density component $J_{\alpha,1}$ [Eq. (14)], which is consistent with the commutation algebra and with the approximate Hamiltonian to be used. The approximation consists in assuming that the density fluctuations in space are small compared with the average density ρ_{av}, i.e.,

$$\tilde{\rho}(\mathbf{x}) = \rho(\mathbf{x}) - \rho_{\text{av}} \ll \rho_{\text{av}} \tag{15}$$

with

$$\int d^3x\, \tilde{\rho}(\mathbf{x}) = 0$$

as

$$N = \int d^3x\, \rho(\mathbf{x}) = \int d^3x\, \rho_{\text{av}}. \tag{16}$$

Using the Fourier series expansion for $\rho(\mathbf{x})$ according to (6), we have

$$\rho(\mathbf{x}) = \frac{1}{\Omega^{\frac{1}{2}}}\left(\rho_0 + \sum_{\mathbf{k}\neq 0} \rho_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}\right) = \rho_{\text{av}} + \frac{1}{\Omega^{\frac{1}{2}}} \sum_{\mathbf{k}\neq 0} \rho_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{17}$$

so that condition (15) is just

$$\tilde{\rho}(\mathbf{x}) = \frac{1}{\Omega^{\frac{1}{2}}} \sum_{\mathbf{k}\neq 0} \rho_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} \ll \rho_{\text{av}}. \tag{18}$$

By aid of a power series expansion,

$$1/\rho(\mathbf{x}) = (1/\rho_{\text{av}})\{1 - \tilde{\rho}(\mathbf{x})/\rho_{\text{av}} + [\tilde{\rho}(\mathbf{x})/\rho_{\text{av}}]^2 - \cdots\} \tag{19}$$

and $\rho(\mathbf{x}) = \rho_{\text{av}} + \tilde{\rho}(\mathbf{x})$, one expands the Hamiltonian (2) in terms of $\tilde{\rho}(\mathbf{x})$, and retains terms up to the second order in $\tilde{\rho}(\mathbf{x})$, $\nabla\rho(\mathbf{x}) = \nabla\tilde{\rho}(\mathbf{x})$, and $\mathbf{j}(\mathbf{x})$, thus obtaining an approximate Hamiltonian

$$H_1 = \frac{1}{8\rho_{\text{av}}} \int d^3x\, [\nabla\tilde{\rho}(\mathbf{x}) - 2i\mathbf{j}(\mathbf{x})] \cdot [\nabla\tilde{\rho}(\mathbf{x}) + 2i\mathbf{j}(\mathbf{x})] + \tfrac{1}{2}\iint d^3x\, d^3y\, \rho_{\text{av}}^2 V(|\mathbf{x}-\mathbf{y}|) + \tfrac{1}{2}\iint d^3x\, d^3y\, \tilde{\rho}(\mathbf{x}) V(|\mathbf{x}-\mathbf{y}|)\tilde{\rho}(\mathbf{y}) - \tfrac{1}{2}NV(0), \tag{20}$$

where the terms linear in $\tilde{\rho}(\mathbf{x})$ in the potential energy have dropped out due to (16). Expressed in terms of Fourier components of the variables, H_1 in (20) becomes

$$H_1' = \frac{1}{8\rho_{\text{av}}} \sum_{\mathbf{k}\neq 0} \Big(\mathbf{k}^2 \rho_{\mathbf{k}}\rho_{-\mathbf{k}} + \sum_{\alpha} (-2k_\alpha j_{\alpha,\mathbf{k}}\rho_{-\mathbf{k}} + 2k_\alpha \rho_{-\mathbf{k}} j_{\alpha,\mathbf{k}} + 4 j_{\alpha,\mathbf{k}} j_{\alpha,-\mathbf{k}})\Big) + \tfrac{1}{2} \sum_{\mathbf{k}\neq 0} V_{\mathbf{k}}\rho_{\mathbf{k}}\rho_{-\mathbf{k}} + \tfrac{1}{2}V_0\rho_{\text{av}}^2\Omega - \tfrac{1}{2}NV(0), \tag{21}$$

when $1/2\rho_{\text{av}}\mathbf{j}_0^2$ is zero or can be neglected (this is so, in particular, for the low-lying states of the system, in which we are chiefly interested here), with V_0 being the $k = 0$ Fourier component of the interaction potential $V(|\mathbf{x}-\mathbf{y}|)$. In writing out (21), we have assumed that $V(|\mathbf{x}|)$ has a Fourier series expansion

$$V(|\mathbf{x}|) = \frac{1}{\Omega} \sum_{\mathbf{k}} V_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} \tag{22}$$

and

$$V_{\mathbf{k}} = \int d^3x V(|\mathbf{x}|) e^{-i\mathbf{k}\cdot\mathbf{x}}, \tag{23}$$

and neglected a kinetic energy term $(\frac{1}{2}/\rho_{\text{av}})\mathbf{j}_0^2$ of small magnitude, $\sim k^2/N$, for large N, $\mathbf{j}_0 = (1/\Omega^{\frac{1}{2}}) \int d^3x \times \mathbf{j}(\mathbf{x})$, for those low-lying excited states with a finite nonzero total linear momentum or total current $\sim \pm k$.

An approximate functional representation for the current density component can be obtained by applying the condition for approximation (18) to the commutation relation between $\rho(\mathbf{x})$ and $\mathbf{j}_\alpha(\mathbf{y})$ in (10). Keeping on the right-hand side only the terms associated with the large density Fourier component ρ_0, one finds, instead of (11),

$$[\rho_{\mathbf{k}}, j_{\alpha,-\mathbf{k}}] = k_\alpha \rho_{\text{av}}. \tag{24}$$

Since $j_{\alpha,-\mathbf{k}\neq 0}$ is of the order of magnitude as $u_s\rho_{-\mathbf{k}\neq 0}$ (u_s is the sound velocity), $\hbar k_\alpha/m$ would have a magnitude about $u_s|\rho_{\mathbf{k}}|^2/\rho_{\text{av}}$. As in very long wavelength density oscillations, k_α may be quite small, and accordingly $|\rho_{\mathbf{k}}|$ would be very small compared with ρ_{av}. We note also that (24) is just one special case of the relation (11), when $\mathbf{l} = -\mathbf{k}$.[13] From (24) one easily obtains an approximate functional representation for $j_{\alpha,-\mathbf{k}}$,

$$j_{\alpha,-\mathbf{k}} = -k_\alpha \rho_{\text{av}} \frac{\delta}{\delta\rho_{\mathbf{k}}},$$

which is a special case of (12) and an approximation to the expression (13). Corresponding to (14), the proper approximate expression is

$$j_{\alpha,-\mathbf{k}} = -k_\alpha\left(\rho_{\text{av}} \frac{\delta}{\delta\rho_{\mathbf{k}}} - \tfrac{1}{2}\rho_{-\mathbf{k}}\right). \tag{25}$$

To use expression (25), in the present approximation, implies that we treat the pairs of variables $\rho_{\mathbf{k}\neq 0}$ and $j_{\alpha,1\neq -\mathbf{k},0}$, and $j_{\alpha,\mathbf{k}\neq 0}$ and $j_{\beta,1\neq 0}$ as commuting variables, as is seen to be true. This leads physically to the independence of each mode of density oscillations associated with a wave vector $\mathbf{k}$.[14] To see whether (25) is consistent with the commutation relations among current density components, one can find an affirmative answer approximately. This will be given in Appendix A. It is noted that (25) gives zero for $j_{\alpha,0}$, which is true for the ground state but not generally right for excited states. Thus we have to use the exact expression (14) for $j_{\alpha,0}$ in general. If the expression (14) were used for all $j_{\alpha,1}$, higher-order terms in $|\rho_{\mathbf{k}\neq 0}|/\rho_{\mathrm{av}}$ with smaller magnitudes would be produced in H_1' (21).

Now by aid of relations (24) and (25), or (25) alone, one can write the approximate Hamiltonian (21) in the following form:

$$H_1' = \sum_{k\neq 0}\left[k^2\left(\tfrac{1}{2}\rho_{\mathbf{k}}\frac{\delta}{\delta\rho_{\mathbf{k}}} - \tfrac{1}{2}\rho_{\mathrm{av}}\frac{\delta}{\delta\rho_{-\mathbf{k}}}\frac{\delta}{\delta\rho_{\mathbf{k}}}\right) + \tfrac{1}{2}V_{\mathbf{k}}\rho_{\mathbf{k}}\rho_{-\mathbf{k}}\right] + \tfrac{1}{2}V_0\rho_{\mathrm{av}}^2\Omega - \tfrac{1}{2}NV(0). \quad (26)$$

In the ρ_k-representation used here, the state vectors of the quantum field for the system will be represented by wavefunctionals of $\rho_{k\neq 0}$ and denoted by $\Psi'\{\rho_{\mathbf{k}}\}$. Then one has a Schrödinger equation of the form

$$H_1'\Psi'\{\rho_{\mathbf{k}}\} = E\Psi'\{\rho_{\mathbf{k}}\}, \quad (27)$$

where E denotes an approximate energy eigenvalue of the system. By inspection of the terms in H_1' (26), it is easy to see that one eigenfunctional is of the Gaussian form usually used for the ground state of a simple harmonic oscillator. Thus we employ as a trial wavefunctional for the ground state here the following:

$$\Psi_0'\{\rho_{\mathbf{k}}\} = A\exp\left(-\sum_{\mathbf{k}\neq 0}\lambda_{\mathbf{k}}\rho_{\mathbf{k}}\rho_{-\mathbf{k}}\right). \quad (28)$$

In (28) A is a normalization constant and $\lambda_{\mathbf{k}}$ denotes an unknown function of the wave vector $\mathbf{k}$, to be determined through Eq. (27). We find, for Ψ_0' to be an eigenfunctional,

$$\begin{aligned} H_1'\Psi_0'\{\rho_{\mathbf{k}}\} &= \left\{\sum_{\mathbf{k}\neq 0}k^2\left[\rho_{\mathrm{av}}\lambda_{\mathbf{k}} - \left(2\rho_{\mathrm{av}}\lambda_{\mathbf{k}}^2 + \lambda_{\mathbf{k}} - \frac{1}{2k^2}V_{\mathbf{k}}\right)\rho_{\mathbf{k}}\rho_{-\mathbf{k}}\right]\right. \\ &\quad \left. + \tfrac{1}{2}V_0\rho_{\mathrm{av}}^2\Omega - \tfrac{1}{2}NV(0)\right\}\Psi_0'\{\rho_{\mathbf{k}}\} \\ &= E_0\Psi_0'\{\rho_{\mathbf{k}}\}, \end{aligned} \quad (29)$$

with

$$E_0 = \sum k^2\rho_{\mathrm{av}}\lambda_{\mathbf{k}} + \tfrac{1}{2}V_0\rho_{\mathrm{av}}^2\Omega - \tfrac{1}{2}NV(0), \quad (30)$$

if the coefficient of $\rho_{\mathbf{k}}\rho_{-\mathbf{k}}\Psi_0'$ vanishes, i.e.,

$$2\lambda_{\mathbf{k}}^2\rho_{\mathrm{av}} + \lambda_{\mathbf{k}} - \frac{1}{2k^2}V_{\mathbf{k}} = 0. \quad (31)$$

This equation (31) gives

$$\lambda_{\mathbf{k}} = \frac{-1 \pm (1 + 4\rho_{\mathrm{av}}V_{\mathbf{k}}/k^2)^{\frac{1}{2}}}{4\rho_{\mathrm{av}}}. \quad (32)$$

We choose

$$\lambda_{\mathbf{k}} = \frac{-1 + (1 + 4\rho_{\mathrm{av}}V_{\mathbf{k}}/k^2)^{\frac{1}{2}}}{4\rho_{\mathrm{av}}} > 0, \quad (33)$$

in order for the wavefunctional $\Psi_0'\{\rho_{\mathbf{k}}\}$ [(28)] to have the meaning of a probability amplitude for each $\rho_{k\neq 0}$ variable. ($1 + 4\rho_{\mathrm{av}}V_k/k^2$ is assumed to be a positive real number for a central repulsive interaction potential here.) With this value for $\lambda_{\mathbf{k}}$, the corresponding energy is, by (30),

$$E_0 = \tfrac{1}{2}V_0\rho_{\mathrm{av}}^2\Omega - \tfrac{1}{2}\sum_{\mathbf{k}\neq 0}[\tfrac{1}{2}k^2 + \rho_{\mathrm{av}}V_{\mathbf{k}} - k(\tfrac{1}{4}k^2 + \rho_{\mathrm{av}}V_{\mathbf{k}})^{\frac{1}{2}}], \quad (34)$$

where we have written $\frac{1}{2}\sum_{k\neq 0}\rho_{\mathrm{av}}V_{\mathbf{k}}$ for $\frac{1}{2}NV(0)$, thereby neglecting a term $\frac{1}{2}\rho_{\mathrm{av}}V_0$, small by a factor $1/N$ compared with the first term. As will be seen later, the energy for any low-lying state of the approximate Hamiltonian H_1' is greater than E_0 in (34) and so is the approximate ground state energy for the system. This result agrees with Bogoliubov's.

We would just like to mention in passing that if one tried to determine the above-mentioned $\lambda_{\mathbf{k}}$ by minimizing E_0,[15] one would find

$$\lambda_{\mathbf{k}}' = \tfrac{1}{2}\rho_{\mathrm{av}}^{-1}(\rho_{\mathrm{av}}V_{\mathbf{k}}/k^2)^{\frac{1}{2}}.$$

This would lead to a higher value for E_0, since, assuming $V_k > 0$ for every k, we have

$$\frac{-1 + (1 + 4\rho_{\mathrm{av}}V_{\mathbf{k}}/k^2)^{\frac{1}{2}}}{4\rho_{\mathrm{av}}} < \frac{1}{2\rho_{\mathrm{av}}}(\rho_{\mathrm{av}}V_{\mathbf{k}}/k^2)^{\frac{1}{2}}.$$

It is to be noted that, in obtaining E_0 in (34), we have never assumed condensation of particles in the free-particle zero-momentum state, nor is it apparent in our treatment.

To get the excitation spectrum formally, one simple way is to use the Heisenberg equation of motion for $\rho(x, t)$. Here, for clarity's sake, we indicate explicitly the t parameter for operators.

One has

$$\dot{\rho}(\mathbf{x}, t) = i[H, \rho(\mathbf{x}, t)] \quad (35)$$

$$= -\mathrm{div}\,\mathbf{j}(\mathbf{x}, t), \quad (36)$$

which is the equation of continuity, obtained through using the equal-time commutation relations (3) and (4).

The H appearing above is the exact Hamiltonian of the system. Since we have been dealing with the approximate Hamiltonian H_1 and thus wish to find an approximate excitation spectrum, we shall replace H by H_1 [Eq. (20)]. Then (35) becomes an approximate relation

$$\dot{\rho}(\mathbf{x}, t) = i[H_1, \rho(\mathbf{x}, t)], \tag{37}$$

which yields the continuity equation

$$\dot{\rho}(\mathbf{x}, t) = -\operatorname{div} \mathbf{j}(\mathbf{x}, t) \tag{38}$$

approximately, if one neglects smaller terms nonlinear in $\mathbf{j}$ and $\tilde{\rho} = \rho - \rho_{\mathrm{av}}$, the density fluctuation. [Note $\dot{\rho}(\mathbf{x}, t) = \dot{\tilde{\rho}}(\mathbf{x}, t)$.] Fourier-analyzing (37) and (38) with respect to $\mathbf{x}$ and assuming that $\rho_{k\neq 0}$ and $j_{\alpha,\mathbf{k}\neq 0}$ have a t dependence like $e^{i\omega t}$, one finds

$$\omega\rho_{\mathbf{k}}(t) = [H_1', \rho_{\mathbf{k}}(t)], \quad \mathbf{k} \neq 0, \tag{39}$$

$$\omega\rho_{\mathbf{k}}(t) = -\sum_{\alpha} k_\alpha j_{\alpha,\mathbf{k}}(t), \tag{40}$$

where ω denotes a frequency and H_1' takes the form (21).[16] Insertion into (39) of the approximate ground state wavefunctional $\Psi_0\{\rho_{\mathbf{k}}(t)\}$ [Eq. (28)], with $\lambda_{\mathbf{k}}$ given by (33), leads to

$$\omega\rho_{\mathbf{k}}\Psi_0\{\rho_{\mathbf{k}}\} = (H_1' - E_0)\rho_{\mathbf{k}}\Psi_0\{\rho_{\mathbf{k}}\}, \tag{41}$$

when H_1' assumes the approximate form (26). So $\rho_{\mathbf{k}}\Psi_0\{\rho_{\mathbf{k}}\}$ is the approximate wavefunctional for an excited state of the system, with the corresponding excitation energy denoted by

$$\epsilon_{\mathbf{k}} = \omega. \tag{42}$$

[It is easy to verify that this excited state has a total linear momentum $= \Omega^{\frac{1}{2}}\langle \mathbf{j}_0 \rangle = -\mathbf{k}$, using (14).] This excitation energy is determined by aid of (40). Inserting $\Psi_0\{\rho_{\mathbf{k}}\}$ also into (40) and using the approximate functional representation for $j_{\alpha,\mathbf{k}\neq 0}$ like (25), we get

$$\begin{aligned}\omega\rho_{\mathbf{k}}\Psi_0\{\rho_{\mathbf{k}}\} &= -\sum_{\alpha} k_\alpha^2\left(\rho_{\mathrm{av}}\frac{\delta}{\delta\rho_{-\mathbf{k}}} - \tfrac{1}{2}\rho_{\mathbf{k}}\right)\Psi_0\{\rho_{\mathbf{k}}\} \\ &= 2k^2(\rho_{\mathrm{av}}\lambda_{\mathbf{k}}\rho_{\mathbf{k}} + \tfrac{1}{4}\rho_{\mathbf{k}})\Psi_0\{\rho_{\mathbf{k}}\}. \end{aligned}\tag{43}$$

Hence the excitation energy is given by

$$\epsilon_{\mathbf{k}} = 2k^2(\rho_{\mathrm{av}}\lambda_{\mathbf{k}} + \tfrac{1}{4}) = k(\tfrac{1}{4}k^2 + \rho_{\mathrm{av}}V_{\mathbf{k}})^{\frac{1}{2}}. \tag{44}$$

Of course, it is also easy to get this excitation spectrum here by noting that $\rho_{\mathbf{k}}\Psi_0\{\rho_{\mathbf{k}}\}$ is an eigenfunctional of the Schrödinger equation (27) with an energy $E_{\mathbf{k}}$, so that

$$E_{\mathbf{k}} = E_0 + \epsilon_{\mathbf{k}} = \frac{\int \rho_{\mathbf{k}}^*\Psi_0^*\{\rho_{\mathbf{k}'}\}H_1'\rho_{\mathbf{k}}\Psi_0\{\rho_{\mathbf{k}'}\}\prod_{\mathbf{k}'\neq 0} d\rho_{\mathbf{k}'}}{\int |\rho_{\mathbf{k}}\Psi_0\{\rho_{\mathbf{k}'}\}|^2\prod_{\mathbf{k}'\neq 0} d\rho_{\mathbf{k}'}}.$$

This $\epsilon_{\mathbf{k}}$ agrees with Bogoliubov's result[3] for the excitation spectrum which he obtained, however, for his quasiparticles. We note also that the spectrum (44) and the set of eigenfunctionals $\Psi_0\{\rho_{\mathbf{k}}\}$, $\rho_{\mathbf{k}}\Psi_0$, $\rho_{\mathbf{k}}\rho_{\mathbf{l}\neq\mathbf{k}}\Psi_0$, etc., for the low states of the system are essentially compatible with the results of Bohm and Salt,[8] using their collective coordinates ($\rho_{\mathbf{k}}\rho_{-\mathbf{k}}\Psi_0$ for any k being not an eigenfunctional).

4. CONNECTION WITH THE QUANTUM HYDRODYNAMIC APPROACH

As is well known, the above density fluctuation excitation spectrum agrees also essentially (only for very small k) with results of the quantum hydrodynamic approach to the irrotational motions of a Bose liquid, as used by Kronig and Thellung[10] and London.[10] If we could take the quantum field Hamiltonian expressed in terms of ρ and $\mathbf{j}$ in (2) to describe the hydrodynamic system of the Bose liquid, as done by Yee,[11] and tried to find the energy spectrum for the irrotational flow within the same kind of approximation, using the commutation rules common to the hydrodynamic methods, i.e.,

$$\begin{aligned}&[\rho(\mathbf{x}), \rho(\mathbf{y})] = 0, \\ &[\phi(\mathbf{x}), \phi(\mathbf{y})] = 0, \\ \text{and}\quad &[\rho(\mathbf{x}), \phi(\mathbf{y})] = -i\delta(\mathbf{x} - \mathbf{y}),\end{aligned}\tag{45}$$

where ϕ is the velocity scalar potential, the results[17] would be identical to ours. The main difference between our spectrum and those of Kronig and Thellung, and London, then comes from the fact that they started with a classical Lagrangian for the liquid while we used a quantum field Hamiltonian, which contains explicitly the two-body interaction potential and some terms of quantum origin. This point has also been noted by Yee. To see more clearly why our present microscopic treatment is equivalent to the quantum hydrodynamic approach to the irrotational flow, we observe two points. The first point is obvious: that both treatments employ the same kind of approximation—that the amplitudes of density oscillations are small compared with the average density. Secondly, the commutation relations used here to derive the above results can be easily shown to be equivalent to those relations (45), provided we assume also the

existence of a velocity potential $\phi(\mathbf{x})$, such that

$$j_\alpha(\mathbf{x}) = -\rho_{\mathrm{av}} \frac{\partial}{\partial x_\alpha} \phi(\mathbf{x}). \tag{46}$$

Let

$$\phi(\mathbf{x}) = \frac{1}{\Omega^{\frac{1}{2}}} \sum_{\mathbf{q}} \phi_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{x}}.$$

So in terms of Fourier components, (46) becomes

$$j_{\alpha,\mathbf{k}} = -ik_\alpha \rho_{\mathrm{av}} \phi_{\mathbf{k}}, \quad \mathbf{k} \neq 0, \tag{47}$$

which actually corresponds to the approximate functional expression for $j_{\alpha,\mathbf{k}}$, i.e.,

$$j_{\alpha,\mathbf{k}} = \left(k_\alpha \rho_{\mathrm{av}} \frac{\delta}{\delta \rho_{-\mathbf{k}}} - \tfrac{1}{2} k_\alpha \rho_{\mathbf{k}} \right). \tag{48}$$

Using Fourier components, we change the relations (45) into

$$\begin{aligned} [\rho_{\mathbf{k}}, \rho_{\mathbf{l}}] &= 0, \\ [\phi_{\mathbf{k}}, \phi_{\mathbf{l}}] &= 0, \\ [\rho_{\mathbf{k}}, \phi_{-\mathbf{k}}] &= -i. \end{aligned} \tag{49}$$

and $\mathbf{k} \neq 0$

As is seen easily, the first relation in (49) holds in both treatments. This is also true for the second relation: As we mentioned before, the use of the approximate functional expression like (48) for $j_{\alpha,\mathbf{k}}$ implies the relation

$$[j_{\alpha,\mathbf{k}\neq 0}, j_{\beta,\mathbf{l}\neq 0}] = 0, \tag{50}$$

which yields by virtue of (47) just the second relation in (49). As to the third relation, we have in our case the commutation relation (24), i.e.,

$$[\rho_{\mathbf{k}\neq 0}, j_{\alpha,-\mathbf{k}}] = k_\alpha \rho_{\mathrm{av}}.$$

Due to (47), this leads to the third relation we want. In accordance with the commutation algebra (49), if we give $\phi_{-\mathbf{k}}$ a functional representation in our $\rho_{\mathbf{k}}$-representation, this will be, after adding a term corresponding to $-\frac{1}{2}k_\alpha \beta_{\mathbf{k}}$ in $j_{\alpha,\mathbf{k}}$,

$$\phi_{-\mathbf{k}} = i\left(\frac{\delta}{\delta \rho_{\mathbf{k}}} - \tfrac{1}{2} \rho_{-\mathbf{k}}/\rho_{\mathrm{av}} \right), \quad \mathbf{k} \neq 0 \tag{51}$$

[although $\delta/\delta\rho_{\mathbf{k}}$ is not well defined (Ref. 12)], so that (47) goes to (48), as previously stated. This also indicates a formal connection between the two methods.

It appears that the above connection cannot be made if general expressions like

$$j_{\alpha,\mathbf{k}} = -\frac{1}{\Omega^{\frac{1}{2}}} \sum_{\mathbf{m}} m_\alpha \rho_{\mathbf{m}+\mathbf{k}} \frac{\delta}{\delta \rho_{\mathbf{m}}} - \tfrac{1}{2} k_\alpha \rho_{\mathbf{k}}$$

are used, for then obviously the relation (50) is no longer valid. However, a formal connection still exists even in such a case if we use Yee's hydrodynamic approach and define the velocity $\mathbf{V}$ by

$$\mathbf{j}(\mathbf{x}) = \tfrac{1}{2}[\rho(\mathbf{x})\mathbf{V}(\mathbf{x}) + \mathbf{V}(\mathbf{x})\rho(\mathbf{x})].^{11}$$

We shall prove the last statement in Appendix B.

No matter whether there exists the above connection or not, these approximate or general functional expressions for $j_{\alpha,\mathbf{k}}$ together with the $\rho_{\mathbf{k}\neq 0}$ functions for $\rho_{\mathbf{k}\neq 0}$ operators in the $\rho_{\mathbf{k}}$-representation can be used to investigate some hydrodynamic motions of a quantum liquid, since they depend only on the basic commutation relations among the density and the current density (3)–(5), which are valid for both microscopic and macroscopic descriptions of the liquid.

5. DISCUSSION

Let us now examine the validity of the present treatment. The following considerations show that our results hold for the low density limit with a short-range repulsive potential. Consider the main approximation involved in the condition (18),

$$\tilde{\rho}(\mathbf{x}) \ll \rho_{\mathrm{av}}$$

for the ground state. That is equivalent to

$$\langle (\tilde{\rho}(\mathbf{x}))^2 \rangle \ll \rho_{\mathrm{av}}^2, \tag{52}$$

where the average $\langle\ \rangle$ is taken over the approximate ground state wavefunctional $\Psi_0\{\rho_{\mathbf{k}}\}$ [Eq. (28)]. Equation (52) may be expressed as

$$\frac{1}{\Omega} \sum_{\mathbf{k},\mathbf{l}\neq 0} \langle \rho_{\mathbf{k}} \rho_{\mathbf{l}} e^{i(\mathbf{k}+\mathbf{l})\cdot\mathbf{x}} \rangle = \frac{1}{\Omega} \sum_{\mathbf{k}\neq 0} \langle \rho_{\mathbf{k}} \rho_{-\mathbf{k}} \rangle \ll \rho_{\mathrm{av}}^2 \tag{53}$$

since

$$\langle \rho_{\mathbf{k}} \rho_{\mathbf{l}} \rangle = \delta_{-\mathbf{k},\mathbf{l}} \langle \rho_{\mathbf{k}} \rho_{-\mathbf{k}} \rangle.^{15}$$

Changing the summation into integration by $\sum_{\mathbf{k}} \to [\Omega/(2\pi)^3] \int d^3k$, we have ($\epsilon$ being an arbitrarily small vector)

$$\frac{1}{(2\pi)^3} \int_{\epsilon\to 0} d^3k \langle \rho_{\mathbf{k}} \rho_{-\mathbf{k}} \rangle = \frac{1}{2\pi^2} \int_{\epsilon\to 0} \frac{k^2\, dk}{4\lambda_{\mathbf{k}}} \ll \rho_{\mathrm{av}}^2 \tag{54}$$

with (33) for $\lambda_{\mathbf{k}}$,

$$\lambda_{\mathbf{k}} = \frac{1}{2\rho_{\mathrm{av}} k} \left[\left(\frac{k^2}{4} + \rho_{\mathrm{av}} V_{\mathbf{k}} \right)^{\frac{1}{2}} - \frac{k}{2} \right].$$

We observe that physically there exists a minimum wavelength of the density oscillations with a magnitude about the average interparticle distance, r_0 ($\sim \rho_{\mathrm{av}}^{-\frac{1}{3}}$), so we shall take a cutoff k_c for the upper integration limit. The actual value of k_c will be estimated from the above inequality. We shall assume that, for those small wave vectors $\mathbf{k} \leq \mathbf{k}_c$ of interest here, the Fourier

component of the interaction potential $V_{\mathbf{k}}$ may be replaced by V_0, the constant $\mathbf{k} = 0$ Fourier component. As usual, this V_0 may be related to a scattering amplitude a by

$$V_0 = 4\pi a\hbar^2/m, \tag{55}$$

where m is understood to be the mass of a Bose particle. With the factors in $\hbar$ and m inserted and (55) substituted into $\lambda_{\mathbf{k}}$, one finds

$$\lambda_{\mathbf{k}} = \frac{1}{2\rho_{\mathrm{av}}k}\left[\left(\frac{k^2}{4} + 4\pi a\rho_{\mathrm{av}}\right)^{\frac{1}{2}} - \frac{k}{2}\right].$$

Then (54) becomes

$$\frac{1}{2\rho_{\mathrm{av}}\pi^2}\int_0^{k_c} \frac{k^3\,dk}{-k + 2(\frac{1}{4}k^2 + 4\pi a\rho_{\mathrm{av}})^{\frac{1}{2}}} \ll 1.$$

For our purpose here, it seems legitimate to consider, instead, a simplified integral

$$\frac{1}{2\pi^2}\rho_{\mathrm{av}}^{-1}\int_0^{k_c} \frac{k^3\,dk}{(k^2 + 16\pi a\rho_{\mathrm{av}})^{\frac{1}{2}} - k_c} \ll 1,$$

or just

$$\frac{1}{2\pi^2}\rho_{\mathrm{av}}^{-1}\int_0^{k_c} \frac{k^3\,dk}{(k^2 + 16\pi a\rho_{\mathrm{av}})^{\frac{1}{2}} - 4(\pi a\rho_{\mathrm{av}})^{\frac{1}{2}}} \ll 1, \tag{56}$$

if

$$k_c^2 \sim 16\pi a\rho_{\mathrm{av}} \quad \text{or} \quad r = 16\pi a\rho_{\mathrm{av}}/k_c^2 \sim 1,$$

and, if $r \ll 1$,

$$\frac{1}{2\pi^2}\rho_{\mathrm{av}}^{-1}\left(\int_0^{4(\pi a\rho_{\mathrm{av}})^{\frac{1}{2}}} \frac{k^3\,dk}{(k^2 + 16\pi a\rho_{\mathrm{av}})^{\frac{1}{2}} - 4(\pi a\rho_{\mathrm{av}})^{\frac{1}{2}}} + \int_{4(\pi a\rho_{\mathrm{av}})^{\frac{1}{2}}}^{k_c} \frac{k^4\,dk}{8\pi a\rho_{\mathrm{av}}}\right) \ll 1. \tag{57}$$

In the first case, $r \sim 1$, (56) can be written as

$$\frac{32}{\pi^{\frac{1}{2}}}(a^3\rho_{\mathrm{av}})^{\frac{1}{2}}\left[\frac{2 + 3\cos\Phi_c}{6\cos^3\Phi_c} - \frac{5}{6}\right] \ll 1,$$

where

$$\cos\Phi_c = r^{\frac{1}{2}}/(1 + r). \tag{58}$$

(58) is easily seen to hold if $(a^3\rho_{\mathrm{av}})^{\frac{1}{2}} \ll 1$, which is just a usual condition for the low density limit with a short-range interaction potential.[4,18] In the second case, $r \ll 1$, (57) becomes after integration

$$\frac{32}{\pi^{\frac{1}{2}}}(a^3\rho_{\mathrm{av}})^{\frac{1}{2}}\left[\frac{2\sqrt{2}}{3} - \frac{1}{6}\right] + \frac{k_c^5}{80\pi^3 a\rho_{\mathrm{av}}^2}[1 - r^{\frac{5}{2}}] \ll 1. \tag{59}$$

Let us consider the second term first. It would be $\ll 1$, if

$$(16\pi a\rho_{\mathrm{av}})^{\frac{5}{2}} \ll k_c^5 \ll 80\pi^3 a\rho_{\mathrm{av}}^2,$$

or

$$(a\rho_{\mathrm{av}})^{\frac{5}{2}} \ll k_c^5/10^4 \ll a\rho_{\mathrm{av}}^2. \tag{60}$$

This latter inequality expression implies also the condition $(a^3\rho_{\mathrm{av}})^{\frac{1}{2}} \ll 1$, which makes the first term in (59) again small compared with unity. Thus with (60) fullfilled, our main approximation would be valid. [E.g., one could choose $k_c^5 = 10^2 a\rho_{\mathrm{av}}^2$ with $(a^3\rho_{\mathrm{av}})^{\frac{1}{2}} \ll 10^{-2}$.]

Under the condition $(a^3\rho_{\mathrm{av}})^{\frac{1}{2}} \ll 10^{-2}$ with the chosen cutoff value for k_c, one can express the approximate ground state energy in (34) as a series in $(a^3\rho_{\mathrm{av}})^{\frac{1}{2}}$. However, in order to show which terms in the series do not depend on the chosen value of k_c, we expand the terms as a power series in the parameter r, leaving k_c as if not fixed. We have from (34)

$$E_0/\Omega = (\hbar^2/2m) \times \left\{4\pi a\rho_{\mathrm{av}}^2 - \frac{1}{(2\pi)^3}\int_0^{\mathbf{k}_c} d^3k\left[\frac{k^2}{2} + 4\pi a\rho_{\mathrm{av}} - \left(\frac{k^2}{4} + 4\pi a\rho_{\mathrm{av}}k^2\right)^{\frac{1}{2}}\right]\right\}. \tag{61}$$

After simple integrations, the second and the third terms combine to give

$$(-k_c^5/2\pi^2)(1/10 + r/12). \tag{62}$$

The last term in (61) can also be easily integrated out to be

$$\frac{1}{2\pi^2} \times 512(\pi a\rho_{\mathrm{av}})^{\frac{5}{2}}\left[\frac{3 - 5\cos^2\Phi_c}{15\cos^5\Phi_c} + \frac{2}{15}\right], \tag{63}$$

where $\cos\Phi_c$ has been given in (58). If we expand (63) with respect to r and keep terms up to the $\frac{5}{2}$th power, we find as an approximation to (63)

$$+k_c^5/2\pi^2(1/10 + r/12 - r^2/16 + r^{\frac{5}{2}}/15\cdots). \tag{64}$$

Substituting (62) and (64) into (61) and replacing r by $16\pi a\rho_{\mathrm{av}}/k_c^2$, one obtains

$$\frac{E_0}{\Omega} = \frac{2\pi\hbar^2}{m}a\rho_{\mathrm{av}}^2 \times \left[1 - \frac{2}{\pi}ak_c + \frac{128}{15\pi^{\frac{1}{2}}}(a^3\rho_{\mathrm{av}})^{\frac{1}{2}} - O\left(\frac{a^2\rho_{\mathrm{av}}}{k_c}\right)\right]. \tag{65}$$

We have noticed from the above that the cancellations of the second and the third terms by part of the last integral in (61), as well as the first and the third terms in (65), do not depend on the chosen value of k_c and thus that they are reliable. However, the second term and the remaining terms in the parentheses of (65), being of order $(a^3\rho_{\mathrm{av}})^{\frac{2}{5}}$ and at least

$O[(a^3\rho_{av})^{\frac{3}{5}}]$ respectively, do depend on the chosen k_c value, so that the coefficients associated with them are inaccurate. Actually this second term is reminiscent of the divergent term when $k_c \to \infty$ as discussed by Lee, Huang, and Yang,[4] and has to be dropped out for a correct treatment of the interaction potential. As to the inaccuracy of the coefficients just mentioned, there is another source here, i.e., that we have not used the general functional expression for $j_{\alpha,-\mathbf{k}}$, which would give some corrections. These corrections will be worked out and published elsewhere. In the present treatment, we can not get the logarithmic term of $(a^3\rho_{av})$, as calculated by Hugenholtz and Pines [7] and by Wu.[19] As far as the reliable terms in (65) are concerned, they are in agreement with those of Lee and Yang[20] and Ref. 4.

It is to be emphasized that our main approximation $\langle[\tilde{\rho}(\mathbf{x})]^2\rangle \ll \rho_{av}^2$ would break down for $k \gg (10^2 \times a\rho_{av}^2)^{\frac{1}{5}}$ as previously estimated.[21] If the k_c value could serve as an order-of-magnitude mark where the collective oscillations end, as in the case with $k_c^2 \sim 16\pi a\rho_{av}$, then, for $k \gg k_c$, $\rho_{\mathbf{k}}$ refers essentially to the individual particle behaviors, and the present approximation is not suitable for studying those with $k^5 \gg 10^2 \times a\rho_{av}^2$. However such a meaning for k_c is inapplicable to the case where $16\pi a\rho_{av} \ll k_c^2$; for those k values such that $16\pi a\rho_{av} \ll k^2 \leq k_c^2$, the density excitation spectrum is already particlelike. It is not proper then to talk about collective oscillations.

Finally, the present method of employing the functional representations of $\rho_{\mathbf{k}\neq 0}$ and $j_{\alpha,\mathbf{k}}$ in the $\rho_{\mathbf{k}}$-representation is useful for both Bose and Fermi systems since the same set of commutation rules for density and current density components exist. So it can be applied to study interacting Fermi systems, if we extend it to take care of the spin and Fermi statistics of the particles, as done, e.g., by Grodnik and Sharp[22] recently. This application will be made later on.

ACKNOWLEDGMENTS

I am deeply indebted to Professor D. M. Dennison for his continuous interest and encouragement throughout this work. It is a great pleasure to thank Professor O. Laporte for helpful discussions and his interest; also Professor H. R. Crane, Chairman, Professor K. T. Hecht, Professor G. W. Ford, and other staff members of the Physics Department for their kind hospitality.

APPENDIX A

We have wanted to see whether the approximate functional representation, $j_{\alpha,-\mathbf{k}\neq 0} = -k_\alpha(\rho_{av}(\delta/\delta\rho_{\mathbf{k}}) - \frac{1}{2}\rho_{-\mathbf{k}})$ [Eq. (25)], is approximately consistent with the commutation relation among the current density components, i.e., Eq. (5),

$$[j_\alpha(\mathbf{x}), j_\beta(\mathbf{y})] = -ij_\beta(\mathbf{x})\left(\frac{\partial}{\partial x_\alpha}\delta(\mathbf{x}-\mathbf{y})\right) + ij_\alpha(\mathbf{y})\left(\frac{\partial}{\partial y_\beta}\delta(\mathbf{x}-\mathbf{y})\right)$$

or its Fourier transform

$$[j_{\alpha,\mathbf{k}}, j_{\beta,\mathbf{l}}] = \frac{1}{\Omega^{\frac{1}{2}}}(-l_\alpha j_{\beta,\mathbf{k}+\mathbf{l}} + k_\beta j_{\alpha,\mathbf{k}+\mathbf{l}}). \quad \text{(A1)}$$

Applying (25) to the left-hand side of (A1), one easily finds zero for any nonzero $\mathbf{k}$ and $\mathbf{l}$ values. But the direct use of (25) for the right-hand side will give nonzero generally if $\mathbf{k} \neq \pm\mathbf{l}$, i.e.,

$$\frac{1}{\Omega^{\frac{1}{2}}}(k_\alpha k_\beta - l_\alpha l_\beta)\left(\rho_{av}\frac{\delta}{\delta\rho_{-\mathbf{k}-\mathbf{l}}} - \tfrac{1}{2}\rho_{\mathbf{k}+\mathbf{l}}\right). \quad \text{(A2)}$$

So there seems to be an inconsistency with the commutation relation (A1). However, those nonzero terms will contribute small magnitudes when operating on a state functional and may be neglected through the following considerations. Let us find out how (A2) comes about by employing the general functional expression (14) for the current density components in the commutator $[j_{\alpha,\mathbf{k}}, j_{\beta,\mathbf{l}}]$. We find, by separating the large terms from the small terms in $j_{\alpha,\mathbf{k}}$ and $j_{\beta,\mathbf{l}}$, for $\mathbf{k} \neq \pm\mathbf{l}$,

$$\begin{aligned} j_{\alpha,\mathbf{k}}j_{\beta,\mathbf{l}} = {} & \left[\left(k_\alpha\rho_{av}\frac{\delta}{\delta\rho_{-\mathbf{k}}} - \tfrac{1}{2}k_\alpha\rho_{\mathbf{k}}\right) - \frac{1}{\Omega^{\frac{1}{2}}}\sum_{\mathbf{m}\neq-\mathbf{k}} m_\alpha\rho_{\mathbf{m}+\mathbf{k}}\frac{\delta}{\delta\rho_{\mathbf{m}}}\right] \\ & \times\left[\left(l_\beta\rho_{av}\frac{\delta}{\delta\rho_{-\mathbf{l}}} - \tfrac{1}{2}l_\beta\rho_{\mathbf{l}}\right) - \frac{1}{\Omega^{\frac{1}{2}}}\sum_{\mathbf{n}\neq-\mathbf{l}} n_\beta\rho_{\mathbf{n}+\mathbf{l}}\frac{\delta}{\delta\rho_{\mathbf{n}}}\right] \quad \text{(A3)} \\ = {} & k_\alpha l_\beta\rho_{av}^2\frac{\delta}{\delta\rho_{-\mathbf{k}}}\frac{\delta}{\delta\rho_{-\mathbf{l}}} - \tfrac{1}{2}k_\alpha l_\beta\rho_{av}\rho_{\mathbf{l}}\frac{\delta}{\delta\rho_{-\mathbf{k}}} \\ & - \frac{1}{\Omega^{\frac{1}{2}}}k_\alpha\rho_{av}\sum_{\mathbf{n}\neq-\mathbf{l}} n_\beta \\ & \times\left(\delta_{-\mathbf{k},\mathbf{n}+\mathbf{l}}\frac{\delta}{\delta\rho_{\mathbf{n}}} + \rho_{\mathbf{n}+\mathbf{l}}\frac{\delta}{\delta\rho_{-\mathbf{k}}}\frac{\delta}{\delta\rho_{\mathbf{n}}}\right) \\ & - \tfrac{1}{2}k_\alpha l_\beta\rho_{av}\rho_{\mathbf{k}}\frac{\delta}{\delta\rho_{-\mathbf{l}}} + \tfrac{1}{4}k_\alpha l_\beta\rho_{\mathbf{k}}\rho_{\mathbf{l}} \\ & + \frac{1}{2\Omega^{\frac{1}{2}}}k_\alpha\rho_{\mathbf{k}}\sum_{\mathbf{n}\neq-\mathbf{l}} n_\beta\rho_{\mathbf{n}+\mathbf{l}}\frac{\delta}{\delta\rho_{\mathbf{n}}} \\ & - \frac{1}{\Omega^{\frac{1}{2}}}l_\beta\rho_{av}\sum_{\mathbf{m}\neq-\mathbf{k}} m_\alpha\rho_{\mathbf{m}+\mathbf{k}}\frac{\delta}{\delta\rho_{\mathbf{m}}}\frac{\delta}{\delta\rho_{-\mathbf{l}}} \\ & + \frac{1}{2\Omega^{\frac{1}{2}}}\sum_{\mathbf{m}\neq-\mathbf{k}} m_\alpha\rho_{\mathbf{m}+\mathbf{k}}\left((\delta_{\mathbf{m},\mathbf{l}}l_\beta) + l_\beta\rho_{\mathbf{l}}\frac{\delta}{\delta\rho_{\mathbf{m}}}\right) \\ & + \frac{1}{\Omega^{\frac{1}{2}}}\sum_{\mathbf{m}\neq-\mathbf{k}} m_\alpha\rho_{\mathbf{m}+\mathbf{k}}\frac{\delta}{\delta\rho_{\mathbf{m}}}\left(\sum_{\mathbf{n}\neq-\mathbf{l}} n_\beta\rho_{\mathbf{n}+\mathbf{l}}\frac{\delta}{\delta\rho_{\mathbf{n}}}\right). \quad \text{(A4)} \end{aligned}$$

It is noted here that the third and the eighth terms in (A4) give

$$\frac{1}{\Omega^{\frac{1}{2}}}\left(k_\alpha(k_\beta + l_\beta)\rho_{\mathrm{av}}\frac{\delta}{\delta\rho_{-\mathbf{k}-\mathbf{l}}} + \tfrac{1}{2}l_\alpha l_\beta\rho_{\mathbf{k}+\mathbf{l}}\right). \quad \text{(A5)}$$

Similarly,

$$\begin{aligned} j_{\beta,\mathbf{l}}j_{\alpha,\mathbf{k}} = {} & l_\beta k_\alpha\rho_{\mathrm{av}}^2\frac{\delta}{\delta\rho_{-\mathbf{l}}}\frac{\delta}{\delta\rho_{-\mathbf{k}}} - \tfrac{1}{2}l_\beta k_\alpha\rho_{\mathrm{av}}\rho_{\mathbf{k}}\frac{\delta}{\delta\rho_{-\mathbf{l}}} \\ & - \frac{1}{\Omega^{\frac{1}{2}}}l_\beta\rho_{\mathrm{av}}\sum_{\mathbf{m}\neq-\mathbf{k}}m_\alpha \\ & \times\left(\delta_{-\mathbf{l},\mathbf{m}+\mathbf{k}}\frac{\delta}{\delta\rho_{\mathbf{m}}} + \rho_{\mathbf{m}+\mathbf{k}}\frac{\delta}{\delta\rho_{-\mathbf{l}}}\frac{\delta}{\delta\rho_{\mathbf{m}}}\right) \\ & - \tfrac{1}{2}l_\beta k_\alpha\rho_{\mathrm{av}}\rho_{\mathbf{l}}\frac{\delta}{\delta\rho_{-\mathbf{k}}} + \tfrac{1}{4}l_\beta k_\alpha\rho_{\mathbf{l}}\rho_{\mathbf{k}} \\ & + \frac{1}{2\Omega^{\frac{1}{2}}}l_\beta\rho_{\mathbf{l}}\sum_{\mathbf{m}\neq-\mathbf{k}}m_\alpha\rho_{\mathbf{m}+\mathbf{k}}\frac{\delta}{\delta\rho_{\mathbf{m}}} \\ & - \frac{1}{\Omega^{\frac{1}{2}}}k_\alpha\rho_{\mathrm{av}}\sum_{\mathbf{n}\neq-\mathbf{l}}n_\beta\rho_{\mathbf{n}+\mathbf{l}}\frac{\delta}{\delta\rho_{\mathbf{n}}}\frac{\delta}{\delta\rho_{-\mathbf{k}}} \\ & + \frac{1}{2\Omega^{\frac{1}{2}}}\sum_{\mathbf{n}\neq-\mathbf{l}}n_\beta\rho_{\mathbf{n}+\mathbf{l}}\left(\delta_{\mathbf{n},\mathbf{k}}k_\alpha + k_\alpha\rho_{\mathbf{k}}\frac{\delta}{\delta\rho_{\mathbf{n}}}\right) \\ & + \frac{1}{\Omega^{\frac{1}{2}}}\sum_{\mathbf{n}\neq-\mathbf{l}}n_\beta\rho_{\mathbf{n}+\mathbf{l}}\frac{\delta}{\delta\rho_{\mathbf{n}}}\left(\sum_{\mathbf{m}\neq-\mathbf{k}}m_\alpha\rho_{\mathbf{m}+\mathbf{k}}\frac{\delta}{\delta\rho_{\mathbf{m}}}\right). \end{aligned} \quad \text{(A6)}$$

Again the third and the eighth terms in (A6) are of interest, and give

$$\frac{1}{\Omega^{\frac{1}{2}}}\left(l_\beta(l_\alpha + k_\alpha)\rho_{\mathrm{av}}\frac{\delta}{\delta\rho_{-\mathbf{k}-\mathbf{l}}} + \tfrac{1}{2}k_\alpha k_\beta\rho_{\mathbf{k}+\mathbf{l}}\right). \quad \text{(A7)}$$

We see then that it is the difference (A5) − (A7) which leads to the nonzero terms in (A2). Since these terms (A5) and (A7) are both small compared with the leading terms in the product

$$\left(k_\alpha\rho_{\mathrm{av}}\frac{\delta}{\delta\rho_{-\mathbf{k}}} - \tfrac{1}{2}k_\alpha\rho_{\mathbf{k}}\right)\left(\mathbf{l}_\beta\rho_{\mathrm{av}}\frac{\delta}{\delta\rho_{-\mathbf{l}}} - \tfrac{1}{2}\mathbf{l}_\beta\rho_{\mathbf{l}}\right),$$

as seen clear from (A3), and since we have implicitly neglected them when we apply the approximate expression (25) to $[j_{\alpha,\mathbf{k}}, j_{\beta,\mathbf{l}}]$, it is reasonable and also consistent to neglect (A2). One may view this neglection in the same sense of approximation as Bogoliubov took $[a_0, a_0^+] = 0$ for bosons[3] near $T = 0\,^\circ\mathrm{K}$, where a_0 and a_0^+ denote, respectively, destruction and creation operators for the free-particle zero-momentum state. Thus within this approximation the approximate expression (25) leads to $[j_{\alpha,\mathbf{k}}, j_{\beta,\mathbf{l}}] = 0$, which is consistent with the commutation relation (A1).

APPENDIX B

As was shown by Yee, the commutation relation (4) may lead to

$$[\rho(\mathbf{x}), V_\alpha(\mathbf{y})] = i\left(\frac{\partial}{\partial y_\alpha}\delta(\mathbf{x}-\mathbf{y})\right). \quad \text{(B1)}$$

In terms of Fourier components, (B1) becomes

$$[\rho_{\mathbf{k}}, V_{\alpha,-\mathbf{k}}] = k_\alpha \quad \text{(B2)}$$

if one assumes

$$V_\alpha(\mathbf{y}) = \frac{1}{\Omega^{\frac{1}{2}}}\sum_{\mathbf{k}}V_{\alpha,\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{y}},$$

and the relation defining $\mathbf{V}$ yields

$$j_{\alpha,\mathbf{k}} = \frac{1}{2\Omega^{\frac{1}{2}}}\sum_{\mathbf{q}}[\rho_{\mathbf{k}+\mathbf{q}}V_{\alpha,-\mathbf{q}} + V_{\alpha,-\mathbf{q}}\rho_{\mathbf{k}+\mathbf{q}}]. \quad \text{(B3)}$$

For any $\mathbf{k}$, we get, by aid of (B2),

$$j_{\alpha,\mathbf{k}} = \frac{1}{\Omega^{\frac{1}{2}}}\sum_{\mathbf{q}}\rho_{\mathbf{k}+\mathbf{q}}V_{\alpha,-\mathbf{q}}, \quad \text{(B4)}$$

but we also have here

$$j_{\alpha,\mathbf{k}} = -\frac{1}{\Omega^{\frac{1}{2}}}\sum_{\mathbf{m}}m_\alpha\rho_{\mathbf{m}+\mathbf{k}}\frac{\delta}{\delta\rho_{\mathbf{m}}} - \tfrac{1}{2}k_\alpha\rho_{\mathbf{k}}. \quad \text{(B5)}$$

To make these two expressions identical, one can give $V_{\alpha,-\mathbf{q}}$ a formal functional expression in our $\rho_{\mathbf{k}}$-representation as (assuming the zeroth Fourier component of the velocity, $\mathbf{V}_0 = 0$)

$$V_{\alpha,-\mathbf{q}} = -q_\alpha\frac{\delta}{\delta\rho_{\mathbf{q}}} - \frac{1}{2i\Omega^{\frac{1}{2}}}\int d^3x\left(\rho^{-1}(\mathbf{x})\frac{\partial}{\partial x_\alpha}\rho(\mathbf{x})\right)e^{i\mathbf{q}\cdot\mathbf{x}}, \quad q \neq 0, \quad \text{(B6)}$$

with $\rho(\mathbf{x})$ expressed in terms of $\rho_{\mathbf{k}}$, although it involves a not well-defined operator $\delta/\delta\rho_{\mathbf{q}}$ and a singular function $1/\rho(\mathbf{x})$. It can be easily verified that substitution of (B6) into (B4) will give (B5). It is also evident that $V_{\alpha,-\mathbf{q}}$ given by (B6) is consistent with the relation (B2) and the commutation relation among the current density components, since the latter is satisfied by the general expression for $j_{\alpha,\mathbf{k}}$. (B6) implies the existence of a formal velocity scalar potential operator ϕ, for

$$q_\beta V_{\alpha,-\mathbf{q}} - q_\alpha V_{\beta,-\mathbf{q}} = 0. \quad \text{(B7)}$$

To show the truth of (B7), clearly we need only to examine the second term of $V_{\alpha,-\mathbf{q}}$ in (B6). On integrating by parts, this term can be changed to

$$\begin{aligned} & + \frac{1}{2i\Omega^{\frac{1}{2}}}\int d^3x\log\rho(\mathbf{x})\frac{\partial}{\partial x_\alpha}e^{i\mathbf{q}\cdot\mathbf{x}} \\ & = \frac{q_\alpha}{2\Omega^{\frac{1}{2}}}\int d^3x\log\rho(\mathbf{x})e^{i\mathbf{q}\cdot\mathbf{x}}, \end{aligned} \quad \text{(B8)}$$

as the other integral vanishes due to the periodic boundary conditions used. Thus terms like (B8) will satisfy (B7). One may then define $\phi_{\mathbf{q}}$ formally by

$$V_{\alpha,-\mathbf{q}} = iq_\alpha \phi_{-\mathbf{q}}, \quad \mathbf{q} \neq 0, \tag{B9}$$

i.e.,

$$\phi_{-\mathbf{q}} = i\frac{\delta}{\delta\rho_{\mathbf{q}}} + \frac{1}{2\Omega^{\frac{1}{2}}q_\alpha}\int d^3x\left(\rho^{-1}(\mathbf{x})\frac{\partial}{\partial x_\alpha}\rho(\mathbf{x})\right)e^{i\mathbf{q}\cdot\mathbf{x}} \tag{B10}$$

or

$$= i\left(\frac{\delta}{\delta\rho_{\mathbf{q}}} - \frac{1}{2\Omega^{\frac{1}{2}}}\int d^3x \log\rho(\mathbf{x})e^{i\mathbf{q}\cdot\mathbf{x}}\right). \tag{B11}$$

[When $1/\rho(\mathbf{x})$ is approximated by $1/\rho_{\mathrm{av}}$, (B10) becomes $\phi_{-\mathbf{q}} = i(\delta/\delta\rho_{\mathbf{q}} - \frac{1}{2}\rho_{-\mathbf{q}}/\rho_{\mathrm{av}})$ as given by (51) in the text.]

Substituting (B9) into (B2), one finds

$$[\rho_{\mathbf{k}}, \phi_{-\mathbf{k}}] = -i, \quad \mathbf{k} \neq 0. \tag{B12}$$

(B10) or (B11) leads to

$$(\mathbf{k}, \mathbf{l} \neq 0) \quad [\phi_{\mathbf{k}}, \phi_{\mathbf{l}}] = 0. \tag{B13}$$

It is not difficult to check the consistency of (B13) with the commutation relation among $j_{\alpha,\mathbf{k}}$ and $j_{\beta,\mathbf{l}}$, i.e., (A1).

We see that (B12) and (B13) are just two of the relations in (49) in the text, which need to be proved.

* Visiting Scholar.

[1] L. D. Landau, J. Phys. (USSR) **5**, 71 (1941); **11**, 91 (1941).

[2] R. P. Feynman, Phys. Rev. **91**, 1291 (1953); **94**, 262 (1954). Experimentally, see D. G. Henshaw and A. D. B. Woods, Phys. Rev. **121**, 1266 (1961).

[3] N. N. Bogoliubov, J. Phys. (USSR) **11**, 23 (1947).

[4] T. D. Lee, K. Huang, and C. N. Yang, Phys. Rev. **106**, 1135 (1957).

[5] K. A. Brueckner and K. Sawada, Phys. Rev. **106**, 1117 (1957).

[6] S. T. Beliaev, Zh. Eksp. Teor. Fiz. **34**, 417 (1958) [Sov. Phys. JETP **7**, 289 (1958)].

[7] N. M. Hugenholtz and D. Pines, Phys. Rev. **116**, 489 (1959).

[8] D. Bohm and B. Salt, Rev. Mod. Phys. **39**, 894 (1967).

[9] R. F. Dashen and D. H. Sharp, Phys. Rev. **165**, 1857 (1968); see also D. H. Sharp, *ibid.* 1867 (1968).

[10] R. Kronig and A. Thellung, Physica **18**, 749 (1952). See also F. London, *Superfluids* (Dover, New York, 1964), Vol. II, pp. 114–18.

[11] D. D. H. Yee, Phys. Rev. **184**, 196 (1969).

[12] Only after reading with much appreciation the referee's comments about the present work, did the author realize that such a functional representation has been independently obtained by J. Grodnik and D. H. Sharp, and already appeared in their paper, Phys. Rev. D **1**, 1531 (1970). While it provides a somewhat different and more formal treatment, that paper supports and amplifies our results in this section except our original functional representation for $j_{\alpha,1}$, which did not contain the term $\frac{1}{2}l_\alpha\rho_{\mathbf{l}}$. As was expected by the referee and will be seen in later sections, inclusion of this term does not affect appreciably our original calculations obtained without it in the lowest approximation; nor does it affect the formal connection of our method with the quantum hydrodynamic approach to the irrotational flow of the Bose liquid.

[13] So the approximation (18) does not affect the validity of (24); it just enables us to ignore the other cases of the general relation (12).

[14] See the Heisenberg equation of motion for $\rho_{\mathbf{k}}(t)$, as put in the form (43), Sec. 3.

[15] The minimization of E_0 with respect to $\lambda_{\mathbf{k}}$ involves calculating the average of $\rho_{\mathbf{k}}\rho_{-\mathbf{k}}$ over the trial wavefunctional (28). This average $\langle\rho_{\mathbf{k}}\rho_{-\mathbf{k}}\rangle$ can be easily calculated, as done by D. Bohm and D. Pines, Phys. Rev. **85**, 338 (1952), through introducing a pair of real variables $r_{\mathbf{k}}$ and $\theta_{\mathbf{k}}$ to replace the pair of complex variables $\rho_{\mathbf{k}}$ and $\rho_{-\mathbf{k}}$ with the relations $\rho_{\mathbf{k}} = r_{\mathbf{k}}e^{i\theta_{\mathbf{k}}}$, $\rho_{-\mathbf{k}} = r_{\mathbf{k}}e^{-i\theta_{\mathbf{k}}}$. Thus, noting $r^2_{-\mathbf{k}} = r^2_{\mathbf{k}}$, we have

$$\langle\rho_{\mathbf{k}}\rho_{-\mathbf{k}}\rangle = \frac{\iint r_{\mathbf{k}}^2\exp(-4\lambda_{\mathbf{k}}r_{\mathbf{k}}^2)|J|\,dr_{\mathbf{k}}\,d\theta_{\mathbf{k}}}{\iint\exp(-4\lambda_{\mathbf{k}}r_{\mathbf{k}}^2)|J|\,dr_{\mathbf{k}}\,d\theta_{\mathbf{k}}}.$$

$|J|$ here denotes the absolute value of the Jacobian for the transformation from the pair $\rho_{\mathbf{k}}$ and $\rho_{-\mathbf{k}}$ to the new variables, and is $2r_{\mathbf{k}}$. In this way, one finds $\langle\rho_{\mathbf{k}}\rho_{-\mathbf{k}}\rangle = \dfrac{1}{4\lambda_{\mathbf{k}}}$. Similar kinds of calculations for $\langle\rho_{\mathbf{k}}^n\rho_{-\mathbf{k}}^n\rangle$ (n = any positive integer) will be of frequent use.

[16] The relation (40) can also be directly obtained from (39) by aid of the two commutation relations $[\rho_{\mathbf{k}}, \rho_{\mathbf{l}}] = 0$ and $[\rho_{\mathbf{k}}, j_{\alpha,\mathbf{l}\neq 0}] = k_\alpha\rho_{\mathrm{av}}\delta_{-\mathbf{k},\mathbf{l}}$. The latter relation is only approximately valid as pointed out previously in Sec. 3. It is connected with using the approximate functional representation for $j_{\alpha,-\mathbf{k}\neq 0}$, (25).

[17] Although, in the hydrodynamic case, the two-body interaction potential has to be interpreted as the average potential among particles, as mentioned by E. P. Gross. See J. Math. Phys. **4**, 195 (1963), where the author used the conventional quantum field Hamiltonian in terms of ψ and ψ^+ to discuss the hydrodynamics of a Bose liquid.

[18] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, New York, 1963), p. 229.

[19] T. T. Wu, Phys. Rev. **115**, 1390 (1959).

[20] T. D. Lee and C. N. Yang, Phys. Rev. **105**, 1119 (1957).

[21] Although the range of validity of our approximation was considered before only for the ground state, it holds also for the low-lying excited states when N is large.

[22] J. Grodnik and D. H. Sharp, Phys. Rev. D **1**, 1546 (1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Lattice Green's Functions for the Cubic Lattices in Terms of the Complete Elliptic Integral

TOHRU MORITA*
Department of Applied Science, Faculty of Engineering, Tohoku University, Sendai, Japan
AND
TSUYOSHI HORIGUCHI*
Department of Physics, Faculty of Science, Tohoku University, Sendai, Japan

(Received 20 July 1970)

The real and imaginary parts of the lattice Green's functions for the simple cubic (actually the tetragonal), body-centered cubic, and face-centered cubic lattices, at the variable from $-\infty$ to $+\infty$, are expressed as a sum of simple integrals of the complete elliptic integral of the first kind. The results of the numerical calculations obtained with the aid of the formulas are shown by graphs.

1. INTRODUCTION[1]

In the present part of this series of works, we present formulas useful for the numerical calculations of the lattice Green's functions for the cubic lattices:

$$G(t) = \frac{1}{\pi^3}\int_0^\pi dx \int_0^\pi dy \int_0^\pi dz \frac{1}{t - \omega(x, y, z)}, \quad (1.1)$$

where $\omega(x, y, z)$ is

$$\omega(x, y, z) = \gamma \cos x + \cos y + \cos z, \quad (1.2)$$

$$\omega(x, y, z) = \cos x \cos y \cos z, \quad (1.3)$$

$$\omega(x, y, z) = \cos y \cos z + \cos z \cos x + \cos x \cos y, \quad (1.4)$$

for the simple cubic (sc), body-centered cubic (bcc), and face-centered cubic (fcc) lattices, respectively. The parameter γ is introduced in (1.2) for generality, as it causes no additional trouble in the formulation. The function defined by (1.1) is real at $t \geq 2 + \gamma$, $t \geq 1$, and $t \geq 3$, respectively, for the sc, bcc, and fcc lattices. For this region, the function $G(t)$ was expressed as an integral of the complete elliptic integral of the first kind,[2,3] and the expressions were used to make an extensive table of the function.[3]

We shall consider the function defined by (1.1) as a complex function of the complex variable t. The function is analytic on the whole complex t plane, excluding the real axis from $-(2 + \gamma)$ to $2 + \gamma$ (sc), from -1 to 1 (bcc), and from -1 to 3 (fcc). Hence an expression of the function $G(t)$ on the whole t plane can be attained by the procedure of analytic continuation from the above-mentioned expressions in terms of the complete elliptic integral. We first prepare the expressions for the complete elliptic integral of the first kind and its analytic continuation at the values of modulus on the real and imaginary axis in terms of the complete elliptic integral between zero and unity. Then we use them to express the real and imaginary parts of $G(s - i\epsilon)$ at all the real values of s from $-\infty$ to $+\infty$, where ϵ is an infinitesimal positive number. The resulting expressions have the form of a sum of simple integrals of the complete elliptic integral of the first kind.

An extensive table of the real and imaginary parts of $G(s - i\epsilon)$ for $-\infty < s < \infty$ is in preparation. The graphs of the curves are given.

2. THE COMPLETE ELLIPTIC INTEGRAL OF THE FIRST KIND AS A COMPLEX FUNCTION OF THE MODULUS

The complete elliptic integral of the first kind $\mathbf{K}(k)$ as a complex function of the complex modulus k is defined by[4]

$$\mathbf{K}(k) = \int_0^{\frac{1}{2}\pi} d\theta (1 - k^2 \sin^2\theta)^{-\frac{1}{2}}. \quad (2.1)$$

This function is an even function of k and $\mathbf{K}(k^*) = \mathbf{K}(k)^*$. As a consequence, $\mathbf{K}(k)$ is real when k is pure imaginary. The expansion in powers of k and $k' = (1 - k^2)^{\frac{1}{2}}$ are given, respectively, as follows:

$$\mathbf{K}(k) = \frac{\pi}{2}\sum_{n=0}^{\infty} \frac{(\frac{1}{2})_n (\frac{1}{2})_n}{n!\, n!} k^{2n}, \quad (2.2)$$

$$\mathbf{K}(k) = -\frac{2}{\pi}\mathbf{K}(k') \ln k' + \sum_{n=0}^{\infty} \frac{(\frac{1}{2})_n (\frac{1}{2})_n}{n!\, n!} [\psi(n + 1) - \psi(n + \tfrac{1}{2})] k'^{2n}, \quad (2.3)$$

where

$$(\tfrac{1}{2})_n \equiv \Gamma(n + \tfrac{1}{2})/\Gamma(\tfrac{1}{2}), \quad \psi(z) = \Gamma'(z)/\Gamma(z).$$

The function $\mathbf{K}(k)$ has branch points at $k = \pm 1$. The expressions (2.1)–(2.3) are analytic on the Riemann surface excluding the branch cuts connecting $+1$ and $+\infty$ and -1 and $-\infty$, respectively, on the real axis. We call this part of the Riemann surface sheet I. In the following calculation, we need the function above

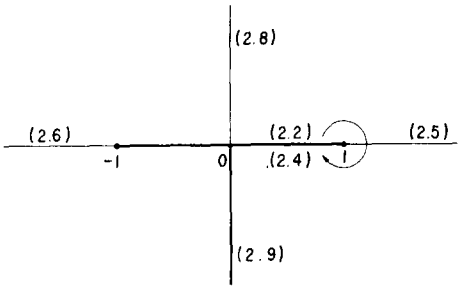

FIG. 1. Complex k plane. The bold solid line between -1 and 1 denotes the branch cut of the complete elliptic integral of the first kind. The numbers of the equation useful at each portion near the axes are referred. In this and the following figures, the upper half-plane represents sheet I and the right lower quarter of the plane, sheet II.

the real axis and its analytic continuation to the lower half-plane through the cut connecting $+1$ and $+\infty$. The latter part of the Riemann surface is denoted as II. Sheet I above the real axis and sheet II are written in the same figure by drawing the branch cut connecting -1 and $+1$, as in Fig. 1. The expressions (2.1)–(2.3) are valid on sheet I as stated above. The analytic continuation of the function $\mathbf{K}(k)$ to sheet II is attained with the aid of the expansion (2.3). One can reach from a point k on sheet I to the same point k on sheet II by encircling the unity clockwise. If $|k'| < 1$, the expression (2.3) is used. By this process, k' changes to $k' \exp(-i\pi)$. As a result, one obtains

$$\mathbf{K}^{(\mathrm{II})}(k) = \mathbf{K}(k) + 2i\mathbf{K}((1-k^2)^{\frac{1}{2}}), \tag{2.4}$$

when $|k'| < 1$. Here superscript (II) denotes sheet II. This expression is analytically continued to the whole sheet II. The function on the real axis $k = k_{\mathrm{R}} + i\epsilon$, where $k_{\mathrm{R}} > 1$ and $\epsilon \gtrsim 0$, is given by the one at $k < 1$ with the aid of the formula[4]

$$\mathbf{K}(k) = k^{-1}[\mathbf{K}(k^{-1}) + i\mathbf{K}(k^{-1}(k^2-1)^{\frac{1}{2}})], \tag{2.5}$$

which is valid when $\mathrm{Im}\, k^2 > 0$. From the definition (2.1) or (2.2), one sees that the value of $\mathbf{K}(k)$ at $k = -k_{\mathrm{R}} + ik_{\mathrm{I}}$ is the complex conjugate of the value at $k = k_{\mathrm{R}} + ik_{\mathrm{I}}$. This implies that the value at $k = k_{\mathrm{R}} + i\epsilon$, where $k_{\mathrm{R}} < -1$ and $\epsilon \gtrsim 0$, is given by

$$\mathbf{K}(k) = |k|^{-1}[\mathbf{K}(k^{-1}) - i\mathbf{K}(k^{-1}(k^2-1)^{\frac{1}{2}})]. \tag{2.6}$$

The values on the imaginary axis is achieved with the aid of the formula[4]

$$\mathbf{K}(ik_1/(1-k_1^2)^{\frac{1}{2}}) = (1-k_1^2)^{\frac{1}{4}}\mathbf{K}(k_1). \tag{2.7}$$

By substituting

$$k_1 = k_{\mathrm{I}}/(1+k_{\mathrm{I}}^2)^{\frac{1}{2}},$$

one obtains

$$\mathbf{K}(ik_{\mathrm{I}}) = (1+k_{\mathrm{I}}^2)^{-\frac{1}{4}}\mathbf{K}(k_{\mathrm{I}}/(1+k_{\mathrm{I}}^2)^{\frac{1}{2}}). \tag{2.8}$$

The corresponding expression for sheet II is obtained by putting $k = -ik_{\mathrm{I}} + \epsilon$, $\epsilon \gtrsim 0$, in (2.4) and then using (2.8) and (2.5):

$$\begin{aligned}\mathbf{K}^{(\mathrm{II})}&(-ik_{\mathrm{I}}+\epsilon)\\ &= \mathbf{K}(-ik_{\mathrm{I}}+\epsilon) + 2i\mathbf{K}((1+k_{\mathrm{I}}^2)^{\frac{1}{2}} + i\epsilon)\\ &= -(1+k_{\mathrm{I}}^2)^{-\frac{1}{4}}[\mathbf{K}(k_{\mathrm{I}}/(1+k_{\mathrm{I}}^2)^{\frac{1}{2}}) - 2i\mathbf{K}((1+k_{\mathrm{I}}^2)^{-\frac{1}{2}})].\end{aligned} \tag{2.9}$$

It should be remembered that the left-hand sides of (2.4) and (2.9) refer to the analytic continuation of the function $\mathbf{K}(k)$. But all the functions on the right-hand sides are referred to the original function $\mathbf{K}(k)$. One sees from (2.4) and (2.9) that the analytically continued function has a branch point at the origin. All equations (2.5), (2.6), (2.8), and (2.9) show that the infinity is a branch point. In Fig. 1 and in the following figures, these branch points are connected along the negative portion of the imaginary axis. Hence sheet II in these figures is restricted to the right lower quarter of the complex k plane.

3. SIMPLE CUBIC OR TETRAGONAL LATTICE

The Green's function defined by (1.1) with (1.2) is real at $t > 2 + \gamma$ and has been expressed in terms of the complete elliptic integral of the first kind $\mathbf{K}(k)$ as follows[3]:

$$G(t) = \frac{1}{\pi^2}\int_0^\pi dx\, k\mathbf{K}(k), \tag{3.1}$$

where

$$k = 2/(t - \gamma\cos x). \tag{3.2}$$

We note that this relation (3.1) is valid on the whole complex t plane.

When $t > 2 + \gamma$, we have $0 < k < 1$, and Eq. (3.1) is used for the numerical calculations [cf., Fig. 2(i)].

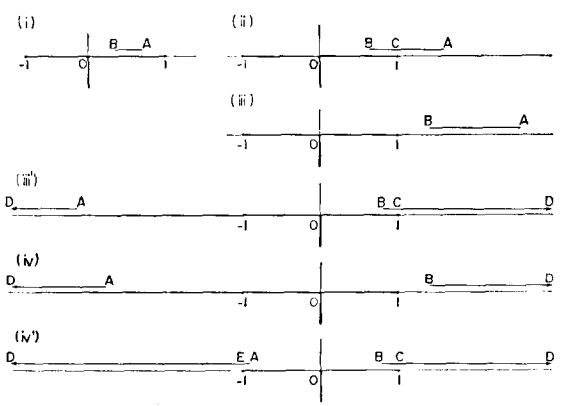

FIG. 2. Complex k plane for the simple cubic lattice. The k occurring in the integral (3.1) takes on values on the line connecting the points labeled by A and B. A corresponds to the value at $x = 0$, where $k = 2/(s-\gamma) + i\epsilon$. B is the value at $x = \pi$, where $k = 2/(s+\gamma) + i\epsilon$. (i) corresponds to the case $s > 2 + \gamma$. (ii), (iii), (iii'), (iv), and (iv') correspond to the subdivisions for the cases of $0 \le \gamma \le 1$, $1 \le \gamma \le 2$, and $2 \le \gamma \le \infty$. In case (ii), (iii'), or (iv'), C, where $k = 1 + i\epsilon$, is taken at $x = \cos^{-1}[(s-2)/\gamma]$. In case (iii'), (iv), or (iv'), D, where $k = +i\infty$, is taken at $x = \cos^{-1}(s/\gamma)$. In case (iv'), E, where $k = -1 + i\epsilon$, is taken at $x = \cos^{-1}[(s+2)/\gamma]$.

We put $t = s - i\epsilon$, where s is a real variable and ϵ is an infinitesimal positive number. $G(s - i\epsilon)$ is complex at $-(2 + \gamma) < s < 2 + \gamma$. We denote the real and imaginary parts as $G_R(s)$ and $G_I(s)$, respectively:

$$G(s - i\epsilon) = G_R(s) + iG_I(s). \tag{3.3}$$

For the sc case, $G_R(s)$ is an odd function of s and $G_I(s)$ is an even function:

$$G_R(-s) = -G_R(s), \quad G_I(-s) = G_I(s). \tag{3.4}$$

This fact can, for instance, be confirmed easily by the definition (1.1) with (1.2). Hence we have only to consider the range $0 < s < 2 + \gamma$ in the following. The formulas are given for the three cases $0 \leq \gamma \leq 1$, $1 \leq \gamma \leq 2$, and $2 \leq \gamma$, separately.

A. $0 \leq \gamma \leq 1$

We shall first assume that $0 \leq \gamma \leq 1$. Then the range $0 \leq s \leq 2 + \gamma$ is subdivided into three regions as follows:

(ii) $2 - \gamma < s < 2 + \gamma$,
(iii) $\gamma < s < 2 - \gamma$,
(iv) $0 < s < \gamma$.

The values of k occurring in the integral (3.1) are shown in Fig. 2 for the respective cases. For the region $k > 1$, one uses the formula (2.5). Namely, it is used for $0 < x < \cos^{-1}[(s - 2)/\gamma]$ in case (ii). As a result, one obtains

$$G_R(s) = \frac{1}{\pi^2}\int_0^{\cos^{-1}[(s-2)/\gamma]} dx\, \mathbf{K}\left(\frac{1}{k}\right) + \frac{1}{\pi^2}\int_{\cos^{-1}[(s-2)/\gamma]}^{\pi} dx\, k\mathbf{K}(k), \tag{3.5}$$

$$G_I(s) = \frac{1}{\pi^2}\int_0^{\cos^{-1}[(s-2)/\gamma]} dx\, \mathbf{K}\left(\frac{(k^2-1)^{\frac{1}{2}}}{k}\right). \tag{3.6}$$

For case (iii), one simply has

$$G_R(s) = \frac{1}{\pi^2}\int_0^{\pi} dx\, \mathbf{K}\left(\frac{1}{k}\right), \tag{3.7}$$

$$G_I(s) = \frac{1}{\pi^2}\int_0^{\pi} dx\, \mathbf{K}\left(\frac{(k^2-1)^{\frac{1}{2}}}{k}\right), \tag{3.8}$$

by (2.5). For case (iv), one uses the formulas (2.6) when $k < -1$, namely when $0 < x < \cos^{-1}(s/\gamma)$, and obtains

$$G_R(s) = -\frac{1}{\pi^2}\int_0^{\cos^{-1}(s/\gamma)} dx\, \mathbf{K}\left(\frac{1}{|k|}\right) + \frac{1}{\pi^2}\int_{\cos^{-1}(s/\gamma)}^{\pi} dx\, \mathbf{K}\left(\frac{1}{k}\right), \tag{3.9}$$

$$G_I(s) = \frac{1}{\pi^2}\int_0^{\pi} dx\, \mathbf{K}\left(\frac{(k^2-1)^{\frac{1}{2}}}{k}\right). \tag{3.10}$$

B. $1 \leq \gamma \leq 2$

When $1 \leq \gamma \leq 2$, the range $0 \leq s \leq 2 + \gamma$ is subdivided as follows:

(ii) $\gamma < s < 2 + \gamma$,
(iii') $2 - \gamma < s < \gamma$,
(iv) $0 < s < 2 - \gamma$.

For the cases (ii) and (iv), the values of k are restricted to the same regions as the corresponding cases of $0 \leq \gamma \leq 1$, and one obtains the same expressions (3.5), (3.6) and (3.9), (3.10), respectively. The values of k for the case (iii') are shown in Fig. 2(iii'). Now we have

$$G_R(s) = -\frac{1}{\pi^2}\int_0^{\cos^{-1}(s/\gamma)} dx\, \mathbf{K}\left(\frac{1}{|k|}\right) + \frac{1}{\pi^2}\int_{\cos^{-1}(s/\gamma)}^{\cos^{-1}[(s-2)/\gamma]} dx\, \mathbf{K}\left(\frac{1}{k}\right) + \frac{1}{\pi^2}\int_{\cos^{-1}[(s-2)/\gamma]}^{\pi} dx\, k\mathbf{K}(k), \tag{3.11}$$

$$G_I(s) = \frac{1}{\pi^2}\int_0^{\cos^{-1}[(s-2)/\gamma]} dx\, \mathbf{K}\left(\frac{(k^2-1)^{\frac{1}{2}}}{k}\right). \tag{3.12}$$

C. $2 \leq \gamma$

The subdivision for this case is

(ii) $\gamma < s < 2 + \gamma$,
(iii') $\gamma - 2 < s < \gamma$,
(iv') $0 < s < \gamma - 2$.

For the cases (ii) and (iii'), we have Eqs. (3.5) and (3.6) and Eqs. (3.11) and (3.12), respectively. The values of k for the case (iv') are shown by Fig. 2(iv'). The function $\mathbf{K}(k + i\epsilon)$ at $-1 < k < 0$ is equal to $\mathbf{K}(|k| + i\epsilon)$, and hence one has

$$G_R(s) = -\frac{1}{\pi^2}\int_0^{\cos^{-1}[(s+2)/\gamma]} dx\, |k|\, \mathbf{K}(|k|) - \frac{1}{\pi^2}\int_{\cos^{-1}[(s+2)/\gamma]}^{\cos^{-1}(s/\gamma)} dx\, \mathbf{K}\left(\frac{1}{|k|}\right) + \frac{1}{\pi^2}\int_{\cos^{-1}(s/\gamma)}^{\cos^{-1}[(s-2)/\gamma]} dx\, \mathbf{K}\left(\frac{1}{k}\right) + \frac{1}{\pi^2}\int_{\cos^{-1}[(s-2)/\gamma]}^{\pi} dx\, k\mathbf{K}(k), \tag{3.13}$$

$$G_I(s) = \frac{1}{\pi^2}\int_{\cos^{-1}[(s+2)/\gamma]}^{\cos^{-1}[(s-2)/\gamma]} dx\, \mathbf{K}\left(\frac{(k^2-1)^{\frac{1}{2}}}{k}\right). \tag{3.14}$$

4. BODY-CENTERED CUBIC LATTICE

The Green's function at $t > 1$ is real and is given by[3]

$$G(t) = \frac{4}{\pi^2 t}\int_0^{\frac{1}{2}\pi} dx\, \mathbf{K}(k), \tag{4.1}$$

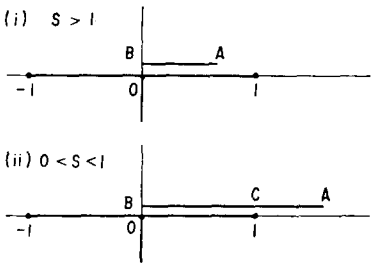

FIG. 3. Complex k plane for the bcc lattice. Point A corresponds to $x = 0$, where $k = (1/s) + i\epsilon$. Point B corresponds to $x = \pi/2$, where $k = i\epsilon$. In case (ii), point C, where $k = 1 + i\epsilon$, is taken at $x = \cos^{-1}s$.

where

$$k = \cos(x)/t. \tag{4.2}$$

We put $t = s - i\epsilon$ and divide $G(t)$ into the real and imaginary parts:

$$G(s - i\epsilon) = G_R(s) + iG_I(s). \tag{4.3}$$

We notice again that $G_R(s)$ is an odd and $G_I(s)$ is an even function of s:

$$G_R(-s) = -G_R(s), \quad G_I(-s) = G_I(s); \tag{4.4}$$

e.g., cf. (1.1) with (1.3).

We now consider the region $0 < s < 1$. The values of k occurring in the integral (4.1) are shown in Fig. 3. By using (2.5), one obtains

$$G_R(s) = \frac{4}{\pi^2 s}\int_0^{\cos^{-1}s} dx\,\frac{1}{k}\mathbf{K}\left(\frac{1}{k}\right) + \frac{4}{\pi^2 s}\int_{\cos^{-1}s}^{\frac{1}{2}\pi} dx\,\mathbf{K}(k), \tag{4.5}$$

$$G_I(s) = \frac{4}{\pi^2 s}\int_0^{\cos^{-1}s} dx\,\frac{1}{k}\mathbf{K}\left(\frac{(k^2-1)^{\frac{1}{2}}}{k}\right). \tag{4.6}$$

5. FACE-CENTERED CUBIC LATTICE

The Green's function at $t > 3$ is expressed as[3]

$$G(t) = \frac{4}{\pi^2(t+1)}\int_0^{\frac{1}{2}\pi} dx\,\mathbf{K}(k), \tag{5.1}$$

where

$$k = 2(t + \cos^2 x)^{\frac{1}{2}}/(t+1). \tag{5.2}$$

We put $t = s - i\epsilon$ and denote the real and imaginary parts $G(s - i\epsilon)$ as $G_R(s)$ and $G_I(s)$:

$$G(s - i\epsilon) = G_R(s) + iG_I(s). \tag{5.3}$$

For the fcc lattice, no symmetric behaviors between positive and negative values of s. We have to divide the range $-\infty < s < \infty$ into five regions as follows:

(i) $3 < s$,
(ii) $1 < s < 3$,
(iii) $0 < s < 1$,
(iv) $-1 < s < 0$,
(v) $s < -1$.

For case (i), the values of k occurring in the integral (5.1) are between 0 and 1, and we can use (5.1) by putting $t = s$.

In order to derive the expressions of $G_R(s)$ and $G_I(s)$ for cases (ii)–(v), we draw the figures to show the values of k occurring in the integral (5.1). They are shown in Fig. 4. In Fig. 1, the numbers of the equations which are useful on the various regions near the real and imaginary axis are shown. By comparing with this figure and then the referred equations, we easily obtain the results for cases (ii)–(v) as follows.

Case (ii): $1 < s < 3$

Comparing Fig. 4(ii) and Fig. 1 and using (5.1) and (2.5), we have

$$G_R(s) = \frac{4}{\pi^2(s+1)}\left(\int_0^{\cos^{-1}[(s-1)/2]} dx\,\frac{1}{k}\mathbf{K}(k) + \int_{\cos^{-1}[(s-1)/2]}^{\frac{1}{2}\pi} dx\,\mathbf{K}(k)\right), \tag{5.4}$$

$$G_I(s) = \frac{4}{\pi^2(s+1)}\int_0^{\cos^{-1}[(s-1)/2]} \frac{1}{k}\mathbf{K}\left(\frac{(k^2-1)^{\frac{1}{2}}}{k}\right). \tag{5.5}$$

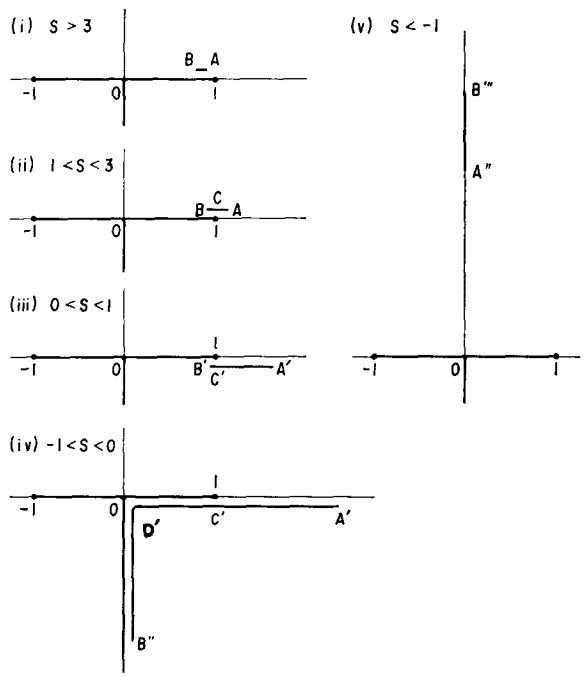

FIG. 4. Complex k plane for the fcc lattice, points A, A', and A'' correspond to $x = 0$, where $k = 2/(s+1)^{1/2} + i\epsilon$, $2/(1+s)^{1/2} - i\epsilon$, and $2i/(-1-s)$, respectively. Points B, B', B'', and B''' correspond to $x = \frac{1}{2}\pi$, where $k = 2s^{1/2}/(s+1) + i\epsilon$, $2s^{1/2}/(s+1) - i\epsilon$, $-2i(-s)^{1/2}/(1+s) + \epsilon$, and $2i(-s)^{1/2}/(-s-1)$, respectively. In case (ii), $x = \cos^{-1}[(s-1)/2]$ at C, where $k = 1 + i\epsilon$. In cases (iii) and (iv), $x = \cos^{-1}[(1-s)/2]$ at C', where $k = 1 - i\epsilon$. In case (iv), $x = \cos^{-1}[(-s)^{1/2}]$ at D', where $k = (-1+i)\epsilon$.

Case (iii): $0 < s < 1$

We compare Fig. 4(iii) and Fig. 1. By using (2.4) and (2.5), one has

$$G_R(s) = \frac{4}{\pi^2(s+1)}\left(\int_0^{\cos^{-1}[(s-1)/2]} dx\,\frac{1}{k}\,\mathbf{K}(k) + \int_{\cos^{-1}[(1-s)/2]}^{\frac{1}{2}\pi} dx\,\mathbf{K}(k)\right), \tag{5.6}$$

$$G_I(s) = \frac{4}{\pi^2(s+1)}\left(\int_0^{\cos^{-1}[(1-s)/2]} dx\,\frac{1}{k}\,\mathbf{K}\left(\frac{(k^2-1)^{\frac{1}{2}}}{k}\right) + 2\int_{\cos^{-1}[\frac{1}{2}(1-s)]}^{\frac{1}{2}\pi} dx\,\mathbf{K}((1-k^2)^{\frac{1}{2}})\right). \tag{5.7}$$

Case (iv): $-1 < s < 0$

Figure 4(iv) and Fig. 1 are compared. With the aid of (2.4), (2.5), and (2.8), we have

$$G_R(s) = \frac{4}{\pi^2(s+1)}\left[\int_0^{\cos^{-1}[\frac{1}{2}(1-s)]} dx\,\frac{1}{k}\,\mathbf{K}\left(\frac{1}{k}\right) + \int_{\cos^{-1}[\frac{1}{2}(1-s)]}^{\cos^{-1}\sqrt{-s}} dx\,\mathbf{K}(k) - \int_{\cos^{-1}\sqrt{-s}}^{\frac{1}{2}\pi} dx\,\frac{1}{(1-k^2)^{\frac{1}{2}}}\,\mathbf{K}\left(\frac{k/i}{(1-k^2)^{\frac{1}{2}}}\right)\right], \tag{5.8}$$

$$G_I(s) = \frac{4}{\pi^2(s+1)}\left[\int_0^{\cos^{-1}[\frac{1}{2}(1-s)]} dx\,\frac{1}{k}\,\mathbf{K}\left(\frac{(k^2-1)^{\frac{1}{2}}}{k}\right) + 2\int_{\cos^{-1}[\frac{1}{2}(1-s)]}^{\cos^{-1}\sqrt{-s}} dx\,\mathbf{K}((1-k^2)^{\frac{1}{2}}) + 2\int_{\cos^{-1}\sqrt{-s}}^{\frac{1}{2}\pi} dx\,\frac{1}{(1-k^2)^{\frac{1}{2}}}\,\mathbf{K}\left(\frac{1}{(1-k^2)^{\frac{1}{2}}}\right)\right]. \tag{5.9}$$

Case (v): $s < -1$

Figure 4(v) and Fig. 1 are compared. Now (2.7) is used to give

$$G_R(s) = \frac{4}{\pi^2(s+1)}\int_0^{\frac{1}{2}\pi} dx\,\frac{1}{(1-k^2)^{\frac{1}{2}}}\,\mathbf{K}\left(\frac{k/i}{(1-k^2)^{\frac{1}{2}}}\right), \tag{5.10}$$

$$G_I(s) = 0. \tag{5.11}$$

6. RESULTS OF NUMERICAL CALCULATIONS

The formulas presented in the preceding sections are used for the numerical evaluation of the lattice Green's functions for the cubic lattices. In the calculation for the sc lattice, the parameter γ is put equal to unity. The results obtained are shown in Figs. 5–7. The results for $s \geq 3$ (sc and fcc) and $s > 1$ (bcc) are in agreement with the previous one.[3,5] The results for $0 \leq s \leq 1$ and $3 < s < \infty$ (sc) are found

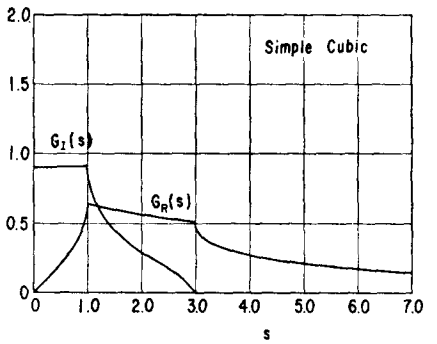

FIG. 5. The lattice Green's function for the simple cubic lattice. $G_R(s)$ and $G_I(s)$ are the real and imaginary parts, respectively.

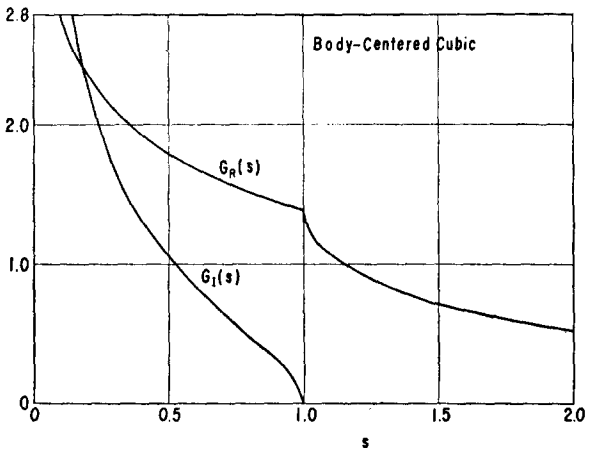

FIG. 6. The lattice Green's function for the body-centered cubic lattice. $G_R(s)$ and $G_I(s)$ are the real and imaginary parts, respectively.

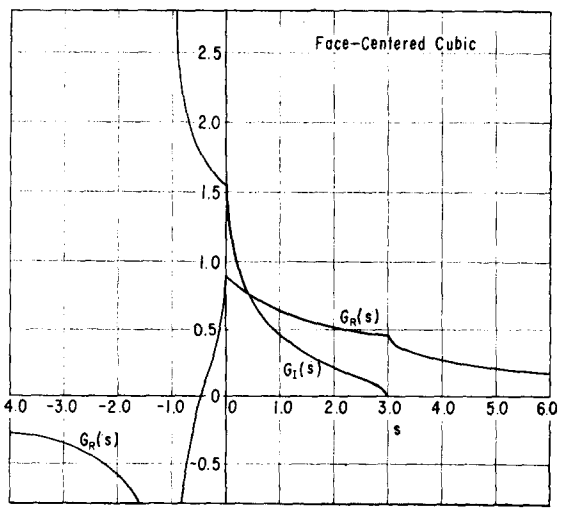

FIG. 7. The lattice Green's function for the face-centered cubic lattice. $G_R(s)$ and $G_I(s)$ are the real and imaginary parts, respectively.

in complete agreement as those obtained by Katsura, Inawashiro, and Abe with the aid of the expansion formulas.[6]

7. CONCLUDING REMARKS

The formulas useful for the calculation of the lattice Green's function $G(s - i\epsilon)$ for the cubic lattices are derived for the whole range, $-\infty < s < \infty$, by the method of the analytic continuation from the expressions which have been derived for the range $s > 2 + \gamma$ (sc or tetragonal), $s \geq 3$ (fcc), or $s \geq 1$ (bcc). We mention here that we have derived all the formulas for $-\infty < s < \infty$ by introducing a suitable change of variables, after the integration over z is performed on (1.1). That derivation is more elementary but laborious and will be published on a suitable occasion.

In the derived expressions for the lattice Green's functions, the integration over one of the coordinates x is left to be performed. Hence if the integrand of (1.1) is multiplied by an arbitrary function of x, each integrand in the final expressions must involve the same function. If the function $\cos lx$ with an integer l is used, one can calculate the lattice Green's function denoted by $I_{\text{tetra}}(t; l, 0, 0; \gamma)$ in the general introduction.[1] Such a calculation is now in progress and $I_{\text{tetra}}(t; l, m, n; \gamma)$ for some sets of l, m, and n are known to be evaluated with the aid of the results (see Ref. 5).

ACKNOWLEDGMENTS

The authors are grateful to Professor S. Katsura for valuable discussions. The numerical computations were performed with the aid of the computer NEAC 2200 of the Computer Center, Tohoku University.

Note added in proof: An alternative derivation of the formulas given in the present article is given by the present authors in J. Phys. Soc. Japan **30,** No. 4 (1971). The formulas for the imaginary part $G_I(s)$ have been given by Jelitto [J. Phys. Chem. Solids **30,** 609 (1969)]. If one is interested only in the values at the origin of the lattice Green's function for the bcc or fcc lattice, a more efficient method has been found, presented in the following paper of the present authors [J. Math. Phys. **12,** 986 (1971)].

* Present address: Dept. of Physics, Ohio University, Athens, Ohio 45701.

[1] A general introduction of the present series is given in a preceding paper by S. Katsura, T. Morita, S. Inawashiro, T. Horiguchi, and Y. Abe, J. Math. Phys. **12,** 892 (1971).

[2] E. W. Montroll, *Proceedings of the Third Berkley Symposium on Mathematical Statistics and Probability* (U. of California Press, Berkeley, Calif., 1955), Vol. III, p. 209.

[3] I. Mannari and C. Kawabata, Department of Physics, Okayama University, Research Notes, No. 15, 1964.

[4] Bateman Manuscript Project, *Higher Transcendental Functions*, edited by A. Erdélyi (McGraw-Hill, New York, 1953), Vol. II, p. 317ff., and Vol. I, p. 105ff.; W. Magnus and F. Oberhettinger, *Formeln und Sätze für die Mathematische Physik* (Springer-Verlag, Berlin, 1948).

[5] A. A. Maradudin, E. W. Montroll, G. H. Weiss, R. Herman, and H. W. Milnes, Acad. Roy. Belg., Classe Sci., **9,** No. 7 (1960).

[6] S. Katsura, S. Inawashiro, Y. Abe, J. Math. Phys. **12,** 895 (1971).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Calculation of the Lattice Green's Function for the bcc, fcc, and Rectangular Lattices

TOHRU MORITA*
Department of Applied Science, Faculty of Engineering, Tohoku University, Sendai, Japan
AND
TSUYOSHI HORIGUCHI*
Department of Physics, Faculty of Science, Tohoku University, Sendai, Japan

(Received 3 August 1970)

Formulas are provided which are convenient for the evaluation of the lattice Green's functions for the bcc, fcc, and rectangular lattices, at an arbitrary complex variable. The formulas involve the complete elliptic integral of the first kind with complex modulus; the integral has been found to be evaluated efficiently by the method of the arithmetic–geometric mean, generalized for the case with complex modulus. The expansions of the lattice Green's functions around the singular points are given for the bcc and fcc lattices. These lattice Green's functions diverge at a variable. The singular points responsible for the divergences are found to form one-dimensional lines.

1. INTRODUCTION

The lattice Green's functions for the cubic lattices are defined by

$$G(t) = \frac{1}{\pi^3}\int_0^\pi dx \int_0^\pi dy \int_0^\pi dz \frac{1}{t - \omega(x, y, z)}, \quad (1.1)$$

where $\omega(x, y, z)$ is

$$\omega(x, y, z) = \cos x + \cos y + \cos z, \quad (1.2)$$

$$\omega(x, y, z) = \cos x \cos y \cos z, \quad (1.3)$$

$$\omega(x, y, z) = \cos y \cos z + \cos z \cos x + \cos x \cos y \quad (1.4)$$

in complete agreement as those obtained by Katsura, Inawashiro, and Abe with the aid of the expansion formulas.[6]

7. CONCLUDING REMARKS

The formulas useful for the calculation of the lattice Green's function $G(s - i\epsilon)$ for the cubic lattices are derived for the whole range, $-\infty < s < \infty$, by the method of the analytic continuation from the expressions which have been derived for the range $s > 2 + \gamma$ (sc or tetragonal), $s \geq 3$ (fcc), or $s \geq 1$ (bcc). We mention here that we have derived all the formulas for $-\infty < s < \infty$ by introducing a suitable change of variables, after the integration over z is performed on (1.1). That derivation is more elementary but laborious and will be published on a suitable occasion.

In the derived expressions for the lattice Green's functions, the integration over one of the coordinates x is left to be performed. Hence if the integrand of (1.1) is multiplied by an arbitrary function of x, each integrand in the final expressions must involve the same function. If the function $\cos lx$ with an integer l is used, one can calculate the lattice Green's function denoted by $I_{\text{tetra}}(t; l, 0, 0; \gamma)$ in the general introduction.[1] Such a calculation is now in progress and $I_{\text{tetra}}(t; l, m, n; \gamma)$ for some sets of l, m, and n are known to be evaluated with the aid of the results (see Ref. 5).

ACKNOWLEDGMENTS

The authors are grateful to Professor S. Katsura for valuable discussions. The numerical computations were performed with the aid of the computer NEAC 2200 of the Computer Center, Tohoku University.

Note added in proof: An alternative derivation of the formulas given in the present article is given by the present authors in J. Phys. Soc. Japan **30,** No. 4 (1971). The formulas for the imaginary part $G_I(s)$ have been given by Jelitto [J. Phys. Chem. Solids **30,** 609 (1969)]. If one is interested only in the values at the origin of the lattice Green's function for the bcc or fcc lattice, a more efficient method has been found, presented in the following paper of the present authors [J. Math. Phys. **12,** 986 (1971)].

* Present address: Dept. of Physics, Ohio University, Athens, Ohio 45701.

[1] A general introduction of the present series is given in a preceding paper by S. Katsura, T. Morita, S. Inawashiro, T. Horiguchi, and Y. Abe, J. Math. Phys. **12,** 892 (1971).

[2] E. W. Montroll, *Proceedings of the Third Berkley Symposium on Mathematical Statistics and Probability* (U. of California Press, Berkeley, Calif., 1955), Vol. III, p. 209.

[3] I. Mannari and C. Kawabata, Department of Physics, Okayama University, Research Notes, No. 15, 1964.

[4] Bateman Manuscript Project, *Higher Transcendental Functions*, edited by A. Erdélyi (McGraw-Hill, New York, 1953), Vol. II, p. 317ff., and Vol. I, p. 105ff.; W. Magnus and F. Oberhettinger, *Formeln und Sätze für die Mathematische Physik* (Springer-Verlag, Berlin, 1948).

[5] A. A. Maradudin, E. W. Montroll, G. H. Weiss, R. Herman, and H. W. Milnes, Acad. Roy. Belg., Classe Sci., **9,** No. 7 (1960).

[6] S. Katsura, S. Inawashiro, Y. Abe, J. Math. Phys. **12,** 895 (1971).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Calculation of the Lattice Green's Function for the bcc, fcc, and Rectangular Lattices

TOHRU MORITA*
Department of Applied Science, Faculty of Engineering, Tohoku University, Sendai, Japan
AND
TSUYOSHI HORIGUCHI*
Department of Physics, Faculty of Science, Tohoku University, Sendai, Japan

(Received 3 August 1970)

Formulas are provided which are convenient for the evaluation of the lattice Green's functions for the bcc, fcc, and rectangular lattices, at an arbitrary complex variable. The formulas involve the complete elliptic integral of the first kind with complex modulus; the integral has been found to be evaluated efficiently by the method of the arithmetic–geometric mean, generalized for the case with complex modulus. The expansions of the lattice Green's functions around the singular points are given for the bcc and fcc lattices. These lattice Green's functions diverge at a variable. The singular points responsible for the divergences are found to form one-dimensional lines.

1. INTRODUCTION

The lattice Green's functions for the cubic lattices are defined by

$$G(t) = \frac{1}{\pi^3}\int_0^\pi dx \int_0^\pi dy \int_0^\pi dz \frac{1}{t - \omega(x, y, z)}, \quad (1.1)$$

where $\omega(x, y, z)$ is

$$\omega(x, y, z) = \cos x + \cos y + \cos z, \quad (1.2)$$

$$\omega(x, y, z) = \cos x \cos y \cos z, \quad (1.3)$$

$$\omega(x, y, z) = \cos y \cos z + \cos z \cos x + \cos x \cos y \quad (1.4)$$

for the simple cubic (sc), the body-centered cubic (bcc),and the face-centered cubic (fcc) lattices, respectively. The function is an analytic function of the complex variable t except on the real axis from -3 to 3 (sc), from -1 to 1 (bcc), and from -1 to 3 (fcc). Therefore, if one has an analytic expression which represents the function $G(t)$ in some region or on a line or on a converging series of points on the complex t plane, where $G(t)$ is analytic, one can in principle achieve the analytic function which represents $G(t)$ on the whole complex t plane by the process of analytic continuation. This procedure has been successfully applied to obtain the expression of $G(t)$ for $t = s - i\epsilon$ ($-\infty < s < \infty$, $\epsilon \geqslant 0$) from the known expressions for $t > 3$ (sc and fcc) or for $t > 1$ (bcc).[1-4]

For the bcc lattice for $t > 1$ and for the fcc lattice for $t < -1$ or $t > 3$, compact expressions have been given for the lattice Green's function $G(t)$ in terms of the complete elliptic integral of the first kind.[5,6] The corresponding expression for the rectangular lattice is easily derived. The process of analytic continuation has been applied to the expression for the bcc lattice by Katsura and Horiguchi.[4] They obtained expressions for the real and imaginary parts separately in terms of the hypergeometric functions. In the present paper, alternative expressions for $G(t)$, useful for an arbitrary complex value t, are obtained for the bcc as well as for the fcc and rectangular lattices in terms of the complete elliptic integral of the first kind with complex modulus.

For real values of the modulus, the method of the arithmetic–geometric mean is a very powerful method of calculating the elliptic integrals and the Jacobian elliptic functions (e.g., see Refs. 7 and 8). In a separate paper[9] the present authors showed that the method is powerful also for the evaluation of these quantities with complex modulus. Then the lattice Green's functions $G(t)$ for the bcc and fcc lattices for the complex values of t in general are calculated with the aid of that method for an arbitrary complex variable t, in particular for $t = s - i\epsilon$, where $-\infty < s < \infty$ and ϵ is an infinitesimal positive number.

Discussions in the following sections are restricted to the calculation of $G(t)$ for t in the lower half of the complex t plane. The results for the upper half of the plane can be obtained through the equation

$$G(t^*) = G(t)^*. \tag{1.5}$$

Special attention is focused on the calculation of $G(t)$ just below the real axis:

$$\lim_{\epsilon \to 0} G(s - i\epsilon) = G_{\mathrm{R}}(s) + iG_{\mathrm{I}}(s) \tag{1.6}$$

for s from $-\infty$ to $+\infty$. The imaginary part $G_{\mathrm{I}}(s)$ is related to the level density of the Bloch electrons and lattice vibration of the lattice.

The numerical calculations of the lattice Green's functions have been attempted by using the expressions which are a sum of definite integrals of the complete elliptic integral of the first kind for the sc, bcc, and fcc lattices.[3] The results for the latter two lattices are reproduced by the present method with far less labor to higher accuracy. The curves obtained for these lattices are reproduced in Figs. 2 and 7. They show a divergence. According to van Hove,[10] the divergence of the level density, which is the imaginary part of the Green's function, $G_{\mathrm{I}}(s)$, occurs for the 3-dimensional lattice only for a special interaction. It is shown that the special relations predicted by van Hove for the occurrence of the divergence are satisfied for the present cases.[11] The expansions of the Green's functions at the van Hove singular points are given with the aid of the expansions of the complete elliptic integral of the first kind.

The formulas for the bcc and fcc lattice are given in Sec. 2 and are analyzed in Sec. 3. The results of Sec. 2 are summarized in Sec. 4. The corresponding discussions for the rectangular lattice are given in the Appendix.

2. EXPRESSIONS FOR THE LATTICE GREEN'S FUNCTIONS

A. bcc Lattice

The bcc lattice Green's function $G(t)$ for $t > 1$ is known to be expressed as follows:

$$G(t) = (4/\pi^2)t^{-1}\mathbf{K}(k)^2, \tag{2.1}$$

where

$$k = [\tfrac{1}{2} - \tfrac{1}{2}(1 - t^{-2})^{\frac{1}{2}}]^{\frac{1}{2}}. \tag{2.2}$$

The value of $G(t)$ for $t < -1$ is evaluated by the relation

$$G(-t) = -G(t), \tag{2.3}$$

which is a consequence of the definition (1.1) with (1.3).

When $t > 1$, k defined by (2.2) is real and between 0 and $1/\sqrt{2}$. We draw the values of k for $t = s - i\epsilon$ in Fig. 1, where $-\infty < s < \infty$ and ϵ is an infinitesimal positive number. In drawing the figure, we first rewrite (2.2) as follows:

$$k = [(t + 1)^{\frac{1}{2}} - (t - 1)^{\frac{1}{2}}]/2t^{\frac{1}{2}}. \tag{2.4}$$

We note that the factor $(s - i\epsilon - a)^{\frac{1}{2}}$ for $s > a$ changes to $-i(a - s + i\epsilon)^{\frac{1}{2}}$ for $s < a$. We denote that limiting value of $(s - i\epsilon - a)^{\frac{1}{2}}$ at $\epsilon \to 0$ ($\epsilon > 0$)

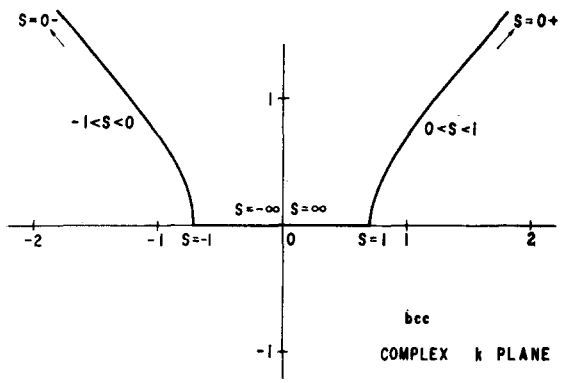

FIG. 1. The values of the argument k of the complete elliptic integral for $t = s - i\epsilon$, $-\infty < s < \infty$, for the bcc lattice.

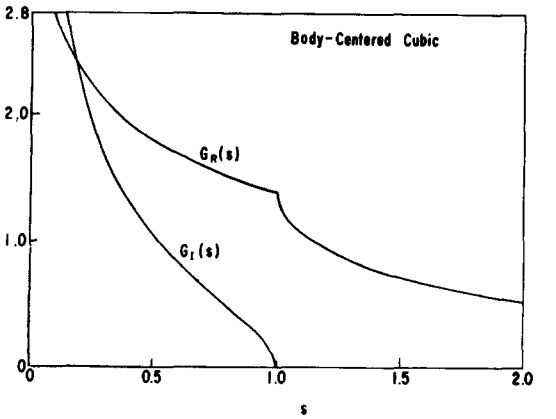

FIG. 2. The Green's function for the bcc lattice. $G_R(s)$ and $G_I(s)$ represent the real and imaginary parts, respectively. They are an odd and an even function of s, respectively, for this lattice.

by $(s-a)^{\frac{1}{2}}$ so that

$$(s-a)^{\frac{1}{2}} = (s-a)^{\frac{1}{2}}, \quad s-a \geq 0, \qquad = -i(a-s)^{\frac{1}{2}}, \quad s-a < 0. \tag{2.5}$$

Then k, for $t = s - i\epsilon$, $\epsilon \gtrsim 0$, is denoted by

$$k = [(s+1)^{\frac{1}{2}} - (s-1)^{\frac{1}{2}}]/2s^{\frac{1}{2}}, \quad -\infty < s < \infty. \tag{2.6}$$

At $s < 1$, this expression represents

$$k = [(s+1)^{\frac{1}{2}} + i(1-s)^{\frac{1}{2}}]/2s^{\frac{1}{2}}, \quad 0 < s < 1, \tag{2.7}$$

$$k = [-(1-s)^{\frac{1}{2}} + i(1+s)^{\frac{1}{2}}]/2(-s)^{\frac{1}{2}}, \quad -1 < s < 0, \tag{2.8}$$

$$k = [-(1-s)^{\frac{1}{2}} + (-1-s)^{\frac{1}{2}}]/2(-s)^{\frac{1}{2}}, \quad s < -1. \tag{2.9}$$

The lower half of the complex t plane is mapped to the region above the line showing k for $s - i\epsilon$. The lattice Green's function $G(t)$ for $t = s - i\epsilon$ and for an arbitrary complex variable t with negative imaginary part is evaluated with the aid of the formula (2.1) with (2.2) or (2.4). Evaluation of $\mathbf{K}(k)$ is performed with the aid of the method of the arithmetic–geometric mean. Its program is attained by changing the declaration of some of the variables and functions occurring in the programs[7,8] prepared for real modulus from *real* to *complex*.[9] The result of the numerical computation is shown in Fig. 2.

B. fcc Lattice

The fcc lattice Green's function $G(t)$ for $t < -1$ and $t > 3$ has recently been expressed as follows[6]:

$$G(t) = [4/\pi^2(t+1)]\mathbf{K}(k_1^-)\mathbf{K}(k_1^+) \tag{2.10}$$

for $t > 3$, where

$$k_1^{\pm} = \frac{1}{2^{\frac{1}{2}}}\left(1 \pm \frac{4t^{\frac{1}{2}}(t+1)^{\frac{1}{2}}}{(t+1)^2} - \frac{(t-1)(t+1)^{\frac{1}{2}}(t-3)^{\frac{1}{2}}}{(t+1)^2}\right)^{\frac{1}{2}}, \tag{2.11}$$

and

$$G(t) = -(8/\pi^2)[(3-t)^{\frac{1}{2}}(-1-t)^{\frac{1}{2}} + 1 - t]^{-1} \times \mathbf{K}(k_2^-)\mathbf{K}(k_2^+) \tag{2.12}$$

for $t < -1$, where

$$k_2^{\pm} = [2(-t)^{\frac{1}{2}} \pm 2(-1-t)^{\frac{1}{2}}]/[(3-t)^{\frac{1}{2}}(-1-t)^{\frac{1}{2}} + 1 - t]. \tag{2.13}$$

The expressions of k_1^- and k_1^+ for $t = s - i\epsilon$ are obtained by using $(s-a)^{\frac{1}{2}}$, defined by (2.5), in place of $(t-a)^{\frac{1}{2}}$ occurring in (2.11). Figures 3 and 4 show the curve of them for $-\infty < s < \infty$. The lower half of the complex t plane is mapped to the region which is to the left of the imaginary axis, and is bounded by the line for $s - i\epsilon$, in Fig. 3. In Fig. 4, it is mapped to the right of the line for $s - i\epsilon$. Now for $-1 < s < 1$, the values of k_1^+ appear below the real axis across the real axis between 1 and $+\infty$. For k_1^+ on this part of the curve and to the right of the curve below the real axis, the analytic continuation $\mathbf{K}^{(\mathrm{II})}(k_1^+)$ defined

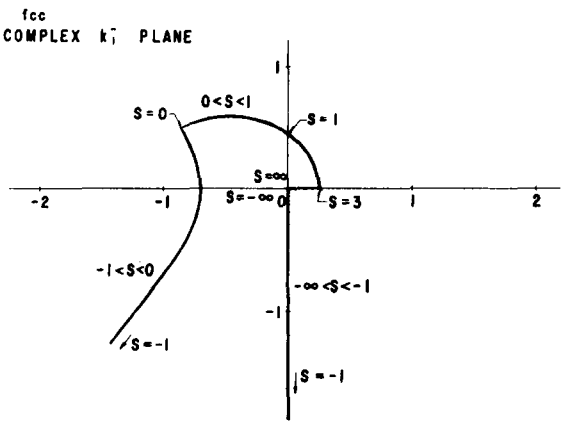

FIG. 3. The values of k_1^- for $t = s - i\epsilon$, $-\infty < s < \infty$.

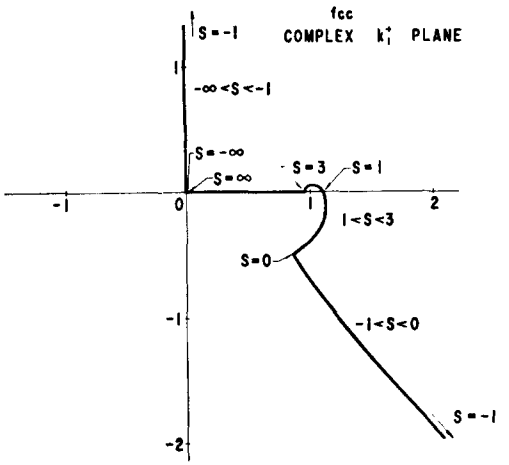

FIG. 4. The values of k^+ for $t = s - i\epsilon$, $-\infty < s < \infty$.

as follows must be used in place of $\mathbf{K}(k_1^+)$:

$$\mathbf{K}^{(\mathrm{II})}(k_1^+) = \mathbf{K}(k_1^+) + 2i\mathbf{K}'(k_1^+), \tag{2.14}$$

where

$$\mathbf{K}'(k) \equiv \mathbf{K}(k'), \quad k' \equiv (1 - k^2)^{\frac{1}{2}}. \tag{2.15}$$

[This analytic continuation is obtained by considering the expansion of $\mathbf{K}(k)$ around $k = 1$ as given by (3.4) in the following section.] For the other values of k_1^+ and k_1^-, we can use the complete elliptic integral, and we have

$$G(t) = [1/\pi^2(t + 1)]\mathbf{K}(k_1^-)\mathbf{K}^{(\mathrm{II})}(k_1^+), \tag{2.16}$$

where

$$\begin{aligned}\mathbf{K}^{(\mathrm{II})}(k_1^+) &= \mathbf{K}(k_1^+), \quad \operatorname{Im} k_1^+ > 0, \quad \operatorname{Im} t < 0,\\ &= \mathbf{K}(k_1^+) + 2i\mathbf{K}'(k_1^+),\\ &\qquad \operatorname{Im} k_1^+ < 0, \quad \operatorname{Im} t < 0.\end{aligned} \tag{2.17}$$

For $t = s - i\epsilon$, $k_2^\pm$ are given by

$$\begin{aligned}k_2^\pm = \tfrac{1}{2}[&(-s)^{\frac{1}{2}}(1 - s) \mp (3 - s)^{\frac{1}{2}}(-1 - s)\\ &\pm (-1 - s)^{\frac{1}{2}}(1 - s)\\ &- (-s)^{\frac{1}{2}}(-1 - s)^{\frac{1}{2}}(3 - s)^{\frac{1}{2}}],\end{aligned} \tag{2.18}$$

where $(a - s)^{\frac{1}{2}}$ denotes

$$\begin{aligned}(a - s)^{\frac{1}{2}} &\equiv (a - s)^{\frac{1}{2}}, \quad a - s > 0,\\ &\equiv i(s - a)^{\frac{1}{2}}, \quad a - s < 0,\end{aligned} \tag{2.19}$$

namely

$$(a - s)^{\frac{1}{2}} = i(s - a)^{\frac{1}{2}} \tag{2.20}$$

[cf. (2.5)].

The values of k_2^- and k_2^+ for $t = s - i\epsilon$, where $-\infty < s < \infty$ and $\epsilon \geqslant 0$, are shown in Figs. 5 and 6. Equation (2.18) with the lower and upper signs map the lower half of the complex t plane to the inside of the curve shown in Figs. 5 and 6, respectively.

For $-1 < s < 1$ and $\epsilon \geqslant 0$, the curve given by Fig. 6 for k_2^+ is above the real axis across the part from

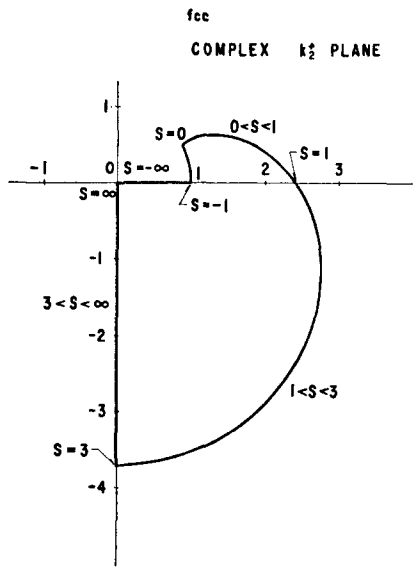

FIG. 5. The values of k_2^- for $t = s - i\epsilon$, $-\infty < s < \infty$.

1 to $+\infty$. Hence, $\mathbf{K}(k_2^+)$ for k_2^+, which is on and inside of this part of the curve above the real axis, must be the analytic continuation of $\mathbf{K}(k_2^+)$:

$$\mathbf{K}^{(\mathrm{III})}(k_2^+) = \mathbf{K}(k_2^+) - 2i\mathbf{K}'(k_2^+). \tag{2.21}$$

For the other parts, $\mathbf{K}(k_2^+)$ and $\mathbf{K}(k_2^-)$ are the complete elliptic integral for complex modulus. Now we have

$$\begin{aligned}G(t) = &-(8/\pi^2)[(3 - t)^{\frac{1}{2}}(-1 - t)^{\frac{1}{2}} + 1 - t]^{-1}\\ &\times \mathbf{K}(k_2^-)\mathbf{K}^{(\mathrm{III})}(k_2^+),\end{aligned} \tag{2.22}$$

where $\mathbf{K}^{(\mathrm{III})}(k_2^+)$ is given by

$$\begin{aligned}\mathbf{K}^{(\mathrm{III})}(k_2^+) &= \mathbf{K}(k_2^+), \quad \operatorname{Im} k_2^+ < 0, \quad \operatorname{Im} t < 0,\\ &= \mathbf{K}(k_2^+) - 2i\mathbf{K}'(k_2^+),\\ &\qquad \operatorname{Im} k_2^+ > 0, \quad \operatorname{Im} t > 0.\end{aligned} \tag{2.23}$$

The lattice Green's function $G(t)$ for the fcc lattice is evaluated with the aid either of (2.16) or (2.22) for an arbitrary complex t or for $s - i\epsilon$, $-\infty < s < \infty$ and $\epsilon \geqslant 0$. The fact that the expressions (2.16) and (2.22) are equivalent is checked by substituting the

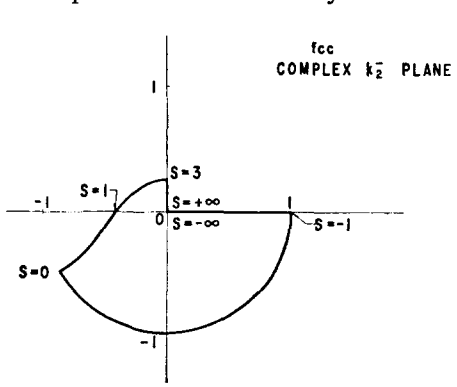

FIG. 6. The values of k_2^+ for $t = s - i\epsilon$, $-\infty < s < \infty$.

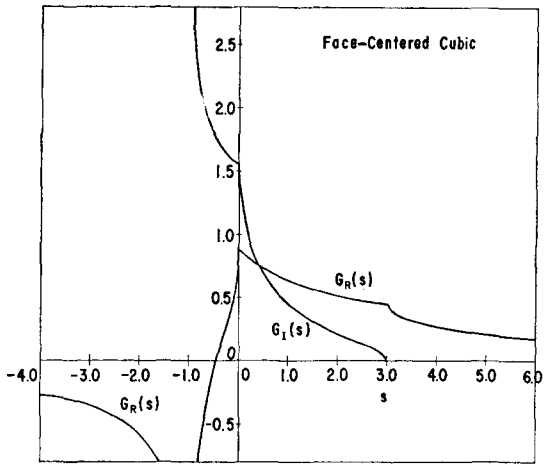

FIG. 7. The Green's function for the fcc lattice. $G_R(s)$ and $G_I(s)$ represent the real and imaginary parts, respectively.

relations connecting $k_1^{\pm}$ and $k_2^{\pm}$,

$$k_2^{\pm} = ik_1^{\pm}/(k_1^{\pm})' \quad \text{or} \quad k_1^{\pm} = ik_2^{\pm}/(k_2^{\pm})', \tag{2.24}$$

into the following formula of the complete elliptic integral:

$$\mathbf{K}(k_1^{\pm}) = \mathbf{K}(ik_2^{\pm}/(k_2^{\pm})') = (k_2^{\pm})'\mathbf{K}(k_2^{\pm}), \tag{2.25}$$

where

$$(k_j^{\pm})' = [1 - (k_j^{\pm})^2]^{\frac{1}{2}}, \quad j = 1, 2. \tag{2.26}$$

The result of the numerical computations is shown in Fig. 7. Both of the formulas (2.15) and (2.21) gives the identical result. But when one uses (2.21), he must take special care in the calculation because for that formula k becomes unity and k' is zero at $s = -1$. The accuracy of the arithmetic-geometric mean depends on the accuracy of k'.[8,9] In fact, we have obtained better results by the aid of (2.21) without special cares in the program for the computer.

3. ANALYTIC BEHAVIORS OF THE LATTICE GREEN'S FUNCTIONS

A. bcc Lattice

We present the expansions of the Green's function $G(t)$ which is given by (2.1) and (2.2) or (2.4) at $t = \infty$, $t = 0$, and $t = 1$. The expansion at $t = \infty$ is obtained by substituting (2.2) or (2.4) into the expansion[12,13]

$$\mathbf{K}(k) = \tfrac{1}{2}\pi \sum_{n=0}^{\infty} \left[\frac{(\frac{1}{2})_n}{n!}\right]^2 k^{2n} \tag{3.1}$$

and then into (2.1). Here $(\frac{1}{2})_n = \Gamma(n + \frac{1}{2})/\Gamma(\frac{1}{2})$. The result is

$$G(t) = \frac{1}{t} + \frac{1}{8}\frac{1}{t^3} + \frac{27}{512}\frac{1}{t^5} + \frac{125}{4096}\frac{1}{t^7} + O\left(\frac{1}{t^9}\right). \tag{3.2}$$

The expansion at $t = 0$ is obtained by using the formula[12,13]

$$\mathbf{K}(k) = (1/k)[\mathbf{K}(1/k) + i\mathbf{K}'(1/k)]. \tag{3.3}$$

For the first term of the right-hand side, (3.1) for $1/k$ is used. For the second term we use[12,13]

$$\mathbf{K}(k) = -\frac{2}{\pi}\mathbf{K}(k')\ln k' + \sum_{n=0}^{\infty}\left(\frac{(\frac{1}{2})_n}{n!}\right)^2 [\psi(n+1) - \psi(n+\tfrac{1}{2})]k'^{2n}, \tag{3.4}$$

where $k' = (1 - k^2)^{\frac{1}{2}}$, and $\psi(z) = \Gamma'(z)/\Gamma(z)$. Substituting (2.2) into (3.3) and then into (2.1), we obtain

$$G(s - i\epsilon) = -(2/\pi)\ln \tfrac{1}{8}s + O(s^2 \ln s) + i[(2/\pi^2)(\ln \tfrac{1}{8}s)^2 - \tfrac{1}{2} + O(s^2(\ln s)^2)]. \tag{3.5}$$

This expression is given in a preliminary report by the present authors.[11]

The expansion at $s = 1$ is obtained simply by expanding k and hence $\mathbf{K}(k)$ and $G(t)$ in powers of $(1 - 1/t^2)^{\frac{1}{2}}$.

More detailed analysis of the lattice Green's function for the bcc lattice is presented by Katsura and Horiguchi.[4]

B. fcc Lattice

The expansion of the lattice Green's function for the fc lattice is derived around the singular points $t = \infty$ and $t = -1$.

The expansion at $t = \infty$ is obtained by expanding $\mathbf{K}(k_1^{\pm})$ in (2.10) or $\mathbf{K}(k_2^{\pm})$ in (2.12) in powers of $k_1^{\pm}$ or $k_2^{\pm}$, where the formula (3.1) is used, and then by expanding $k_1^{\pm}$ or $k_2^{\pm}$ in powers of $1/t$. The result is

$$G(t) = t^{-1} + \tfrac{3}{4}t^{-3} + \tfrac{3}{4}t^{-4} + O(t^{-5}). \tag{3.6}$$

The expansion at $t = -1$ is obtained by using (3.4) into (2.12) and then expanding $k_2^{\pm}$ in powers of $(-1 - t)^{\frac{1}{2}}$. As a result, one obtains

$$G(s - i\epsilon) = -\frac{3}{4\pi^2}\left[\ln\left(\frac{1+s}{16}\right)\right]^2 + \frac{3}{4} + O((1+s)[\ln(1+s)]^2) + i\left(-\frac{3}{2\pi}\ln\left(\frac{1+s}{16}\right) + O((1+s)[\ln(1+s)]^2)\right) \tag{3.7}$$

for $s \geqslant -1$ and

$$G(s - i\epsilon) = -\frac{3}{4\pi^2}\left[\ln\left(\frac{|1+s|}{16}\right)\right]^2 + O((1+s)[\ln|1+s|]^2) \tag{3.8}$$

for $s \leqslant -1$. The same result is attained with the aid of (2.10) with (2.12). In that case, one uses the formula (3.3).

The expansions at $s = 0$ and 3 are obtained simply by expanding $\mathbf{K}(k_1^{\pm})$ or $\mathbf{K}(k_2^{\pm})$ at the respective points, in powers of $t^{\frac{1}{2}}$ and $(t-3)^{\frac{1}{2}}$, respectively.

C. Discussions of the Divergences

As shown in Figs. 2 and 7 and Eqs. (3.5), (3.7), and (3.8), divergences are observed at $s = 0$ for the bcc lattice and at $s = -1$ for the fcc lattice, respectively.

According to van Hove,[10] the divergences in the 3-dimensional lattice must be related to the critical point where the first derivative of $\omega(\mathbf{k})$ vanishes and also where the determinant of the second derivative, that is the Hessian, is zero:

$$\frac{\partial\omega(\mathbf{k})}{\partial\mathbf{k}} = 0, \quad \left|\frac{\partial^2\omega(\mathbf{k})}{\partial k_i \partial k_j}\right| = 0, \tag{3.9}$$

where $\mathbf{k} = (x, y, z)$, $k_1 = x$, $k_2 = y$, and $k_3 = z$.

In fact, for the present choice of $\omega(\mathbf{k})$, we find that both of these relations are satisfied at an arbitrary point on the singular lines on which two of the three components of $\mathbf{k}$ are $\frac{1}{2}\pi$ for the bcc lattice and at an arbitrary point on the singular lines on which two of the components are π and 0, respectively, for the fcc lattice. We confirm that $\omega(\mathbf{k})$ on these lines are equal to zero and minus unity, respectively; these values of $\omega(\mathbf{k})$ correspond to the center of the band of the bcc lattice and the top of the band of the fcc lattice. For the bcc lattice, $\omega(\mathbf{k})$ is equal to zero on the planes where one of the components is $\frac{1}{2}\pi$, and the cross lines constitute the singular lines. For the fcc case, the plane where $\omega(\mathbf{k})$ is a little more than -1 encloses the singular lines.

We shall now consider the case where the next-neighbor interaction is also not zero. Then we find that $\partial\omega(\mathbf{k})/\partial\mathbf{k}$ still vanishes at some points on the singular lines but the Hessian becomes finite at these critical points. Thus the divergent behaviors disappear from the curves for the Green's functions. In this sense, the nearest-neighbor interaction is, in fact, a very special form of the interaction for which the singular points form the singular lines and a divergence occurs, for the cases of the bcc and fcc lattices.

4. SUMMARY

The formulas for calculating the lattice Green's function $G(t)$ are provided for the bcc and fcc lattices. For the bcc lattice, we have (2.1) with (2.4). For the fcc lattice, we have two formulas: (i) (2.16), (2.17), and (2.11) and (ii) (2.22), (2.23), and (2.13); the former one is the more advisable. The formulas are applicable to arbitrary complex values of t for which $\mathrm{Im}\, t < 0$. If $\mathrm{Im}\, t > 0$, $G(t)$ is given by $G(t^*)^*$. When t is just below the real axis, $t = s - i\epsilon$, $-\infty < s < \infty$, $(t-a)^{\frac{1}{2}}$ in (2.4), or (2.11) must be replaced by the expressions given on the right-hand side of (2.5) and $(a-t)^{\frac{1}{2}}$ in (2.13) by the ones on the right-hand side of (2.19).

The formulas for the rectangular lattice are provided in the Appendix.

ACKNOWLEDGMENTS

The authors express their sincere thanks to Professor S. Katsura for valuable discussions. The numerical computations were performed with the aid of the computer NEAC 2200 of the Computer Center Tohoku University.

APPENDIX: THE LATTICE GREEN'S FUNCTION FOR THE RECTANGULAR LATTICE

In this appendix, we consider the rectangular lattice, which includes the square lattice as a special case. For this lattice, function $G(t)$ is defined by

$$G(t) = \frac{1}{\pi^2}\int_0^{\pi} dx \int_0^{\pi} dy \frac{1}{t - \gamma\cos x - \cos y}. \tag{A1}$$

When $t > 1 + \gamma$, we have

$$G(t) = \frac{1}{\pi}\int_0^{\pi} dx[(t - \gamma\cos x)^2 - 1]^{-\frac{1}{2}} \tag{A2}$$

$$= (\pi\gamma^{\frac{1}{2}})^{-1} k_1 \mathbf{K}(k_1), \tag{A3}$$

where

$$k_1 = \{4\gamma/[t^2 - (\gamma-1)^2]\}^{\frac{1}{2}}; \tag{A4}$$

the transformation which leads (A2) to (A3) is

$$\frac{t - \gamma\cos x - 1}{t - \gamma\cos x + 1} = \frac{t + \gamma - 1}{t + \gamma + 1}(1 - k_1^2 \sin^2\theta). \tag{A5}$$

Here γ is assumed to be greater than unity without loss of generality. Now we can apply the discussion in the text directly to the above expression (A3). By (A4), the lower half of the t plane is mapped to the upper half of the k_1 plane, and $G(t)$ for an arbitrary complex t is evaluated by (A3) with (A4). $G(t)$ for $t = s - i\epsilon$, $-(\gamma-1) < s < \gamma - 1$, is calculated by replacing $(t-a)^{\frac{1}{2}}$ by the right-hand side of (2.5). Then, by using the standard formulas for the complete elliptic integral of the first kind, one obtains

$$G(t) = \frac{1}{\pi\gamma^{\frac{1}{2}}}[\mathbf{K}(1/k_1) + i\mathbf{K}(1/k_2)], \quad \gamma - 1 < s < \gamma + 1, \tag{A6}$$

and

$$G(t) = \frac{i}{\pi\gamma^{\frac{1}{2}}} k_2 \mathbf{K}(k_2), \quad 0 < s < \gamma - 1, \tag{A7}$$

where

$$k_2 = \{4\gamma[(\gamma+1)^2 - t^2]\}^{\frac{1}{2}}. \tag{A8}$$

The imaginary parts of the expressions (A6) and (A7) have already been given by Montroll.[14] The expansions of (A3), (A6), and (A7) at $s = \gamma - 1 - i\epsilon$, $\gamma + 1 - i\epsilon$, and ∞ are easily obtained with the aid of the expansion formulas (3.1) and (3.4).

From the general theory of the elliptic integrals, we find that we can, in principle, express the integrals which involve $\cos mx \cos ny$, $m, n = 0, 1, 2, \cdots$, in the numerator of (A1) in terms of the complete elliptic integrals and that then the present method is applicable to these integrals. The investigation of this problem is left as a future problem.

* Present address: Department of Physics, Ohio University, Athens, Ohio 45701.

[1] S. Katsura, T. Morita, S. Inawashiro, T. Horiguchi, and Y. Abe, J. Math. Phys. **12**, 892 (1971).
[2] S. Katsura, S. Inawashiro, and Y. Abe, J. Math. Phys. **12**, 895 (1971).
[3] T. Morita and T. Horiguchi, J. Math. Phys. **12**, 981 (1971).
[4] S. Katsura and T. Horiguchi, J. Math. Phys. **12**, 230 (1971).
[5] See A. A. Maradudin *et al.*, Acad. Roy. Belg., Classe Sci., **9**, No. 7 (1960).
[6] G. Iwata, Natural Science Rept., Ochanomizu Univ., **20**, 13 (1969).
[7] D. J. Hofsommer and R. P. van de Riet, Num. Math. **5**, 291 (1963).
[8] R. Bulirsch, Num. Math. **7**, 78 (1965).
[9] T. Morita and T. Horiguchi, unpublished.
[10] L. van Hove, Phys. Rev. **89**, 1189 (1953).
[11] T. Horiguchi and T. Morita, Phys. Letters A **32**, 191 (1970).
[12] W. Magnus and F. Oberhettinger, *Formeln und Sätze für die Mathematische Physik* (Springer-Verlag, Berlin, 1948).
[13] Bateman Manuscript Project, *Higher Transcendental Functions*, A. Erdélyi Ed. (McGraw-Hill, New York, 1953), Vol. II.
[14] E. W. Montroll, in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, edited by J. Neyman (U. of California Press, Berkeley, Calif., 1955), Vol. III, p. 209.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

AC Susceptibility of Biased One-Dimensional Stochastic Ising Model*

V. Hugo Schmidt
Montana State University, Bozeman, Montana 59715

(Received 27 September 1970)

The ac susceptibility for the one-dimensional Ising model is obtained for arbitrary coupling strength in the presence of a dc bias field strong enough to align most of the dipoles in one direction. The dipole flip probability is assumed proportional to the Boltzmann factor corresponding to half the energy change resulting from the flip. The general expression for ac susceptibility is analyzed in three limiting cases: weak coupling with strong bias, strong coupling with strong bias, and strong coupling with weak bias. In the latter case, relatively long chains of anti-aligned dipoles exist and give rise to large susceptibility.

The dynamic behavior of the one-dimensional stochastic Ising model has been studied previously by Meijer, Tanaka, and Barry[1] in the limit of weak spin coupling, and by Glauber,[2] who found the ac susceptibility for the case of zero bias field. In the present work the isothermal ac susceptibility is found for arbitrary coupling strength in the presence of a dc bias field sufficiently large to align most dipoles in one direction. The region of applicability for each of these calculations is indicated in Fig. 1. Glauber's results are exact for zero bias. The present results are complementary to his in that they become exact as the bias field becomes strong enough to align all dipoles.

The one-dimensional Ising model is applicable to materials in which significant coupling of neighboring electric or magnetic dipoles exists in only one direction. The present results are potentially useful for such materials with large bias fields applied, but of particular interest is possible application to crystals of noncentric structure in which the structure rather than an externally applied field favors alignment of electric dipoles in one direction. The relation of this model to such a crystal, lithium hydrazinium sulfate, is briefly discussed.

This model assumes electric or magnetic dipoles of moment μ aligned parallel (up; $\sigma = 1$) or antiparallel (down; $\sigma = -1$) to the dc bias field E_0. The Hamiltonian in the presence of E_0 and a time-varying field E_t is

$$H = -J \sum_l \sigma_l \sigma_{l+1} - \mu(E_0 + E_t) \sum_l \sigma_l, \qquad (1)$$

where the nearest-neighbor interaction energy parameter J is positive, tending to align adjacent dipoles.

The polarization resulting from N dipoles in a

The imaginary parts of the expressions (A6) and (A7) have already been given by Montroll.[14] The expansions of (A3), (A6), and (A7) at $s = \gamma - 1 - i\epsilon$, $\gamma + 1 - i\epsilon$, and ∞ are easily obtained with the aid of the expansion formulas (3.1) and (3.4).

From the general theory of the elliptic integrals, we find that we can, in principle, express the integrals which involve $\cos mx \cos ny$, $m, n = 0, 1, 2, \cdots$, in the numerator of (A1) in terms of the complete elliptic integrals and that then the present method is applicable to these integrals. The investigation of this problem is left as a future problem.

* Present address: Department of Physics, Ohio University, Athens, Ohio 45701.

[1] S. Katsura, T. Morita, S. Inawashiro, T. Horiguchi, and Y. Abe, J. Math. Phys. **12**, 892 (1971).
[2] S. Katsura, S. Inawashiro, and Y. Abe, J. Math. Phys. **12**, 895 (1971).
[3] T. Morita and T. Horiguchi, J. Math. Phys. **12**, 981 (1971).
[4] S. Katsura and T. Horiguchi, J. Math. Phys. **12**, 230 (1971).
[5] See A. A. Maradudin *et al.*, Acad. Roy. Belg., Classe Sci., **9**, No. 7 (1960).
[6] G. Iwata, Natural Science Rept., Ochanomizu Univ., **20**, 13 (1969).
[7] D. J. Hofsommer and R. P. van de Riet, Num. Math. **5**, 291 (1963).
[8] R. Bulirsch, Num. Math. **7**, 78 (1965).
[9] T. Morita and T. Horiguchi, unpublished.
[10] L. van Hove, Phys. Rev. **89**, 1189 (1953).
[11] T. Horiguchi and T. Morita, Phys. Letters A **32**, 191 (1970).
[12] W. Magnus and F. Oberhettinger, *Formeln und Sätze für die Mathematische Physik* (Springer-Verlag, Berlin, 1948).
[13] Bateman Manuscript Project, *Higher Transcendental Functions*, A. Erdélyi Ed. (McGraw-Hill, New York, 1953), Vol. II.
[14] E. W. Montroll, in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, edited by J. Neyman (U. of California Press, Berkeley, Calif., 1955), Vol. III, p. 209.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

AC Susceptibility of Biased One-Dimensional Stochastic Ising Model*

V. Hugo Schmidt
Montana State University, Bozeman, Montana 59715

(Received 27 September 1970)

The ac susceptibility for the one-dimensional Ising model is obtained for arbitrary coupling strength in the presence of a dc bias field strong enough to align most of the dipoles in one direction. The dipole flip probability is assumed proportional to the Boltzmann factor corresponding to half the energy change resulting from the flip. The general expression for ac susceptibility is analyzed in three limiting cases: weak coupling with strong bias, strong coupling with strong bias, and strong coupling with weak bias. In the latter case, relatively long chains of anti-aligned dipoles exist and give rise to large susceptibility.

The dynamic behavior of the one-dimensional stochastic Ising model has been studied previously by Meijer, Tanaka, and Barry[1] in the limit of weak spin coupling, and by Glauber,[2] who found the ac susceptibility for the case of zero bias field. In the present work the isothermal ac susceptibility is found for arbitrary coupling strength in the presence of a dc bias field sufficiently large to align most dipoles in one direction. The region of applicability for each of these calculations is indicated in Fig. 1. Glauber's results are exact for zero bias. The present results are complementary to his in that they become exact as the bias field becomes strong enough to align all dipoles.

The one-dimensional Ising model is applicable to materials in which significant coupling of neighboring electric or magnetic dipoles exists in only one direction. The present results are potentially useful for such materials with large bias fields applied, but of particular interest is possible application to crystals of noncentric structure in which the structure rather than an externally applied field favors alignment of electric dipoles in one direction. The relation of this model to such a crystal, lithium hydrazinium sulfate, is briefly discussed.

This model assumes electric or magnetic dipoles of moment μ aligned parallel (up; $\sigma = 1$) or antiparallel (down; $\sigma = -1$) to the dc bias field E_0. The Hamiltonian in the presence of E_0 and a time-varying field E_t is

$$H = -J \sum_l \sigma_l \sigma_{l+1} - \mu(E_0 + E_t) \sum_l \sigma_l, \qquad (1)$$

where the nearest-neighbor interaction energy parameter J is positive, tending to align adjacent dipoles.

The polarization resulting from N dipoles in a

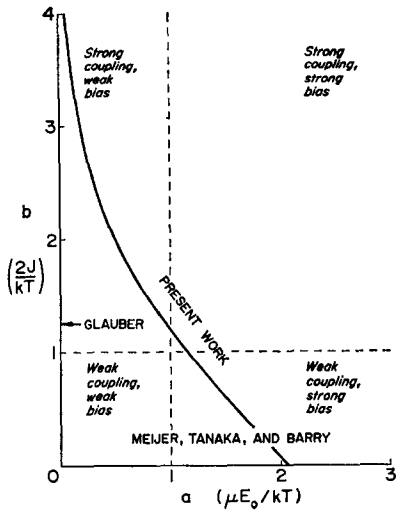

FIG. 1. Regions of applicability for studies of dynamic behavior of one-dimensional stochastic Ising model. Meijer, Tanaka, and Barry (Ref. 1) studied the weak coupling case ($b \ll 1$). Glauber (Ref. 2) found the ac susceptibility for the zero bias case ($a = 0$). In the present work the ac susceptibility given in Eq. (17) should approximate the unknown exact value in the region to the right of the curved line, because in this region the dc limit of Eq. (17) as given in Eq. (18) is within 10% of the exact dc susceptibility (Ref. 3) given in Eq. (19).

volume V is given by

$$P = \frac{\mu}{V}\left[N - 2\sum_m m(N_m + n_m)\right], \tag{2}$$

where N_m is the equilibrium number of chains each composed of m adjacent down dipoles and n_m is the deviation caused by E_t. The energy required to create a chain is $4J + 2\mu(E_0 + E_t)m$. The large-polarization approximation used herein is that the equilibrium probability of such a chain beginning at any one of the N sites is simply the Boltzmann factor $e^{-\beta(4J+2\mu E_0 m)}$, where $\beta = (kT)^{-1}$. In this approximation the relations

$$N_m = Ne^{-\beta(4J+2\mu E_0 m)}, \tag{3}$$

$$\sum mN_m = Ne^{-4\beta J}/4\sinh^2(\beta\mu E_0) \tag{4}$$

are obtained. The equilibrium polarization obtained by combining Eqs. (2) and (4) approaches the exact expression[3]

$$P = \frac{N\mu\sinh(\beta\mu E_0)}{V[\sinh^2(\beta\mu E_0) + e^{-4\beta J}]^{\frac{1}{2}}} \tag{5}$$

if the large-polarization condition $\sinh(\mu\beta E_0) \gg e^{-2\beta J}$ is fulfilled.

The equilibrium N_m in Eq. (3) can also be obtained by assuming that the flip probability per unit time for a given dipole is a thermally induced basic flip rate ν multiplied by the Boltzmann factor corresponding to half the energy change resulting from the flip. This probability was chosen with electric dipoles in mind. Even if their initial and final states have equal energy, they usually must surmount a large barrier in order to flip. Then a change U in the final state energy will change the barrier height by approximately $\frac{1}{2}U$, giving a flip rate of $\nu e^{-\beta U/2}$. Glauber[2] uses different flip probabilities, $\nu[1 - \frac{1}{2}\tanh(2\beta J)\sigma_l(\sigma_{l-1} + \sigma_{l+1})]$, which were chosen for their simplicity. As he points out, there are infinitely many choices of flip probabilities which will yield the correct equilibrium populations.

In equilibrium, the rates of creation and annihilation of isolated down dipoles must be equal, as indicated below,

$$\left.\frac{dN_1}{dt}\right|_{N_2} = N\nu e^{-\beta(2J+\mu E_0)} - N_1\nu e^{\beta(2J+\mu E_0)} = 0, \tag{6}$$

and a similar relation governs equilibrium between numbers of chains containing $m - 1$ and m down dipoles:

$$\left.\frac{dN_m}{dt}\right|_{N_{m+1}} = 2N_{m-1}\nu e^{-\beta\mu E_0} - 2N_m\nu e^{\beta\mu E_0} = 0. \tag{7}$$

(The 2 occurs because these chains can grow or shrink at either end.) Both flip probabilities in either Eq. (6) or Eq. (7) could be multiplied by the same parameter without upsetting the equilibrium, but the symmetric expressions used seem physically the most reasonable. In Eqs. (6) and (7) the use of N, rather than the more exact value $N - \sum(m + 2)N_m$, and the neglect of the effects of isolated up dipoles, are large-polarization approximations which allow these equations to be satisfied by the approximate N_m of Eq. (3).

Application of a time-dependent field E_t multiplies the transition rates in Eqs. (6) and (7) by the factors of $e^{\pm\beta\mu E_t}$, and these modified transition rates provide differential equations governing the population changes:

$$\begin{aligned}\dot{n}_1 = {} & N\nu e^{-\beta(2J+\mu E_0+\mu E_t)} - (N_1 + n_1)\nu e^{\beta(2J+\mu E_0+\mu E_t)} \\ & + 2(N_2 + n_2)\nu e^{\beta\mu(E_0+E_t)} \\ & - 2(N_1 + n_1)\nu e^{-\beta\mu(E_0+E_t)},\end{aligned} \tag{8}$$

$$\begin{aligned}\dot{n}_m = {} & 2(N_{m-1} + n_{m-1})\nu e^{-\beta\mu(E_0+E_t)} \\ & - 2(N_m + n_m)\nu e^{\beta\mu(E_0+E_t)} \\ & + 2(N_{m+1} + n_{m+1})\nu e^{\beta\mu(E_0+E_t)} \\ & - 2(N_m + n_m)\nu e^{-\beta\mu(E_0+E_t)}, \quad m > 1.\end{aligned} \tag{9}$$

Upon setting $a = \beta\mu E_0$, $b = 2\beta J$, $c = N\beta\mu e^{-4\beta J}$, making the small-field approximation $e^{\pm\beta\mu E_t} = 1 + \beta\mu E_t$, eliminating terms which cancel according to Eqs. (6) and (7), and terms of the form $n_m\beta\mu E_t$, which are of second order in E_t, and using the expression in Eq. (3) for N_m, these equations become

$$\begin{aligned}\dot{n}_1 + \nu n_1(e^{b+a} + 2e^{-a}) - 2\nu n_2 e^a \\ = -2\nu E_t c e^{-2a}(e^{b+a} - 2e^{-a}),\end{aligned} \tag{10}$$

$$\begin{aligned}\dot{n}_m + 4\nu n_m\cosh a - 2\nu n_{m-1}e^{-a} - 2\nu n_{m+1}e^a \\ = -8\nu E_t c e^{-2am}\sinh a, \quad m > 1.\end{aligned} \tag{11}$$

For a small sinusoidal applied field $E_t = E_{t0}e^{i\omega t}$, these equations become

$$(2i\Omega + \tfrac{1}{2}e^{b+a} + e^{-a})S_1 - e^{-a}S_2 = -\tfrac{1}{2}e^{b+a} + e^{-a}, \quad (12)$$

$$(2i\Omega + 2\cosh a)S_m - e^{a}S_{m-1} - e^{-a}S_{m+1} = -2\sinh a, \quad m > 1, \quad (13)$$

where $\Omega = \omega/4\nu$ and $S_m = (n_m/2c)e^{-2am}E_t$. If S_m is assumed to be of the form

$$S_m = S_0G^m - (i\Omega)^{-1}\sinh a, \quad (14)$$

then Eq. (13) is satisfied if G has the values

$$G = e^{a}\{\cosh a + i\Omega \pm [(\cosh a + i\Omega)^2 - 1]^{\frac{1}{2}}\}. \quad (15)$$

The positive sign gives S_m and n_m which increase without limit for large m and Ω, so the negative sign must be chosen for G. The constant S_0 is determined by requiring that Eq. (12) be satisfied by S_1 and S_2 as given in Eq. (14). This value of S_0 inserted into Eq. (14) provides the following simultaneous solution for Eqs. (12) and (13):

$$S_m = -\frac{\sinh a}{i\Omega} - \frac{1 - 2e^{-b} - (i\Omega)^{-1}\sinh a}{(1 - 2e^{-b})G + 2e^{-b}}G^m. \quad (16)$$

The ac susceptibility in the presence of the dc bias field E_0 is $\chi = dP/dE$, which for small E_t reduces to $(P - P_{eq})/E_t$, and from Eq. (2) $P - P_{eq} = 2\mu V^{-1} \times \sum_{m=1}^{N} mn_m$. For large N the upper limit can be allowed to become infinite. Then, from Eq. (16), the susceptibility is found to be

$$\chi = -\frac{2\mu}{VE_t}\sum_{m=1}^{\infty} mn_m$$

$$= -\frac{4c\mu}{V}\sum_{m=1}^{\infty} mS_m e^{-2am}$$

$$= \frac{4N\mu^2\beta e^{-2b}}{V}\left(\frac{1}{4i\Omega\sinh a} + \frac{1 - 2e^{-b} - (i\Omega)^{-1}\sinh a}{[1 + 2e^{-b}(G^{-1} - 1)](e^{a} - e^{-a}G)^2}\right). \quad (17)$$

This expression can be better understood by considering various limiting cases. In the low-frequency limit, $G = (1 - i\Omega/\sinh a - e^{a}\Omega^2/2\sinh^3 a + \cdots)$ to second order in Ω, and substitution of this G gives χ to first order in Ω:

$$\chi \underset{\Omega\to 0}{=} \frac{N\mu^2\beta e^{-2b}\cosh a}{V\sinh^3 a} \times \left(1 - i\Omega\frac{4 + e^{-2a} + 4e^{a-b}\sinh a}{4\sinh^2 a\cosh a} + \cdots\right). \quad (18)$$

The first term is the dc susceptibility in the presence of a bias field $E_0 = a/\beta\mu$. It agrees in the large-polarization limit ($\sinh a \gg e^{-b}$) with the exact value

$$\chi_{\text{exact}}(\Omega = 0) = \frac{N\mu^2\beta e^{-2b}\cosh a}{V(\sinh^2 a + e^{-2b})^{\frac{3}{2}}} \quad (19)$$

obtained by evaluating dP/dE using Eq. (5). The agreement is within 10% in the region to the right of the curved line in Fig. 1. This region within which the large polarization condition $\sinh(\mu E_0/kT) \gg e^{-2J/kT}$ is valid includes three cases, as indicated in Fig. 1. The weak-coupling case has $2J \ll kT$ ($b \ll 1$), so a strong bias field is required for which $\mu E_0 \gg kT$ ($a \gg 1$). In this case there are many chains of down dipoles, but most of the chains are only one dipole long. The strong-coupling, strong-bias case ($b \gg 1$, $a \gg 1$) shows the smallest deviation from maximum polarization. The strong-coupling, weak-bias case ($b \gg 1$, $e^{-b} \ll a \ll 1$) results in relatively few chains of down dipoles having long average chain length since

$$\bar{m} = \sum mN_m/\sum N_m = (e^{a}/2)\sinh a$$

varies from $1 + e^{-a}$ for $a \gg 1$, to $1/2a$ for $a \ll 1$.

To a good approximation, the susceptibility in Eq. (17) can be represented by two components χ_f and χ_s corresponding to fast and slow modes:

$$\chi \simeq \chi_{f0}/(1 + i\omega\tau_f) + \chi_{s0}/(1 + i\omega\tau_s). \quad (20)$$

The fast mode is governed by the frequency dependence of the factor $(1 - 2e^{-b} + 2e^{-b}G^{-1})$ in Eq. (17). The approximation $e^{a}G^{-1} \simeq 2(\cosh a + i\Omega)$ is valid under conditions for which this mode is active. This mode results from the field-induced shift of equilibrium between creation and annihilation of isolated down dipoles ($m = 1$ "chains"). The relaxation times for this mode form a narrow distribution around $\tau_f = (\nu e^{a+b})^{-1}$ for each of the three cases. This is the only mode which exists for the weak-coupling case and is the dominant mode for the $b \gg 1$, $a \gg 1$ case. For both cases, $\chi_{f0} \simeq \chi_0$, where χ_0 is the dc susceptibility given by the first term in Eq. (18). Because $a \gg 1$ for these cases, $\cosh a \simeq \sinh a$, and the large-polarization condition $\sinh a \gg e^{-b}$ requires that $\chi_{f0} \ll N\mu^2\beta/V$.

A comparison can be made with the results of Glauber's[2] study of the zero-bias ($a = 0$) case, which of course does not satisfy the large-polarization condition. His flip probabilities are not exponential in the flip energy, but reduce to those used herein for $b \ll 1$. In this limit he finds $\chi \simeq N\mu^2\beta/V(1 + i\omega/2\nu)$, so his correlation time is $(2\nu)^{-1}$.

For the $b \gg 1$, $a \ll 1$ case, χ_{f0} falls off to $4a^3\chi_0$. The susceptibility for this weak-bias case is dominated by

the slow mode, which is caused by the shifts in populations n_m of existing chains of down dipoles. This mode has a narrow range of time constants near $(4va^2)^{-1}$ for this case, and $\chi_{s0} \simeq \chi_0 \simeq N\mu^2\beta e^{-2b}/Va^3$. The large-polarization condition requires that $e^{-2b}/a^2 \ll 1$, but if a is in the lower part of its allowed range $e^{-b} \ll a \ll 1$, it is possible to have $\chi_{s0} \gg N\mu^2\beta/V$. If the coupling were turned off ($b = 0$), the exact expressions in Eqs. (19) and (5) show that χ_{s0} would drop to $N\mu^2\beta/V$, and that the polarization would be near zero rather than near maximum. Accordingly, even in the presence of a dc field which aligns most of the dipoles, the ac susceptibility can show what Wannier[3] has termed "enhanced paramagnetism."

This work was undertaken in order to explain the unusual dielectric properties[4,5] of lithium hydrazinum sulfate, which has generally been considered to be a ferroelectric. The structure[6] contains ordered N–H · · · N–H · · · chains running along the "ferroelectric" c axis. This biased one-dimensional Ising model seems quite applicable to the N–H dipole system if E_0 is the effective field with which the noncentric structure tends to align the N–H dipoles. This model allows for polarization reversal of the N–H dipole system if E_0 can be overcome by an externally applied field smaller than the breakdown field; but it predicts no hysteresis in the dc limit and so is inconsistent with ferroelectric behavior. We have since determined that this crystal is not ferroelectric but that a mechanism other than dipolar reorientation dominates the dielectric behavior, at least below 10 MHz. This mechanism appears to be protonic conduction along the N–H · · · N–H · · · chains, with extrinsic barriers of random height partially blocking the flow. A detailed account of this mechanism will be presented elsewhere. There remains a significant difference between the susceptibilities at 10 MHz and 9.3 GHz. Experimental study of the intervening frequency region could determine whether the effective bias field E_0 and other parameters of this biased Ising model have the correct magnitudes to cause dielectric relaxation in this frequency range.

ACKNOWLEDGMENTS

Thanks are offered to Dr. J. A. Ball and Dr. R. J. Swenson for several helpful discussions, to Dr. J. E. Drumheller for reading the manuscript, to Dr. S. J. Glass for pointing out related work, and to the referee for helpful comments.

* Work supported by the National Science Foundation.

[1] P. H. E. Meijer, T. Tanaka, and J. Barry, J. Math. Phys. **3**, 793 (1962).
[2] R. J. Glauber, J. Math. Phys. **4**, 294 (1963).
[3] G. H. Wannier, *Statistical Physics* (Wiley, New York, 1966), p. 355.
[4] F. L. Howell and V. H. Schmidt, Bull. Am. Phys. Soc. **13**, 876 (1968).
[5] P. Wan, F. L. Howell, V. H. Schmidt, and J. E. Drumheller, Bull. Am. Phys. Soc. **15**, 606 (1970).
[6] V. M. Padmanabhan and R. Balasubramanian, Acta Cryst. **22**, 532 (1967).

Note added in proof: The frequency-wavenumber dependent susceptibility $\chi(q, \omega)$ has been derived for the zero bias case by M. Suzuki and R. Kubo, J. Phys. Soc. Japan **24**, 51 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Scaling Behavior of Gravitational Energy

DAVID C. ROBINSON* AND JEFFREY WINICOUR
Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio 45433

(Received 23 November 1970)

The transformation property of gravitational energy under scale transformations is presented in the context of a general asymptotically flat space-time. The relevance of scale transformations to the sign of gravitational energy is discussed. Arguments for the positive-definiteness of gravitational energy are presented and criticized.

1. INTRODUCTION

Einstein,[1] in one of his less well-known papers entitled "Demonstration of the Non-Existence of (vacuum) Gravitational Fields with a Non-Vanishing Total Mass Free of Singularities," pointed out a certain scaling property of the energy of asymptotically Schwarzschild space-times. Einstein's model of radiative space-times was oversimplified, and at present the claim made in the title of his paper is widely thought to be wrong. The energy scaling property, however, does have a general validity and, as we show in Sec. 2, can be extended to energy expressions formulated in terms of more recent concepts of asymptotic flatness.

the slow mode, which is caused by the shifts in populations n_m of existing chains of down dipoles. This mode has a narrow range of time constants near $(4va^2)^{-1}$ for this case, and $\chi_{s0} \simeq \chi_0 \simeq N\mu^2\beta e^{-2b}/Va^3$. The large-polarization condition requires that $e^{-2b}/a^2 \ll 1$, but if a is in the lower part of its allowed range $e^{-b} \ll a \ll 1$, it is possible to have $\chi_{s0} \gg N\mu^2\beta/V$. If the coupling were turned off ($b = 0$), the exact expressions in Eqs. (19) and (5) show that χ_{s0} would drop to $N\mu^2\beta/V$, and that the polarization would be near zero rather than near maximum. Accordingly, even in the presence of a dc field which aligns most of the dipoles, the ac susceptibility can show what Wannier[3] has termed "enhanced paramagnetism."

This work was undertaken in order to explain the unusual dielectric properties[4,5] of lithium hydrazinum sulfate, which has generally been considered to be a ferroelectric. The structure[6] contains ordered N–H · · · N–H · · · chains running along the "ferroelectric" c axis. This biased one-dimensional Ising model seems quite applicable to the N–H dipole system if E_0 is the effective field with which the noncentric structure tends to align the N–H dipoles. This model allows for polarization reversal of the N–H dipole system if E_0 can be overcome by an externally applied field smaller than the breakdown field; but it predicts no hysteresis in the dc limit and so is inconsistent with ferroelectric behavior. We have since determined that this crystal is not ferroelectric but that a mechanism other than dipolar reorientation dominates the dielectric behavior, at least below 10 MHz. This mechanism appears to be protonic conduction along the N–H · · · N–H · · · chains, with extrinsic barriers of random height partially blocking the flow. A detailed account of this mechanism will be presented elsewhere. There remains a significant difference between the susceptibilities at 10 MHz and 9.3 GHz. Experimental study of the intervening frequency region could determine whether the effective bias field E_0 and other parameters of this biased Ising model have the correct magnitudes to cause dielectric relaxation in this frequency range.

ACKNOWLEDGMENTS

Thanks are offered to Dr. J. A. Ball and Dr. R. J. Swenson for several helpful discussions, to Dr. J. E. Drumheller for reading the manuscript, to Dr. S. J. Glass for pointing out related work, and to the referee for helpful comments.

* Work supported by the National Science Foundation.

[1] P. H. E. Meijer, T. Tanaka, and J. Barry, J. Math. Phys. **3**, 793 (1962).
[2] R. J. Glauber, J. Math. Phys. **4**, 294 (1963).
[3] G. H. Wannier, *Statistical Physics* (Wiley, New York, 1966), p. 355.
[4] F. L. Howell and V. H. Schmidt, Bull. Am. Phys. Soc. **13**, 876 (1968).
[5] P. Wan, F. L. Howell, V. H. Schmidt, and J. E. Drumheller, Bull. Am. Phys. Soc. **15**, 606 (1970).
[6] V. M. Padmanabhan and R. Balasubramanian, Acta Cryst. **22**, 532 (1967).

Note added in proof: The frequency-wavenumber dependent susceptibility $\chi(q, \omega)$ has been derived for the zero bias case by M. Suzuki and R. Kubo, J. Phys. Soc. Japan **24**, 51 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Scaling Behavior of Gravitational Energy

DAVID C. ROBINSON* AND JEFFREY WINICOUR
Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio 45433

(Received 23 November 1970)

The transformation property of gravitational energy under scale transformations is presented in the context of a general asymptotically flat space-time. The relevance of scale transformations to the sign of gravitational energy is discussed. Arguments for the positive-definiteness of gravitational energy are presented and criticized.

1. INTRODUCTION

Einstein,[1] in one of his less well-known papers entitled "Demonstration of the Non-Existence of (vacuum) Gravitational Fields with a Non-Vanishing Total Mass Free of Singularities," pointed out a certain scaling property of the energy of asymptotically Schwarzschild space-times. Einstein's model of radiative space-times was oversimplified, and at present the claim made in the title of his paper is widely thought to be wrong. The energy scaling property, however, does have a general validity and, as we show in Sec. 2, can be extended to energy expressions formulated in terms of more recent concepts of asymptotic flatness.

Einstein used this scaling property as part of a variational argument which has many features in common with recent arguments of Brill, Deser, and Faddeev[2-4] for the positive-definiteness of gravitational energy. These latter arguments, however, are also incomplete. In Secs. 3 and 4, we discuss the implications of energy scaling for the positive-definiteness problem. We show that the arguments of Brill, Deser, and Faddeev can be strengthened, but that the final step still involves an apparently unjustified application of finite-dimensional theorems to function space.

For the purpose of brevity, we concentrate on vacuum space-times. Although all our results have straightforward generalizations to the nonvacuum case, it is in the vacuum case that the possibility of negative energy is best isolated. Physically reasonable macroscopic sources have positive-definite energy densities, but there is not even any clear-cut local concept of gravitational energy density. This paper deals with the global concept of total gravitational energy.

2. ENERGY SCALING

Einstein's derivation of the energy scaling property can be restated in the following way. Consider two Schwarzschild line elements ds_1 and ds_2 with masses m_1 and m_2, such that

$$ds_1^2 = (1 - 2m_1/r)\,dt^2 - (1 - 2m_1/r)^{-1}\,dr^2 - r^2(d\theta^2 + \sin^2\theta\,d\varphi^2)$$

and

$$ds_2^2 = (1 - 2m_2/r)\,dt^2 - (1 - 2m_2/r)^{-1}\,dr^2 - r^2(d\theta^2 + \sin^2\theta\,d\varphi^2).$$

By means of the transformation

$$t = (m_1/m_2)T, \quad r = (m_1/m_2)R,$$

we arrive at the conformal correspondence

$$ds_1 = (m_1/m_2)\,ds_2. \tag{2.1}$$

That is to say, two Schwarzschild geometries with different masses are conformally related by a constant scale factor equal to the ratio of their masses. It follows, when the asymptotic concepts involved can be legitimately defined, that the same result applies in an asymptotic sense to two asymptotically Schwarzschild space-times. Furthermore, asymptotic structure and Einstein's vacuum field equations are preserved by scale transformations such as in Eq. (2.1). Hence, given one solution with the appropriate asymptotic structure, we may construct a one-parameter family of solutions with similar asymptotic structure but with varying mass. Einstein used this result in an infinitesimal form to establish the nonvanishing of the variation of mass with geometry:

$$\frac{\delta m}{\delta g_{\mu\nu}} \neq 0, \quad \text{unless} \quad m = 0, \tag{2.2}$$

where the variation is with respect to changes in geometry preserving Einstein's equations and the asymptotic Schwarzschild structure. Clearly, the one-parameter family of scale related geometries associated with any solution is sufficient to establish this result. The exception to Eq. (2.2) occurs when $m = 0$, in which case Eq. (2.1) is no longer applicable.

More generally, consider a space-time which is asymptotically flat in the limit of null infinity, as described in the treatments of Sachs,[5] of Bondi, van der Burg, and Metzner,[6] and of Newman and Penrose.[7] Each outgoing null hypersurface with topologically spherical cross-sections determines a retarded time Σ^+ at future null infinity. According to Penrose's conformal picture of null infinity,[8] we may visualize Σ^+ to be a sphere at future null infinity. To each generator of the Bondi-Metzner-Sachs[6,9] asymptotic symmetry group at null infinity there corresponds a descriptor field ξ^μ defined on Σ^+. These can be used to form the asymptotic symmetry linkages[10,11]

$$L_\xi(\Sigma^+) = \oint_{\Sigma^+} (\xi^{[\mu;\nu]} - \xi^\rho;\, {}_\rho N^{\mu\nu})\,dS_{\mu\nu}, \tag{2.3}$$

where $N^{\mu\nu}$ is the unit bivector normal to Σ^+. In particular, for time translational generators this procedure correctly determines the total energy $E(\Sigma^+)$ at retarded time Σ^+ to within the usual freedom of a Lorentz transformation. Complete details may be found in Refs. 10 and 11. Energy here corresponds to active gravitational mass in the same sense as the Schwarzschild mass. However, whereas, in the asymptotically Schwarzschild case envisaged by Einstein, the Schwarzschild mass describes the energy measured at spatial infinity, $E(\Sigma^+)$ describes the energy measured at spatial infinity minus the energy which has been radiated away to null infinity by zero-rest-mass fields prior to the retarded time Σ^+.

With each asymptotically flat solution $g_{\mu\nu}(x^\alpha)$ of the vacuum Einstein equations with energy $E(\Sigma^+)$, we now associate the one-parameter family of asymptotically flat solutions

$$g_{\mu\nu}(x^\alpha, \lambda) = \lambda^2 g_{\mu\nu}(x^\alpha) \tag{2.4}$$

with energies $E_\lambda(\Sigma^+)$. Under this scale transformation the metrical quantities involved in Eq. (2.3) transform as is customary under conformal transformations. Special consideration must be given to the scale transformation properties of the descriptor field ξ^μ.

The generators of symmetry transformations themselves possess scale transformation freedom. Conventionally, in special relativity the descriptor of, say, a unit time translation is normalized by $\xi^\mu = (1, 0, 0, 0)$ in the coordinate system x^α, in which the metric takes the standard Minkowski form $\eta_{\mu\nu}$. Then in the coordinate system $y^\alpha = \lambda^{-1}x^\alpha$ the metric takes the form $\lambda^2\eta_{\mu\nu}$, and the descriptor of a unit time translation takes the form $(\lambda^{-1}, 0, 0, 0)$. To consistently apply this convention, we must then demand that a unit time translation in the geometry $\lambda^2\eta_{\mu\nu}$ be given by $\xi^\mu = (\lambda^{-1}, 0, 0, 0)$. Applying exactly the same considerations at null infinity to asymptotically flat geometries leads us to require

$$\xi^\mu(x^\alpha, \lambda) = \lambda^{-1}\xi^\mu(x^\alpha)$$

for the descriptors of unit symmetry transformations associated with the two geometries in Eq. (2.4). This is the appropriate scale transformation law for transforming the linkage integrals if physical quantities such as energy are to remain associated with generators of unit translations in the proper way. By carrying out the scale transformation of Eq. (2.3), we now obtain

$$E_\lambda(\Sigma^+) = \lambda E(\Sigma^+), \tag{2.5}$$

which becomes, for infinitesimal scale transformations,

$$\lambda\frac{\delta E_\lambda}{\delta\lambda} = E_\lambda. \tag{2.6}$$

In a manner analogous to Eq. (2.2), this establishes the nonvanishing of the variation of energy with respect to variations in geometry which preserve Einstein's equations and asymptotic flatness,

$$\frac{\delta E}{\delta g_{\mu\nu}} \neq 0, \quad \text{unless} \quad E = 0. \tag{2.7}$$

By considering either infinitely past retarded times at future null infinity or infinitely future advanced times at past null infinity, Eqs. (2.5)–(2.7) can be extended to the energy evaluated at spatial infinity. This can also be done in terms of a spacelike approach to spatial infinity by means of the treatment of Arnowitt, Deser, and Misner.[12] However, while all these limiting processes lead to the same definition of energy at spatial infinity in stationary space-times, the conditions for their equivalence in radiative space-times has not been clearly established. Equations (2.5)–(2.7) also apply to the momenta, angular momenta, and supermomenta associated with the remaining generators of the Bondi–Metzner–Sachs group.[10,11]

3. THE ZERO ENERGY LIMIT

Brill and Deser[2,3] have shown that the energy measured at spatial infinity has a strict minimum of zero energy at flat space with respect to variations of the geometry. A similar result holds for the energy measured at null infinity.[13] Brill and Deser[2,3] and Brill, Deser, and Faddeev[4] attempted to extend this result to prove the positive-definiteness of the energy of nonsingular vacuum gravitational fields without any weak field qualifications. In this effort, they derived variational results equivalent to the application of Eq. (2.7) to the constraint problem of a spacelike Cauchy hypersurface.[14] They established that the energy has no critical points as a functional of the geometry except a strict minimum of zero energy at flat space. In one dimension, a smooth function with a strict minimum at the value zero and no other critical points is strictly positive except at the minimum. It was conjectured that this result could be extended to the function space of the energy functional. However, even in two dimensions, counterexamples to this conjecture were constructed by Geroch (see Appendix A).

On the other hand, Eq. (2.6) does provide a one-parameter path of geometries along which the energy increases monotonically for increasing λ. This feature could be used to establish the positive-definiteness of the energy provided that this path approached the minimum at flat space in a suitable way in the limit $\lambda \to 0$. The limit obtained directly from Eq. (2.4) by keeping the space–time points fixed as $\lambda \to 0$ is clearly pathological. However, as Geroch[15] has emphasized, a great deal of ambiguity exists in the possible limits of a one-parameter family of space–times. This ambiguity corresponds to the freedom in the way points in neighboring space–times are associated as the parameter varies. By such considerations, Geroch has shown that one possible limit for the Schwarzschild geometry as $m \to 0$ is Minkowski space with no singular points. This is achieved by displacing the singular regions of the Schwarzschild manifold to infinity as the limit is taken. Since varying the Schwarzschild mass corresponds to a scale transformation as in Eq. (2.1), it is conceivable that for some choice of limiting operation all one-parameter families of nonsingular vacuum geometries generated by scale transformations approach flat space as $\lambda \to 0$.

However, even the validity of this plausible conjecture would not be sufficient to complete the argument for positive-definiteness. The proof that $E(\Sigma^+)$ has a minimum at flat space relies upon Σ^+ being a topologically spherical surface at null or spatial infinity which

surrounds a nonsingular interior. This property of Σ^+ is necessary in order to apply the Gauss theorem to convert the two-dimensional energy integral at infinity into a three-dimensional integral over a null or spacelike hypersurface spanning Σ^+. In order to apply the minimum at flat space to the energy of a particular one-parameter family of geometries approaching flat space, this property of Σ^+ must be maintained in the limit. This is not true of the above-mentioned flat space limit of the Schwarzschild geometry, in which Σ^+ is missing points at infinity, to which the singularity is displaced in the limit $m \to 0$.

We now show more generally that this property of Σ^+ is not maintained in the scaling limit $\lambda \to 0$. Consider the Weyl tensor scalar

$$S = C^{\mu\nu\rho\sigma}C_{\mu\nu\rho\sigma}$$

associated with the asymptotically flat vacuum metric $g_{\mu\nu}(x^\alpha)$. To each scaled metric $g_{\mu\nu}(x^\alpha, \lambda)$ given by Eq. (2.4) corresponds the analogous invariant S_λ, which satisfies

$$S_\lambda = \lambda^{-4}S.$$

In the manifold M with geometry $g_{\mu\nu}$, let σ be any spacelike hypersurface spanning Σ^+. We form the invariant integral

$$I = \int_\sigma S d^3V,$$

where d^3V is the usual scalar volume element. This integral converges for any nonsingular asymptotically flat space–time. In each member of the one-parameter family of scale related manifolds M_λ, we repeat this construction on the hypersurface σ_λ corresponding conformally to the original hypersurface σ. From the scale behavior of volume elements

$$d^3V_\lambda = \lambda^3 d^3V,$$

we find

$$I_\lambda = \lambda^{-1}I.$$

In the limit $\lambda \to 0$, the invariant diverges and hence all spanning spacelike hypersurfaces develop singularities. The null hypersurface N spanning Σ^+ (when Σ^+ is at null infinity) can be treated in a similar way by introducing the scalar measure

$$d^3M = \rho m^\mu \, dV_\mu,$$

where dV_μ is the usual vector volume element on N, ρ is the divergence of the null vector field k_μ generating N, and m^μ is an arbitrary vector field pointing out of N whose extension is fixed by

$$k_\mu m^\mu = 1. \tag{3.1}$$

The definition of d^3M is insensitive to the remaining freedom in m^μ. We construct the invariant integral

$$J = \int_N S d^3M,$$

where the integration over N extends from the outermost caustic along each generator out to Σ^+. Again it is easy to check that the integral converges for any nonsingular asymptotically flat space–time. We repeat this construction in the one-parameter family of manifolds M_λ. The normalization given by Eq. (3.1) is maintained in this process. This leads to the scale behavior

$$d^3M_\lambda = \lambda^2 d^3M$$

and

$$J_\lambda = \int_{N_\lambda} S_\lambda d^3M_\lambda = \lambda^{-2}J.$$

Again in the limit $\lambda \to 0$ the invariant diverges, and hence all spanning hypersurfaces develop singularities. We conclude that the one-parameter family of scale related geometries cannot approach flat space in a way which can be used to establish the positive-definiteness of the energy.

4. A FINITE-DIMENSIONAL ARGUMENT

Although the one-parameter family of scale-related geometries cannot be used to construct a one-dimensional variational argument for the positive-definiteness of gravitational energy, it does provide the basis for an argument resting upon multidimensional techniques. This would establish positive-definiteness provided certain finite-dimensional results were applicable to the function space of geometries. We first proceed as if this function space were finite dimensional and then analyze the shortcomings of this assumption.

Suppose G were an asymptotically flat vacuum geometry with negative energy

$$E(G) < 0. \tag{4.1}$$

Let F denote flat space, so that

$$E(F) = 0. \tag{4.2}$$

Consider the set of paths in function space connecting F to G. Along each path the energy attains a maximum. A lower bound of zero for the maximum is provided by flat space. Let m be the greatest lower bound of the energy maximum on the set of all paths from F to G. Because there is a strict local minimum at flat space, we have (in the finite dimensional case) $m > 0$. Although this greatest lower bound may not be attained by any particular path, for arbitrarily small positive ϵ there exists a path whose maximum energy is $(1 + \epsilon)m$. According to the assumption in Eq. (4.1) this maximum must be an interior point of the path.

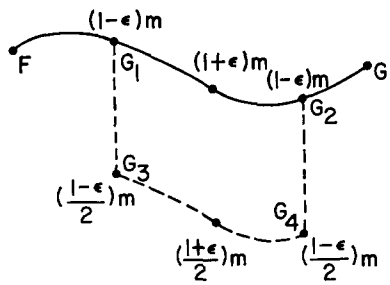

FIG. 1. The deformed path is obtained by replacing the segment $[G_1, G_2]$ by the segments $[G_1, G_3]$ along which λ varies from 1 to $\frac{1}{2}$, $[G_3, G_4]$ along which $\lambda = \frac{1}{2}$, and $[G_4, G_2]$ along which λ varies from $\frac{1}{2}$ to 1.

We can surround it with points G_1 and G_2 whose energy is $(1 - \epsilon)m$ such that $(1 - \epsilon)m$ is the maximum energy in the end segments $[F, G_1]$ and $[G_2, G]$. We now construct a new path from F to G by scale transforming the interior segment $[G_1, G_2]$, with the scale factor λ ranging from 1 to $\frac{1}{2}$. The deformed path is illustrated in Fig. 1. It has maximum energy $(1 - \epsilon)m$. But m was the greatest lower bound on the energy maxima so that we have reached a contradiction.

If our problem were really a finite-dimensional one, we could conclude that the initial assumption of negative energy in Eq. (3.1) is incompatible with the properties of a nonsingular asymptotically flat vacuum space–time. However, there is one critical aspect of the above argument which depends upon finite-dimensional analysis and cannot be extended to function space in any straightforward way. This is the intermediate conclusion $m > 0$ drawn from the existence of the strict local minimum at F. Let (t, α) coordinatize points in the neighborhood of F, with F given by the parameter $t = 0$ and the set α representing the directions emanating from F. Along each path $\alpha = \text{const}$ there is a number t_α such that the energy is positive for $t \leq t_\alpha$. In the finite-dimensional case, the set α is compact and consequently the set $\{t_\alpha\}$ possesses a nonzero minimum. But, in function space, α is a noncompact set and the possibility arises that zero is the greatest lower bound of the set $\{t_\alpha\}$.

The possibility $m = 0$ prohibits drawing any conclusions from our finite-dimensional argument. The importance of this argument is that it reduces the question of positive-definiteness to an examination of the properties of some neighborhood of flat space. The results of this paper are quite insensitive to any particular choice of topological structure on the function space of geometries. It remains to be seen whether some proper choice leads to a nonzero lower bound for the energy on a small surface surrounding flat space.

ACKNOWLEDGMENT

We wish to thank R. Geroch for informing us of the relevance of Ref. 1 as well as for many helpful ideas.

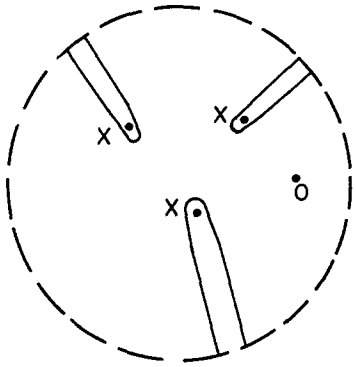

FIG. 2. The infinite plane is represented as the interior of a circle. The critical points X are removed by deleting the closed regions bounded by solid lines from the plane.

APPENDIX

Suppose a smooth function in n-dimensional Euclidean space satisfies the following four properties:

(1) $f(0) = 0$;
(2) $f_{,i}(0) = 0$;
(3) $f_{,ij}(0)x^i x^j > 0$ unless $x^j = 0$;
(4) $f_{,i}(x^j) \neq 0$ unless $x^i = 0$.

Then only in the case $n = 1$ can one conclude $f(x^i) \geq 0$. The generic method of constructing counterexamples due to Geroch is best illustrated in two dimensions. The Euclidean plane can be represented by the interior of the unit circle. It is easy to construct functions with negative regions if we allow other critical points X besides the strict local minimum at 0. The points X can be surrounded by closed regions of arbitrarily small size which can be removed from the plane without changing its topology (see Fig. 2). The remaining manifold contains no critical points except 0 but still contains negative regions. Since the remaining manifold is diffeomorphic to the infinite Euclidean plane, this establishes the counterexample.

Note that if we replace property (4) by the condition that $f(x^i)$ be homogeneous of degree 2 in the variables x^i, then the result $f(x^i) \geq 0$ holds in n dimensions. This is the essence of the finite-dimensional argument in Sec. 4.

* National Academy of Sciences Postdoctoral Research Associate.

[1] A. Einstein, Univ. Nacl. Tucuman Rev. **A2**, 11 (1941).
[2] D. Brill and S. Deser, Phys. Rev. Letters **20**, 75 (1968).
[3] D. Brill and S. Deser, Ann. Phys. **50**, 548 (1968).
[4] D. Brill, S. Deser, and L. Faddeev, Phys. Letters **26A**, 538 (1968).
[5] R. K. Sachs, Proc. Roy. Soc. (London) **A264**, 309 (1961).
[6] H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner, Proc. Roy. Soc. (London) **A270**, 103 (1962).
[7] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).
[8] R. Penrose, Proc. Roy. Soc. (London) **A284**, 159 (1965).
[9] R. K. Sachs, Phys. Rev. **128**, 2851 (1962).
[10] L. Tamburino and J. Winicour, Phys. Rev. **150**, 1039 (1966).
[11] J. Winicour, J. Math. Phys. **9**, 861 (1968).
[12] R. Arnowitt, S. Deser, and C. W. Misner, Phys. Rev. **116**, 1322 (1959).
[13] D. C. Robinson and J. Winicour, preprint. Preliminary results reported at the 1969 Relativity Conference in the Midwest appear in *Relativity*, M. Carmeli, S. I. Fickler, and L. Witten, Eds. (Plenum, New York, 1970).
[14] In Refs. 2 and 3, the qualification "unless $E = 0$" in Eq. (2.7) is reduced to "except at flat space" by assuming minimal surface properties for the Cauchy hypersurface.
[15] R. Geroch, Commun. Math. Phys. **13**, 180 (1969).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Infinite-Spin Limit of the Quantum Heisenberg Model*

KENNETH MILLARD† AND HARVEY S. LEFF
Department of Physics, Case Western Reserve University, Cleveland, Ohio 44106

(Received 16 October 1970)

The canonical partition function $Q(\mathbf{H}, N, s)$ for a spin-s anisotropic Heisenberg model with N sites, in external magnetic field $\mathbf{H}$, is examined in the limit $s \to \infty$. The coupling coefficients and the magnetic moment per spin are taken to be inversely proportional to s^2 and s, respectively. It is proved that

$$\lim_{s\to\infty} s^{-N} Q(\mathbf{H}, N, s) = (2\pi)^{-N} Q(\mathbf{H}, N),$$

where $Q(\mathbf{H}, N)$ is the partition function for a corresponding classical Heisenberg model with spins of unit magnitude. This theorem makes precise the widely believed, but heretofore unproved equivalence between the classical Heisenberg model and the infinite-spin limit of the quantum Heisenberg model. No appeal to the thermodynamic limit is necessary.

I. INTRODUCTION

Classical Ising and Heisenberg systems have been the subject of numerous investigations in recent years.[1-3] It is commonly believed that these continuum models represent, in some sense, the infinite spin limits of their quantum mechanical counterparts. Fisher[1] has given a precise definition of the infinite-spin limit, and has discussed the limiting forms of the spin operator commutation relations and of the Hamiltonian operator for an isotropic Heisenberg model. Using a similar definition, Griffiths[2] has pointed out that the infinite-spin limit of the canonical partition function for the quantum Ising model (with certain necessary multiplicative factors) is equal to the classical Ising partition function. This observation is based on the definition of a Riemann integral as the limit of its approximating series. The proof is direct because the eigenvalues of the Ising Hamiltonian are known explicitly, so that the partition function can be written immediately in a useful series form.

For the anisotropic Heisenberg model, the eigenvalues are not known explicitly, and an immediate useful series representation of the partition function does not seem possible. Thus, while it is clearly suggestive that the classical, continuum model is related to the infinite-spin limit of the corresponding quantum mechanical model, a simple proof is not apparent. To our knowledge, no rigorous proof, simple or complex, has been published with regard to the general, anisotropic Heisenberg model. The objective of this paper is to present such a proof.

II. STATEMENT OF THE THEOREM

Consider the Hamiltonian

$$\mathcal{H}_N = -\tfrac{1}{2} \sum_{i=x,y,z} \sum_{k\neq l}^{N} J_{i,kl} \frac{\mathcal{A}_{i,k}}{s} \frac{\mathcal{A}_{i,l}}{s} - \mu \sum_{i=x,y,z} \sum_{k=1}^{N} H_i \frac{\mathcal{A}_{i,k}}{s}, \tag{1}$$

for a system containing N spins, with spin quantum number $s \geq \frac{1}{2}$, on some lattice, in one, two, or three dimensions. $\mathcal{A}_{i,k}$ represents the ith component of the spin operator for site k, and $\mathbf{H} = (H_x, H_y, H_z)$ represents a uniform external magnetic field. $(J_{i,kl}/s^2)$ and (μ/s) are, in effect, a typical interspin coupling constant and the magnetic moment per spin, respectively. The inverse s dependence was introduced first by Fisher[1] and was used subsequently by Griffiths[2] with regard to the Ising Model. It assures the existence of $s^{-N}Q(\mathbf{H}, N, s)$ in the infinite-spin limit, where $Q(\mathbf{H}, N, s)$ is the canonical partition function, defined by

$$Q(\mathbf{H}, N, s) = \mathrm{Tr} \exp(-\beta \mathcal{H}_N). \tag{2}$$

By convention, the spin label s in $Q(\mathbf{H}, N, s)$ denotes that this is the *quantum* partition function.

For the corresponding classical Heisenberg model, the object $\mathcal{A}_{i,k}/s$ in Eq. (1) is interpreted as the ith component of the spin vector for site k. For convenience, each such vector is chosen to have unit magnitude; i.e., $\mathcal{A}_{i,k}/s \to \mathrm{s}_{i,k}$, with

$$|\mathbf{s}_k| = 1. \tag{3}$$

The canonical partition function for the classical case is

$$Q(\mathbf{H}, N) = \int_\Omega d\mathbf{s}_1 \cdots \int_\Omega d\mathbf{s}_N \exp(-\beta \mathfrak{H}_N), \tag{4}$$

where each integral runs over the solid angle $\Omega = 4\pi$ steradians. Introducing spherical coordinates at each site,

$$\mathrm{s}_{x,k} = \sin\theta_k \cos\phi_k, \tag{5a}$$

$$\mathrm{s}_{y,k} = \sin\theta_k \sin\phi_k, \tag{5b}$$

$$\mathrm{s}_{z,k} = \cos\theta_k, \tag{5c}$$

we see that the classical Hamiltonian $\mathfrak{H}_N$ is simply

$$\mathfrak{H}_N = -\tfrac{1}{2} \sum_{i=x,y,z} \sum_{k\neq l}^{N} J_{i,kl} \mathrm{s}_{i,k} \mathrm{s}_{i,l} - \mu \sum_{k=1}^{N} \mathbf{H} \cdot \mathbf{s}_k. \tag{6}$$

Each integral in Eq. (4) is of the form

$$\int_\Omega d\mathbf{s}_k(\,) = \int_0^{2\pi} d\phi_k \int_0^{\pi} d\theta_k \sin\theta_k(\,). \tag{7}$$

Theorem: For the quantum and classical anisotropic Heisenberg systems defined by Eqs. (1)–(6),

$$\lim_{s\to\infty} s^{-N} Q(\mathbf{H}, N, s) = (2\pi)^{-N} Q(\mathbf{H}, N). \tag{8}$$

III. PROOF OF THE THEOREM

We choose to evaluate the trace in (2) using the basis whose elements are the direct products of single-site eigenvectors of $\mathcal{S}_z$ and $(\mathcal{S}_x^2 + \mathcal{S}_y^2 + \mathcal{S}_z^2)$. Let $|s_{z,k}\rangle$ be the normalized eigenvector corresponding to site k. It has simultaneous eigenvalues $s_{z,k}$ and $s(s+1)$, but the s label is suppressed for simplicity. The chosen basis for our N site system is the orthonormal set of vectors

$$|\{s_{z,k}\}\rangle = \prod_k |s_{z,k}\rangle, \tag{9}$$

for all possible choices of the eigenvalues $\{s_{z,k}\}$ and fixed s.

Raising and lowering operators are now introduced as follows:

$$\mathcal{S}_{\pm,k} = \mathcal{S}_{x,k} \pm i\mathcal{S}_{y,k} \tag{10}$$

or, equivalently,

$$\mathcal{S}_{x,k} = \tfrac{1}{2}(\mathcal{S}_{+,k} + \mathcal{S}_{-,k}), \tag{11a}$$

$$\mathcal{S}_{y,k} = (2i)^{-1}(\mathcal{S}_{+,k} - \mathcal{S}_{-,k}). \tag{11b}$$

The operation of these operators on the basis elements gives the following results:

$$\mathcal{S}_{z,k}|s_{z,k}\rangle = s_{z,k}|s_{z,k}\rangle, \tag{12a}$$

$$\mathcal{S}_{\pm,k}|s_{z,k}\rangle = [s(s+1) - s_{z,k}(s_{z,k}\pm 1)]^{\frac{1}{2}}|s_{z,k}\pm 1\rangle. \tag{12b}$$

Note that $s_{z,k}$ can take on the $(2s+1)$ values s, $s-1, \cdots, -s$. The Hamiltonian, defined by (1), can now be written entirely in terms of these three operators:

$$\begin{aligned}\mathcal{H}_N = -\frac{1}{2}\sum_{k\neq l}^{N}\Bigg[&\frac{1}{4}J_{x,kl}\left(\frac{\mathcal{S}_{+,k}}{s}+\frac{\mathcal{S}_{-,k}}{s}\right)\left(\frac{\mathcal{S}_{+,l}}{s}+\frac{\mathcal{S}_{-,l}}{s}\right)\\ &-\frac{1}{4}J_{y,kl}\left(\frac{\mathcal{S}_{+,k}}{s}-\frac{\mathcal{S}_{-,k}}{s}\right)\left(\frac{\mathcal{S}_{+,l}}{s}-\frac{\mathcal{S}_{-,l}}{s}\right)\\ &+J_{z,kl}\frac{\mathcal{S}_{z,k}}{s}\frac{\mathcal{S}_{z,l}}{s}\Bigg]-\mu\sum_{k=1}^{N}\Bigg[\frac{1}{2}H_x\left(\frac{\mathcal{S}_{+,k}}{s}+\frac{\mathcal{S}_{-,k}}{s}\right)\\ &-\frac{i}{2}H_y\left(\frac{\mathcal{S}_{+,k}}{s}-\frac{\mathcal{S}_{-,k}}{s}\right)+H_z\frac{\mathcal{S}_{z,k}}{s}\Bigg].\end{aligned} \tag{13}$$

The operators in (13) satisfy the following commutation relations:

$$\left[\frac{\mathcal{S}_z}{s}, \frac{\mathcal{S}_\pm}{s}\right] = \pm\frac{1}{s}\frac{\mathcal{S}_\pm}{s}, \tag{14a}$$

$$\left[\frac{\mathcal{S}_+}{s}, \frac{\mathcal{S}_-}{s}\right] = \frac{2}{s}\frac{\mathcal{S}_z}{s}. \tag{14b}$$

We now note the following property: For *any* product Π_l of the operators $\mathcal{S}_z/s$ and $\mathcal{S}_\pm/s$, $k = 1, \cdots, N$, containing *at most* a total of $2l$ factors we have

$$|\langle\{s_{z,k}\}|\,\Pi_l\,|\{s_{z,k}\}\rangle| \leq [2(s+1)/s]^l \leq 6^l. \tag{15}$$

The validity of (15) follows immediately from (12) for $s \geq \frac{1}{2}$. From (13) it is clear that the highest-order product operator contained in the Hamiltonian is $2l = 2$. Therefore, (13) and (15) together imply that

$$|\langle\{s_{z,k}\}|\,\mathcal{H}_N\,|\{s_{z,k}\}\rangle| \leq 6W(N), \tag{16}$$

where

$$W(N) = \tfrac{1}{2}\sum_{i=x,y,z}\sum_{k\neq l}^{N}|J_{i,kl}| + \mu\sum_{i=x,y,z}\sum_{k=1}^{N}|H_{i,k}|. \tag{17}$$

The quantity $W(N)$ is useful in obtaining upper and lower bounds for $\langle\{s_{z,k}\}|\,(\mathcal{H}_N)^l\,|\{s_{z,k}\}\rangle$, as developed below.

We assume all the coupling constants and the magnetic field in (1) are finite. Then, for finite N, the quantity $W(N)$ must also be finite and, of course, independent of s. This assumption is the only restriction placed on the coupling constants and magnetic field in the proof of the stated equivalence theorem.

We now obtain upper and lower bounds for $s^{-N}Q(\mathbf{H}, N, s)$. First, we expand the exponential[4] in (2):

$$\begin{aligned}&s^{-N}Q(\mathbf{H}, N, s)\\ &\quad= s^{-N}\sum_{\{s_{z,k}\}}\langle\{s_{z,k}\}|\exp(-\beta\mathcal{H}_N)|\{s_{z,k}\}\rangle\\ &\quad= s^{-N}\sum_{\{s_{z,k}\}}\sum_{l=0}^{\infty}\frac{1}{l!}\langle\{s_{z,k}\}|\,(-\beta\mathcal{H}_N)^l\,|\{s_{z,k}\}\rangle.\end{aligned} \tag{18}$$

The first summation in (18) is over all possible values of the eigenvalue set $\{s_{z,k}\}$. There are $(2s+1)^N$ terms in this sum, where each $s_{z,k}$ can take on values $s, s-1, \cdots, -s$.

We now examine a term of the form

$$\langle\{s_{z,k}\}|\,(-\beta\mathcal{H}_N)^l\,|\{s_{z,k}\}\rangle.$$

We can think of expanding $(\mathcal{H}_N)^l$, using (13), so that we obtain a sum of terms consisting of a purely real coefficient $\Gamma_r^{(j)}$ (or a purely imaginary coefficient, $\Gamma_i^{(j)}$) multiplied by a product operator $\Pi_l^{(j)}$.

Symbolically, the expansion of $(\mathcal{H}_N)^l$ is of the form

$$(\mathcal{H}_N)^l = \sum_j \Gamma_r^{(j)} \Pi_l^{(j)} + \sum_k \Gamma_i^{(k)} \Pi_l^{(k)}. \tag{19}$$

The product operator [as in (15)] is a product of the operators $\mathcal{S}_{z,k}/s, \mathcal{S}_{+,k}/s, \mathcal{S}_{-,k}/s, k = 1, 2, \cdots, N$, consisting of at most $2l$ factors. Note that if in the above expression each operator product is replaced by unity and $\Gamma_r^{(j)}$ and $\Gamma_i^{(k)}$ are replaced by their magnitudes, the resulting expression is

$$\sum_j |\Gamma_r^{(j)}| + \sum_k |\Gamma_i^{(k)}| = [W(N)]^l, \tag{20}$$

where $W(N)$ is defined by (17). This identity is useful in what follows.

We observe two properties of the expectation value, $\langle\{s_{z,k}\}| (-\beta\mathcal{H}_N)^l |\{s_{z,k}\}\rangle$, where $(\mathcal{H}_N)^l$ is given by (19):

(i) Only product operators with equal numbers of raising and lowering operators for *each* site can be nonzero.

(ii) The expectation value of the second term on the right-hand side of (19) is zero.

Property (i) follows from (12) plus the orthonormality of the basis vectors. Property (ii) is obtained from the following argument. Since $\mathcal{H}_N$ [and therefore $(\mathcal{H}_N)^l$] is Hermitian, the expectation value of $\mathcal{H}_N^l$ must be real. But, by (12), the expectation value of any product operator is real. We thus conclude [property (ii)] that the sum of all terms with purely imaginary coefficients must be zero.

We define an "ordered product" of operators in the following way. Let Π_l be a product of at most $2l$ operators $\mathcal{S}_{\pm,k}/s$ and $\mathcal{S}_{z,k}/s$, $1 \leq k \leq N$, containing an equal number of $\mathcal{S}_{+,k}$ and $\mathcal{S}_{-,k}$ operators for each k. The ordered product $(\Pi_l)_{\text{ord}}$ is defined such that, for each k, a raising operator $\mathcal{S}_{+,k}/s$ must be followed *immediately* to its left by a lowering operator $\mathcal{S}_{-,k}/s$. The order in which these *pairs* appear is, of course, irrelevant. It is always possible to write

$$\Pi_l = (\Pi_l)_{\text{ord}} + \Lambda_l, \tag{21}$$

where Λ_l contains all terms resulting from the use of the commutation relations (14). Since there are at most $2l$ factors making up Π_l, the ordering procedure involves fewer than $(2l)^2$ permutations. Λ_l then consists of fewer than $(2l)^2$ terms, each with a number of factors equal to one less than the number of factors contained in Π_l. Each term in Λ_l has a coefficient equal to $\pm 1/s$ or $\pm 2/s$ due to the commutation relations (14).

Now, suppose that in the expansion of $\mathcal{H}_N^l$ we focus attention on the jth term, of the form $\Gamma_r^{(j)}\Pi_l^{(j)}$, where $\Pi_l^{(j)}$ has equal numbers of raising and lowering operators for each site. We wish to bound the corresponding expectation value $\Gamma_r^{(j)} \langle\{s_{z,k}\}| \Pi_l^{(j)} |\{s_{z,k}\}\rangle$ above and below in terms of the quantity

$$\Gamma_r^{(j)} \langle\{s_{z,k}\}| (\Pi_l^{(j)})_{\text{ord}} |\{s_{z,k}\}\rangle.$$

But from the previous discussion, the magnitude of $\langle\{s_{z,k}\}| \Lambda_l^{(j)} |\{s_{z,k}\}\rangle$ is bounded above by $(2l)^2(2/s)(6)^l$. This bound is obtained since there are at most $(2l)^2$ terms in $\Lambda_l^{(j)}$, each resulting from a single application of the commutation relations (14); thus, the $(2/s)$ factor. Finally, the factor $(6)^l$ comes from (15) since each term in $\Lambda_l^{(j)}$ is a product operator with fewer than $2l$ factors. We conclude that

$$\begin{aligned}(-\beta)^l\Gamma_r^{(j)} &\langle\{s_{z,k}\}| (\Pi_l^{(j)})_{\text{ord}} |\{s_{z,k}\}\rangle - c_{l,j} \\ &\leq (-\beta)^l\Gamma_r^{(j)} \langle\{s_{z,k}\}| \Pi_l^{(j)} |\{s_{z,k}\}\rangle \\ &\leq (-\beta)^l\Gamma_r^{(j)} \langle\{s_{z,k}\}| (\Pi_l^{(j)})_{\text{ord}} |\{s_{z,k}\}\rangle + c_{l,j},\end{aligned} \tag{22a}$$

where

$$c_{l,j} = (2l)^2(2/s)(6\beta)^l |\Gamma_r^{(j)}|. \tag{22b}$$

Such an inequality must hold for each possible "surviving" term [i.e., any term which is not zero *a priori* due to (i) or (ii)] in the expansion (19).

We define the *scalar* quantity $\tilde{\Pi}_l$ which is obtained by making the replacements

$$\mathcal{S}_{+,k}/s \to \sigma_k e^{i\phi_k}, \tag{23a}$$

$$\mathcal{S}_{-,k}/s \to \sigma_k e^{-i\phi_k}, \tag{23b}$$

and

$$\mathcal{S}_{z,k}/s \to s_{z,k}/s, \tag{24}$$

where

$$\sigma_k = s^{-1}[s(s+1) - s_{z,k}(s_{z,k}+1)]^{\frac{1}{2}} \tag{25}$$

in the product operator Π_l. We then note the identity

$$\langle\{s_{z,k}\}| (\Pi_l)_{\text{ord}} |\{s_{z,k}\}\rangle = \frac{1}{(2\pi)^N} \int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N \tilde{\Pi}_l. \tag{26}$$

This equality follows because (a) Eq. (12b) implies that

$$s^{-2} \langle s_{z,k}| \mathcal{S}_{-,k}\mathcal{S}_{+,k} |s_{z,k}\rangle = \sigma_k^2 \tag{27}$$

and (b) the operator $(\Pi_l)_{\text{ord}}$ has equal numbers of raising and lowering operators for each site, and the pair operator becomes

$$(\mathcal{S}_{-,k}/s)(\mathcal{S}_{+,k}/s) \to \sigma_k^2 e^{-i\phi_k} e^{i\phi_k} = \sigma_k^2. \tag{28}$$

The preceding argument shows that expression (22) can be written as

$$\begin{aligned}(-\beta)^l\Gamma_r^{(j)}(2\pi)^{-N} &\int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N \tilde{\Pi}_l^{(j)} - c_{l,j} \\ &\leq (-\beta)^l\Gamma_r^{(j)} \langle\{s_{z,k}\}| \Pi_l^{(j)} |\{s_{z,k}\}\rangle \\ &\leq (-\beta)^l\Gamma_r^{(j)}(2\pi)^{-N} \int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N \tilde{\Pi}_l^{(j)} + c_{l,j},\end{aligned} \tag{29}$$

provided that $\Pi_l^{(j)}$ has equal numbers of raising and lowering operators for each site. In fact, (29) is valid in general. To see this, suppose that $\Pi_l^{(j)}$ does *not* have equal numbers of raising and lowering operators. Then $\langle\{s_{z,k}\}|\,\Pi_l^{(j)}\,|\{s_{z,k}\}\rangle$ is zero by property (i) above. But, if, for such a $\Pi_l^{(j)}$, the replacements (23)–(25) are made, then

$$(2\pi)^{-N}\int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N \tilde{\Pi}_l^{(j)} = 0. \tag{30}$$

This follows, since $\tilde{\Pi}_l^{(j)}$ then contains at least one factor of the form $e^{\pm i\phi_k}$, integrated from 0 to 2π. We thus conclude that expression (29) is valid in general.

Summing over all j in expression (29), we find

$$\begin{aligned}(-\beta)^l(2\pi)^{-N}\int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N\left(\sum_j \Gamma_r^{(j)}\tilde{\Pi}_l^{(j)}\right) - \sum_j c_{l,j} &\le \langle\{s_{z,k}\}|\,(-\beta\mathcal{J}\mathcal{C}_N)^l\,|\{s_{z,k}\}\rangle \\ &\le (-\beta)^l(2\pi)^{-N}\int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N \\ &\quad\times\left(\sum_j \Gamma_r^{(j)}\tilde{\Pi}_l^{(j)}\right) + \sum_j c_{l,j}. \end{aligned} \tag{31}$$

The term $\langle\{s_{z,k}\}|\,(-\beta\mathcal{J}\mathcal{C}_N)^l\,|\{s_{z,k}\}\rangle$ follows from property (ii) above. Using (20), we can write (31) as

$$\begin{aligned}(-\beta)^l(2\pi)^{-N}\int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N\left(\sum_j \Gamma_r^{(j)}\tilde{\Pi}_l^{(j)}\right) - c_l &\le \langle\{s_{z,k}\}|\,(-\beta\mathcal{J}\mathcal{C}_N)^l\,|\{s_{z,k}\}\rangle \\ &\le (-\beta)^l(2\pi)^{-N}\int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N \\ &\quad\times\left(\sum_j \Gamma_r^{(j)}\tilde{\Pi}_l^{(j)}\right) + c_l, \end{aligned} \tag{32a}$$

where

$$c_l = (2l)^2(2/s)(6\beta)^l[W(N)]^l. \tag{32b}$$

If the replacements (23)–(25) are made in the Hamiltonian (13), we obtain the scalar quantity

$$\begin{aligned}\tilde{\mathfrak{H}}_N(\{s_{z,k}\}) = &-\tfrac{1}{2}\sum_{k\neq l}^{N}[\tfrac{1}{4}J_{x,kl}\sigma_k\sigma_l(e^{i\phi_k}+e^{-i\phi_k})(e^{i\phi_l}+e^{-i\phi_l}) \\ &-\tfrac{1}{4}J_{y,kl}\sigma_k\sigma_l(e^{i\phi_k}-e^{-i\phi_k})(e^{i\phi_l}-e^{-i\phi_l}) \\ &+J_{z,kl}s^{-2}s_{z,k}s_{z,l}] \\ &-\mu\sum_{k=1}^{N}[\tfrac{1}{2}H_x\sigma_k(e^{i\phi_k}+e^{-i\phi_k}) \\ &-\tfrac{1}{2}iH_y\sigma_k(e^{i\phi_k}-e^{-i\phi_k})+H_z s^{-1}s_{z,k}]. \end{aligned} \tag{33}$$

In a manner similar to (19), the expansion of $(\tilde{\mathfrak{H}}_N)^l$ can be written symbolically as

$$(\tilde{\mathfrak{H}}_N)^l = \sum_j \Gamma_r^{(j)}\tilde{\Pi}_l^{(j)} + \sum_k \Gamma_i^{(k)}\tilde{\Pi}_l^{(k)}. \tag{34}$$

We now note that $\tilde{\mathfrak{H}}_N$ is a real function, i.e.,

$$\tilde{\mathfrak{H}}_N^* = \tilde{\mathfrak{H}}_N. \tag{35}$$

We thus conclude that $(2\pi)^{-N}\int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N(\tilde{\mathfrak{H}}_N)^l$ must be purely real. Further, in (34) the product $\tilde{\Pi}_l^{(k)}$ represents a product of complex exponentials, and the multiple ϕ-integral over each such product is real. It follows that, since $\Gamma_i^{(k)}$ is purely imaginary,

$$(2\pi)^{-N}\int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N\left(\sum_k \Gamma_i^{(k)}\tilde{\Pi}_l^{(k)}\right) = 0. \tag{36}$$

Adding this quantity to the upper and lower bounds in (32) and using (34), we find

$$\begin{aligned}(2\pi)^{-N}\int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N(-\beta\tilde{\mathfrak{H}}_N)^l - c_l &\le \langle\{s_{z,k}\}|\,(-\beta\tilde{\mathfrak{J}}\tilde{\mathcal{C}}_N)^l\,|\{s_{z,k}\}\rangle \\ &\le (2\pi)^{-N}\int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N(-\beta\tilde{\mathfrak{H}}_N)^l + c_l. \end{aligned} \tag{37}$$

Multiplying (37) by $s^{-N}/l!$, summing over l, and using (18), we find

$$\begin{aligned}(2\pi s)^{-N}\sum_{\{s_{z,k}\}}\int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N &\\ \times\exp(-\beta\tilde{\mathfrak{H}}_N) - s^{-N}\sum_{\{s_{z,k}\}}\sum_{l=0}^{\infty}(l!)^{-1}c_l &\le s^{-N}Q(\mathbf{H}, N, s) \\ &\le (2\pi s)^{-N}\sum_{\{s_{z,k}\}}\int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N \\ &\quad\times\exp(-\beta\tilde{\mathfrak{H}}_N) + s^{-N}\sum_{\{s_{z,k}\}}\sum_{l=0}^{\infty}(l!)^{-1}c_l, \end{aligned} \tag{38}$$

where the order of summation and integration have been interchanged.[5] Since there are $(2s+1)^N$ terms in the summation over $\{s_{z,k}\}$, the use of (32b) yields

$$\begin{aligned}s^{-N}\sum_{\{s_{z,k}\}}\sum_{l=0}^{\infty}\frac{1}{l!}c_l &\le \frac{2^3}{s}\left(\frac{2s+1}{s}\right)^N\sum_{l=1}^{\infty}\frac{l^2}{l!}[6\beta W(N)]^l \\ &\le \frac{2^3}{s}(4)^N\sum_{l=1}^{\infty}\frac{l^2}{l!}[6\beta W(N)]^l \\ &= \frac{2^3}{s}(4)^N 6\beta W(N)[6\beta W(N)+1]e^{6\beta W(N)}. \end{aligned} \tag{39}$$

The third line in (39) is obtained by setting $s = \frac{1}{2}$ in the factor $(2s+1)/s$. We can then conclude that, in the classical limit,

$$\lim_{s\to\infty} s^{-N}\sum_{\{s_{z,k}\}}\sum_{l=0}^{\infty}\frac{1}{l!}c_l = 0. \tag{40}$$

Taking the limit $s \to \infty$ of (38), we find that the upper and lower bounds coalesce, so that

$$\begin{aligned}\lim_{s\to\infty} s^{-N}Q(\mathbf{H}, N, s) = (2\pi)^{-N}\lim_{s\to\infty}\int_0^{2\pi} d\phi_1 \cdots \int_0^{2\pi} d\phi_N &\\ \times\sum_{\{s_{z,k}\}} s^{-N}\exp(-\beta\tilde{\mathfrak{H}}_N), \end{aligned} \tag{41}$$

where the order of integration and the finite summation have been interchanged.

We now write (33) in a slightly different form for convenience in the remaining steps:

$$\begin{aligned}\tilde{\mathfrak{H}}_N(\{s_{z,j}\}) = &-\tfrac{1}{2}\sum_{k\neq l}^{N}[J_{x,kl}\sigma_k\sigma_l\cos\phi_k\cos\phi_l \\ &+ J_{y,kl}\sigma_k\sigma_l\sin\phi_k\sin\phi_l + J_{z,kl}s^{-2}s_{z,k}s_{z,l}] \\ &- \mu\sum_{k=1}^{N}[H_x\sigma_k\cos\phi_k \\ &+ H_y\sigma_k\sin\phi_k + H_z s^{-1}s_{z,k}]. \qquad (42)\end{aligned}$$

From (17) and (42) we observe that

$$|\tilde{\mathfrak{H}}_N(\{s_{z,j}\}| \leq 2[(s+1)/s]W(N) \leq 6W(N), \qquad (43)$$

which implies the bound

$$\begin{aligned}\left|\sum_{\{s_{z,k}\}} s^{-N}\exp(-\beta\tilde{\mathfrak{H}}_N)\right| &\leq \left(\frac{2s+1}{s}\right)^N \exp[6\beta W(N)] \\ &\leq 4^N\exp[6\beta W(N)]. \qquad (44)\end{aligned}$$

We thus conclude that any term in the sequence

$$\sum_{\{s_{z,k}\}} s^{-N}\exp(-\beta\tilde{\mathfrak{H}}_N) \quad \text{as} \quad s\to\infty$$

is uniformly bounded.[6]

We now examine the limit of the sequence

$$\sum_{\{s_{z,k}\}} s^{-N}\exp(-\beta\tilde{\mathfrak{H}}_N), \quad \text{for} \quad s\to\infty.$$

To do this, we define the quantities

$$t_k = s_{z,k}/s \qquad (45a)$$

and

$$\Delta = 1/s. \qquad (45b)$$

Then,

$$\begin{aligned}\lim_{s\to\infty}\sum_{\{s_{z,k}\}} s^{-N}\exp[-\beta\tilde{\mathfrak{H}}_N(\{s_{z,k}\})] \\ = \lim_{\Delta\to 0}\sum_{\{t_k\}}\Delta^N\exp[-\beta\tilde{\mathfrak{H}}_N(\{t_k\})], \qquad (46)\end{aligned}$$

where the sum on $\{t_k\}$ is over $(2\Delta^{-1}+1)^N$ terms such that

$$t_k = 1, 1-\Delta, 1-2\Delta, \cdots, -1.$$

But (46) is just the form of an N-dimensional Riemann integral. Thus[7]

$$\begin{aligned}\lim_{s\to\infty}\sum_{\{s_{z,k}\}} s^{-N}\exp[-\beta\tilde{\mathfrak{H}}_N(\{s_{z,k}\})] \\ = \int_{-1}^{1}dt_1\cdots\int_{-1}^{1}dt_N\exp[-\beta\mathfrak{H}_N(\{t_k\})], \qquad (47)\end{aligned}$$

where

$$\begin{aligned}\mathfrak{H}_N(\{t_j\}) = &-\tfrac{1}{2}\sum_{k\neq l}^{N}[J_{x,kl}(1-t_k^2)^{\frac{1}{2}}(1-t_l^2)^{\frac{1}{2}}\cos\phi_k\cos\phi_l \\ &+ J_{y,kl}(1-t_k^2)^{\frac{1}{2}}(1-t_l^2)^{\frac{1}{2}} \\ &\times \sin\phi_k\sin\phi_l + J_{z,kl}t_k t_l] \\ &- \mu\sum_{k=1}^{N}[H_x(1-t_k^2)^{\frac{1}{2}}\cos\phi_k \\ &+ H_y(1-t_k^2)^{\frac{1}{2}}\sin\phi_k + H_z t_k]. \qquad (48)\end{aligned}$$

By defining $t_k = \cos\theta_k$, (47) becomes

$$\begin{aligned}\lim_{s\to\infty}\sum_{\{s_{z,k}\}} s^{-N}\exp[-\beta\tilde{\mathfrak{H}}_N(\{s_{z,k}\})] \\ = \int_0^{\pi}\sin\theta_1\,d\theta_1\cdots\int_0^{\pi}\sin\theta_N\,d\theta_N\exp(-\beta\mathfrak{H}_N). \qquad (49)\end{aligned}$$

We note that $\mathfrak{H}_N$ is equal to the classical Hamiltonian as given by (5) and (6).

The sequence

$$\sum_{\{s_{z,k}\}} s^{-N}\exp(-\beta\tilde{\mathfrak{H}}_N) \quad \text{as} \quad s\to\infty$$

converges [by (47)] and is uniformly bounded [by (44)], and each term in the sequence and the limit of the sequence are both Riemann integrable over the $\{\phi_k\}$ coordinates. We can then use Arzela's theorem[8] to commute the limit as $s\to\infty$ and the $\{\phi_k\}$ integrals in (41). We thus find, using (4),

$$\begin{aligned}\lim_{s\to\infty} s^{-N}&Q(\mathbf{H}, N, s) \\ &= (2\pi)^{-N}\int_\Omega d\mathbf{s}_1\cdots\int_\Omega d\mathbf{s}_N\exp(-\beta\mathfrak{H}_N) \\ &= (2\pi)^{-N}Q(\mathbf{H}, N), \qquad (50)\end{aligned}$$

which completes the proof.

* Supported in part by the United States Atomic Energy Commission.

† This work is based on portions of a dissertation submitted in partial fulfillment of the Ph.D. degree, Case Western Reserve University, 1971.

[1] M. E. Fisher, Am. J. Phys. **32**, 343 (1964).

[2] R. B. Griffiths, J. Math. Phys. **10**, 1559 (1969).

[3] There exists a substantial number of papers dealing with classical spin systems. The following partial list serves as an indication of their scope. The papers are listed chronologically and alphabetically within each year: H. A. Brown and J. M. Luttinger, Phys. Rev. **100**, 685 (1955); N. W. Dalton, Proc. Phys. Soc. **89**, 845 (1966); H. E. Stanley and T. A. Kaplan, Phys. Rev. Letters **16**, 981 (1966); P. J. Wood and G. S. Rushbrooke, *ibid.* **17**, 307 (1966); M. E. Fisher, *ibid.* **19**, 581 (1967); G. S. Joyce, Phys. Rev. **155**, 478 (1967); N. D. Mermin, J. Math. Phys. **8**, 1061 (1967); H. E. Stanley, Phys. Rev. **158**, 537 (1967); **164**, 709 (1967); C. Domb, J. Appl. Phys. **39**, 614 (1968); D. Jasnow and M. A. Moore, Phys. Rev. **176**, 751 (1968); D. Jasnow and M. Wortis, *ibid.*, 739 (1968); H. E. Stanley, Phys. Rev. Letters **20**, 150 (1968); R. L. Stephenson and P. J. Wood, Phys. Rev. **173**, 475 (1968); C. J. Thompson, J. Math. Phys. **9**, 241 (1968); D. Jasnow and M. Wortis, J. Appl. Phys. **40**, 1277 (1969); H. E. Stanley, *ibid.*, 1272 (1969); Phys. Rev. **179**, 570 (1969).

[4] A function $f(\mathcal{A})$ of a Hermitian operator $\mathcal{A}$ may be defined as follows. Let $\{|a_k\rangle\}$ represent the complete, orthonormal set of eigenvectors of $\mathcal{A}$. Then, $\langle a_k|f(\mathcal{A})|a_l\rangle = f(a_k)\delta_{kl}$ defines $f(\mathcal{A})$ in

the eigenvector basis. For an arbitrary vector $|\psi\rangle = \sum_k c_k |a_k\rangle$, one has $\langle\psi| f(\mathcal{A}) |\psi\rangle = \sum_k |c_k|^2 f(a_k)$. Now, choose $f(\mathcal{A}) = \exp(\mathcal{A})$. Then, if $|a_k|$ is finite for all k,

$$\langle\psi| \exp(\mathcal{A}) |\psi\rangle = \sum_k \sum_{l=0}^{\infty} |c_k|^2 \frac{(a_k)^l}{l!}.$$

But the k- and l-summations can be interchanged. [See Theorem 13-19 of T. Apostol, *Mathematical Analysis* (Addison-Wesley, Reading, Mass., 1957), p. 398.] Thus,

$$\langle\psi| \exp(\mathcal{A}) |\psi\rangle = \sum_{l=0}^{\infty} \frac{1}{l!} \langle\psi| (\mathcal{A})^l |\psi\rangle.$$

Choosing $\mathcal{A} = -\beta\mathcal{K}_N$, and choosing $\{|a_k\rangle\}$ to be the eigenvector set for $\{\mathcal{A}_{s,k}\}$, one obtains (18) for finite N.

[5] The interchange is valid since the series converges uniformly in the set $\{\phi_k\}$, and each term is continuous. See, for example, E. C. Titchmarsh, *The Theory of Functions* (Oxford U.P., London, 1932), p. 36.

[6] A sequence $\{f_n(x)\}$ is said to be *uniformly bounded* on a domain D if there exists an $M > 0$ such that $|f_n(x)| \leq M$ for all x in D and all $n = 1, 2, \cdots$.

[7] To obtain Eq. (47), $\tilde{\mathfrak{H}}_N$ is bounded above and below in terms of $\mathfrak{H}_N$ and a term which depends upon Δ, but which is independent of the summation variables $\{s_{z,k}\}$. From Eqs. (25) and (45) and the inequality

$$a^{\frac{1}{2}} - b^{\frac{1}{2}} \leq (a+b)^{\frac{1}{2}} \leq a^{\frac{1}{2}} + b^{\frac{1}{2}}, \quad \text{for} \quad a \geq 0, b \geq 0,$$

we obtain the bounds

$$(1 + t_k^2)^{\frac{1}{2}} - (2\Delta)^{\frac{1}{2}} \leq \sigma_k \leq (1 + t_k^2)^{\frac{1}{2}} + (2\Delta)^{\frac{1}{2}},$$

$$[(1 + t_k^2)(1 + t_l^2)]^{\frac{1}{2}} - 2\Delta^{\frac{1}{2}}(2 + \Delta^{\frac{1}{2}}) \leq \sigma_k\sigma_l \leq [(1 + t_k^2)(1 + t_l^2)]^{\frac{1}{2}} + 2\Delta^{\frac{1}{2}}(2 + \Delta^{\frac{1}{2}}).$$

Using Eqs. (17) and (48) and these bounds in Eq. (42), we obtain

$$\mathfrak{H}_N - 2\Delta^{\frac{1}{2}}(2 + \Delta^{\frac{1}{2}})W(N) \leq \tilde{\mathfrak{H}}_N \leq \mathfrak{H}_N + 2\Delta^{\frac{1}{2}}(2 + \Delta^{\frac{1}{2}})W(N).$$

The right-hand side of Eq. (46) is then bounded by

$$\lim_{\Delta\to 0} \{\exp[-2\beta\Delta^{\frac{1}{2}}(2 + \Delta^{\frac{1}{2}})W(N)]\} \Big(\sum_{\{t_k\}} \Delta^N \exp(-\beta\tilde{\mathfrak{H}}_N) \Big) \leq \lim_{\Delta\to 0} \sum_{\{t_k\}} \Delta^N \exp(-\beta\tilde{\mathfrak{H}}_N) \leq \lim_{\Delta\to 0} \{\exp[+2\beta\Delta^{\frac{1}{2}}(2 + \Delta^{\frac{1}{2}})W(N)]\} \Big(\sum_{\{t_k\}} \Delta^N \exp(-\beta\mathfrak{H}_N) \Big).$$

Proceeding to the limit, we obtain Eq. (47).

[8] *Arzela's Theorem:* Assume $\{f_n(x)\}$ is boundedly convergent on $[a, b]$ and suppose each $f_n(x)$ is Riemann-integrable on $[a, b]$. Assume also that the limit function $f(x)$ is Riemann-integrable on $[a, b]$. Then

$$\lim_{n\to\infty} \int_a^b f_n(x)\, dx = \int_a^b \lim_{n\to\infty} f_n(x)\, dx = \int_a^b f(x)\, dx.$$

See, for example, T. Apostol, Ref. 4, p. 405. See also W. W. Rogosinski, *Volume and Integral* (Interscience, New York, 1962), Sec. 5.8.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Generating Function of the N-fold Photoelectric Counting Distribution for Gaussian Light*

C. D. CANTRELL
Physics Department, Swarthmore College, Swarthmore, Pennsylvania 19081

(Received 27 August 1970)

An exact expression is obtained for the generating function of the N-fold (i.e., N-detector) photoelectron counting distribution for cross-spectrally pure Gaussian light, in terms of the onefold generating function. It is assumed that the counting times at the different detectors are all equal, but not necessarily short compared to the coherence time of the light. In cases where the onefold generating function can be calculated, a closed formula for the N-fold generating function can be obtained.

1. INTRODUCTION

Considerable experimental and theoretical effort has recently been focused on determining the joint N-fold photoelectric counting distribution for various incident optical fields.[1–8] In an N-fold (i.e., N-interval) counting experiment, one measures (or can obtain directly from the data) the probability $P_N(n_1, t_1, T_1; \cdots; n_N, t_N, T_N)$ that n_1 counts will be recorded by the first detector in a counting interval of length T_1 beginning at time t_1; $\cdots$; n_N counts by the Nth detector in the interval $[t_N, t_N + T_N]$. (Some or all of the detectors may be identical.) Multiple coincidence experiments[9–18] are essentially measurements of P_N for the special case $n_1 = n_2 = \cdots = n_N = 1$.

For one detector ($N = 1$), both the probability distribution and the factorial moments of $n = n_1$ can be calculated from the generating function

$$\mathfrak{G}_1(n; s) = \langle(1 - s)^n\rangle.$$

An N-fold generating function can be similarly defined[3,4]:

$$\mathfrak{G}_N(n_1, \cdots, n_N; s_1, \cdots, s_N) = \langle(1 - s_1)^{n_1} \cdots (1 - s_N)^{n_N}\rangle.$$

In this paper we will derive an exact expression for $\mathfrak{G}_N$ in terms of $\mathfrak{G}_1$, assuming that: (i) The light incident on the N detectors is Gaussian; (ii) the light is cross-spectrally pure; (iii) all the counting intervals are

the eigenvector basis. For an arbitrary vector $|\psi\rangle = \sum_k c_k |a_k\rangle$, one has $\langle\psi| f(\mathcal{A}) |\psi\rangle = \sum_k |c_k|^2 f(a_k)$. Now, choose $f(\mathcal{A}) = \exp(\mathcal{A})$. Then, if $|a_k|$ is finite for all k,

$$\langle\psi| \exp(\mathcal{A}) |\psi\rangle = \sum_k \sum_{l=0}^{\infty} |c_k|^2 \frac{(a_k)^l}{l!}.$$

But the k- and l-summations can be interchanged. [See Theorem 13-19 of T. Apostol, *Mathematical Analysis* (Addison-Wesley, Reading, Mass., 1957), p. 398.] Thus,

$$\langle\psi| \exp(\mathcal{A}) |\psi\rangle = \sum_{l=0}^{\infty} \frac{1}{l!} \langle\psi| (\mathcal{A})^l |\psi\rangle.$$

Choosing $\mathcal{A} = -\beta\mathcal{K}_N$, and choosing $\{|a_k\rangle\}$ to be the eigenvector set for $\{\mathcal{A}_{s,k}\}$, one obtains (18) for finite N.

[5] The interchange is valid since the series converges uniformly in the set $\{\phi_k\}$, and each term is continuous. See, for example, E. C. Titchmarsh, *The Theory of Functions* (Oxford U.P., London, 1932), p. 36.

[6] A sequence $\{f_n(x)\}$ is said to be *uniformly bounded* on a domain D if there exists an $M > 0$ such that $|f_n(x)| \leq M$ for all x in D and all $n = 1, 2, \cdots$.

[7] To obtain Eq. (47), $\tilde{\mathfrak{H}}_N$ is bounded above and below in terms of $\mathfrak{H}_N$ and a term which depends upon Δ, but which is independent of the summation variables $\{s_{z,k}\}$. From Eqs. (25) and (45) and the inequality

$$a^{\frac{1}{2}} - b^{\frac{1}{2}} \leq (a+b)^{\frac{1}{2}} \leq a^{\frac{1}{2}} + b^{\frac{1}{2}}, \quad \text{for} \quad a \geq 0, b \geq 0,$$

we obtain the bounds

$$(1 + t_k^2)^{\frac{1}{2}} - (2\Delta)^{\frac{1}{2}} \leq \sigma_k \leq (1 + t_k^2)^{\frac{1}{2}} + (2\Delta)^{\frac{1}{2}},$$

$$[(1 + t_k^2)(1 + t_l^2)]^{\frac{1}{2}} - 2\Delta^{\frac{1}{2}}(2 + \Delta^{\frac{1}{2}}) \leq \sigma_k\sigma_l \leq [(1 + t_k^2)(1 + t_l^2)]^{\frac{1}{2}} + 2\Delta^{\frac{1}{2}}(2 + \Delta^{\frac{1}{2}}).$$

Using Eqs. (17) and (48) and these bounds in Eq. (42), we obtain

$$\mathfrak{H}_N - 2\Delta^{\frac{1}{2}}(2 + \Delta^{\frac{1}{2}})W(N) \leq \tilde{\mathfrak{H}}_N \leq \mathfrak{H}_N + 2\Delta^{\frac{1}{2}}(2 + \Delta^{\frac{1}{2}})W(N).$$

The right-hand side of Eq. (46) is then bounded by

$$\lim_{\Delta\to 0} \{\exp[-2\beta\Delta^{\frac{1}{2}}(2 + \Delta^{\frac{1}{2}})W(N)]\} \Big(\sum_{\{t_k\}} \Delta^N \exp(-\beta\tilde{\mathfrak{H}}_N) \Big) \leq \lim_{\Delta\to 0} \sum_{\{t_k\}} \Delta^N \exp(-\beta\tilde{\mathfrak{H}}_N) \leq \lim_{\Delta\to 0} \{\exp[+2\beta\Delta^{\frac{1}{2}}(2 + \Delta^{\frac{1}{2}})W(N)]\} \Big(\sum_{\{t_k\}} \Delta^N \exp(-\beta\mathfrak{H}_N) \Big).$$

Proceeding to the limit, we obtain Eq. (47).

[8] *Arzela's Theorem:* Assume $\{f_n(x)\}$ is boundedly convergent on $[a, b]$ and suppose each $f_n(x)$ is Riemann-integrable on $[a, b]$. Assume also that the limit function $f(x)$ is Riemann-integrable on $[a, b]$. Then

$$\lim_{n\to\infty} \int_a^b f_n(x)\, dx = \int_a^b \lim_{n\to\infty} f_n(x)\, dx = \int_a^b f(x)\, dx.$$

See, for example, T. Apostol, Ref. 4, p. 405. See also W. W. Rogosinski, *Volume and Integral* (Interscience, New York, 1962), Sec. 5.8.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Generating Function of the N-fold Photoelectric Counting Distribution for Gaussian Light*

C. D. CANTRELL
Physics Department, Swarthmore College, Swarthmore, Pennsylvania 19081

(Received 27 August 1970)

An exact expression is obtained for the generating function of the N-fold (i.e., N-detector) photoelectron counting distribution for cross-spectrally pure Gaussian light, in terms of the onefold generating function. It is assumed that the counting times at the different detectors are all equal, but not necessarily short compared to the coherence time of the light. In cases where the onefold generating function can be calculated, a closed formula for the N-fold generating function can be obtained.

1. INTRODUCTION

Considerable experimental and theoretical effort has recently been focused on determining the joint N-fold photoelectric counting distribution for various incident optical fields.[1–8] In an N-fold (i.e., N-interval) counting experiment, one measures (or can obtain directly from the data) the probability $P_N(n_1, t_1, T_1; \cdots; n_N, t_N, T_N)$ that n_1 counts will be recorded by the first detector in a counting interval of length T_1 beginning at time t_1; $\cdots$; n_N counts by the Nth detector in the interval $[t_N, t_N + T_N]$. (Some or all of the detectors may be identical.) Multiple coincidence experiments[9–18] are essentially measurements of P_N for the special case $n_1 = n_2 = \cdots = n_N = 1$.

For one detector ($N = 1$), both the probability distribution and the factorial moments of $n = n_1$ can be calculated from the generating function

$$\mathfrak{G}_1(n; s) = \langle(1 - s)^n\rangle.$$

An N-fold generating function can be similarly defined[3,4]:

$$\mathfrak{G}_N(n_1, \cdots, n_N; s_1, \cdots, s_N) = \langle(1 - s_1)^{n_1} \cdots (1 - s_N)^{n_N}\rangle.$$

In this paper we will derive an exact expression for $\mathfrak{G}_N$ in terms of $\mathfrak{G}_1$, assuming that: (i) The light incident on the N detectors is Gaussian; (ii) the light is cross-spectrally pure; (iii) all the counting intervals are

equal ($T_j = T$, $j = 1, 2, \cdots, N$). As usual, it is assumed that the detectors are located at single points in space. Physically, this means that the area of each detector is assumed to be small compared to a coherence area of the optical field.

Approximate expressions for this N-fold generating function have been derived by Arecchi *et al.*[3] and Bédard,[4] for the limit in which the counting time is short compared to the coherence time of the light. No such approximation is made here. Dialetis[7] has shown that $\mathcal{G}_N$ for the special case of N counting intervals at a single detector can be expressed as an infinite product involving the eigenvalues of a certain integral equation. The new contribution of this paper is, therefore, a procedure for calculating the joint N-fold photocounting distribution in terms of the simpler onefold distribution when the effects of finite counting time cannot be neglected.

Note added in proof: The N-fold generating function has been expressed as an infinite product involving the eigenvalues of an $N \times N$ matrix integral equation by A. K. Jaiswal and C. L. Mehta, Phys. Rev. A **2**, 2570 (1970), and A. Zardecki, Can. J. Phys. (to be published).

2. DERIVATION

Let

$$V(\mathbf{r}_j, t) = V_j(t) = V_j, \quad j = 1, 2, \cdots, N,$$

be the values of the (complex) scalar optical field at the N detectors. We shall assume that these are Gaussian variables with zero mean value and with a given covariance matrix Γ:

$$(\Gamma)_{jk} = \langle V_j^* V_k \rangle.$$

The calculation of the N-fold photocount generating function[3,4]

$$\mathcal{G}_N(n_1, \cdots, n_N; s_1, \cdots, s_N) = \left\langle \prod_{j=1}^{N} (1 - s_j)^{n_j} \right\rangle$$

depends on the identity[4]

$$\mathcal{G}_N(n_1, \cdots, n_N; s_1, \cdots, s_N) = \mathcal{M}(W_1, \cdots, W_N; -s_1, \cdots, -s_N), \quad (1)$$

which relates $\mathcal{G}_N$ to the N-fold moment-generating function

$$\mathcal{M}(W_1, \cdots, W_N; -s_1, \cdots, -s_N) = \left\langle \prod_{j=1}^{N} e^{-s_j W_j} \right\rangle$$

for the integrated intensities at the N detectors:

$$W_j = \alpha_j \int_{t_j}^{t_j+T} I_j(t)\, dt,$$

where $I_j(t) = |V_j(t)|^2$. (In the expression for W_j, α_j is proportional to the quantum efficiency of the jth detector, so that $\langle W_j \rangle = \alpha_j \langle I_j \rangle T = \langle n_j \rangle$, the mean number of counts recorded at the jth detector.) The moment generating function for the W's will now be calculated by an extension of the method used in the onefold case.[15,19–21]

The functions $V_j(t)$ are sample functions of N different (but statistically correlated) Gaussian random processes. Each $V_j(t)$ can be represented, on the interval $[t_j, t_j + T]$, as a Karhunen–Loéve series[19–22]

$$V_j(t) = \langle I_j \rangle^{\frac{1}{2}} \sum_{l=0}^{\infty} [\lambda(l, j)]^{\frac{1}{2}} G(l, j) f(l, j; t). \quad (2)$$

For a given j the quantities $G(l, j)$ are independent complex Gaussian variables with mean zero and variance unity,

$$\langle G(l, j)^* G(l', j) \rangle = \delta_{ll'}. \quad (3)$$

The functions $f(l, j; t)$ are solutions of the homogeneous Fredholm equations

$$\int_{t_j}^{t_j+T} \gamma_{jj}(s - t) f(l, j; t)\, dt = \lambda(l, j) f(l, j; s) \quad (4)$$

and are orthonormal on the interval $[t_j, t_j + T]$:

$$\int_{t_j}^{t_j+T} f(l, j; t)^* f(l', j; t)\, dt = \delta_{ll'}. \quad (5)$$

In (4),

$$\gamma_{jk}(s - t) = \gamma_{jk}(s, t) = (\langle I_j \rangle \langle I_k \rangle)^{-\frac{1}{2}} \langle V_j(t)^* V_k(s) \rangle$$

is the usual normalized optical coherence function. In statistical terminology, $\gamma_{jj}(t)$ is the normalized covariance of the random process $V_j(t)$.

For light which is cross-spectrally pure,[23]

$$\gamma_{jk}(t + \tau) = \gamma_{jk}(\tau) \gamma_{jj}(t) \quad (6)$$

for any τ and for all j, k. This implies that the function $\gamma_{jj}(t)$, which is the Fourier transform of the spectral density, is the same for all the detectors. Then (4) is the same integral equation for every $j = 1, 2, \cdots, N$, except for a translation of the time variables. The eigenfunctions and eigenvalues are therefore

$$f(l, j; t) = f_l(t - t_j),$$
$$\lambda(l, j) = \lambda_l,$$

where the functions $f_l(t)$ satisfy the integral equation

$$\int_0^T \gamma_{jj}(s - t) f_l(t)\, dt = \lambda_l f_l(s). \quad (7)$$

The average $\langle G(l, j)^* G(l', j') \rangle$, which is needed for the evaluation of (1), can be calculated as follows[24]: Solving (2) for

$$G(l, j) = (\langle I_j \rangle \lambda_l)^{-\frac{1}{2}} \int_0^T f_l(t)^* V_j(t + t_j)\, dt,$$

we find from (5) and (6) that

$$\begin{aligned}\langle G(l,j)^*G(l',j')\rangle &= (\lambda_l\lambda_{l'})^{-\frac{1}{2}}\int_0^T\int_0^T f_l(t)f_{l'}(t')^* \\ &\quad\times\gamma_{jj'}(t'-t-t_j+t_{j'})\,dt\,dt' \\ &= (\lambda_l\lambda_{l'})^{-\frac{1}{2}}\gamma_{jj'}(t_{j'}-t_j) \\ &\quad\times\int_0^T\int_0^T f_l(t)f_{l'}(t')^*\gamma_{jj}(t'-t)\,dt\,dt' \\ &= \gamma_{jj'}(t_{j'}-t_j)\delta_{ll'}. \qquad (8)\end{aligned}$$

In the representation (2),

$$W_j = \alpha_j\langle I_j\rangle\sum_{l=0}^{\infty}\lambda_l|G(l,j)|^2,$$

and, according to (8), $|G(l,j)|^2$ and $|G(l',j')|^2$ are statistically independent for $l\neq l'$. Thus the moment-generating function for the integrated intensities is

$$\mathcal{M}(W_1,\cdots,W_N;-s_1,\cdots,-s_N) = \prod_{l=0}^{\infty}\left\langle\prod_{j=1}^{N}e^{-s_j\alpha_j\langle I_j\rangle\lambda_l|G(l,j)|^2}\right\rangle. \qquad (9)$$

Each factor in (9) can be evaluated by simple calculations involving the N-variable complex Gaussian distribution,[25] with the result

$$\left\langle\prod_{j=1}^{N}e^{-s_j\alpha_j\langle I_j\rangle\lambda_j|G(l,j)|^2}\right\rangle = \frac{|A|}{|A+2\lambda_lS|}. \qquad (10)$$

The vertical bars on the right-hand side indicate the determinant of the enclosed $N\times N$ matrix, and

$$2A^{-1} = \tilde{\gamma},\quad (\gamma)_{jk} = \gamma_{jk}(t_k-t_j),$$
$$(S)_{jk} = \delta_{jk}s_j\alpha_j\langle I_j\rangle = \delta_{jk}s_j\langle n_j\rangle/T.$$

In order to effect another interchange of products over j and l in (9), it is convenient to express the right-hand side of (10) in terms of the eigenvalues of an $N\times N$ matrix. Since $\gamma_{jk}=\gamma_{kj}^*$, the matrix A is Hermitian and can be diagonalized by a similarity transformation:

$$U^{-1}AU = D. \qquad (11)$$

Because the eigenvalues of the matrix γ are all strictly positive,[26] the elements of the diagonal matrix D are finite and strictly positive. Hence it is possible to define the diagonal matrix $D^{\frac{1}{2}}$, with elements which are the positive square roots of the elements of D. We now have

$$A+2\lambda_lS = UD^{\frac{1}{2}}[1+\lambda_lS']D^{\frac{1}{2}}U^{-1},$$

where

$$S' = 2D^{-\frac{1}{2}}U^{-1}SUD^{-\frac{1}{2}}, \qquad (12)$$

and therefore

$$|A|/|A+2\lambda_lS| = 1/|1+\lambda_lS'|. \qquad (13)$$

The matrix S' is Hermitian (for real $s_1,\cdots,s_N$) and hence can be diagonalized by another similarity transformation. If $\sigma_1',\cdots,\sigma_N'$ are the eigenvalues of S', then, from (10) and (13),

$$\begin{aligned}\mathcal{M}(W_1,\cdots,W_N;-s_1,\cdots,-s_N) &= \prod_{l=0}^{\infty}\prod_{j=1}^{N}(1+\lambda_l\sigma_j')^{-1} \\ &= \prod_{j=1}^{N}[d(-\sigma_j')]^{-1}, \qquad (14)\end{aligned}$$

where

$$d(z) = \prod_{l=0}^{\infty}(1-\lambda_lz).$$

The function $d(z)$ is the Fredholm determinant of the kernel $\gamma_{jj}(s-t)$ of the integral equation (7).[27] Equation (14) is the generalization to N-fold counting statistics of the well-known theorem that the generating function of the onefold photocounting distribution for Gaussian light is

$$\mathcal{G}_1(n;s) = [d(-s\langle n\rangle/T)]^{-1}.$$

A physically more interesting interpretation is that

$$[d(\xi/T)]^{-1} = \mathcal{M}(w;\xi),$$

the moment-generating function for a single variable

$$w = \int_0^T|v(t)|^2\,dt,$$

where $v(t)$ is a complex Gaussian random process with zero mean and covariance

$$\langle v(t)^*v(s)\rangle = \gamma_{jj}(s-t),\quad j=1,2,\cdots,N.$$

Within a scale factor, $v(t)$ has the same statistical properties as the optical field incident on any one of the detectors.

Equations (1) and (14) imply the main result of this paper,

$$\begin{aligned}\mathcal{G}_N(n_1,\cdots,n_N;s_1,\cdots,s_N) &= \prod_{j=1}^{N}[d(-\sigma_j')]^{-1} \qquad (15)\\ &= |d(-S')|^{-1} \qquad (16)\\ &= |d(-\tilde{\gamma}S)|^{-1}. \qquad (17)\end{aligned}$$

Since the Fredholm determinant $d(z)$ is essentially the reciprocal of the onefold generating function, (15) shows how the N-fold photocounting statistics are determined by the onefold statistics and the mutual coherence functions γ_{jk}. The $N\times N$ matrix $d(-S')$ in (16) is defined to be the same function of the matrix $-S'$ as $d(z)$ is of the single variable z. Equation (16) follows from (14) and (15) by the observation that the numbers $d(-\sigma_j')$ are the eigenvalues of the matrix $d(-S')$. Equation (17) follows from (16) by expanding

$d(-S')$ in a power series in[27] S' and performing simple matrix manipulations.

To compare this calculation with Bédard's results for short counting times, we note that (13) can be written in the form

$$|A|/|A + 2\lambda_l S| = 1/|1 + \lambda_l \tilde{\gamma} S|.$$

For short counting times only one eigenvalue, $\lambda_0 \approx T$, makes an appreciable contribution to the infinite product over l. Equation (16) then reduces to the expression derived by Bédard.[4]

If $d(z)$ can be calculated in closed form, then (15)–(17) can be used to calculate both P_N and the joint factorial moments without any approximation involving the length of the counting interval. An exact result for these quantities can, in principle, be useful in at least two ways:

(i) There are several approximate formulas for the onefold probability distribution P_1,[15,28,29] which, in the future, it may prove possible to generalize to the N-fold case. Equation (15) allows one to evaluate these approximations.

(ii) In a photocounting experiment, where the object is to gain information on the statistics of the incident optical field, it is customary to extrapolate the measured P_N to zero counting times. Exact results for P_N or its moments facilitate this correction.

In practice (15)–(17) are not likely to prove particularly convenient for either purpose. Obtaining the eigenvalues of the $N \times N$ matrix S' and then differentiating (15) is a considerable task. Evaluating (16) or (17) is no easier. In addition, it is not possible to calculate $d(z)$ in closed form for a general spectral distribution of the incident light. Alternatively, there are formulas which express the N-fold cumulants of the integrated intensities W_j as linear combinations of the cumulants of the photocounts n_j.[8] For finite counting times the cumulants are related [through integrals involving the function $\gamma_{jj}(t)$] to the cumulants for zero counting times, in a way which is valid for any spectral distribution. In the onefold case the cumulants have proven to be convenient for evaluating approximate probability distributions[28] and for making the correction for finite counting times[29,30]; the same will probably be true in N-fold experiments. The interest and the usefulness of (15)–(17) are expected to be mainly theoretical, in showing compactly how the photo-counting statistics of Gaussian light at a single detector and in a single counting interval determine the counting statistics at N detectors and N counting intervals of equal length.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the hospitality of the Aspen Center for Physics, where the first draft of this paper was written, and the support of an institutional grant from the Alfred P. Sloan Foundation. The author also wishes to thank Dr. E. A. Klotz for a helpful discussion and to express his appreciation to Dr. C. L. Mehta for several essential comments on a preliminary version of this paper.

* Research partially supported by an institutional grant from the Alfred P. Sloan Foundation.

[1] M. L. Goldberger and K. M. Watson, Phys. Rev. **134**, B919 (1964). See especially Eq. (2.19).

[2] L. Mandel and E. Wolf, Rev. Mod. Phys. **37**, 231 (1965). See especially Sec. 6.4.

[3] F. T. Arecchi, A. Berné, and A. Sona, Phys. Rev. Letters **17**, 260 (1966).

[4] G. Bédard, Phys. Rev. **161**, 1304 (1967).

[5] J. R. Klauder and E. C. G. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1968), pp. 177–78, 253–55.

[6] G. L. Fillmore, Phys. Rev. **182**, 1384 (1969).

[7] D. Dialetis, J. Phys. A **2**, 229 (1969).

[8] C. D. Cantrell, Phys. Rev. A **1**, 672 (1970).

[9] R. Q. Twiss, A. G. Little, and R. H. Brown, Nature **180**, 324 (1957).

[10] G. A. Rebka, Jr., and R. V. Pound, Nature **180**, 1035 (1957).

[11] E. Brannen, H. I. S. Ferguson, and W. Wehlau, Can. J. Phys. **36**, 871 (1958).

[12] R. Q. Twiss and A. G. Little, Australian J. Phys. **12**, 77 (1959).

[13] L. Mandel, in *Progress in Optics*, edited by E. Wolf (North-Holland, Amsterdam, 1962), Vol. 2, pp. 181–248.

[14] W. Martienssen and E. Spiller, Am. J. Phys. **32**, 919 (1964).

[15] R. J. Glauber, in *Quantum Optics and Electronics*, edited by C. DeWitt, A. Blandin, and C. Cohen-Tannoudji (Gordon and Breach, New York, 1965), pp. 65–185.

[16] B. L. Morgan and L. Mandel, Phys. Rev. Letters **16**, 1012 (1966).

[17] D. B. Scarl, Phys. Rev. Letters **17**, 663 (1966).

[18] F. Davidson and L. Mandel, J. Appl. Phys. **39**, 62 (1968); Phys. Letters **27A**, 579 (1968).

[19] M. Kac and A. J. F. Siegert, Ann. Math. Stat. **18**, 438 (1947).

[20] D. Slepian, Bell System Tech. J. **37**, 163 (1958).

[21] G. Bédard, Phys. Rev. **151**, 1038 (1966).

[22] W. B. Davenport, Jr., and W. L. Root, *An Introduction to the Theory of Random Signals and Noise* (McGraw-Hill, New York, 1958), pp. 96–101.

[23] L. Mandel, J. Opt. Soc. Am. **51**, 1342 (1961).

[24] C. L. Mehta, private communication.

[25] C. L. Mehta, in *Lectures in Theoretical Physics*, edited by W. E. Brittin (U. of Colorado Press, Boulder, 1965), Vol. 7, pp. 345–401. See especially p. 398.

[26] R. J. Glauber, Phys. Rev. **130**, 2529 (1963), Eq. (A8). Note that since the mean value of a product of N intensities is always nonzero in the present case, the symbol $\geq$ in Glauber's inequality may be replaced by $>$.

[27] F. Smithies, *Integral Equations* (Cambridge U.P., Cambridge, 1962), Chap. V.

[28] G. Bédard, J. C. Chang, and L. Mandel, Phys. Rev. **160**, 1496 (1967).

[29] E. Jakeman and E. R. Pike, J. Phys. A **1**, 128 (1968).

[30] R. F. Chang, V. Korenman, C. O. Alley, and R. W. Detenbeck, Phys. Rev. **178**, 612 (1969).

Lorentz Basis for the 3 + 2 de Sitter Group*

S. K. Bose and R. Parker†
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

(Received 24 July, 1970)

Matrix elements of the Lie algebra generators of the de Sitter group $O(3, 2)$ are explicitly constructed on a basis in which the Casimir operators of the homogeneous Lorentz group are diagonal. As an application, a class of Hermitian representations of the algebra are obtained.

1. INTRODUCTION

In the following, we give an explicit construction of the matrix elements of the "translation like" generators Π_μ of the Lie algebra of the (universal covering group of) de Sitter group $O(3, 2)$ on the "Lorentz basis." By "Lorentz basis" we mean a basis in which the Casimir operators of the homogeneous Lorentz group are diagonal. It is felt that, apart from its intrinsic mathematical interest, a knowledge of these matrix elements may be useful in possible physical applications[1] involving the de Sitter group $O(3, 2)$.

The representation of translationlike generators on the "Lorentz basis," in general, does not correspond to unitary representations. If one starts with a pure state corresponding to an eigenvalue λ (real for principal series representation) of the Casimir operators of homogeneous Lorentz group, the action of the "translationlike" generators results formally in states with $(\lambda \pm i)$. The corresponding matrix elements are not Hermitic. This situation is a special example of a more general malaise, which appears in the reduction of the representation of a noncompact group with respect to its noncompact subgroups.[2] The corresponding problem for the Poincaré group on "Lorentz basis" has been discussed recently by Chakrabarti *et al.*[3] (Indeed, a perusal of this paper prompted us to the present work.)

Somewhat remarkably, it turns out that a class of Hermitian representations of the de Sitter algebra exists even on the "Lorentz basis." This happens when either (i) λ is restricted to some suitable fixed value or (ii) is allowed to take up only two values $\lambda = i/2$ and $\lambda = -i/2$ (together with other restrictions!). These representations are discussed in Sec. 4.

2. MATRIX ELEMENTS OF "TRANSLATION-LIKE" OPERATORS ON THE LORENTZ BASIS

The Lie algebra of the de Sitter group $O(3, 2)$ is generated by 10 operators $M_{\mu\nu}$ and Π_μ having the following commutation relations:

$$[M_{\mu\nu}, M_{\lambda\sigma}] = i(g_{\nu\lambda}M_{\mu\sigma} - g_{\mu\lambda}M_{\nu\sigma} + g_{\mu\sigma}M_{\nu\lambda} - g_{\nu\sigma}M_{\mu\lambda}), \quad (2.1)$$

$$[M_{\mu\nu}, \Pi_\lambda] = i(g_{\nu\lambda}\Pi_\mu - g_{\mu\lambda}\Pi_\nu), \quad (2.2)$$

$$[\Pi_\mu, \Pi_\nu] = -iM_{\mu\nu}. \quad (2.3)$$

The operators $M_{\mu\nu}\,(= -M_{\nu\mu})$ generate the Lie algebra of homogeneous Lorentz group (HLG) and we call Π_μ "translationlike" as the latter go over to the translation operators of Poincaré group upon contraction.[4] We propose to determine explicitly the matrix elements of Π_μ with respect to the canonical basis of HLG. The matrix elements of the generators of HLG are given by[5]

$$\begin{aligned}
M_3\,|jm\rangle_{j_0\lambda} &= m\,|jm\rangle_{j_0\lambda}, \quad (M_1 \pm iM_2)\,|j\cdot m\rangle_{j_0\lambda} = [(j \mp m)(j \pm m + 1)]^{\frac{1}{2}}\,|jm \pm 1\rangle_{j_0\lambda}, \\
N_3\,|jm\rangle_{j_0\lambda} &= \frac{1}{j}\left[\frac{(j^2 - m^2)(j^2 - j_0^2)(j^2 + \lambda^2)}{(2j + 1)(2j - 1)}\right]^{\frac{1}{2}}|j - 1, m\rangle_{j_0\lambda} + \frac{j_0\lambda}{j(j + 1)}\,m\,|jm\rangle_{j_0\lambda} \\
&\quad + \frac{1}{j + 1}\left[\frac{[(j + 1)^2 - m^2][(j + 1)^2 + \lambda^2][(j + 1)^2 - j_0^2]}{(2j + 1)(2j + 3)}\right]^{\frac{1}{2}}|j + 1\,m\rangle_{j_0\lambda}, \\
(N_1 \pm iN_2)\,|jm\rangle_{j_0\lambda} &= \pm\frac{1}{j}\left(\frac{(j \mp m)(j \mp m - 1)(j^2 - j_0^2)(j^2 + \lambda^2)}{(2j + 1)(2j - 1)}\right)^{\frac{1}{2}}|j - 1, m \pm 1\rangle_{j_0\lambda} \\
&\quad + j_0\lambda\left(\frac{(j \pm m + 1)(j \mp m)}{[j(j + 1)]^2}\right)^{\frac{1}{2}}|j, m \pm 1\rangle_{j_0\lambda} \\
&\quad \mp \frac{1}{j + 1}\left(\frac{[(j \pm m + 1)(j \pm m + 2)][(j + 1)^2 - j_0^2][(j + 1)^2 + \lambda^2]}{(2j + 1)(2j + 3)}\right)^{\frac{1}{2}}|j + 1, m \pm 1\rangle_{j_0\lambda}.
\end{aligned} \quad (2.4)$$

The two Casimir operators of HLG are given by

$$(M^2 - N^2)\,|jm\rangle_{j_0\lambda} = (j_0^2 - \lambda^2 - 1)\,|jm\rangle_{j_0\lambda}, \quad (2.5)$$

$$\mathbf{M}\cdot\mathbf{N}\,|jm\rangle_{j_0\lambda} = j_0\lambda\,|jm\rangle_{j_0\lambda}. \quad (2.6)$$

The unitary representations of HLG are classified as

$$\left.\begin{array}{ll}(1)\ j_0 = 0, 1, 2, 3, \cdots, & -\infty < \lambda < \infty \\ (2)\ j_0 = \frac{1}{2}, \frac{3}{2}, \cdots, & -\infty < \lambda < \infty\end{array}\right\} \text{principal series,}$$

$$(3)\ j_0 = 0, \quad -1 < -i\lambda < 1 \quad \text{supplementary series.} \quad (2.7)$$

The normalization corresponding to the above matrix elements is

$${}_{j_0'\lambda'}\langle j'm'\,|\,jm\rangle_{j_0\lambda} = \delta_{jj'}\delta_{mm'}\delta_{j_0j_0'}\delta(\lambda - \lambda'). \quad (2.8)$$

Let us now proceed to construct matrix elements of Π_μ. Evidently it is sufficient to consider Π_0, as the rest Π_i, $i = 1, 2, 3$, can be written down from Eqs. (2.2) and (2.4) once Π_0 is known. The general form of Π_0 follows from Eq. (2.2) and has been derived before.[6]

$$\begin{aligned}\Pi_0\,|jm\rangle_{j_0\lambda} &= c_\lambda^{j_0-1,j_0}[(j + j_0)(j - j_0 + 1)]^{\frac{1}{2}}\,|jm\rangle_{j_0-1,\lambda} \\ &+ c_\lambda^{j_0+1,j_0}[(j - j_0)(j + j_0 + 1)]^{\frac{1}{2}}\,|jm\rangle_{j_0+1,\lambda} \\ &+ c_{j_0}^{\lambda-i,\lambda}[(j - i\lambda)(j + i\lambda + 1)]^{\frac{1}{2}}\,|jm\rangle_{j_0,\lambda-i} \\ &+ c_{j_0}^{\lambda+i,\lambda}[(j + i\lambda)(j - i\lambda + 1)]^{\frac{1}{2}}\,|jm\rangle_{j_0,\lambda+i}. \quad (2.9)\end{aligned}$$

It remains for us to determine the coefficients $c_\lambda^{j_0\pm1,j_0}$ and $c_{j_0}^{\lambda\pm i,\lambda}$. For this purpose it is convenient to use two commutation relations which follow from Eqs. (2.2) and (2.3):

$$[\Pi_0, [\Pi_0, \mathbf{M}\cdot\mathbf{N}]] = \mathbf{M}\cdot\mathbf{N}, \quad (2.10)$$

$$[\Pi_0, [\Pi_0, M^2 - N^2]] = 2(\mathbf{\Pi}^2 - \mathbf{N}^2). \quad (2.11)$$

The above two relations in conjunction with Eq. (2.9) yield the following recursion relations:

$$\begin{aligned}&\lambda(c_\lambda^{j_0-1,j_0}c_\lambda^{j_0,j_0-1} - c_\lambda^{j_0+1,j_0}c_\lambda^{j_0,j_0+1}) \\ &\quad + ij_0(c_{j_0}^{\lambda-i,\lambda}c_{j_0}^{\lambda,\lambda-i} - c_{j_0}^{\lambda+i,\lambda}c_{j_0}^{\lambda,\lambda+i}) = 0, \quad (2.12)\end{aligned}$$

$$\begin{aligned}&(j_0 + 1)c_\lambda^{j_0+1,j_0}c_\lambda^{j_0,j_0+1} - (j_0 - 1)c_\lambda^{j_0-1,j_0}c_\lambda^{j_0,j_0-1} \\ &\quad + (1 + i\lambda)c_{j_0}^{\lambda-i,\lambda}c_{j_0}^{\lambda,\lambda-i} + (1 - i\lambda)c_{j_0}^{\lambda+i,\lambda}c_{j_0}^{\lambda,\lambda+i} = \tfrac{1}{2}, \quad (2.13)\end{aligned}$$

$$\begin{aligned}&j_0[(j_0 + 1)^2c_\lambda^{j_0+1,j_0}c_\lambda^{j_0,j_0+1} - (j_0 - 1)^2c_\lambda^{j_0-1,j_0}c_\lambda^{j_0,j_0-1}] \\ &\quad + i\lambda[(1 + i\lambda)^2c_{j_0}^{\lambda-i,\lambda}c_{j_0}^{\lambda,\lambda-i} - (1 - i\lambda)^2c_{j_0}^{\lambda+i,\lambda}c_{j_0}^{\lambda,\lambda+i}] \\ &\quad = \tfrac{1}{2}\xi + j_0^2 - \lambda^2 - 1. \quad (2.14)\end{aligned}$$

The parameter ξ occurring in Eq. (2.14) denotes the eigenvalue of the first Casimir operator c_1 of de Sitter algebra, i.e.,

$$c_1\,|jm\rangle_{j_0\lambda,\xi} = \xi\,|jm\rangle_{j_0\lambda\xi},$$

$$c_1 = -(\tfrac{1}{2}M_{\mu\nu}M^{\mu\nu} + \Pi_\mu\Pi^\mu). \quad (2.15)$$

To solve the above set of recursion relations,we proceed as follows. First, combine Eqs. (2.12)–(2.14) to obtain

$$\begin{aligned}&[\lambda^4 + \lambda^2(2j_0^2 + 2j_0 + 1) \\ &\quad + j_0^2(j_0^2 + 2j_0 + 1)]c_\lambda^{j_0+1,j_0}c_\lambda^{j_0,j_0+1} \\ &\quad - [\lambda^4 + \lambda^2(2j_0^2 - 2j_0 + 1) \\ &\quad + j_0^2(j_0^2 - 2j_0 + 1)]c_\lambda^{j_0-1,j_0}c_\lambda^{j_0,j_0-1} \\ &= (\tfrac{1}{2}\xi + j_0^2 - 1)j_0. \quad (2.16)\end{aligned}$$

Now let us put

$$\begin{aligned}&[\lambda^4 + \lambda^2(2j_0^2 + 2j_0 + 1) \\ &\quad + j_0^2(j_0 + 1)^2]c_\lambda^{j_0+1,j_0}c_\lambda^{j_0,j_0+1} = \beta_{j_0+1}, \quad (2.17)\end{aligned}$$

so that Eq. (2.16) reduces to

$$\beta_{j_0+1} - \beta_{j_0} = j_0^3 + (\tfrac{1}{2}\xi - 1)j_0. \quad (2.18)$$

To solve Eq. (2.18),we notice that, in an irreducible representation of the de Sitter algebra, there must evidently exist a minimum value of j_0. Let us call this j_{0L}, so that $\beta_{j_{0L}} = 0$. Hence, we get from Eq. (2.18)

$$\begin{aligned}\beta_{j_0+1} &= \beta_{j_0+1} - \beta_{j_{0L}} \\ &= \sum_{k_0=j_{0L}}^{j_0}(\beta_{k_0+1} - \beta_{k_0}) \\ &= \sum_{k_0=j_{0L}}^{j_0} k_0^3 + (\tfrac{1}{2}\xi - 1)\sum_{k_0=j_{0L}}^{j_0} k_0 \\ &= \tfrac{1}{4}[j_0(j_0 + 1) - j_{0L}(j_{0L} - 1)] \\ &\quad \times [j_0(j_0 + 1) + j_{0L}(j_{0L} - 1) + \xi - 2]. \quad (2.19)\end{aligned}$$

From Eqs. (2.17) and (2.19) we obtain finally

$$\begin{aligned}c_\lambda^{j_0-1,j_0}c_\lambda^{j_0,j_0-1} &= \frac{[j_0(j_0 - 1) - j_{0L}(j_{0L} - 1)][j_0(j_0 - 1) + j_{0L}(j_{0L} - 1) + \xi - 2]}{4[\lambda^4 + \lambda^2(2j_0^2 - 2j_0 + 1) + j_0^2(j_0 - 1)^2]} \\ &= \frac{j_0^2(j_0 - 1)^2 + (\xi - 2)[j_0(j_0 - 1) - j_{0L}(j_{0L} - 1)] - j_{0L}^2(j_{0L} - 1)^2}{4(j_0 + i\lambda)(j_0 - i\lambda)(j_0 - i\lambda - 1)(j_0 + i\lambda - 1)} \quad (2.20)\end{aligned}$$

The determination of the remaining coefficient $c_{j_0}^{\lambda-i,\lambda}$ is now straightforward. We substitute Eq. (2.20) into Eq. (2.12) and solve the resulting recursion relation. We obtain

$$c_{j_0}^{\lambda-i,\lambda}c_{j_0}^{\lambda,\lambda-i} = \frac{\lambda^2(\lambda-i)^2 + (2-\xi)(\lambda^2 - i\lambda + j_{0L}^2 - j_{0L}) - j_{0L}^2(j_{0L}-1)^2}{4(j_0 + i\lambda)(j_0 - i\lambda)(j_0 - i\lambda - 1)(j_0 + i\lambda + 1)} \tag{2.21}$$

If we want Π_0 to be represented by a Hermitian matrix, then we must have $c_\lambda^{j_0-1,j_0} = (c_\lambda^{j_0,j_0-1})^*$ and $c_{j_0}^{\lambda-i,\lambda} = (c_{j_0}^{\lambda,\lambda-i})^*$. Therefore

$$\begin{aligned} c_\lambda^{j_0-1,j_0} &= |c_\lambda^{j_0-1,j_0}|\, e^{i\Phi}, \quad c_\lambda^{j_0,j_0-1} = |c_\lambda^{j_0-1,j_0}|\, e^{-i\Phi}, \\ c_{j_0}^{\lambda-i,\lambda} &= |c_{j_0}^{\lambda-i,\lambda}|\, e^{i\chi}, \qquad c_{j_0}^{\lambda,\lambda-i} = |c_{j_0}^{\lambda-i,\lambda}|\, e^{-i\chi}, \end{aligned} \tag{2.22}$$

when Φ and χ are phase functions to be fixed according to some convention and we have

$$|c_\lambda^{j_0-1,j_0}| = \tfrac{1}{2}\left(\frac{j_0^2(j_0-1)^2 + (\xi-2)[j_0(j_0-1) - j_{0L}(j_{0L}-1)] - j_{0L}^2(j_{0L}-1)^2}{(j_0 + i\lambda)(j_0 - i\lambda)(j_0 - i\lambda - 1)(j_0 + i\lambda - 1)}\right)^{\frac{1}{2}}, \tag{2.23}$$

$$|c_{j_0}^{\lambda-i,\lambda}| = \tfrac{1}{2}\left(\frac{\lambda^2(\lambda-i)^2 + (2-\xi)[\lambda(\lambda-i) + j_{0L}(j_{0L}-1)] - j_{0L}^2(j_{0L}-1)^2}{(j_0 + i\lambda)(j_0 - i\lambda)(j_0 - i\lambda - 1)(j_0 + i\lambda + 1)}\right)^{\frac{1}{2}}. \tag{2.24}$$

Equations (2.23) and (2.24) give the desired matrix elements of the operator Π_0 on the "Lorentz basis."

Finally, let us state the connection between j_{0L} and the second Casimir operator c_2 of DeSitter algebra. By direct calculation we find

$$\begin{aligned} &c_2\,|jm\rangle_{j_0\lambda\xi j_{0L}} \\ &\quad = j_{0L}(j_{0L}-1)[j_{0L}(j_{0L}-1) + \xi - 2]\,|jm\rangle_{j_0\lambda\xi j_{0L}}, \end{aligned} \tag{2.25}$$

when

$$\begin{aligned} c_2 &= W_\mu W^\mu + d^2, \quad W_\mu = \tfrac{1}{2}\epsilon_{\mu\nu\lambda\sigma}M^{\nu\lambda}\Pi^\sigma, \\ d &= \tfrac{1}{8}\epsilon_{\mu\nu\lambda\sigma}M^{\mu\nu}M^{\lambda\sigma}. \end{aligned} \tag{2.26}$$

Thus we may rewrite Eqs. (2.23) and (2.24) as follows:

$$\begin{aligned} &|c_\lambda^{j_0-1,j_0}| \\ &\quad = \tfrac{1}{2}\left(\frac{j_0(j_0-1)[j_0(j_0-1) + \xi - 2] - \eta}{(j_0 + i\lambda)(j_0 - i\lambda)(j_0 - i\lambda - 1)(j_0 + i\lambda - 1)}\right)^{\frac{1}{2}}, \end{aligned} \tag{2.27}$$

$$\begin{aligned} &|c_{j_0}^{\lambda-i,\lambda}| \\ &\quad = \tfrac{1}{2}\left(\frac{\lambda(\lambda-i)[\lambda(\lambda-i) + 2 - \xi] - \eta}{(j_0 + i\lambda)(j_0 - i\lambda)(j_0 - i\lambda - 1)(j_0 + i\lambda + 1)}\right)^{\frac{1}{2}}. \end{aligned} \tag{2.28}$$

In the above, ξ and η are, respectively, the eigenvalues of the Casimir operators $c_1[-(\frac{1}{2}M_{\mu\nu}M^{\mu\nu} + \Pi_\mu\Pi^\mu)]$ and $c_2(W_\mu W^\mu + d^2)$ of the 3 + 2 deSitter group.

3. COMMENTS ON THE MATRIX ELEMENTS

The matrix elements given by Eqs. (2.23) and (2.24) can be used to study irreducible representations of de Sitter algebra on the Lorentz basis. First, a brief remark on an interesting class of nonunitary representations in which the operator Π_0 is Hermitian (but not N). This happens when $\lambda = in$, n being an integral (for j_{0L} an integer) or half-integral (for j_{0L} a half integer) number. The Dirac representation of deSitter algebra provided by the 4 × 4 Hermitian gamma matrices is a special case of this type of representation. We do not discuss nonunitary representations any further and pass on to the physically interesting case of unitary representations.

From the discussion of Sec. 2, it is quite clear that in general one does not get unitary representations of de Sitter group on the "Lorentz basis." To see this, it is sufficient to notice that the action of the generators Π_μ on a pure state belonging to the eigenvalue λ (real or pure imaginary lying between $-i$ and i for principal or supplementary series unitary representations of HLG) gives us formally the states corresponding to $(\lambda \pm i)$. These matrix elements are not Hermitic. This situation is not, of course, a special feature of the de Sitter algebra but exists, in general, in the reduction of the representations of any noncompact group with respect to those of one of its noncompact subgroups.[2] In particular, the same problem occurs in the construction of representations of Poincaré group on the Lorentz basis. In the present case there seems to be two ways of constructing unitary representations. First is to use suitably "smeared" basis vectors and obtain unitary representations as a continuous superposition of nonunitary ones. (In the corresponding case of Poincaré group, this has been done for the special case $P_\mu P^\mu = 0$ in Ref. 3.) However, this procedure amounts to carrying out a change of basis, from the "Lorentz basis" to that in which the maximal compact subgroup $O(3) \times O(2)$ is diagonal. In this case, the original motivation of constructing representations on the "Lorentz basis" is completely lost

as one can directly reduce[7] representations of de Sitter group $O(3,2)$ with respect to its maximal compact subgroup $O(3) \times O(2)$. Hence, we refrain from discussing this possibility any further, in spite of its intrinsic mathematical interest.

A second possibility of constructing Hermitian representations of the de Sitter algebra is to impose additional restrictions on the parameters which occur in Eqs. (2.27) and (2.28). As we will see in the next section, it is possible to consistently impose such restrictions, and the resulting Hermitian representations of the de Sitter algebra $O(3,2)$ on the "Lorentz basis" have the property that (i) λ can only acquire a single fixed value and (ii) λ can acquire two fixed values, $\lambda = i/2$ and $-i/2$. For this reason we may call these "discrete Lorentz" representations.

4. DISCRETE LORENTZ REPRESENTATIONS[8]

From the discussion of Sec. 3, it is obvious that in order to get Hermitian representations of the algebra, it is necessary to forbid the appearance of "undesired" states, as a result of the action of Π_0 on the canonical basis vectors of HLG. Thus we must prevent two types of transitions: (a) $\lambda \to \lambda \pm i$ if it leads to outside the interval $-i$ to $+i$ for supplementary series of HLG and (b) $\lambda \to \lambda \pm i$ for the principal series representations of HLG. We now investigate these two cases separately.

Case (a)

Here we seek representations of the de Sitter algebra which contain only supplementary series representations of HLG. To get this, we notice that, for $\lambda \neq 0$, $\xi \neq 2, j_{0L} = 0$, and $\eta = 0$, Eq. (2.27) gives $|c_\lambda^{0,1}| = 0$. This means that for $j_{0L} = 0$, the vector corresponding to $j_0 = 1$ is not present in the representation. Consequently, $j_0 = 2, 3$ are also absent, and $j_0 = 0$ is the only allowed value in this case. Equation (2.28) now reduces to

$$|c_0^{\lambda-i,\lambda}| = \tfrac{1}{2}\left(\frac{\lambda(\lambda - i) + 2 - \xi}{\lambda(\lambda - i)}\right)^{\frac{1}{2}}. \tag{4.1}$$

Hence, we get the desired solution for $\xi = \frac{5}{4}$. In this case $|c_0^{\frac{3}{2}i,\frac{1}{2}i}| = 0$, $|c_0^{-\frac{3}{2}i,-\frac{1}{2}i}| = 0$, as they should be, and $|c_0^{-\frac{1}{2}i,\frac{1}{2}i}| = 1$. Thus this representation is characterized by $\xi = \frac{5}{4}$, $\eta = 0$, $j_0 = 0$, and λ takes up the two values $-\frac{1}{2}i$ and $\frac{1}{2}i$.

Case (b)

Here we seek special values of λ, with λ real, for which $\lambda \to \lambda \pm i$ transitions are forbidden. Let λ_0 be such a value. Hence, we must have

$$c_{j_0}^{\lambda_0-i,\lambda_0} = 0, \quad c_{j_0}^{\lambda_0+i,\lambda_0} = 0. \tag{4.2}$$

From Eqs. (4.2) and (2.28) we get a pair of equations

$$\lambda_0(2\lambda_0^2 + 2 - \xi) = 0, \quad \lambda_0^2(\lambda_0^2 + 1 - \xi) - \eta = 0. \tag{4.3}$$

Equation (4.3) admits of two classes of solutions:

$$(1) \qquad \lambda_0 = 0, \quad \eta = 0, \tag{4.4}$$

$$(2)\ \lambda_0 \neq 0, \quad \lambda_0^2 = \tfrac{1}{2}\xi - 1, \quad \eta = \tfrac{1}{2}\xi(1 - \tfrac{1}{2}\xi). \tag{4.5}$$

The above relations can also be obtained directly (and more simply) if, from the very beginning, we seek a solution of the form

$$\begin{aligned}\Pi_0 |jm\rangle_{j_0,\lambda} &= c_\lambda^{j_0-1,j_0}[(j + j_0)(j - j_0 + 1)]^{\frac{1}{2}} |jm\rangle_{j_0-1,\lambda} \\ &\quad + c_\lambda^{j_0+1,j_0}[(j - j_0)(j + j_0 + 1)]^{\frac{1}{2}} |jm\rangle_{j_1+1,\lambda}.\end{aligned} \tag{4.6}$$

Equations (4.6), (2.10), and (2.11) give the recursion relations

$$\lambda(c_\lambda^{j_0-1,j_0} c_\lambda^{j_0,j_0-1} - c_\lambda^{j_0+1,j_0} c_\lambda^{j_0,j_0+1}) = 0, \tag{4.7}$$

$$(j_0 + 1)c_\lambda^{j_0+1,j_0} c^{j_0,j_0+1} - (j_0 - 1)c_\lambda^{j_0-1,j_0} c_\lambda^{j_0,j_0-1} = \tfrac{1}{2}, \tag{4.8}$$

$$\begin{aligned} j_0[(1 - j_0)^2 c_\lambda^{j_0-1,j_0} c_\lambda^{j_0,j_0-1} - (1 + j_0)^2 c_\lambda^{j_0+1,j_0} c_\lambda^{j_0,j_0+1}] \\ = 1 + \lambda^2 - j_0^2 - \tfrac{1}{2}\xi. \end{aligned} \tag{4.9}$$

Solving recursion relations (4.7)–(4.9), we get once again Eqs. (4.4) and (4.5). These are now used to classify unitary irreducible representations:

(b1) $\lambda_0 = 0, \quad \xi = 2, \quad \eta = 0$:

$$|c_0^{j_0+1,j_0}| = \tfrac{1}{2} \quad \text{for all } j_0.$$

j_0 takes up all values $j_0 = 0, 1, 2, \cdots$.

(b2)

$$\lambda_0 = 0, \quad \eta = 0, \quad \xi = 2 - j_{0L}(j_{0L} - 1), \quad \xi \neq 2:$$

$$|c_0^{j_0+1,j_0}| = \tfrac{1}{2}\left(\frac{j_0(j_0 + 1) - j_{0L}(j_{0L} - 1)}{j_0(j_0 + 1)}\right)^{\frac{1}{2}}.$$

Here j_{0L} can take up any half-integral value ($j_{0L} = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \cdots$) or any integral value ≥ 2 ($j_{0L} = 2, 3, 4, \cdots$). For each choice of j_{0L}, the allowed values of j_0 are $j_{0L}, j_{0L} + 1, j_{0L} + 2, \cdots$ Each choice of j_{0L}, together with the correspondingly determined value of ξ, provides a unitary, irreducible representation:

(b3) $\lambda_0 \neq 0, \quad \lambda_0^2 = \xi/2 - 1, \quad \eta = (\xi/2)(1 - \xi/2), \quad \xi > 2, \quad \eta \neq 0$:

$$|c_{\lambda_0}^{j_0+1,j_0}| = \tfrac{1}{2} \quad \text{for all } j_0.$$

Here every value of $\xi > 2$, together with corresponding values of η, provides a pair of representations,

depending on whether j_0 takes up all integral ($j_0 = 0, 1, 2, \cdots$) or all half-integral ($j_0 = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \cdots$) values.

5. CONCLUDING REMARKS

We gave an explicit construction of the Lie algebra generators of the universal covering group of the 3 + 2 deSitter group on the Lorentz basis. The representations thus obtained are not, in general, Hermitian. Imposing additional restrictions, we obtained a class of Hermitian representations on the Lorentz basis (the discrete Lorentz representations of Sec. 4). It should be emphasized that we did *not* obtain the complete system of Hermitian representations of the 3 + 2 deSitter algebra. There exists additional Hermitian representations (however, not on the Lorentz basis) which cannot be derived by the procedure followed in this paper. Finally, no attempt has been made in this paper to determine if the "discrete Lorentz" representations of the 3 + 2 de Sitter algebra are also global representations of the corresponding group.

* Research supported in part by the U.S. Atomic Energy Commission under Contract No. AT(11-1)-427.

† This work is based on a dissertation submitted in partial fulfillment of the requirements for the Ph.D. degree at the University of Notre Dame.

[1] For possible applications of de Sitter group $O(3, 2)$, see P. A. M. Dirac, Ann. Math. **36**, 657 (1935); and more recently, E. C. G. Sudarshan and N. Mukunda, Phys. Rev. D **1**, 571 (1970). Further references are cited in the latter.

[2] See, for instance, E. C. G. Sudarshan, "Noncompact Groups in Particle Physics," *Coral Gables Conference, 1966* (Freeman, San Francisco, 1966), p. 254.

[3] A. Chakrabarti, M. Levy-Nahas, and R. Seneor, J. Math. Phys. **9**, 1274 (1968).

[4] E. Inönü and E. P. Wigner, Proc. Natl. Acad. Sci.(U.S.) **39**,510 (1953).

[5] The definition is standard. $M_{ij} = \epsilon_{ijk}M_k$, $N_i = M_{0i}$. ijk take up values 1, 2, 3. The metric used in this paper is $g_{00} = 1$, $g_{11} = g_{22} = g_{33} = -1$.

[6] I. M. Gel'fand, R. Milnos, and Z. Shapiro, *Representations of the Rotation and Lorentz Groups and Their Applications* (Pergamon, New York, 1963), p. 275.

[7] J. B. Ehrman, Proc. Cambridge Phil. Soc. **53**, 290 (1957).

[8] Since submitting this manuscript, a paper by L. Jaffe, "The Discrete Lorentz and Singleton Representations of the Universal Covering Group of the 3 + 2 de Sitter Group," University of Texas Preprint CPT-31, March 1970 [J. Math. Phys. **12**,882 (1971)] has come to our attention. This paper contains all the results of Sec. 4 of the present manuscript. Several additional "discrete Lorentz" UIR's are obtained in Jaffe's paper. These may also be obtained from Eq. (4.3) of the present text. For full details, the reader is referred to Jaffe's paper where further properties of discrete Lorentz representations are discussed.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Generalized Hidden Variables Theorem

PLATON C. DELIYANNIS
Department of Mathematics, Illinois Institute of Technology, Chicago, Illinois 60616

(Received 6 October 1970)

Two versions of the hidden variables theorem are presented with minimum hypotheses. The second is intended to be an answer to a question raised by Jauch and Mackey regarding the existence of approximately dispersion-free states.

INTRODUCTION

All the hidden variables theorems that have appeared in the literature have the following form: If a system admits "sufficiently many" states in which all observables are measured exactly, then it is "classical." [1-6] The concept of system varies from author to author, as well as that of state; in most papers the term "sufficiently many" means that the states in this distinguished set determine in some specific (but not always the same) way all other states,[2,3,5] but there are some exceptions; finally the term "classical system" is given various meanings, not all equivalent. In this work we shall present two versions of the hidden variables theorem (both in the "if and only if" form) under as weak conditions as possible, and we shall indicate the role of certain further assumptions made by other authors. Our second theorem is intended to give an answer to a question raised by Jauch and Mackey[3] as to whether hidden variables can exist in systems for which the above-mentioned "sufficiently many" states are not exactly, but approximately, dispersion free.

In the next section we shall define the various terms and state the axioms used; the reader will note that they are strictly weaker than all others related to the same basic ideas. We shall also try to defend our point of view by pointing out the physical relevance of each term and axiom. In the remaining sections we

depending on whether j_0 takes up all integral ($j_0 = 0, 1, 2, \cdots$) or all half-integral ($j_0 = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \cdots$) values.

5. CONCLUDING REMARKS

We gave an explicit construction of the Lie algebra generators of the universal covering group of the 3 + 2 deSitter group on the Lorentz basis. The representations thus obtained are not, in general, Hermitian. Imposing additional restrictions, we obtained a class of Hermitian representations on the Lorentz basis (the discrete Lorentz representations of Sec. 4). It should be emphasized that we did *not* obtain the complete system of Hermitian representations of the 3 + 2 deSitter algebra. There exists additional Hermitian representations (however, not on the Lorentz basis) which cannot be derived by the procedure followed in this paper. Finally, no attempt has been made in this paper to determine if the "discrete Lorentz" representations of the 3 + 2 de Sitter algebra are also global representations of the corresponding group.

* Research supported in part by the U.S. Atomic Energy Commission under Contract No. AT(11-1)-427.

† This work is based on a dissertation submitted in partial fulfillment of the requirements for the Ph.D. degree at the University of Notre Dame.

[1] For possible applications of de Sitter group $O(3, 2)$, see P. A. M. Dirac, Ann. Math. **36**, 657 (1935); and more recently, E. C. G. Sudarshan and N. Mukunda, Phys. Rev. D **1**, 571 (1970). Further references are cited in the latter.

[2] See, for instance, E. C. G. Sudarshan, "Noncompact Groups in Particle Physics," *Coral Gables Conference, 1966* (Freeman, San Francisco, 1966), p. 254.

[3] A. Chakrabarti, M. Levy-Nahas, and R. Seneor, J. Math. Phys. **9**, 1274 (1968).

[4] E. Inönü and E. P. Wigner, Proc. Natl. Acad. Sci.(U.S.) **39**,510 (1953).

[5] The definition is standard. $M_{ij} = \epsilon_{ijk}M_k$, $N_i = M_{0i}$. ijk take up values 1, 2, 3. The metric used in this paper is $g_{00} = 1$, $g_{11} = g_{22} = g_{33} = -1$.

[6] I. M. Gel'fand, R. Milnos, and Z. Shapiro, *Representations of the Rotation and Lorentz Groups and Their Applications* (Pergamon, New York, 1963), p. 275.

[7] J. B. Ehrman, Proc. Cambridge Phil. Soc. **53**, 290 (1957).

[8] Since submitting this manuscript, a paper by L. Jaffe, "The Discrete Lorentz and Singleton Representations of the Universal Covering Group of the 3 + 2 de Sitter Group," University of Texas Preprint CPT-31, March 1970 [J. Math. Phys. **12**,882 (1971)] has come to our attention. This paper contains all the results of Sec. 4 of the present manuscript. Several additional "discrete Lorentz" UIR's are obtained in Jaffe's paper. These may also be obtained from Eq. (4.3) of the present text. For full details, the reader is referred to Jaffe's paper where further properties of discrete Lorentz representations are discussed.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Generalized Hidden Variables Theorem

PLATON C. DELIYANNIS
Department of Mathematics, Illinois Institute of Technology, Chicago, Illinois 60616

(Received 6 October 1970)

Two versions of the hidden variables theorem are presented with minimum hypotheses. The second is intended to be an answer to a question raised by Jauch and Mackey regarding the existence of approximately dispersion-free states.

INTRODUCTION

All the hidden variables theorems that have appeared in the literature have the following form: If a system admits "sufficiently many" states in which all observables are measured exactly, then it is "classical." [1-6] The concept of system varies from author to author, as well as that of state; in most papers the term "sufficiently many" means that the states in this distinguished set determine in some specific (but not always the same) way all other states,[2,3,5] but there are some exceptions; finally the term "classical system" is given various meanings, not all equivalent. In this work we shall present two versions of the hidden variables theorem (both in the "if and only if" form) under as weak conditions as possible, and we shall indicate the role of certain further assumptions made by other authors. Our second theorem is intended to give an answer to a question raised by Jauch and Mackey[3] as to whether hidden variables can exist in systems for which the above-mentioned "sufficiently many" states are not exactly, but approximately, dispersion free.

In the next section we shall define the various terms and state the axioms used; the reader will note that they are strictly weaker than all others related to the same basic ideas. We shall also try to defend our point of view by pointing out the physical relevance of each term and axiom. In the remaining sections we

state and prove our results; the method of proof is analogous to that of Stone, and has already been used by Zierler and Schlessinger[6] for similar purposes.

1. DEFINITIONS AND AXIOMS

We shall work with the concept of a system as introduced by Mackey.[7]

Events

This is the basic undefined concept. It is assumed that the set $\mathfrak{L}$ of all events carries a *partial order* $\leq$, which represents implication; to each event A there corresponds a unique event A', representing the negation of A in such a way that $(A')' = A$ and $A \leq B$ implies $B' \leq A'$ for all $A, B \in \mathfrak{L}$. Infima $A \wedge B$ and suprema $A \vee B$ relative to $\leq$, representing conjunction and disjunction, are not assumed to exist universally. We shall also assume the existence of elements $0, I \in \mathfrak{L}$ (impossible and certain event) such that $0 \leq A \leq I$, $A \wedge A' = 0$, $A \vee A' = I$ for all $A \in \mathfrak{L}$, and $0' = I$, $I' = 0$. We say that A and B are *disjoint* (mutually exclusive) if $A \leq B'$, which is the same as $B \leq A'$; write $A \perp B$. We shall assume that any finite or infinite sequence of pairwise disjoint events A_k ($A_i \perp A_j$ for $i \neq j$) admits a supremum written as $\sum A_k$, or $A_1 + A_2 + \cdots + A_n$ for the finite case; this is an essential physical condition for the formation of observables which are defined later. Thus this operation $\sum$ forms a basic part of the algebra of $\mathfrak{L}$, although not studied by all authors.[6] The final assumption on $\mathfrak{L}$ is this: If $A \leq B$, then $A' \wedge B$ exists and $B = (A' \wedge B) + A$.

We shall use the term *orthomodular* to describe such an algebraic system.

States

A state of the system is considered to assign to each event its probability of occurrence and is, in fact, considered to be uniquely determined by this map. Probabilities are required to be additive on disjoint events in order to conform to statistical interpretations. Here, however, a distinction must be made, according to whether infinite sequences of events are allowed or not. Not all authors agree on this. We shall use the term *state* for maps $m: \mathfrak{L} \to$ interval $[0, 1]$ such that $m(\sum A_k) = \sum m(A_k)$ for all (finite or infinite) sequences of pairwise disjoint events; a map for which this condition is satisfied only for finite sequences will be called a *quasistate*. In a sense quasistates occur in limiting situations; it is conceivable that as a system passes successively through a sequence of states $m_1, m_2, \cdots$, the probabilities $m_1(A), m_2(A), \cdots$ converge to a value $q(A)$ for all events A; the map q will, in general, not be a state, but a quasistate. In Ref. 8 we have shown that under a certain physically plausible condition the converse is also true, although instead of simple sequences more general convergence elements such as nets may be needed. We shall assume that the set of all states of $\mathfrak{L}$ is *full*, i.e., that if $m(A) \leq m(B)$ for all m, then $A \leq B$. A convex combination of states $\sum_{i=1}^{\infty} \lambda_i m_i$ (where $\lambda_i \geq 0$, $\sum \lambda_i = 1$) is always a state corresponding to the statistical mixture of the m_i. There exist, however, more general situations. We shall say that the state m is the μ *mixture* of the family $(m_x)_{x \in X}$ if, for all $A \in \mathfrak{L}$, $m(A) = \int_X m_x(A)\, d\mu(x)$; here μ is a probability measure on X, and each m_x is assumed to be a quasistate. If each m_x is a state, then any μ mixture will be a state too, the converse being false in general. We shall say that a family $(m_x)_{x \in X}$ is *generating* if there exists a σ-algebra $\mathfrak{X}$ of sets in X such that every state m of the system is a μ mixture of this family for a suitable probability measure μ on $\mathfrak{X}$.

Observables

An *observable* is defined as a map u from the Borel sets in the reals R to $\mathfrak{L}$ such that $u(\varnothing) = 0$, $u(R) = I$, $E \subseteq F$ implies $u(E) \leq u(F)$ and $E_i \cap E_j = \varnothing$ implies $u(E_i) \perp u(E_j)$ and $u(\bigcup E_i) = \sum u(E_i)$. The interpretation is that $u(E)$ is the event of u having a value in the set E. For any Borel function $f: R \to R$, the observable $f(u)$ is the map $u \circ f^{-1}$ (composition). The observable u is *bounded* if $u(E) = 0$ for all E disjoint from some bounded interval. If m is a state of the system, then $p_{m,u}: E \to m(u(E))$ is the *probability distribution* of u in this state. The *expectation* of u in the state m is $\int_{-\infty}^{+\infty} \lambda\, dp_{m,u}(\lambda)$, which we shall write as $m(u)$; this exists if u is bounded. Thus the *dispersion* of the values of u in the state m is $m(u^2) - m(u)^2$. Now the events are in a one-to-one correspondence with the observables u for which $u = u^2$, and from the above we have that a state m is such that all observables are measured exactly in it if and only if m assumes on $\mathfrak{L}$ the two values 0 and 1; such a state is called *dispersion free*. The above procedures can, of course, be formally applied to quasistates, and the main difference will be that the "quasiprobability" distributions obtained will be finitely additive. In case the quasistate q is the pointwise (or weak) limit of the states m_i, the expectation of a bounded observable will be the limit of its expectations in the states m_i; such a quasistate which is also dispersion free appears to be the natural object to study in the absence of dispersion-free states.

Classical Systems

Everybody agrees that this concept is somehow associated with a Boolean algebra; there are, however, some variations. It is generally accepted that one can

say that two events A and B are commuting (or simultaneous or compatible) if there exist pairwise disjoint elements A_1, B_1, and C such that $A = A_1 + C$ and $B = B_1 + C$[5,7,9,10] (several equivalent conditions are known). The center $\mathfrak{C}$ of $\mathfrak{L}$ is then defined as the set of all events which commute with everything, and is a Boolean subalgebra of $\mathfrak{L}$ in the sense that the Boolean operations of $\mathfrak{C}$ are actually performed in $\mathfrak{L}$. The standard definition of a classical system is the condition $\mathfrak{C} = \mathfrak{L}$, so that $\mathfrak{L}$ becomes a Boolean σ algebra. Some authors go beyond this in requiring that $\mathfrak{L}$ be a Boolean algebra of sets,[5] and the requirement of atomicity is occasionally thrown in.[3,5] In some cases[5] the extra condition of completeness is required, which is certainly too much, both for physical and for mathematical reasons (for example, no countably additive measures can exist on the algebra of all subsets of certain sets, except the discrete ones, thus excluding from consideration many physically important cases). We wish to give a definition which will not force $\mathfrak{L}$ to be a lattice, but which agrees with the standard one in case it is. A system will be *classical* if it is a subsystem of one for which $\mathfrak{L}$ is a Boolean σ-algebra; it is *strictly classical* if this $\mathfrak{L}$ is a Boolean σ-algebra of sets. We use the term *subsystem* here in the following sense: Let $\mathfrak{L}_1$ and $\mathfrak{L}_2$ be the orthomodular sets of events associated with the two systems $\mathcal{S}_1$ and $\mathcal{S}_2$; we say that $\mathcal{S}_1$ is a subsystem of $\mathcal{S}_2$ if $\mathfrak{L}_1 \subseteq \mathfrak{L}_2$, all operations (including $\sum$) in $\mathfrak{L}_1$ are restrictions of those in $\mathfrak{L}_2$, and every state of $\mathcal{S}_1$ is the restriction of a state of $\mathcal{S}_2$. It is well known that, even for Boolean algebras, infinite suprema need not coincide whenever finite suprema do so and that, even if they do, the third condition need not hold. This third condition has the extremely important interpretation that in order to study some observables only, one need not consider the whole system—which is what is actually done in practice.

It should, perhaps, be pointed out that physically a system behaves classically if all the observables are simultaneously measurable with absolute accuracy (at least in a pure state). Thus one can view the hidden variables theorem as the link between the physically relevant conditions and the algebraic structure of $\mathfrak{L}$.

2. THE STANDARD THEOREM

Consider any generating family $(m_x)_{x\in X}$ of states. To each $A \in \mathfrak{L}$ there corresponds a function $\hat{A}$ on X which takes x into $m_x(A)$.

Lemma: The map $\hat{}$ is one to one; the order on $\mathfrak{L}$ is transformed to pointwise order on the range of $\hat{}$, A' is transformed into $1 - \hat{A}$, and disjoint suprema in $\mathfrak{L}$ are transformed into pointwise sums.

Proof: If $A \leq B$, then $m(A) \leq m(B)$ for all states m, so that in particular $m_x(A) \leq m_x(B)$ for $x \in X$, i.e., $\hat{A}(x) \leq \hat{B}(x)$. Conversely, if this last relation holds, we integrate to obtain $m(A) \leq m(B)$ for all m, so that $A \leq B$. Trivially $\hat{A'}(x) = m_x(A') = 1 - m_x(A) = (1 - \hat{A})(x)$. Now, if $A = \sum A_i$ with A_i pairwise disjoint, then $m_x(A) = \sum m_x(A_i)$, or $\hat{A}(x) = \sum \hat{A}_i(x)$, and the converse follows again by integration which is termwise permitted since all functions are nonnegative.

Now suppose that each m_x is dispersion free, so that $m_x(A)$ is either 0 or 1 for all $A \in \mathfrak{L}$. Then $\hat{A}$ is the characteristic function of some set $\bar{A} \in \mathfrak{X}$, and by the lemma we have that the map $^-$ sends $\mathfrak{L}$ onto a collection of subsets of X so that the operations are transformed into set-theoretic ones. Finally, by hypothesis, each state m gives rise to a measure μ on $\mathfrak{X}$ so that $m(A) = \mu(\bar{A})$. So we have proved

Theorem 1: If $\mathfrak{L}$ admits a generating family of dispersion-free states, then $\mathfrak{L}$ is strictly classical.

So far we have ignored any lattice operations that may exist in $\mathfrak{L}$. Does the above isomorphism preserve such an extra structure? The necessary and sufficient condition is this:

(J–P–Z) For any state m, the relations $m(A) = m(B) = 1$ imply $m(A \wedge B) = 1$ [or, equivalently, $m(A) = m(B) = 0$ imply $m(A \vee B) = 0$ in case both $A \wedge B$, $A \vee B$ exist].

This is included as an axiom (all A, $B \in \mathfrak{L}$) in the system which was studied by Jauch and Piron,[2] but had already been studied by Zierler.[11]

Theorem 2: Under the hypotheses of Theorem 1, the necessary and sufficient condition for the isomorphism to preserve an existing infimum (or supremum) is the J–P–Z condition for this particular pair.

Proof: Let $C = A \wedge B$; we anyway have $\hat{C}(x) \leq \hat{A}(x)$, $\hat{C}(x) \leq \hat{B}(x)$ for all $x \in X$, so that $\bar{C} \subseteq \bar{A} \cap \bar{B}$. If J–P–Z holds, then for $x \in \bar{A} \cap \bar{B}$ we have $\hat{A}(x) = \hat{B}(x) = 1$, or $m_x(A) = m_x(B) = 1$, so that $m_x(C) = 1$, i.e., $x \in \bar{C}$. Conversely, suppose that $\bar{C} = \bar{A} \cap \bar{B}$, so that $\hat{C}(x) = \hat{A}(x)\hat{B}(x)$. If $m(A) = m(B) = 1$, then $\hat{A}(x) = \hat{B}(x) = 1$ a.e. relative to the measure μ corresponding to m; since $\hat{A}, \hat{B} \leq 1$, we obtain $\hat{C}(x) = 1$ a.e. relative to μ, and so $m(C) = \mu(\bar{C}) = 1$. A similar argument works for $A \vee B$.

In connection with the preservation of the lattice operations under an order and complement preserving map of $\mathfrak{L}$ in a Boolean algebra $\mathfrak{B}$, we note that

Zierler and Schlessinger[6] have proved that if all existing suprema are preserved, then $A \vee B$ exists in $\mathfrak{L}$ iff A and B commute; in such a case, condition J-P-Z follows easily. Here we have proved more, since our condition involves only the preservation of suprema for the two involved elements. Naturally, our hypotheses are stronger.

The converse to Theorem 1 is valid and its proof trivial; so we omit it.

Theorem 3: If a system is strictly classical, then it admits a generating system of dispersion-free states.

It is clear that the states in question are just the evaluation maps associated with the points of the space X.

3. THE GENERALIZED THEOREM

We shall now investigate the case where no dispersion-free states are known *a priori* to exist, only approximately such. As already mentioned, the proper objects to use in such a case are the dispersion-free quasistates. In the next section we shall defend this point.

Now assume that $(m_x)_{x\in X}$ is a generating system of quasistates, and let $\mathcal{N}$ be the σ-ideal of $\mathfrak{X}$ consisting of those sets on which all measures associated with the states vanish. We define the map $\hat{}$ as before, and note that the first two statements in the lemma are still valid, as well as their proofs. The third statement will change. Let $A = \sum A_i$, so that we have $m(A) = \sum m(A_i)$ for all states; since each m_x is only finitely additive, we have $\sum \hat{A}_i(x) \leq \hat{A}(x)$; but $\int_X (\hat{A}(x) - \sum \hat{A}_i(x))\, d\mu(x) = 0$ for all μ, together with this inequality, implies that $\hat{A} = \sum \hat{A}_i$ modulo $\mathcal{N}$. So we obtain:

Lemma: Let A^* be the class of $\hat{A}$ modulo $\mathcal{N}$; then * maps $\mathfrak{L}$ isomorphically into the space of classes of bounded measurable (relative to $\mathfrak{X}$) functions modulo $\mathcal{N}$.

If we further assume that each m_x is dispersion free, so that the $\hat{A}$ are characteristic functions of sets in $\mathfrak{X}$, we see that the above isomorphism sends $\mathfrak{L}$ into the Boolean σ-algebra $\mathfrak{B} = \mathfrak{X}/\mathcal{N}$. Clearly all algebraic operations are transformed by the lemma to operations in $\mathfrak{B}$, and, since all measures which correspond to states vanish on $\mathcal{N}$, they produce measures on $\mathfrak{B}$ so that $m(A) = \mu(A^*)$. We have proved

Theorem 4: If the system admits a generating family of dispersion-free quasistates, then it is classical.

Again the preservation of arbitrary lattice operations must be discussed.

Theorem 5: Under the hypotheses of Theorem 4, the necessary and sufficient condition for the isomorphism to preserve an existing infimum (or supremum) is the J-P-Z condition for this particular pair.

Proof: First note that the second part of the proof is valid in this case, so that we consider the first part, letting $C = A \wedge B$. Again we have $\bar{C} \subseteq \bar{A} \cap \bar{B} = \{x \mid \hat{A}(x)\hat{B}(x) = 1\}$, and all we need is $\bar{A} \cap \bar{B} - \bar{C}$ to be in $\mathcal{N}$. But for any probability measure μ on $\mathfrak{X}$ the map

$$m : L \to \frac{1}{\mu(\bar{A} \cap \bar{B})} \int_{A\cap B} \hat{L}(x)\, d\mu(x)$$

is a state of the system. Since $\hat{A}$ and $\hat{B}$ are 1 on $\bar{A} \cap \bar{B}$, we have $m(A) = m(B) = 1$, so that $m(C) = 1$ also. But C vanishes outside $\bar{A} \cap \bar{B}$, and thus

$$\mu(\bar{C}) = \int \hat{C}(x)\, d\mu(x) = \int_{\bar{A}\cap\bar{B}} \hat{C}(x)\, d\mu(x) = \mu(\bar{A} \cap \bar{B}).$$

Therefore, $\mu(\bar{A} \cap \bar{B} - \bar{C}) = 0$ for all μ.

Finally we prove the converse to Theorem 4.

Theorem 6: Any classical system admits a generating family of dispersion-free quasistates.

Proof: Consider the locally convex space of all bounded real-valued finitely additive measures on the given Boolean σ-algebra $\mathfrak{B}$ (we impose the weak topology). The probability measures on $\mathfrak{B}$ form a convex set with compact closure, which consists of quasistates of $\mathfrak{B}$. Let X be the set of extreme points in this compact set; since $\mathfrak{B}$ is a Boolean algebra, the elements of X assume values 0 and 1 only, i.e., are dispersion-free quasistates of $\mathfrak{B}$. Therefore, for each $x \in X$ the functional $m_x = x \mid \mathfrak{L}$ is a dispersion-free quasistate of $\mathfrak{L}$. Now by the Choquet–Bishop–de Leeuw theorem[12] any probability measure p on $\mathfrak{B}$ will have the form

$$p(A) = \int_X x(A)\, d\mu(x)$$

for a suitable probability measure μ on X, and, since every state of $\mathfrak{L}$ extends to some such p by our hypothesis, we have the desired generating family.

Corollary: A system is classical if and only if all pure quasistates are dispersion free.

4. REMARKS

In this section we shall make more precise our previous vague remarks on the relation between

approximately dispersion-free states and dispersion-free quasistates.

Theorem 7: Let m_k, $k = 1, 2, \cdots$, be states such that for each $A \in \mathfrak{L}$ we have

$$\lim_{k\to\infty} [m_k(A) - m_k(A)^2] = 0.$$

Then there exists a dispersion-free quasistate.

Proof: On the set N of natural numbers we consider an ultrafilter $\mathfrak{F}$ finer than the Frechet filter, following an idea due to Dieudonne. Consider the space T of all maps $\mathfrak{L} \to [0, 1]$ which is compact for the pointwise convergence and the map $\hat{}$ sending each $k \in N$ to the function $\hat{k}: A \to m_k(A)$ in T. The image of $\mathfrak{F}$ under this map will converge to a limit q. Since each m_k is finitely additive, so is q; i.e., we have obtained a quasistate of $\mathfrak{L}$. On the other hand, $q(A) - q(A)^2$ is the limit of $m_k(A) - m_k(A)^2$ along $\mathfrak{F}$; but this is the same as its limit over the Frechet filter, which is 0 by hypothesis. Thus the values of q are 0 and 1 only, so that q is dispersion free.

By following this line of thought we arrive at a result showing how a generating family of dispersion-free quasistates can be obtained from an "approximately generating" and "approximately dispersion-free" family of states.

Theorem 8: Suppose that a family $(m_{k,x})_{x\in X}$, $k = 1, 2, \cdots$, of states is given such that for each $x \in X$ and each $A \in \mathfrak{L}$ the dispersion of A in $m_{k,x}$ tends to zero. Suppose that there exists in X a σ-algebra of sets $\mathfrak{X}$ and that with each state m of the system a probability measure μ on $\mathfrak{X}$ is associated, so that the integrals $\int_X m_{k,x}(A)\, d\mu(x) = m_k(A)$ exist uniformly in k and converge to $m(A)$ as $k \to \infty$. Then there exists a generating family of dispersion-free quasistates.

Proof: We consider the ultrafilter $\mathfrak{F}$ as in the proof of Theorem 7, so that

$$\lim_{\mathfrak{F}} m_{k,x}(A) = m_x(A)$$

exists for each $x \in X$ and $A \in \mathfrak{L}$; we have that m_x is a dispersion-free quasistate. Thus we have $|m_{k,x}(A) - m_x(A)| < \epsilon$ provided that $k \in$ some $U_x \in \mathfrak{F}$ (A is fixed). Given any $a_1, a_2, \cdots, a_n \geq 0$ with $\sum a_i = 1$ and $x_1, x_2, \cdots, x_n \in X$, we shall then have

$$\left| \sum_{i=1}^n a_i m_{x_i}(A) - \sum_{i=1}^n a_i m_{k,x_i}(A) \right| < \epsilon \quad \text{for all} \quad k \in \bigcap_{i=1}^n U_{x_i}.$$

On the other hand, by the uniform integrability of the functions $x \to m_{k,x}(A)$, there exists a partition $X_1, \cdots, X_r \in \mathfrak{X}$ such that, for any other finer partition $E_1, \cdots, E_n$ of X and any choice of $x_i \in E_i$, we shall have

$$\left| m_k(A) - \sum_{i=1}^n \mu(E_i) m_{k,x_i}(A) \right| < \epsilon \quad \text{for all} \quad k = 1, 2, \cdots.$$

We also have $|m(A) - m_k(A)| < \epsilon$ for $k > k(\epsilon)$. It follows that if $k \in \bigcap_{i=1}^n U_{x_i}$ and $x_i \in E_i$, we then have

$$\left| \sum_{i=1}^n \mu(E_i) m_{x_i}(A) - \sum_{i=1}^n \mu(E_i) m_{k,x_i}(A) \right| < \epsilon; \quad \text{but} \quad \bigcap_{i=1}^n U_{x_i} \in \mathfrak{F},$$

which is finer than the Frechet filter and hence contains a $k > k(\epsilon)$. Therefore, we have for any $x_i \in E_i$ the relation

$$\left| m(A) - \sum_{i=1}^n \mu(E_i) m_{x_i}(A) \right| < 2\epsilon,$$

provided that the partition (E_i) is finer than the partition (X_j). This means that

$$m(A) = \int_X m_x(A)\, d\mu(x).$$

As a concluding remark, we shall compare the whole situation to what occurs within the totally different axiomatic scheme proposed by Segal.[13] In this context a classical system is represented in a natural way by a commutative (associative) Banach $*$-algebra consisting of the observables. Segal has proved[14] the hidden variables theorem in the following form: If the (pure) states of the system in which all observables are measured exactly (have zero dispersion) separates the observables, then the system is classical. Note that a state in this context is identified with the expectation functional on the observables. Although the hypothesis here appears to be much weaker, it is not hard to see that it is actually equivalent to ours. It would be interesting to investigate this equivalence in the system we have been studying.

[1] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer-Verlag, Berlin, 1932).
[2] J. Jauch and C. Piron, Helv. Phys. Acta **36**, 827 (1963).
[3] J. Jauch, *Foundations of Quantum Mechanics* (Addison-Wesley, Reading, Mass., 1968).
[4] S. Kochen and E. Specker, J. Math. Mech. **17**, 59 (1967).
[5] S. Gudder, Proc. Am. Math. Soc. **19**, 319 (1968).
[6] N. Zierler and M. Schlessinger, Duke Math. J. **32**, 251 (1965).
[7] G. Mackey, *The Mathematical Foundations of Quantum Mechanics* (Benjamin, New York, 1963).
[8] P. C. Deliyannis, J. Math. Phys. **12**, 860 (1971).
[9] A. Ramsey, J. Math. Mech. **15**, 227 (1966).
[10] V. Varadarajan, *Geometry of Quantum Theory* (Van Nostrand, Princeton, N.J., 1968).
[11] N. Zierler, Pacific J. Math. **11**, 1151 (1961).
[12] R. Phelps, *Lectures on Choquet's Theorem* (Van Nostrand, Princeton, N. J., 1966).
[13] I. E. Segal, Ann. Math. **48**, 930 (1947).
[14] I. E. Segal, "A Mathematical Approach to Elementary Particles and Their Fields" (University of Chicago Lecture Notes, 1955) (unpublished).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Generalized Rajnak–Wybourne Identity

J. OREG*
Department of Theoretical Physics, The Hebrew University of Jerusalem, Israel

(Received 15 October 1970)

The representation of the electrostatic interaction by unit operators makes it possible to derive a general equation which provides many identities. The Rajnak–Wybourne identity is obtained as a special case. Some interesting examples are considered. The application of the identities to the construction of the electrostatic energy matrices by computer is discussed.

I. DERIVATION OF THE GENERALIZED IDENTITY

The electrostatic interaction between the electrons in the "central field approximation" can be decomposed in the following way:

$$\sum_{i<j} e^2/r_{ij} = e^2 \sum_k \sum_{n_a l_a n_b l_b n_a' l_a' n_b' l_b'} X^k(n_a l_a n_b l_b, n_a' l_a' n_b' l_b') \times \sum_{i<j} Z_i^{(k)}(n_a l_a, n_a' l_a') \cdot Z_j^{(k)}(n_b l_b, n_b' l_b'). \quad (1)$$

The quantities $X^k(n_a l_a n_b l_b, n_a' l_a' n_b' l_b')$ and the operators $Z_i^{(k)}(nl, n'l')$ are defined in Ref. 1 (see also the Appendix). Since

$$X^k(n_a l_a n_b l_b, n_a' l_a' n_b' l_b') = X^k(n_b l_b n_a l_a, n_b' l_b' n_a' l_a') \quad (2a)$$

and

$$[Z_i^{(k)}(nl, n'l'), Z_j^{(k)}(n''l'', n'''l''')] = 0 \quad \text{for} \quad i \neq j, \quad (2b)$$

when combining terms in (1), the coefficient of $X^k(n_a l_a n_b l_b, n_a' l_a' n_b' l_b')$ becomes

$$[1 + \delta(n_a, n_b)\delta(l_a, l_b)\delta(n_a', n_b')\delta(l_a', l_b')]^{-1} \times S^k(n_a l_a n_b l_b, n_a' l_a' n_b' l_b'), \quad (3)$$

where

$$S^k(n_a l_a n_b l_b, n_a' l_a' n_b' l_b') = \sum_{i \neq j} Z_i^{(k)}(n_a l_a, n_a' l_a') \cdot Z_j^{(k)}(n_b l_b, n_b' l_b').$$

In the "Racah–Slater method" one considers the quantities $X^k(n_a l_a n_b l_b, n_a' l_a' n_b' l_b')$ as parameters which multiply the matrices of the operators (3).

The operators $S^k(n_a l_a n_b l_b, n_a' l_a' n_b' l_b')$, which are essentially two-particle operators, can be expressed by one-particle operators in the usual way[2]:

$$\begin{aligned} &S^k(n_a l_a n_b l_b, n_a' l_a' n_b' l_b') \\ &= \sum_{i \neq j} Z_i^{(k)}(n_a l_a, n_a' l_a') \cdot Z_j^{(k)}(n_b l_b, n_b' l_b') \\ &= \sum_{i,j} Z_i^{(k)}(n_a l_a, n_a' l_a') \cdot Z_j^{(k)}(n_b l_b, n_b' l_b') \\ &\quad - \sum_{i=j} Z_i^{(k)}(n_a l_a, n_a' l_a') \cdot Z_j^{(k)}(n_b l_b, n_b' l_b') \\ &= Z^{(k)}(n_a l_a, n_a' l_a') \cdot Z^{(k)}(n_b l_b, n_b' l_b') \\ &\quad - \delta(l_a, l_b')\delta(n_b, n_a')\delta(l_b, l_a')[l_a]^{-\frac{1}{2}} Z^{(0)}(n_a l_a, n_b' l_a'), \end{aligned} \quad (4)$$

where

$$Z^{(k)}(nl, n'l') = \sum_i Z_i^{(k)}(nl, n'l').$$

The term containing $Z^{(0)}(n_a l_a, n_a' l_a)$ is obtained by using equations (15.16) and (14.5) of Fano and Racah.[3] (This reference is referred to as F.R. in the sequel.)

Using (2b) again, one may also write

$$S^k(n_a l_a n_b l_b, n_a' l_a' n_b' l_b')$$

in the form

$$\begin{aligned} &S^k(n_a l_a n_b l_b, n_a' l_a' n_b' l_b') \\ &= \sum_{i \neq j} Z_i^{(k)}(n_b l_b, n_b' l_b') \cdot Z_j^{(k)}(n_a l_a, n_a' l_a') \\ &= Z^{(k)}(n_b l_b n_b' l_b') \cdot Z^{(k)}(n_a l_a, n_a' l_a') \\ &\quad - \delta(l_b, l_a')\delta(n_a, n_b')\delta(l_a, l_b')[l_b]^{-\frac{1}{2}} Z^{(0)}(n_b l_b, n_a' l_b'). \end{aligned} \quad (5)$$

Equating expressions (4) and (5), one obtains the identity

$$\begin{aligned} &Z^{(k)}(n_a l_a, n_a' l_a') \cdot Z^{(k)}(n_b l_b, n_b' l_b') \\ &\quad - \delta(l_a, l_b')\delta(n_b, n_a')\delta(l_b, l_a')[l_a]^{-\frac{1}{2}} Z^{(0)}(n_a l_a, n_b' l_a') \\ &= Z^{(k)}(n_b l_b, n_b' l_b') \cdot Z^{(k)}(n_a l_a, n_a' l_a') \\ &\quad - \delta(l_b, l_a')\delta(n_a, n_b')\delta(l_a, l_b')[l_b]^{-\frac{1}{2}} Z^{(0)}(n_b l_b, n_a' l_b'). \end{aligned} \quad (6)$$

Taking matrix elements of Eq. (6) between antisymmetrized states ψ and ψ' according to F.R. (15.16), one finally obtains the identity

$$\begin{aligned} &\sum_{\psi''} (-1)^{J-J''-k}[J]^{-1}(\psi \| Z^{(k)}(n_a l_a, n_a' l_a') \| \psi'') \\ &\qquad \times (\psi'' \| Z^{(k)}(n_b l_b, n_b' l_b') \| \psi') \\ &\qquad - \delta(l_a, l_b')\delta(n_b, n_a')\delta(l_b, l_a')[l_a, j]^{-\frac{1}{2}} \\ &\qquad \times (\psi \| Z^{(0)}(n_a l_a, n_b' l_a) \| \psi') \\ &= \sum_{\psi'''} (-1)^{J-J'''-k}[J]^{-1}(\psi \| Z^{(k)}(n_b l_b, n_b' l_b') \| \psi''') \\ &\qquad \times (\psi''' \| Z^{(k)}(n_a l_a, n_a' l_a') \| \psi') \\ &\qquad - \delta(l_b, l_a')\delta(n_a, n_b')\delta(l_a, l_a')[l_b, J]^{-\frac{1}{2}} \\ &\qquad \times (\psi \| Z^{(0)}(n_b l_b, n_a' l_b) \| \psi'), \end{aligned} \quad (7)$$

where J, J'', and J''' are the total angular momenta of the states ψ, ψ'', and ψ''', respectively.

Because of the particular definition of the operators $Z^{(k)}$, the summations over ψ'' and ψ''' are each reduced to states of a single configuration, Γ'' and Γ''' respectively, which will be referred to in the following as intermediate configurations.

Identity (7) is of a general form, which reduces to special cases by specifying the initial and final states and the interaction operators. The Rajnak–Wybourne identity[4,5] is obtained as one such special case. This, as well as some other interesting cases, is described in Sec. II. In Sec. III we discuss the use of Eq. (7) in the construction of the matrices of $S^k(n_a l_a n_b l_b, n_a' l_a' n_b' l_b')$ by computer.

II. EXAMPLES

A. $(nl)^N - (nl)^{N-1}n'l'$ Configuration Interaction—The Rajnak–Wybourne Identity

Limiting the states ψ and ψ' in (7) to the configurations $(nl)^N$ and $(nl)^{N-1}n'l'$ respectively, the intermediate configurations will be $\Gamma'' = (nl)^N$ and $\Gamma''' = (nl)^{N-1}n'l'$, and Eq. (7) will be of the form

$$\begin{aligned}
&\sum_{(nl)^N\alpha''L''} (-1)^{L-L''-k}(L)^{-1} \\
&\quad\times [(nl)^N\alpha SL \,\|Z^{(k)}(nl, nl)\|\, (nl)^N\alpha'' SL''] \\
&\quad\times ((nl)^N\alpha'' SL'' \,\|Z^{(k)}(nl, n'l')\|\, (nl)^{N-1}(\alpha_1' S_1' L_1')n'l'SL) \\
&\quad - \delta(l, l')[l, L]^{-\frac{1}{2}} \\
&\quad\times ((nl)^N\alpha SL \,\|Z^{(0)}(nl, n'l)\|\, (nl)^{N-1}(\alpha_1' S_1' L_1')n'l'SL) \\
&= \sum_{(nl)^{N-1}\alpha_1'' S_1'' L_1'' n'l'L''} (-1)^{L-L''-1}[L]^{-1} \\
&\quad\times ((nl)^N\alpha SL \,\|Z^{(k)}(nl, n'l')\|\, (nl)^{N-1}(\alpha_1'' S_1'' L_1'')n'l'SL'') \\
&\quad\times ((nl)^{N-1}(\alpha_1'' S_1'' L_1'')n'l'SL'' \,\|Z^{(k)}(nl, nl)\| \\
&\quad\times (nl)^{N-1}(\alpha_1' S_1' L_1')n'l'SL). \qquad (7')
\end{aligned}$$

Using the relations

$$\begin{aligned}
&((nl)^N\alpha SL \,\|Z^{(k)}(nl, n'l')\|\, (nl)^{N-1}(\alpha_1' S_1' L_1')n'l'SL') \\
&\quad = N^{\frac{1}{2}}(-1)^{L+L_1'+l'+k}[L, L']^{\frac{1}{2}}\begin{Bmatrix} L & L' & k \\ l' & l & L_1' \end{Bmatrix} \\
&\qquad\qquad\times (l^N\alpha SL \,|\}\, l^{N-1}\alpha_1' S_1' L_1')
\end{aligned}$$

and

$$\begin{aligned}
&((nl)^{N-1}(\alpha_1 S_1 L_1)n'l'SL \,\|Z^{(k)}(nl, nl)\| \\
&\qquad\times (nl)^{N-1}(\alpha_1' S_1' L_1')n'l'SL') \\
&\quad = \delta(S_1, S_1')(-1)^{L'+L_1+l'+k}[L, L']^{\frac{1}{2}}\begin{Bmatrix} L & L' & k \\ L_1' & L_1 & l' \end{Bmatrix} \\
&\qquad\times ((nl)^{N-1}\alpha_1 S_1 L_1 \,\|Z^{(k)}(nl, nl)\|\, (nl)^{N-1}\alpha_1' S_1' L_1')
\end{aligned}$$

and summing over L'', according to F.R. Eq. (I.2), (7') reduces to the Rajnak–Wybourne identity.

B. $(nl)^{4l}n'l' - (nl)^{4l}n''l''$ Exchange Interaction $(l' \neq l'')$

Sometimes, the intermediate configuration on one side of Eq. (7) is much simpler than that on the other side. Its evaluation and the resulting expression are then also much simpler than those of the other side.

As an example we take the exchange interaction between the configurations $l^{4l}l'$ and $l^{4l}l''$,[6] which is represented by the parameters $X^k(ll', l''l)$. In this case identity (6) takes the form

$$Z^{(k)}(l, l'') \cdot Z^{(k)}(l', l) = Z^{(k)}(l', l) \cdot Z^{(k)}(l, l'').$$

Both scalars $Z^{(0)}$ vanish here.[7]

The intermediate configuration on the left-hand side of (7) is $\Gamma'' = l^{4l-1}l''l'$. For this side one obtains specifically

$$\begin{aligned}
&(l^{4l}(S_1L_1)l'SL \,|Z^{(k)}(l, l'') \cdot Z^{(k)}(l', l)|\, l^{4l}(S_1'L_1')l''SL) \\
&\quad = \sum_{l^{4l-1}(\alpha_2S_2L_2)l''S_3L_3, l'L_4} (-1)^{L-L_4-k}[L]^{-1} \\
&\qquad\times (l^{4l}(S_1L_1)l'SL \,\|Z^{(k)}(l, l'')\| \\
&\qquad\times l^{4l-1}(\alpha_2S_2L_2)l''S_3L_3, l'SL_4) \\
&\qquad\times (l^{4l-1}(\alpha_2S_2L_2)l''S_3L_3, l'SL_4 \,\|Z^{(k)}(l', l)\| \\
&\qquad\times l^{4l}(S_1'L_1')l''SL), \qquad (8)
\end{aligned}$$

which, after a long derivation, can be written as

$$\begin{aligned}
&4l \sum_{l^{4l-1}\alpha_2S_2L_2} (-1)^{L+L_1+L_2+l+l'+S_1+1}(L_1, L_1', S_1, S_1')^{\frac{1}{2}} \\
&\qquad\times (l^{4l}S_1L_1 \,|\}\, l^{4l-1}\alpha_2S_2L_2) \\
&\qquad\times (l^{4l-1}\alpha_2S_2L_2 \,\{|\, l^{4l}S_1'L_1') \\
&\qquad\times \begin{Bmatrix} S_2 & \frac{1}{2} & S_1 \\ S & \frac{1}{2} & S_1' \end{Bmatrix}\begin{Bmatrix} L & L_1 & l' \\ l'' & l & k \\ L_1' & L_2 & l \end{Bmatrix}. \qquad (8')
\end{aligned}$$

On the other hand, the intermediate configuration on the right-hand side of (7) is $\Gamma''' = l^{4l+1}$ and the summation therefore reduces to the single term ^2L with $L = l$. Thus this side becomes simply

$$\begin{aligned}
&(l^{4l}(S_1L_1)l'SL \,|Z^{(k)}(l', l) \cdot Z^{(k)}(l, l'')|\, l^{4l}(S_1'L_1')l''SL) \\
&\quad = (-1)^L[L]^{-1}(l^{4l}(S_1L_1)l'SL \,\|Z^{(k)}(l', l)\|\, l^{4l+1}\,{}^2l) \\
&\qquad\times (l^{4l-1}\,{}^2l \,\|Z^{(k)}(l, l'')\|\, l^{4l}(S_1'L_1')l''SL), \qquad (9)
\end{aligned}$$

which can be simply written as

$$\begin{aligned}
&(-1)^{l''+S_1+S_1'}[S_1, L_1, S_1', L_1']^{\frac{1}{2}} \\
&\qquad\times \begin{Bmatrix} L & l & k \\ l & l' & L_1 \end{Bmatrix}\begin{Bmatrix} L & l & k \\ l & l'' & L_1' \end{Bmatrix}, \qquad (9')
\end{aligned}$$

by using Eq. (19) of Ref. 8. This is obviously much simpler than (8'), and its derivation much more straightforward.

C. $(nl)^{4l+1}n'l' - (nl)^{4l+1}n''l''$ Exchange Interaction—Selection Rules

In some cases, one has a closed shell as intermediate configuration on one of the sides of Eq. (7). This can then be used to obtain selection rules whose

derivation by conventional methods requires detailed calculation. As an example we take the exchange interaction between and within configurations containing one hole and one electron.

It is known that the exchange interaction within a configuration of the type $(nl)^{4l+1}n'l'$, which is represented by the parameter $X^k(nln'l', n'l'nl)$, is zero unless $L = k$ and $S = 0$. The same feature occurs also for the exchange part of the interaction between the configurations $(nl)^{4l+1}n'l'$ and $(nl)^{4l+1}n''l''$ (where $l' \neq l''$), which is represented by the parameter $X^k(nln'l', n''l''nl)$, as was found by Goldschmidt.[9]

These facts follow immediately from the right-hand side of (7), where the intermediate configuration becomes the closed shell l^{4l+2}. Specifically, one obtains[7]

$$(l^{4l-1}l'SL\,|Z^{(k)}(l', l)\cdot Z^{(k)}(l, l'')|\,l^{4l-1}l''SL) = (-1)^{L-k}[L]^{-1}(l^{4l+1}l'SL\,\|Z^{(k)}(l', l)\|\,l^{4l+2}\,{}^1S) \times (l^{4l+2}\,{}^1S\,\|Z^{(k)}(l, l'')\|\,l^{4l+1}l''SL).$$

Since the operators $Z^{(k)}$ are diagonal with respect to the total spin, one has $S = 0$. In addition $L = k$, since L, k, and 0 should satisfy the triangular condition.

One may also conclude that for the interactions between $l^{4l+1}l'''^{4l''+1}l' - l^{4l+1}l'''^{4l''+1}l''$ and $l^{4l+1}l'''l' - l^{4l+1}l'''l''$, $S^k(ll', l''l)$ (Ref. 10) vanishes unless $S = \frac{1}{2}$, as was also pointed out by Goldschmidt.[9] This last result is also obtainable by directly using the right-hand side of Eq. (7).

III. APPLICATION TO THE CONSTRUCTION OF THE MATRICES OF S^k BY COMPUTER

The calculation of the matrix elements of the operators S^k is very tedious. It can only be efficiently performed with the aid of a computer. One of the central problems in using a computer for complicated problems is the reliability of the results. It is always difficult to be sure that a complicated and long program is entirely error free. In addition, the input data needed for the calculation (in our case, lists of terms and f.p. tables) might contain mistakes. Finally the computer itself might err.

Since the matrices of the operators S^k are applied in many different calculations, it is essential to have methods for checking them before using them further. The conventional checking methods[11] are based on the previous knowledge of the eigenvalues or the rank of certain combinations of the matrices, and are performed after the matrices have gone through arranging and organizing stages which bring them to an appropriate format for diagonalization. Therefore, when these checks give negative results, it is difficult to know how many mistakes have occurred, at what stage they have appeared, and to which definite matrix they correspond.

On the other hand, identity (7) enables direct checking of individual matrices, in such a way that each matrix element is checked separately. Therefore, any mistake which occurs is detected at once and at an early stage of the construction process of the matrices.

Since identity (7) is trivial when the intermediate configurations are equal to each other, it does not provide a check for the matrices of the parameters F^k within a configuration, nor for the matrices of interaction between configurations which differ from each other by individual quantum numbers of two equivalent electrons. However, checks are provided in all other cases, which are generally more complicated.

A computer program which makes use of identity (7) in calculating and checking both algebraic expressions and numerical values of the matrix elements for electrostatic interactions between any two configurations was written and is now available in the department of theoretical physics of the Hebrew University of Jerusalem.[12] Identity (7) has been found to be an extremely useful tool in all debugging stages of the program.

ACKNOWLEDGMENT

I wish to thank Dr. Z. B. Goldschmidt and Dr. J. Stein for reading the manuscript and for helpful comments.

APPENDIX

1. Definition of $Z_i^{(k)}(nl, n'l')$: $Z_i^{(k)}(nl, n'l')$ is a tensor operator of degree k, operating on the ith electron, whose reduced matrix elements are

$$(n''l''\,\|Z_i^{(k)}(nl, n'l')\|\,n'''l''') = \delta(n''n)\delta(l''l)\delta(n'''n')\delta(l'''l').$$

As was pointed out by Judd,[13] $Z^{(k)}(nl, n'l')$ is actually proportional to coupled creation and destruction operators. $Z_i^{(k)}(nl, n'l')$, operating on the left, destroys an nl electron and creates an $n'l'$ electron. This property is often used.

2. The parameter $X^k(n_al_an_bl_b, n'_al'_an'_bl'_b)$ is defined by

$$X^k(n_al_an_bl_b, n'_al'_an'_bl'_b) = (l_a\,\|c^{(k)}\|\,l'_a)(l_b\,\|c^{(k)}\|\,l'_b)R^k(n_al_an_bl_b, n'_al'_an'_bl'_b),$$

where $R^k(n_al_an_bl_b, n'_al'_an'_bl'_b)$ are the Slater integrals and $C^{(k)}$ are the spherical harmonics normalized according to Eq. (5.19) of Ref. 3.

* This work was supported by the National Bureau of Standards, Washington, D.C.

[1] G. Racah and J. Stein, Phys. Rev. **156**, 58 (1967).

[2] We use the traditional notation $[l_1, l_2, l_3, \cdots] = (2l_1 + 1) \times (2l_2 + 1)(2l_3 + 1) \cdots$.

[3] U. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic, New York, 1959), referred to as F.R. in the sequel.

[4] K. Rajnak and B. G. Wybourne, Phys. Rev. **132**, 280 (1963), Appendix.

[5] Judd obtained[13] the Rajnak–Wybourne identity in a straightforward way making use of the anticommutation relations of creation and destruction operators.

[6] For simplicity of notation we usually omit the quantum numbers n.

[7] $Z^{(0)}(n'l', n''l'') = 0$ for $l' \neq l''$. When $n' = n''$ and $l' = l''$, $Z^{(0)}$ contributes an additive constant which may be added to the coefficient of $F^{(0)}$. Only when $n' \neq n''$ and $l' = l''$, $Z^{(0)}$ must be included.

[8] G. Racah, Phys. Rev. **63**, 367 (1943).

[9] Z. B. Goldschmidt, Ph.D. thesis, The Hebrew University, Jerusalem, 1969; Phys. Rev. A **3**, 1872 (1971)

[10] For the first interaction $S^k(n'''l'''n'l', n''l''n'''l''')$ vanishes as well unless $S = \frac{1}{2}$.

[11] G. Racah, Bull. Res. Council Israel **8F**, (1959).

[12] Description of the program and the method of calculation will be published.

[13] B. R. Judd, *Second Quantization and Atomic Spectroscopy* (Johns Hopkins, Baltimore, Maryland, 1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Abstract Formulation of the Quantum Mechanical Oscillator Phase Problem

EVANGELOS K. IFANTIS
"Democritos" Nuclear Research Center, Aghia Paraskevi Attikis, Athens, Greece

(Received 28 September 1970)

The phase operators "cosine" and "sine" are characterized as a special and peculiar class of tridiagonal operators, defined on an abstract separable Hilbert space. The two disjoint sets of these operators, which lie on the unit sphere of the algebra of bounded operators, are convex and closed in the uniform topology. The whole treatment gives a new and systematic aspect to the quantum mechanical oscillator phase problem.

I. INTRODUCTION

The quantum mechanical phase problem has both practical and tutorial character: Practical, because of its applications to the laser physics, superconductivity, and superfluidity; tutorial, because it is related to the fundamentals of quantum mechanics.

Quantum mechanically the phase problem begins with the definition of the phase operators C and S ("cosine" and "sine"), which satisfy commutation rules analogous to the classical Poisson bracket relations

$$\{\cos\phi, H\} = \omega \sin\phi, \quad \{\sin\phi, H\} = -\omega\cos\phi,$$

where

$$H = (2m)^{-1}[p^2 + (m\omega q)^2]$$

is the harmonic oscillator and $\phi = \arg(m\omega q + ip)$.

In previous work[1–5] on this problem the following operators have been introduced:

$$C = \tfrac{1}{2}(U + U^*), \quad S = \frac{1}{2i}(U - U^*).$$

The operator U is defined as

$$U|n\rangle = a(n)|n-1\rangle,$$

and satisfies the relation

$$[U, N] = UN - NU = U,$$

where N is the oscillator number operator, $|n\rangle$, $n = 0, 1, 2, \cdots$, its normalized eigenstates, and $a(n)$ a real suitable[1] sequence.

Since there exists not only one but a class of phase operators which satisfy the commutation rules

$$[C, N] = iS, \quad [S, N] = -iC,$$

the problem was to choose phase operators C and S leading to reasonable physical results.[2,3] Another problem is to study the general common properties of the phase operators. To this end we give in this paper an abstract formulation of the problem. The problem of characterizing the class of phase operators appears in the present formulation as a peculiar case of the perturbation problem of continuous spectra.

The phase operators appear as a special and peculiar class of tridiagonal operators of the first kind (and the angle-operators as a special class of tridiagonal operators of the second kind), i.e., the phase operators appear as a special and peculiar

* This work was supported by the National Bureau of Standards, Washington, D.C.

[1] G. Racah and J. Stein, Phys. Rev. **156**, 58 (1967).

[2] We use the traditional notation $[l_1, l_2, l_3, \cdots] = (2l_1 + 1) \times (2l_2 + 1)(2l_3 + 1) \cdots$.

[3] U. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic, New York, 1959), referred to as F.R. in the sequel.

[4] K. Rajnak and B. G. Wybourne, Phys. Rev. **132**, 280 (1963), Appendix.

[5] Judd obtained[13] the Rajnak–Wybourne identity in a straightforward way making use of the anticommutation relations of creation and destruction operators.

[6] For simplicity of notation we usually omit the quantum numbers n.

[7] $Z^{(0)}(n'l', n''l'') = 0$ for $l' \neq l''$. When $n' = n''$ and $l' = l''$, $Z^{(0)}$ contributes an additive constant which may be added to the coefficient of $F^{(0)}$. Only when $n' \neq n''$ and $l' = l''$, $Z^{(0)}$ must be included.

[8] G. Racah, Phys. Rev. **63**, 367 (1943).

[9] Z. B. Goldschmidt, Ph.D. thesis, The Hebrew University, Jerusalem, 1969; Phys. Rev. A **3**, 1872 (1971)

[10] For the first interaction $S^k(n'''l'''n'l', n''l''n'''l''')$ vanishes as well unless $S = \frac{1}{2}$.

[11] G. Racah, Bull. Res. Council Israel **8F**, (1959).

[12] Description of the program and the method of calculation will be published.

[13] B. R. Judd, *Second Quantization and Atomic Spectroscopy* (Johns Hopkins, Baltimore, Maryland, 1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Abstract Formulation of the Quantum Mechanical Oscillator Phase Problem

EVANGELOS K. IFANTIS
"Democritos" Nuclear Research Center, Aghia Paraskevi Attikis, Athens, Greece

(Received 28 September 1970)

The phase operators "cosine" and "sine" are characterized as a special and peculiar class of tridiagonal operators, defined on an abstract separable Hilbert space. The two disjoint sets of these operators, which lie on the unit sphere of the algebra of bounded operators, are convex and closed in the uniform topology. The whole treatment gives a new and systematic aspect to the quantum mechanical oscillator phase problem.

I. INTRODUCTION

The quantum mechanical phase problem has both practical and tutorial character: Practical, because of its applications to the laser physics, superconductivity, and superfluidity; tutorial, because it is related to the fundamentals of quantum mechanics.

Quantum mechanically the phase problem begins with the definition of the phase operators C and S ("cosine" and "sine"), which satisfy commutation rules analogous to the classical Poisson bracket relations

$$\{\cos\phi, H\} = \omega \sin\phi, \quad \{\sin\phi, H\} = -\omega\cos\phi,$$

where

$$H = (2m)^{-1}[p^2 + (m\omega q)^2]$$

is the harmonic oscillator and $\phi = \arg(m\omega q + ip)$.

In previous work[1–5] on this problem the following operators have been introduced:

$$C = \tfrac{1}{2}(U + U^*), \quad S = \frac{1}{2i}(U - U^*).$$

The operator U is defined as

$$U|n\rangle = a(n)|n-1\rangle,$$

and satisfies the relation

$$[U, N] = UN - NU = U,$$

where N is the oscillator number operator, $|n\rangle$, $n = 0, 1, 2, \cdots$, its normalized eigenstates, and $a(n)$ a real suitable[1] sequence.

Since there exists not only one but a class of phase operators which satisfy the commutation rules

$$[C, N] = iS, \quad [S, N] = -iC,$$

the problem was to choose phase operators C and S leading to reasonable physical results.[2,3] Another problem is to study the general common properties of the phase operators. To this end we give in this paper an abstract formulation of the problem. The problem of characterizing the class of phase operators appears in the present formulation as a peculiar case of the perturbation problem of continuous spectra.

The phase operators appear as a special and peculiar class of tridiagonal operators of the first kind (and the angle-operators as a special class of tridiagonal operators of the second kind), i.e., the phase operators appear as a special and peculiar

class of operators, the general properties of which were studied from a different point of view.[6-8]

In the present paper we derive easily all the known results on the mathematical properties of phase operators. Moreover, we prove that the phase operators form two disjoint convex sets on the unit sphere of bounded operators, which are closed in the uniform topology. Finally some general properties of all the phase operators are obtained.

II. THE UNILATERAL SHIFT AND WEIGHTED SHIFT OPERATOR

Let $\mathcal{H}$ be a separable Hilbert space with the orthonormal basis $\{e_n\}_1^\infty$. The unilateral shift operator $V: Ve_n = e_{n+1}$ is an isometry from $\mathcal{H}$ onto $\mathcal{H}\theta\{e_1\}$ and its adjoint $V^*: V^*e_n = e_{n-1}$ for $n > 1$ and $V^*e_1 = 0$, a partial isometry from $\mathcal{H}\theta\{e_1\}$ onto $\mathcal{H}$. The spectrum of V is the closed unit disk in the complex plane and it is purely continuous except for the points $z: |z| < 1$ which belong to the residual spectrum. Every point z in the interior of the unit disk is a proper value of V^* with the corresponding normalized proper element

$$f_z = (1 - |z|^2)^{\frac{1}{2}} \sum_{n=1}^{\infty} z^{n-1} e_n.$$

The points on the circumference belong to the continuous spectrum of V^* and the residual spectrum is empty.

On the other hand, the spectrum of the self-adjoint operator $V + V^*$ is purely continuous and it is confined to the closed interval $[-2, 2]$.[7]

The normalized elements $f_z \in \mathcal{H}$, not being proper values of V, have the property

$$Vf_z = z^{-1}[f_z - (1 - |z|^2)^{\frac{1}{2}} e_1],$$

which is convenient for the determination of the expectation values of $V + V^*$ in the states f_z.

If A is a diagonal operator $Ae_n = a(n)e_n$, $n = 1, 2, \cdots$, then a left-weighted[9,10] unilateral shift is by definition the operator $W = AV$. It is easy to see that W is hyponormal, i.e., the relation $WW^* - W^*W \geq 0$ or $WW^* - W^*W \leq 0$ holds, if and only if $a(n)$ is monotone. For hyponormal operators it is well known[11] that the spectral radius is equal to the norm. They also have many similar properties with the normal operators.

The operators V, V^*, W, and W^* play an important role in operator theory. From a purely mathematical point of view, V is characterized[9] as a universal operator. Many counterexamples, some of them very important,[12,13] were constructed in the past with the help of the operators V and V^*. For instance, V is an example of an isometry which is not unitary, and in the case $a(n) = 1/n$, $n = 1, 2, 3, \cdots$, AV is an example of a compact operator for which the point spectrum is empty.[14]

III. THE TRIDIAGONAL OPERATORS IN HILBERT SPACE

Let $a(n)$, $b(n)$, and $d(n)$ be complex sequences. Define the operators A, B, and D as follows:

$$Ae_n = a(n)e_n, \quad Be_n = b(n)e_n, \quad De_n = d(n)e_n, \quad n = 1, 2, \cdots.$$

A tridiagonal operator of the first kind is an operator of the form

$$T = AV + BV^* + D. \tag{1}$$

In case $A = B = I$ the operator (1) is called Schrödinger-type tridiagonal operator of the first kind. (The tridiagonal operators of the second kind are defined in a similar way if instead of the unilateral shift we use the bilateral one; see Appendix A.)

The nature of the spectrum of self-adjoint Schrödinger-type tridiagonal operators of the first kind was studied in Ref. 6. The nature of the spectrum of tridiagonal operators of the first and the second kind in their general form, some typical examples, as well as the advantages of the approach to the study of difference equations were discussed in Ref. 8.

For real sequences $a(n)$, $b(n)$, and $d(n)$ it is easy to see that operator (1) is self-adjoint if and only if $a(n) = b(n + 1)$. We shall see later that the oscillator phase operators C and S are a special case of bounded self-adjoint tridiagonal operators of the first kind.

IV. THE FORM OF THE OPERATORS C AND S

Definition: The oscillator number operator N is defined as follows:

$$N: Ne_n = (n - 1)e_n, \quad n = 1, 2, \cdots.$$

Since $Ne_1 = 0$, it is easy to see that

$$N^{\frac{1}{2}} V V^* N^{\frac{1}{2}} = N,$$

so that the well-known creation and annihilation operators a and a^* are the following:

$$a^* = N^{\frac{1}{2}} V, \quad a = V^* N^{\frac{1}{2}}.$$

In what follows we shall use the commutation relation

$$V^*N - NV^* = V^*. \tag{2}$$

Now we are able to find easily the form of the phase operators C and S from the requirement that they are bounded and satisfy the relations

$$[C, N] = iS, \quad [S, N] = -iC. \tag{3}$$

Proposition 1: For bounded C and S the relations (3) are satisfied if and only if there exists a bounded operator U such that

$$[U, N] = U.$$

Proof: For the sufficiency observe that

$$[N, U^*] = U^*$$

and

$$[U + U^*, N] = U - U^*,$$
$$[U - U^*, N] = U + U^*.$$

Relations (3) follow by taking

$$C = \tfrac{1}{2}(U + U^*), \quad S = (1/2i)(U - U^*).$$

The necessity follows from (3) if we take

$$U = C + iS.$$

Proposition 2: U is the adjoint of a unilateral weighted shift operator. Especially it is of the form $U = V^*A$.

Proof: Set $VU = A$. Due to $V^*V = I$ we then have $U = V^*A$. Moreover, because of (2),

$$V^*(AN - NA) = 0. \tag{4}$$

From (4) it follows that

$$ANf - NAf = \lambda Pf, \tag{5}$$

where P projects on the subspace, spanned by the element e_1. Relation (5) must be satisfied for every $f \in D(N)$, i.e., for every f in the definition domain of N. Thus for $f = e_1$ it follows that $\lambda = 0$. Therefore,

$$ANf = NAf, \quad \forall f \in D(N). \tag{6}$$

From (6) we easily derive that A is a diagonal operator, i.e.,

$$A: Ae_n = a(n)e_n, \quad n = 1, 2, \cdots.$$

In case $A = I$ we have $U = V^*$ and $C = \frac{1}{2}(V^* + V)$, $S = (1/2i)(V^* - V)$. In the general case

$$C = \tfrac{1}{2}(V^*A + A^*V), \quad S = (1/2i)(V^*A - A^*V). \tag{7}$$

Thus, we have the following proposition.

Proposition 3: The phase operators C and S are self-adjoint tridiagonal operators of the first kind.

The set of phase operators must be characterized from the set of sequences $a(n)$ for which the spectrum of C is the entire interval $[-1, 1]$. As far as the spectrum of S is concerned, it is easy to see that S is unitarily equivalent to C. In fact, if we define the unitary operator

$$B: Be_n = i^n e_n, \quad n = 1, 2, \cdots, \quad i^2 = -1,$$

then

$$B^{-1}SBe_n = (1/2i)B^{-1}(V^*A - A^*V)e_n$$
$$= \tfrac{1}{2}(V^*A + A^*V)e_n = Ce_n,$$

and, for every f in $\mathfrak{H}$,

$$B^{-1}SBf = \sum_{n=1}^{\infty}(B^{-1}SBf, e_n)e_n = \sum_{n=1}^{\infty}(f, B^{-1}SBe_n)e_n$$
$$= \sum_{n=1}^{\infty}(f, Ce_n)e_n = Cf, \quad \text{i.e.,} \quad B^{-1}SB = C.$$

Thus, the operators C and S have exactly the same continuous and point spectrum.

It is also well known[1] that, in the special class of the tridiagonal operators (7), the sequence $a(n)$ can be taken, without restriction of the generality, as positive. We write

$$C = \tfrac{1}{2}(V^*A + AV), \tag{8}$$

and we restrict ourselves to the study of the spectrum of the operator (8).

The sequence $a(n)$ because of the classical relation $C^2 + S^2 = 1$ is assumed to converge to unity.[1] This is due to the heuristic quantal rule: "Quantum results must go over into the corresponding classical ones in the limit of large quantum numbers." In Ref. 1, it was assumed that $a(n) \neq 0$ for $n > 1$ and also that it converges monotonically to unity. In the following we assume that $a(n)$ is convergent.

V. CHARACTERIZATION OF THE CLASS OF PHASE OPERATORS

Assume that $\lim a(n)$ as $n \to \infty$ exists and require that the spectrum of C is the entire interval $[-1, 1]$.

Theorem 1: If $\lim a(n) = \alpha$ as $n \to \infty$, then necessarily $\alpha = 1$.

Proof: Consider the two possible cases $\alpha = 0$ and $\alpha \neq 0$.

1st case ($\alpha = 0$): In this case A (and consequently C) is completely continuous. Therefore, the spectrum of C consists only of isolated proper values, i.e., the spectrum of C does not cover the interval $[-1, 1]$.

2nd case ($\alpha \neq 0$): In this case the operator C can be written

$$C = (\alpha/2)(V + V^*) + \tfrac{1}{2}[(A - \alpha I)V + V^*(A - \alpha I)], \tag{9}$$

where the operator

$$\tfrac{1}{2}[(A - \alpha I)V + V^*(A - \alpha I)]$$

is completely continuous and self-adjoint and the purely continuous spectrum of the operator $\alpha/2(V + V^*)$ is the entire interval $[-\alpha, \alpha]$.

If $\alpha > 1$, then it follows from (9) due to the Weyl's theorem[9,15] that the spectrum of C is always extended beyond the interval $[-1, 1]$. If $\alpha < 1$, then the spectrum of C covers a part of the interval $[-1, 1]$, and the residual part may belong to the resolvent set of C or may contain isolated proper values of C. In any case, the spectrum of C does not cover the interval $[-1, 1]$. Thus $\alpha = 1$.

Theorem 2: If $\lim a(n) = 1$ as $n \to \infty$, then a necessary and sufficient condition in order that the spectrum of C is the entire interval $[-1, 1]$ is

$$\|C\| = 1.$$

Proof: Write the operator C as follows:

$$C = \tfrac{1}{2}(V^* + V) + R, \tag{10}$$

where

$$R = \tfrac{1}{2}[(A - I)V + V^*(A - I)]$$

is completely continuous and self-adjoint.

The theorem is a simple consequence of Weyl's theorem because the purely continuous spectrum of $\frac{1}{2}(V^* + V)$ is the entire interval $[-1, 1]$.

Corollary: If $a(n) \leq 1$ and $\lim a(n) = 1$ as $n \to \infty$, then the spectrum of C is the entire interval $[-1, 1]$.

Proof: Since $a(n) \leq 1$ and $\lim a(n) = 1$ as $n \to \infty$, it follows that $\|A\| = \sup a(n) = 1$. Consequently, $\|C\| \leq 1$. This means

$$\text{sp } (C) = \text{spectrum of } C \subseteq [-1, 1]. \tag{11}$$

But from (10) it follows

$$[-1, 1] \subseteq \text{sp } (C). \tag{12}$$

Relations (11) and (12) complete the proof.

We now ask the question: Is the case in which C has a pure point spectrum dense in $[-1, 1]$, as the Weyl–von Neumann theorem[6,15] predicts, possible? This question is very reasonable because the class of operators of the form $V^*A + AV$ does not cover the class of self-adjoint Hilbert–Schmidt type operators. The fact that $A - I$ [and consequently $(A - 1)V + V^*(A - I)$] is an operator of trace class is sufficient to exclude[15] the above case. The exclusion in the general case seems to be difficult with the methods of perturbation theory.[6,15]

Remark: In order to prove that the spectrum of all phase operators is purely continuous, it is sufficient to prove that the operator $AV + V^*A$ has not proper values in the interval $[-2, 2]$. Suppose that there exists an f in $\mathfrak{H}$ such that

$$(V^*A + AV)f = Ef, \quad |E| \leq 2. \tag{13}$$

The realization of Eq. (13) in the Hilbert space $l_2(1, \infty)$ leads to the difference equation

$$a(n + 1)f(n + 1) + a(n)f(n - 1) = Ef(n). \tag{14}$$

Since $\lim a(n) = 1$ as $n \to \infty$, question arises if the solutions of Eq. (14) behave for large n as the solutions of the equation

$$f(n + 1) + f(n - 1) = Ef(n), \tag{15}$$

which are oscillatory for $|E| < 2$ and of the form $f(n) = c_1 n + c_2$ for $E = 2$, i.e., in any case they do not belong in $l_2(1, \infty)$ (see Appendix B).

The corollary of Theorem 2 says that every sequence $a(n)$ convergent to unity from below defines a phase operator C. Note that every sequence convergent to unity from above does not define a phase operator. In fact, from the relation

$$Ce_n = \tfrac{1}{2}[a(n)e_{n-1} + a(n + 1)e_{n+1}]$$

it follows that

$$\|Ce_n\|^2 = \tfrac{1}{4}[a^2(n) + a^2(n + 1)].$$

Thus, if $a^2(n) + a^2(n + 1) > 4$ for at least one n, then $\|C\| > 1$.

Examples of phase operators constructed from sequences convergent from above to unity have been studied by various authors.[1,2] The following theorem is very useful in connection to this problem.

Theorem 3: The class of "cosine" ("sine") phase operators form a convex set.

Proof: We shall prove that if C_1 and C_2 are phase operators and $0 \leq \mu \leq 1$, then

$$T = \mu C_1 + (1 - \mu)C_2 \tag{16}$$

is also a phase operator.

Obviously T is of the form (8). Since C_1 and C_2 are phase operators due to Theorem 2 we have $\|C_1\| = \|C_2\| = 1$. Thus, from (16), it follows that

$$\|T\| \leq 1. \tag{17}$$

On the other hand, T can be written as follows:

$$T = \tfrac{1}{2}(V + V^*) + R, \tag{18}$$

where

$$R = \tfrac{1}{2}V^*[\mu(A_1 - I) + (1 - \mu)(A_2 - I)] + \tfrac{1}{2}[\mu(A_1 - 1) + (1 - \mu)(A_2 - 1)]V,$$

and A_1 and A_2 are the diagonal operators characterizing C_1 and C_2, respectively. Since R is completely continuous and self-adjoint, it follows from (18) that the spectrum of T covers the interval $[-1, 1]$. Finally, from (17) it follows that the spectrum cannot extend beyond $[-1, 1]$.

Theorem 4: The convex sets of "cosine" and "sine" operators are closed in the uniform topology.

Proof: Consider a sequence of "cosine" operators,

$$C_n = \tfrac{1}{2}(V^*A_n + A_nV),$$

such that $\lim \|C_n - C\| = 0$ as $n \to \infty$. Since $\|C_n\| = 1$, it follows that $\|C\| \leq 1 + \|C_n - C\| \leq 1$ and $\|C\| \geq 1 - \|C_n - C\| \geq 1$, i.e.,

$$\|C\| = 1. \tag{19}$$

But

$$C_n = \tfrac{1}{2}(V + V^*) + \tfrac{1}{2}[V^*(A_n - I) + (A_n - I)V],$$

where $A_n - I$ are completely continuous. Therefore, since the set of completely continuous operators is closed in the uniform topology, there exists a completely continuous operator $A - I$ such that

$$C = \tfrac{1}{2}(V + V^*) + \tfrac{1}{2}[V^*(A - I) + (A - I)V], \tag{20}$$

where $\lim \|C_n - C\| = 0$ as $n \to \infty$.

From (19) and (20) it follows that C is a "cosine" operator. This completes the proof that the set of "cosine" operators is closed. The same follows for the set of "sine" operators.

VII. SOME GENERAL PROPERTIES OF THE PHASE OPERATORS

Theorem 5: For every unbounded self-adjoint operator H with a complete system of proper elements and for every phase operator C, there exist normalizable states, which minimize the uncertainty product

$$(\Delta C)^2 \cdot (\Delta H)^2. \tag{21}$$

Proof: The normalizable states, which minimize the uncertainty product (21), are obtained[3,4] as proper elements of the following non-self-adjoint operator

$$C + i\gamma H, \quad \gamma = \text{real}. \tag{22}$$

The operator (22) can be considered, without restriction of the generality, as an operator with compact resolvent (see Ref. 8 or Appendix in Ref. 16). Therefore, operator (22) has a discrete spectrum. Since the spectrum is not empty, it follows that there exists f in $\mathcal{H}$ satisfying the proper value equation

$$(C + i\gamma H)f = \lambda f.$$

Corollary: For every phase operator C, there exist normalizable states which minimize the uncertainty product $(\Delta C)^2 \cdot (\Delta N)^2$.

Theorem 6: There exist normalizable states for which the quantum results of any phase operator C (or S) are not distinguishable from the results of the phase operator

$$\tfrac{1}{2}(V^* + V) \quad [\text{or } (1/2i)(V^* - V)].$$

Proof: All quantum mechanical physical information for C are obtained from the expectation value

$$(Cf, f), \quad f \in \mathcal{H},$$

or

$$(\tfrac{1}{2}(V^* + V)f, f) + (Rf, f),$$

where R is completely continuous and self-adjoint. If f_n ($\|f_n = 1$) is a sequence in $\mathcal{H}$ convergent to null in the weak topology, then, since R is completely continuous, Rf_n is a null sequence in the strong topology. Thus

$$\lim |(Rf_n, f_n)| \leq \lim \|Rf_n\| = 0 \quad \text{as} \quad n \to \infty.$$

Therefore, there exists a k such that, for $n \geq k$, the expectation values of C and $\frac{1}{2}(V^* + V)$ in the states f_n are not distinguishable.

APPENDIX A: THE ANGLE OPERATORS

Let $\{e_n\}_{-\infty}^{+\infty}$ be an orthonormal basis of an abstract separable Hilbert space $\mathcal{H}$. The bilateral shift V_0 is defined as follows:

$$V_0 : V_0 e_n = e_{n+1}, \quad n = 0, \pm 1, \cdots.$$

V_0 is unitary with purely continuous spectrum the entire unit disk. Define the operator N_0 as follows:

$$N_0 : N_0 e_n = n e_n, \quad n = 0, \pm 1, \cdots.$$

The class of bounded operators C_0 and S_0 which satisfy the relations

$$[C_0, N_0] = iS_0, \quad [S_0, N_0] = -iC_0$$

is the following special class of tridiagonal operators of the second kind:

$$C_0 = \tfrac{1}{2}(V_0^*A + AV_0), \quad S_0 = (1/2i)(V_0^*A - AV_0), \tag{A1}$$

where

$$A : Ae_n = a(n)e_n, \quad n = 0, \pm 1, \cdots.$$

The operators (A1) are essentially the angle operators[5] in an abstract form. The results for the phase operators found hold also for the angle operators because they are obtained from general theorems of the perturbation theory. The difference consists of the

unitarity of the operator V_0. The consequences of this difference are well known.[3,5]

APPENDIX B

Equation (14) is a second order difference equation of Poincaré type. The theorem of Poincaré and Perron[17] as it is formulated and proved in the general form is not applicable in the present case because the roots of the characteristic equation of (15) have the same absolute value for $|E| \leq 2$. The assumption $f(n) \in l_2(1, \infty)$ implies the existence of the $\lim f(n+1)/f(n)$, which is the central point in the classical theory of Poincaré and Perron; this, however, does not lead to contradictions.

Assume that there exists $f(n)$ in $l_2(1, \infty)$ such that Eq. (14) is satisfied. Since

$$f(n) \in l_2(1, \infty), \tag{B1}$$

it follows that $\lim f(n) = 0$ as $n \to \infty$ and because of (14) there exists a subsequence $f(n_i)$, $i = 1, 2, \cdots$, such that $f(n_i) \neq 0$ and

$$\lim_{n_i \to \infty} \frac{f(n_i + 1)}{f(n_i)} = \frac{0}{0} = \lambda. \tag{B2}$$

Since $\lim a(n) = 1$ as $n \to \infty$, we have from (14) that

$$\lambda + 1/\lambda = E. \tag{B3}$$

For $E \neq \infty$ it follows from (B3) that $\lambda \neq 0$ and $\lambda \neq \infty$.

From (B2) and (B3) it follows that for large n $f(n+1)/f(n)$ behaves as $f_1(n+1)/f_1(n)$, where $f_1(n)$ is a solution of Eq. (15), for $|E| < 2$. From this it does not follow that $f(n) \sim e^{ikn}$, $k =$ real, and $f(n) \sim c_1 n + c_2$ for $E = 2$, or $f(n) \sim c_1(-1)^n n + c_2$ for $E = -2$. It can, for instance, be $f(n) = (1/n)e^{ikn}$, which does not contradict (B1).

A rigorous proof that the spectrum of all the phase operators is purely continuous has not been achieved yet.

[1] E. C. Lerner, H. W. Huang, and G. E. Walters, J. Math. Phys. **11**, 1679 (1970).
[2] E. C. Lerner, Nuovo Cimento **56B**, 183 (1968).
[3] P. Carruthers and M. Nieto, Rev. Mod. Phys. **40**, 411 (1968).
[4] R. Jackiw, J. Math. Phys. **9**, 339 (1968).
[5] J. Zak, Phys. Rev. **187**, 1803 (1969).
[6] P. C. Deliyannis and E. K. Ifantis, J. Math. Phys. **10**, 421 (1969).
[7] E. K. Ifantis, J. Math. Phys. **11**, 1138 (1970).
[8] E. K. Ifantis, Bull. Greek Math. Soc. **10**, Part 2, 89 (1969).
[9] P. R. Halmos, *A Hilbert Space Problem Book* (Van Nostrand, Princeton, N.J., 1967).
[10] W. C. Ridge, Trans. Am. Math. Soc. **147**, 349 (1970).
[11] J. G. Stampli, Pacific J. Math. **12**, 1453 (1962).
[12] C. E. Rickart, *Banach Algebras* (Van Nostrand, Princeton, N. J. 1960), p. 282.
[13] P. R. Halmos, Acta Sci. Math. **29**, 290 (1968).
[14] A. C. Zannen, *Linear Analysis* (North-Holland, Amsterdam, 1964), p. 307.
[15] T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, 1966).
[16] E. K. Ifantis, J. Math. Phys. (to be submitted).
[17] H. Meschkowski, *Differenzengleichungen* (Van Denhoek und Ruprecht, Göttingen, 1959).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Critical-Point Behavior of a Class of Model Hamiltonians

H. L. Scott
Department of Physics, University of Utah, Salt Lake City, Utah 84112

(Received 12 October 1970)

A rigorous method for obtaining the thermodynamic functions for a class of model Hamiltonians is examined near points where fluctuations in certain order parameters become large. It is shown that if a phase transition occurs at such a point in a model, then the transition is classical in nature. That is, the free energy is analytic in the relevant order parameter.

I. INTRODUCTION

The paucity of exactly soluble models for phase transitions has led workers in this field to consider various ways of obtaining approximate expressions for the thermodynamic behavior of model systems.[1] One such method involves taking the limit of infinite interaction range. For a wide class of local interactions, this limit, if taken carefully, renders the models soluble and phase transitions can occur.[2] However, the detailed behavior of these models near the critical point has generally turned out to be identical to that of the "classical" theories of critical phenomena.[1] Rigorous proofs to this end were obtained by Lieb [3] and Lebowitz and Penrose[4] for systems with local interactions. In a previous paper[5] (hereafter referred to as I), we also obtained this result for a class of

unitarity of the operator V_0. The consequences of this difference are well known.[3,5]

APPENDIX B

Equation (14) is a second order difference equation of Poincaré type. The theorem of Poincaré and Perron[17] as it is formulated and proved in the general form is not applicable in the present case because the roots of the characteristic equation of (15) have the same absolute value for $|E| \leq 2$. The assumption $f(n) \in l_2(1, \infty)$ implies the existence of the $\lim f(n+1)/f(n)$, which is the central point in the classical theory of Poincaré and Perron; this, however, does not lead to contradictions.

Assume that there exists $f(n)$ in $l_2(1, \infty)$ such that Eq. (14) is satisfied. Since

$$f(n) \in l_2(1, \infty), \tag{B1}$$

it follows that $\lim f(n) = 0$ as $n \to \infty$ and because of (14) there exists a subsequence $f(n_i)$, $i = 1, 2, \cdots$, such that $f(n_i) \neq 0$ and

$$\lim_{n_i \to \infty} \frac{f(n_i + 1)}{f(n_i)} = \frac{0}{0} = \lambda. \tag{B2}$$

Since $\lim a(n) = 1$ as $n \to \infty$, we have from (14) that

$$\lambda + 1/\lambda = E. \tag{B3}$$

For $E \neq \infty$ it follows from (B3) that $\lambda \neq 0$ and $\lambda \neq \infty$.

From (B2) and (B3) it follows that for large n $f(n+1)/f(n)$ behaves as $f_1(n+1)/f_1(n)$, where $f_1(n)$ is a solution of Eq. (15), for $|E| < 2$. From this it does not follow that $f(n) \sim e^{ikn}$, $k =$ real, and $f(n) \sim c_1 n + c_2$ for $E = 2$, or $f(n) \sim c_1(-1)^n n + c_2$ for $E = -2$. It can, for instance, be $f(n) = (1/n)e^{ikn}$, which does not contradict (B1).

A rigorous proof that the spectrum of all the phase operators is purely continuous has not been achieved yet.

[1] E. C. Lerner, H. W. Huang, and G. E. Walters, J. Math. Phys. **11**, 1679 (1970).
[2] E. C. Lerner, Nuovo Cimento **56B**, 183 (1968).
[3] P. Carruthers and M. Nieto, Rev. Mod. Phys. **40**, 411 (1968).
[4] R. Jackiw, J. Math. Phys. **9**, 339 (1968).
[5] J. Zak, Phys. Rev. **187**, 1803 (1969).
[6] P. C. Deliyannis and E. K. Ifantis, J. Math. Phys. **10**, 421 (1969).
[7] E. K. Ifantis, J. Math. Phys. **11**, 1138 (1970).
[8] E. K. Ifantis, Bull. Greek Math. Soc. **10**, Part 2, 89 (1969).
[9] P. R. Halmos, *A Hilbert Space Problem Book* (Van Nostrand, Princeton, N.J., 1967).
[10] W. C. Ridge, Trans. Am. Math. Soc. **147**, 349 (1970).
[11] J. G. Stampli, Pacific J. Math. **12**, 1453 (1962).
[12] C. E. Rickart, *Banach Algebras* (Van Nostrand, Princeton, N. J. 1960), p. 282.
[13] P. R. Halmos, Acta Sci. Math. **29**, 290 (1968).
[14] A. C. Zannen, *Linear Analysis* (North-Holland, Amsterdam, 1964), p. 307.
[15] T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, 1966).
[16] E. K. Ifantis, J. Math. Phys. (to be submitted).
[17] H. Meschkowski, *Differenzengleichungen* (Van Denhoek und Ruprecht, Göttingen, 1959).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Critical-Point Behavior of a Class of Model Hamiltonians

H. L. Scott
Department of Physics, University of Utah, Salt Lake City, Utah 84112

(Received 12 October 1970)

A rigorous method for obtaining the thermodynamic functions for a class of model Hamiltonians is examined near points where fluctuations in certain order parameters become large. It is shown that if a phase transition occurs at such a point in a model, then the transition is classical in nature. That is, the free energy is analytic in the relevant order parameter.

I. INTRODUCTION

The paucity of exactly soluble models for phase transitions has led workers in this field to consider various ways of obtaining approximate expressions for the thermodynamic behavior of model systems.[1] One such method involves taking the limit of infinite interaction range. For a wide class of local interactions, this limit, if taken carefully, renders the models soluble and phase transitions can occur.[2] However, the detailed behavior of these models near the critical point has generally turned out to be identical to that of the "classical" theories of critical phenomena.[1] Rigorous proofs to this end were obtained by Lieb [3] and Lebowitz and Penrose[4] for systems with local interactions. In a previous paper[5] (hereafter referred to as I), we also obtained this result for a class of

nonlocal models for liquid–vapor phase transitions in fermion systems.[6,7]

In this paper we wish to show that the results of I can be applied to a more general class of models. The models we shall be concerned with are all exactly soluble in the limit of an infinite system by the thermodynamically equivalent Hamiltonian (TEH) method.[8] We consider only phase transitions of the type which involve, below a certain thermodynamic point, long range ordering in a particular order parameter η associated with the model Hamiltonian. Furthermore, we assume that, at the critical point, fluctuations in η become macroscopically large. Under these assumptions we show that if a phase transition occurs in a given model in the class, then this transition is classical in nature.

In the next section we describe the TEH approach, and Sec. III contains our application of the methods of I to the new class of models. We state our conclusions in Sec. IV.

II. THE TEH METHOD

In this section we outline the TEH approach, following Ref. 8. We consider a many-body system interacting according to the Hamiltonian

$$H = \sum_{\mathbf{k},\sigma} (t_\sigma(\mathbf{k}) b_{\mathbf{k}\sigma} + t_\sigma^*(\mathbf{k}) b_{\mathbf{k}\sigma}^+) + V^{-1} \sum_{\substack{\mathbf{k},\mathbf{q} \\ \sigma,\sigma'}} W_{\sigma,\sigma'}(\mathbf{k},\mathbf{q}) b_{\mathbf{k}\sigma}^+ b_{\mathbf{q}\sigma'}. \quad (1)$$

The operators $b_{\mathbf{k}\sigma}^+$ and $b_{\mathbf{k}\sigma}$ are unspecified bilinear combinations of the single particle creation and destruction operators $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^+$, and the subscript σ denotes the associated spin state. The quantities $t_\sigma(\mathbf{k})$ and $t_\sigma^*(\mathbf{k})$ are the energies for single excitations, and $W_{\sigma\sigma'}(\mathbf{k},\mathbf{q})$ is the interaction between the excitations. The crucial factor in the analysis of Ref. 8 is the inverse volume $(1/V)$ factor in the interaction term. It is this factor which, in the thermodynamic limit ($V \to \infty$, particle density held constant) gives the interaction an infinite range and allows the model partition function to be explicitly calculated.

The TEH solution is obtained by introducing a set of variational parameters $\eta_{\mathbf{k}\sigma}$, $\eta_{\mathbf{k}\sigma}^*$ into Eq. (1), rewriting the Hamiltonian in the form

$$\begin{aligned} H &= H_0 + H', \\ H_0 &= U + \sum_{\mathbf{k},\sigma} (G_\sigma(\mathbf{k}) b_{\mathbf{k}\sigma} + G_\sigma^*(\mathbf{k}) b_{\mathbf{k}\sigma}^+), \quad (2) \\ H' &= V^{-1} \sum_{\mathbf{k},\mathbf{q}\ \sigma,\sigma'} W_{\sigma\sigma'}(\mathbf{k},\mathbf{q}) B_{\mathbf{k}\sigma}^+ B_{\mathbf{q}\sigma'}, \end{aligned}$$

where

$$\begin{aligned} B_{\mathbf{k}\sigma} &= b_{\mathbf{k}\sigma} - \eta_{\mathbf{k}\sigma}, \\ U &= -V^{-1} \sum_{\mathbf{k},\mathbf{q},\sigma,\sigma'} W_{\sigma\sigma'}(\mathbf{k},\mathbf{q}) \eta_{\mathbf{k}\sigma}^* \eta_{\mathbf{q}\sigma'}, \quad (3) \\ G_\sigma(\mathbf{k}) &= t_\sigma(\mathbf{k}) + V^{-1} \sum_{\mathbf{q},\sigma'} W_{\sigma\sigma'}(\mathbf{k},\mathbf{q}) \eta_{\mathbf{q}\sigma'}^*. \end{aligned}$$

Now, in the thermodynamic limit, the "perturbation" H' does not contribute to the thermodynamical behavior of the model,[8,9] providing one chooses the functions $\eta_{\mathbf{k}\sigma}$ to satisfy

$$\eta_{\mathbf{k}\sigma} = \beta^{-1} \frac{\partial}{\partial G_\sigma(\mathbf{k})} [\beta^{-1} \ln (\mathrm{Tr}\, e^{-\beta H_0})], \quad (4)$$

where β is the inverse of Boltzmann's constant times the temperature. The free energy of the model is then given by

$$\lim_{V\to\infty} \beta^{-1} \ln [\mathrm{Tr}\, e^{-\beta(H_0+H')}] = \beta^{-1} \ln [\mathrm{Tr}\, e^{-\beta H_0}] \equiv F_0. \quad (5)$$

For given $b_{\mathbf{k}\sigma}$, $b_{\mathbf{k}\sigma}^+$, this trace can be evaluated using the Bogoliubov transformation method. The resulting expression, coupled with the variational equations

$$\eta_{\mathbf{k}\sigma} = \frac{\partial F_0}{\partial G_\sigma(\mathbf{k})} = \langle b_{\mathbf{k}\sigma} \rangle, \quad (6)$$

contains all the thermodynamic information about the given model.

In the thermodynamic limit it is thus rigorously established that the model defined by Eq. (1) is equivalent to the system described by the nonlocal Hamiltonian H_0. At this point the similarity of the TEH method and the work of Girardeau,[10] which formed the basis for the analysis of I, is evident. The fact that a phase transition can occur in a model of this type was demonstrated by Gartenhaus and Stranahan.[6]

Finally, we note that according to Eq. (6), the Hamiltonian H' contains fluctuation terms in the parameters $\langle b_{\mathbf{k}\sigma} \rangle$, and $\langle b_{\mathbf{k}\nu}^+ \rangle$. We therefore expect that when these fluctuations become macroscopically large, TEH should lead to erroneous results. In the next section we verify this expectation by showing that F_0 behaves analytically even near points of large fluctuations. This analyticity is characteristic of the classical theories of critical phenomena.

III. ANALYSIS OF THE TEH

In this section we study the analytic behavior of the TEH free energy F_0, Eq. (5), near points where fluctuations in the order parameter

$$\eta \equiv \sum_{\mathbf{k},\sigma} \eta_{\mathbf{k}\sigma} \quad (7)$$

become anomalously large. The analysis parallels that used previously in I, so we shall merely indicate here the manner in which the methods of I apply to the present case.

As in I, we consider a class of interactions which may be written as a finite symmetric sum of factorable terms in the form

$$W_{\sigma\sigma'}(\mathbf{k}, \mathbf{q}) = \sum_{i=1}^{N} \alpha_i(\mathbf{k}, \sigma)\beta_i(\mathbf{q}, \sigma'). \tag{8}$$

To accomplish our end, we first determine the conditions under which fluctuations in η diverge. To facilitate our calculations, we include an "external field" term in the single-particle energy by writing

$$t_\sigma(\mathbf{k}) \equiv e_\sigma(\mathbf{k}) - \epsilon,$$
$$t^*_\sigma(\mathbf{k}) \equiv e^*_\sigma(\mathbf{k}) - \epsilon^*. \tag{9}$$

The quantities ϵ and ϵ^* are the fields conjugate to η and η^* respectively, and are introduced here as a convenient manner by which to analyze fluctuations in these order parameters. Then we have

$$\eta = \left\langle \sum_\sigma \int \frac{d\mathbf{k}}{(2\pi)^3} b_{\mathbf{k},\sigma} \right\rangle = \left(\frac{\partial F_0}{\partial \epsilon}\right)_\beta, \tag{10}$$

where we have made use of Eqs. (2), (3), (5), and (6) and changed sums to integrals by the usual prescription (in the limit $V \to \infty$). A similar equation can be written for the conjugate parameter η^*.[11] Furthermore, fluctuations in η are obtained from the formula (analogous to the equation for density fluctuations in a fluid used in I)

$$(\Delta\eta)^2 = \left(\frac{\partial^2 F_0}{\partial \epsilon^2}\right)_\beta = \left(\frac{\partial \eta}{\partial \epsilon}\right)_\beta. \tag{11}$$

Since η is a function of ϵ through the functions $G_0(\mathbf{k})$, we have

$$\left(\frac{\partial \eta}{\partial \epsilon}\right)_\beta = \sum_\sigma \int \frac{d\mathbf{k}}{(2\pi)^3} \left[\frac{\partial \eta}{\partial G_\sigma(\mathbf{k})} \frac{\partial G_\sigma(\mathbf{k})}{\partial \epsilon} + \frac{\partial \eta}{\partial G^*_\sigma(\mathbf{k})} \frac{\partial G^*_\sigma(\mathbf{k})}{\partial \epsilon}\right]. \tag{12}$$

Therefore, assuming the integral over $\mathbf{k}$ always exists,[12] a necessary condition for the onset of critical fluctuations in η is that, at some β_c, ϵ_c,

$$\frac{\partial G_\sigma(\mathbf{k})}{\partial \epsilon} \to \infty \quad \text{or} \quad \frac{\partial G^*_\sigma(\mathbf{k})}{\partial \epsilon} \to \infty.$$

Now according to Eqs. (3) and (6), the functions $G_\sigma(\mathbf{k})$ and their conjugates obey the coupled nonlinear integral equations

$$G_\sigma(\mathbf{k}) = t_\sigma(\mathbf{k}) + \sum_{\sigma'} \int \frac{d\mathbf{q}}{(2\pi)^3} W_{\sigma\sigma'}(\mathbf{k}, \mathbf{q}) \left[\frac{\partial F_0}{\partial G^*_{\sigma'}(\mathbf{q})}\right],$$
$$G^*_\sigma(\mathbf{k}) = t^*_\sigma(\mathbf{k}) + \sum_{\sigma'} \int \frac{d\mathbf{q}}{(2\pi)^3} W_{\sigma\sigma'}(\mathbf{k}, \mathbf{q}) \left[\frac{\partial F_0}{\partial G_{\sigma'}(\mathbf{q})}\right]. \tag{13}$$

It follows then that the derivatives of $G_\sigma(\mathbf{k})$ and $G^*_\sigma(\mathbf{k})$ obey coupled linear integral equations of the form

$$\left(\frac{\partial G_\sigma(\mathbf{k})}{\partial \epsilon}\right)_\beta = -1 + \sum_{\sigma',\sigma''} \int \frac{d\mathbf{q}\, d\mathbf{p}}{(2\pi)^6} W_{\sigma\sigma'}(\mathbf{k}, \mathbf{q}) \times \left\{\left[\frac{\partial^2 F_0}{\partial G_{\sigma''}(\mathbf{p}) \partial G^*_{\sigma'}(\mathbf{q})}\right] \frac{\partial G_{\sigma''}(\mathbf{p})}{\partial \epsilon} + \left[\frac{\partial^2 F_0}{\partial G^*_{\sigma''}(\mathbf{p}) \partial G^*_{\sigma'}(\mathbf{q})}\right] \frac{\partial G^*_{\sigma''}(\mathbf{p})}{\partial \epsilon}\right\}, \tag{14}$$

with an analogous formula for $(\partial G^*_\sigma(\mathbf{k})/\partial \epsilon)$. We simplify these formulas by writing

$$\left(\frac{\partial G_\sigma(\mathbf{k})}{\partial \epsilon}\right)_\beta = -1 + \sum_{\sigma''} \int \frac{d\mathbf{p}}{(2\pi)^3} \left\{K^{11}_{\sigma\sigma''}(\mathbf{k}, \mathbf{p}) \frac{\partial G_{\sigma''}(\mathbf{p})}{\partial \epsilon} + K^{12}_{\sigma\sigma''}(\mathbf{k}, \mathbf{p}) \frac{\partial G^*_{\sigma''}(\mathbf{p})}{\partial \epsilon}\right\},$$
$$\left(\frac{\partial G^*_\sigma(\mathbf{k})}{\partial \epsilon}\right)_\beta = -1 + \sum_{\sigma''} \int \frac{d\mathbf{p}}{(2\pi)^3} \left\{K^{21}_{\sigma\sigma''}(\mathbf{k}, \mathbf{p}) \frac{\partial G_{\sigma''}(\mathbf{p})}{\partial \epsilon} + K^{22}_{\sigma\sigma''}(\mathbf{k}, \mathbf{p}) \frac{\partial G^*_{\sigma''}(\mathbf{p})}{\partial \epsilon}\right\}, \tag{15}$$

where

$$K^{11}_{\sigma\sigma''}(\mathbf{k}, \mathbf{p}) = \sum_{\sigma'} \int \frac{d\mathbf{q}}{(2\pi)^3} W_{\sigma\sigma'}(\mathbf{k}, \mathbf{q}) \left\{\frac{\partial^2 F_0}{\partial G_{\sigma'}(\mathbf{q}) \partial G^*_{\sigma''}(\mathbf{p})}\right\} = K^{22*}_{\sigma\sigma'}(\mathbf{k}, \mathbf{p}),$$
$$K^{12}_{\sigma\sigma'}(\mathbf{k}, \mathbf{p}) = \sum_{\sigma'} \int \frac{d\mathbf{q}}{(2\pi)^3} W_{\sigma\sigma'}(\mathbf{k}, \mathbf{q}) \left\{\frac{\partial^2 F_0}{\partial G^*_{\sigma''}(\mathbf{p}) \partial G^*_{\sigma'}(\mathbf{q})}\right\} = K^{21*}_{\sigma\sigma'}(\mathbf{k}, \mathbf{p}). \tag{16}$$

Upon substitution of the interaction defined in Eq. (8) into the above expressions, each of the kernels can be expressed in the form

$$K^{ij}_{\sigma\sigma'}(\mathbf{k}, \mathbf{p}) = \sum_{r=1}^{N} \alpha_r(\mathbf{k}, \sigma) B^{ij}_r(\mathbf{p}, \sigma'), \quad i, j = 1, 2. \tag{17}$$

As in I, $B^{ij}_r(\mathbf{p}, \sigma)$ is a function of β and ϵ obtained by integrating in Eq. (16) over $\mathbf{q}$.

For kernels of this type the coupled integral equations reduce to a system of $2N$ linear algebraic equations. Following I, these equations can be solved by Cramer's rule, and the result is

$$\frac{\partial G_\sigma(\mathbf{k})}{\partial \epsilon} = \sum_r \alpha_r(\mathbf{k}, \sigma) N_r(\beta, \epsilon)/D(\beta, \epsilon),$$
$$\frac{\partial G^*_\sigma(\mathbf{k})}{\partial \epsilon} = \sum_r \alpha_r(\mathbf{k}, \sigma) N^*_r(\beta, \epsilon)/D(\beta, \epsilon). \tag{18}$$

The function $D(\beta, \epsilon)$ is a $2N \times 2N$ real determinant, the elements of which are functions of β and ϵ given by

$$D(\beta, \epsilon)_{\substack{li \\ mj}} = \delta_{\substack{li \\ mj}} - \sum_\sigma \int \frac{d\mathbf{q}}{(2\pi)^3} \alpha_l(\mathbf{q}, \sigma) B^{ij}_m(\mathbf{q}, \sigma). \tag{19}$$

The functions $N_r(\beta, \mu)$ can be written as a linear combination of the minors of $D(\beta, \mu)$. Thus we obtain for the fluctuations, using Eq. (12),

$$\left(\frac{\partial \eta}{\partial \epsilon}\right)_\beta = M(\beta, \epsilon)/D(\beta, \epsilon). \tag{20}$$

As the number of terms in the interaction is finite and all integrals are assumed to converge,[12] we conclude that $M(\beta, \epsilon)$ and $D(\beta, \epsilon)$ are bounded in β and ϵ. Therefore, if fluctuations in η are to diverge at a critical point, we must have at this point

$$D(\beta_c, \epsilon_c) = 0. \tag{21}$$

We now examine the derivative of η with respect to temperature. By arguments identical to the above, and carried out in detail in I, we find

$$\left(\frac{\partial \eta}{\partial \beta}\right)_\epsilon = \frac{P(\beta, \epsilon)}{D(\beta, \epsilon)}, \tag{22}$$

where $P(\beta, \epsilon)$ is also expressible as a linear combination of the minors of the determinant $D(\beta, \epsilon)$.

In order to determine the analytic properties of the free energy F_0 in η, we examine the derivatives of the denominator function D with respect to β and η, using relations analogous to those used in I;[13]

$$\left(\frac{\partial D}{\partial \eta}\right)_\beta = \left(\frac{\partial D}{\partial \epsilon}\right)_\beta \left(\frac{\partial \epsilon}{\partial \eta}\right)_\beta \tag{23}$$

and

$$\left(\frac{\partial D}{\partial \beta}\right)_\eta = \frac{\left(\frac{\partial D}{\partial \beta}\right)_\epsilon \left(\frac{\partial \eta}{\partial \epsilon}\right)_\beta - \left(\frac{\partial D}{\partial \epsilon}\right)_\beta \left(\frac{\partial \eta}{\partial \beta}\right)_\epsilon}{(\partial \eta/\partial \epsilon)_\beta^2}. \tag{24}$$

At this point we can apply the arguments of I directly to the present case. We write the numerators $M(\beta, \epsilon)$, Eq. (20), and $P(\beta, \epsilon)$, Eq. (22), as sums of minors of $D(\beta, \epsilon)$, and substitute into Eqs. (23) and (24). Then, as in I, all $D^{-1}(\beta, \epsilon)$ terms cancel and we conclude that the derivatives of $D(\beta, \epsilon)$ are bounded at a critical point. Similar arguments apply to the higher-order derivatives, and so $D[\beta, \epsilon(\beta, \eta)] \equiv D(\beta, \eta)$ is analytic at a critical point. From this result, and the fact that $M(\beta, \epsilon)$ must be bounded and nonzero at the critical point (see the Appendix of I), it follows that $(\partial \epsilon/\partial \eta)_\beta$ is analytic in η and β. However,

$$\left(\frac{\partial F_0}{\partial \eta}\right)_\beta = \left(\frac{\partial F_0}{\partial \epsilon}\right)_\beta \left(\frac{\partial \epsilon}{\partial \eta}\right)_\beta, \tag{25}$$

so that $(\partial F_0/\partial \eta)_\beta$ (and therefore F_0 itself) is analytic in η. Consequently, the TEH models are capable of producing only classical phase transitions.

IV. CONCLUSIONS

The results presented above, coupled with the results of I, indicate that the classical theory of phase transitions is inevitably obtained from the study of model Hamiltonians of the form of Eq. (1), when the limit of infinite interaction range is taken. In the present case, this limit was justified as a consequence of the ordinary thermodynamic limit, and the $1/V$ term in the interaction. However, if the $1/V$ is replaced by a general inverse range parameter γ, and we take the limit as $\gamma \to 0$ with the restriction $\gamma \sum_{k,q,\sigma,\sigma'} W_{\sigma\sigma'}(k, q)$ is nonvanishing, then the TEH method is again applicable and the conclusions are unaltered. This fact provides a connection between our work and the work of Lebowitz and Penrose [4] and Lieb.[3] Basically, we have found that their result also applies to a wider class of nonlocal models. Thus we have confirmed, more generally than I, that the limit of infinite interaction range erases any nonanalytic critical-point behavior that might otherwise be predicted by a model. What is still lacking is a detailed understanding of how local, microscopic irregularities contribute in a macroscopic manner to thermodynamic behavior at the critical point.

Finally, we note that by adding an infinite series of three-body and higher interactions to the model Hamiltonian, it is possible in the infinite range limit to obtain a nonclassical phase transition.[14]

[1] For a comprehensive review, see M. E. Fisher, Rept. Progr. Phys. **30**, 615 (1967).

[2] See, for example, M. Kac in *Fundamental Problems in Statistical Mechanics*, E. G. D. Cohen, Ed. (Interscience, N.Y., 1968).

[3] E. H. Lieb, J. Math. Phys. **7**, 1016 (1966).

[4] J. Lebowitz and O. Penrose, J. Math. Phys. **7**, 98 (1966).

[5] H. L. Scott, J. Math. Phys. **11**, 3159 (1970).

[6] S. Gartenhaus and G. Stranahan, Phys. Rev. **A138**, 1346 (1966).

[7] S. Gartenhaus and G. Stranahan, Phys. Rev. **173**, 260 (1968).

[8] G. Wentzel, Phys. Rev. **120**, 1572 (1960).

[9] N. N. Bogoliubov, D. B. Zubarev, and U. A. Tserkovnikov, Dokl. Akad. Nauk SSSR **117**, 788 (1957) [Soviet Phys. Doklady **2**, 535 (1957)].

[10] M. Girardeau, J. Math. Phys. **3**, 131 (1962).

[11] In the following, we will consider only the order parameter η. The conclusions obtained also apply to η^* if one replaces ϵ by ϵ^*.

[12] This assumption holds rigorously for Fermi systems and is valid for Bose systems provided no single excitation state is macroscopically occupied.

[13] As pointed out in I, existence of derivatives of all order is generally not sufficient to insure analyticity. We assume, however, that the additional requirement, which involves the nth Taylor coefficient approaching zero, is met.

[14] S. Gartenhaus and H. L. Scott, Phys. Rev. **A1**, 1270 (1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Approach to the Eulerian–Lagrangian Problem for Sound Propagation in Continuous Stochastic Media*

JEROME A. NEUBERT
Naval Undersea Research and Development Center, Pasadena, California 91107

(Received 7 August 1970)

The formal solution of the Eulerian–Lagrangian problem for sound propagation in continuous stochastic media is reframed so that the emphasis on the need for complete knowledge of the statistical nature of the Lagrangian functional of interest is shifted to the need for knowledge of the asymptotic behavior of certain stochastic Lagrangian integrals which result from the application of a central limit theorem for stochastic functionals. The resultant practical relation for calculating the Eulerian ensemble expectation corresponding to the Lagrangian functional in question is shown to depend in a natural manner upon the Lagrangian path spreading and, in the case of a statistically isotropic medium, to reduce to the proper nonstochastic limit.

INTRODUCTION

In general, the stochastic Eulerian–Lagrangian problem involves obtaining the statistics of a Lagrangian functional $u[\mathbf{X}(t, \xi), t]$ from the known statistics of an Eulerian field $u(\mathbf{x}, t)$, where $\mathbf{X}(t, \xi)$ is the path as a function of the parameter t from the initial point $\xi = \mathbf{X}(0, \xi)$. For example, $u(\mathbf{x}, t)$ could be the Eulerian velocity field specified in laboratory coordinates at time t for a turbulent diffusion problem; see Lumley.[1] In the case of steady-state sound propagation in continuous stochastic media, the complete, rigorous stochastic Eulerian–Lagrangian problem would involve finding the statistics of Lagrangian functions of the refractive index $\mu[\mathbf{X}(s, \xi)]$, where s is the curvilinear distance (or arc length) traveled, from the known, or assumed, statistics of the Eulerian refractive field $\mu(\mathbf{x})$; see, for example, Eq. (39) of Neubert.[2] Unfortunately, this exact approach proves to be horrendous, if not impossible, if a sufficiently general class of physically meaningful functions $\mu[\mathbf{X}(s, \xi)]$ is to be considered; see Lumley[1] and Monin and Yaglom[3] (Sec. 9). Therefore, this study is confined to a more modest approach which involves reframing the rigorous formal solution in a form which is sufficient for, but not restricted to, treating many steady-state sound propagation problems of practical interest in continuous stochastic media. This new form for the solution proves to be convenient for applying the central limit theorem for Lagrangian stochastic functionals discussed in Sec. 3.15 of Lumley[4] and in Sec. 4.5 of Neubert[5] and the methods of asymptotic integral evaluation presented in Sec. 9 of Monin and Yaglom[3] and in Secs. 4.3 and 4.8 of Neubert.[5] It will be found that the general stochastic Eulerian–Lagrangian problem will not have to be solved totally for many sound propagation problems when the equation for the path $\mathbf{X}(s, \xi)$ can be expressed in terms of $\mu[\mathbf{X}(s, \xi)]$, including its derivatives, and the initial conditions and when $\mu(\mathbf{x})$ is statistically isotropic, since the emphasis on the need for complete knowledge of the statistics of $\mu[\mathbf{X}(s, \xi)]$ has been shifted to the need for knowledge of the asymptotic behavior of certain stochastic Lagrangian integrals which result from the application of the central limit theorem for stochastic Lagrangian functionals.

I. FORMAL SOLUTION

Let the continuous Lagrangian functional $F[\mathbf{X}(s, \xi)]$ be the value of a stochastic Lagrangian physical quantity, such as the sound pressure wave $p[\mathbf{X}(s, \xi)]$ of Ref. 2, at arc length s along the continuous path $\mathbf{X}(s, \xi)$ from a point ξ on the continuous initial surface S_0 (e.g., the face of a transducer) in the ensemble realization μ_β. The quantity $F[\mathbf{X}(s, \xi)]$ is a function of s and of ξ and a functional of the field $\mu_\beta(\mathbf{x})$ and, hence, of the path $\mathbf{X}(s, \xi)$ over $\mu_\beta(\mathbf{x})$ as determined by ξ, $\theta_0(\xi)$ and s. It is, of course, assumed that all paths from a single point ξ have the same initial angle $\theta_0(\xi)$ for all μ_β in the total ensemble $\{\mu_\beta\}$ and for all s in the interval $[0, \infty)$. In this paper, an initial surface S_0 and a single terminal point $\mathbf{x}$ are considered. This analysis can be modified to include a terminal surface S. Let $F(\mathbf{x})$ be the corresponding continuous Eulerian field at a known point $\mathbf{x}$. The ensemble expectation of a Lagrangian functional $F[\mathbf{X}(s, \xi)]$ over all μ_β of $\{\mu_\beta\}$ for a given s and ξ will be represented by $E\{F[\mathbf{X}(s, \xi)]\}$, which will be called the *Lagrangian ensemble expectation* of $F[\mathbf{X}(s, \xi)]$. The ensemble expectation of the corresponding Eulerian field $F(\mathbf{x})$ over all μ_β of $\{\mu_\beta\}$ at the field point $\mathbf{x}$ will be represented by $\langle F(\mathbf{x})\rangle$, which will be called the *Eulerian ensemble expectation* of $F(\mathbf{x})$. A different notation is used to distinguish the ensemble expectations of Lagrangian functionals from the ensemble expectations of Eulerian fields for the purpose of clarity

in the following development. It will become apparent that these two stochastic concepts are quite different in their behavior due to the Lagrangian spreading of the terminal location $\mathbf{X}(s, \boldsymbol{\xi})$. Note that $\mathbf{X} = \mathbf{X}(s, \boldsymbol{\xi})$ is the Lagrangian (path) position, while $\mathbf{x}$ is the Eulerian (space) position.

The stochastic Eulerian–Lagrangian problem for steady-state sound propagation in continuous stochastic media can now be stated as follows: Given the statistics of the Eulerian field $\mu(\mathbf{x})$ and assuming the form of the Lagrangian functional $F[\mathbf{X}(s, \boldsymbol{\xi})]$ of $\mu[\mathbf{X}(s, \boldsymbol{\xi})]$ is known, determine $\langle F(\mathbf{x})\rangle$. One senses that a Lagrangian ensemble expectation will have to enter into the formalism, but $E\{F[\mathbf{X}(s, \boldsymbol{\xi})]\rangle$ is too indiscriminate since it contains more of $\{\mu_\beta\}$ than is wanted. Not all μ_β result in paths of length s from $\boldsymbol{\xi}$ that terminate "near" a desired point $\mathbf{x}$, i.e., not all μ_β produce paths such that $\mathbf{X}(s, \boldsymbol{\xi}) \approx \mathbf{x}$, where the rigorous interpretation of "$\mathbf{X}(s, \boldsymbol{\xi}) \approx \mathbf{x}$" is $\mathbf{x} + d\mathbf{x} > \mathbf{X}(s, \boldsymbol{\xi}) \geq \mathbf{x}$. In fact, as s and $\boldsymbol{\xi}$ vary over all their possible values, different subsets of $\{\mu_\beta\}$ contribute paths that satisfy $\mathbf{X}(s, \boldsymbol{\xi}) \approx \mathbf{x}$. Thus, it is necessary to do the ensemble averaging only over those μ_β which result in $\mathbf{X}(s, \boldsymbol{\xi}) \approx \mathbf{x}$ for each possible s and $\boldsymbol{\xi}$. This limited ensemble, called the *Lagrangian subensemble*, is represented mathematically by

$$\{\mu_\beta\}_s \equiv \{\mu_\beta \mid \mathbf{x} + d\mathbf{x} > \mathbf{X}(s, \boldsymbol{\xi}) \geq \mathbf{x}\} \tag{1}$$

$$= \{\mu_\beta \mid \mathbf{X}(s, \boldsymbol{\xi}) \approx \mathbf{x}\} \subset \{\mu_\beta\}, \tag{2}$$

and the conditional ensemble expectation of the Lagrangian functional $F[\mathbf{X}(s, \boldsymbol{\xi})]$ over this limited ensemble, called the *Lagrangian subensemble expectation*, is represented by

$$\tilde{E}\{F[\mathbf{X}(s, \boldsymbol{\xi})]\}$$

$$\equiv E\{F[\mathbf{X}(s, \boldsymbol{\xi})] \mid \mathbf{x} + d\mathbf{x} \geq \mathbf{X}(s, \boldsymbol{\xi}) \geq \mathbf{x}\} \tag{3}$$

$$= E\{F[\mathbf{X}(s, \boldsymbol{\xi})] \mid \mathbf{X}(s, \boldsymbol{\xi}) \approx \mathbf{x}\} \tag{3'}$$

for notational simplicity. Thus, $\tilde{E}\{F[\mathbf{X}(s, \boldsymbol{\xi})]\}$ gives the value of $F[\mathbf{X}(s, \boldsymbol{\xi})]$ averaged over all paths of length s from $\boldsymbol{\xi}$ that do reach $\mathbf{x}$ (there can be at most one such path for each realization μ_β). Infinitely many possible paths may contribute at $\mathbf{x}$ for each s and $\boldsymbol{\xi}$. The next question is how many or, better, what percentage of the μ_β in $\{\mu_\beta\}$ contribute paths to $\tilde{E}\{F[\mathbf{X}(s, \boldsymbol{\xi})]\}$? This latter quantity is just the measure M of $\{\mu_\beta\}_s$, since the measure of $\{\mu_\beta\}$ is unity, and is given by

$$M\{\mu_\beta \mid \mathbf{X}(s, \boldsymbol{\xi}) \approx \mathbf{x}\}$$

$$\equiv M\{\mu_\beta \mid \mathbf{x} + d\mathbf{x} \geq \mathbf{X}(s, \boldsymbol{\xi}) \geq \mathbf{x}\} \tag{4}$$

$$= B(\mathbf{x}, \boldsymbol{\xi} \mid s)\, d\boldsymbol{\xi}\, ds \tag{5}$$

$$= B(x_i - \xi_i \mid s)\, d\boldsymbol{\xi}\, ds \tag{6}$$

[Eq. (5) reduces to Eq. (6) if the medium is statistically homogeneous], where $B(\mathbf{x}, \boldsymbol{\xi} \mid s)$ (if it exists) is the joint probability density for the events $X(s, \boldsymbol{\xi}) \approx x$, $Y(s, \boldsymbol{\xi}) \approx y$, and $Z(s, \boldsymbol{\xi}) \approx z$ and will be called the *Lagrangian measure function*.

Therefore, the Eulerian ensemble expectation of $F(\mathbf{x})$, for a given pair of $\boldsymbol{\xi}$ and s, is

$$\langle F(\mathbf{x}; \boldsymbol{\xi} \mid s)\rangle$$

$$\equiv \langle F(\mathbf{X}) \mid \mathbf{x} + d\mathbf{x} > \mathbf{X}(s, \boldsymbol{\xi}) \geq \mathbf{x}\rangle \tag{7}$$

$$= \langle F(\mathbf{X}) \mid \mathbf{X}(s, \boldsymbol{\xi}) \approx \mathbf{x}\rangle \tag{7'}$$

$$= \tilde{E}\{F[\mathbf{X}(s, \boldsymbol{\xi})]\} M\{\mu_\beta \mid \mathbf{X}(s, \boldsymbol{\xi}) \approx \mathbf{x}\} \tag{8}$$

$$= \tilde{E}\{F[\mathbf{X}(s, \boldsymbol{\xi})]\} B(\mathbf{x}, \boldsymbol{\xi} \mid s)\, d\boldsymbol{\xi}\, ds. \tag{9}$$

It is interesting to note that, although

$$\langle F(\mathbf{X}) \mid \mathbf{X}(s, \boldsymbol{\xi}) \approx \mathbf{x}\rangle$$

is "quasi-Eulerian" in the sense that only a single known terminal point $\mathbf{x}$ is considered, its condition (that the ensemble averaging be done only over $\{\mu_\beta\}_s$) concerns a Lagrangian path selection process. This dual character also occurs in the equivalent representation given by Eq. (9) where $B(\mathbf{x}, \boldsymbol{\xi} \mid s)$ continually measures $\{\mu_\beta\}_s$ as $\mathbf{x}$, $\boldsymbol{\xi}$, and s vary, while $\tilde{E}\{F[\mathbf{X}(s, \boldsymbol{\xi})]\}$ is purely Lagrangian but restricted to $\mathbf{X}(s, \boldsymbol{\xi}) \approx \mathbf{x}$ paths. Thus, the Eulerian ensemble expectation of $F(\mathbf{x})$, for a given s and for all $\boldsymbol{\xi} \in S_0$, is

$$\langle F(\mathbf{x}; S_0 \mid s)\rangle \equiv \langle F(\mathbf{X}) \mid \mathbf{X}(s, \boldsymbol{\xi}) \approx \mathbf{x} \text{ and } \boldsymbol{\xi} \in S_0\rangle \tag{10}$$

$$= ds \int_{S_0} d\boldsymbol{\xi} B(\mathbf{x}, \boldsymbol{\xi} \mid s) \tilde{E}\{F[\mathbf{X}(s, \boldsymbol{\xi})]\} \tag{11}$$

and follows by considering all $\boldsymbol{\xi} \in S_0$ in Eqs. (7′) and (9), respectively. The statement that Eq. (10) equals Eq. (11) is a tautology in the sense that Eqs. (10) and (11) are two ways of saying the same thing. Equation (10) specifies formally the quantity that is desired, and Eq. (11) states how this quantity can be calculated in terms of the factors $B(\mathbf{x}, \boldsymbol{\xi} \mid s)$ and $\tilde{E}\{F[\mathbf{X}(s, \boldsymbol{\xi})]\}$, which are to be determined for each specific physical problem that is considered. Finally, the Eulerian ensemble expectation of $F(\mathbf{x})$ for all $\boldsymbol{\xi} \in S_0$ and all possible s is

$$\langle F(\mathbf{x})\rangle = \langle F(\mathbf{X}) \mid \mathbf{X}(s, \boldsymbol{\xi}) \approx \mathbf{x}, \boldsymbol{\xi} \in S_0 \text{ and } s \in [0, \infty)\rangle \tag{12}$$

$$= \int_0^\infty ds \int_{S_0} d\boldsymbol{\xi} B(\mathbf{x}, \boldsymbol{\xi} \mid s) \tilde{E}\{F[\mathbf{X}(s, \boldsymbol{\xi})]\} \tag{13}$$

$$= \int_0^\infty ds \int_{S_0} d\boldsymbol{\xi} B(x_i - \xi_i \mid s) \tilde{E}\{F[\mathbf{X}(s, \boldsymbol{\xi})]\}; \tag{14}$$

Eq. (14) applies only when the medium is statistically homogeneous. Equations (12) and (13) follow by considering all $s \in [0, \infty)$ in Eqs. (10) and (11), respectively. Note that in the integrand of Eq. (13)

both $\tilde{E}\{F[\mathbf{X}(s, \xi)]\}$ and $B(\mathbf{x}, \xi \mid s)$ are defined for the same Eulerian statistical and path assumptions; only the ensembles differ since $B(\mathbf{x}, \xi \mid s)$ must permit any subensemble to be chosen while $\tilde{E}\{F[\mathbf{X}(s, \xi)]\}$ must utilize only the particular subensemble $\{\mu_\beta\}_s$ chosen.

It should be emphasized that, in obtaining Eq. (13) for determining the Eulerian ensemble expectation $\langle F(\mathbf{x})\rangle$ from the Lagrangian subensemble expectation $\tilde{E}\{F[\mathbf{X}(s, \xi)]\}$, two stochastic aspects of this Eulerian–Lagrangian problem have been explicitly considered together:

(i) the probability density $B(\mathbf{x}, \xi \mid s)$ that, for a given s and ξ, the path $\mathbf{X}(s, \xi)$ terminates "near" the chosen (fixed) point $\mathbf{x}$ [i.e., that $\mathbf{X}(s, \xi) \approx \mathbf{x}$]; symbolically, this density can be denoted by $B(A)$;

(ii) the Lagrangian conditional (subensemble) expectation $\tilde{E}\{F[\mathbf{X}(s, \xi)]\}$ of the Lagrangian functional $F[\mathbf{X}(s, \xi)]$ given that the event $\mathbf{X}(s, \xi) \approx \mathbf{x}$ has occurred; symbolically, this conditional expectation can be associated with the conditional density $B(B \mid A)$.

Thus, Eq. (13) can be expressed symbolically as follows:

$$\begin{aligned}\langle F\rangle &= \iint dA\, dB F(A) B(A, B)\\ &= \iint dA\, dB F(A) B(B \mid A) B(A)\\ &= \int dA B(A) E\{F(A) \mid A\}\\ &= \int dA B(A) \tilde{E}\{F(A)\}, \qquad (13')\end{aligned}$$

where dA represents $d\xi\, ds$ in Eq. (13).

II. A CENTRAL LIMIT THEOREM FOR STOCHASTIC LAGRANGIAN FUNCTIONALS

Section 3.15 of Lumley[4] contains an extensive exposition on the existence of a central limit theorem for random functions and discusses several sufficient conditions for such a theorem. Unfortunately, the direct use of the several proofs that exist is not possible since the sufficient conditions which are imposed on the stochastic functionals under consideration cannot be verified for real processes (see Monin and Yaglom,[3] Sec. 9). For the purposes of this study, it will be assumed that the stochastic Lagrangian functionals $u_i(s) = u_i[\mathbf{X}(s, \xi)]$, of continuous paths of interest, $\mathbf{X}(s, \xi)$, governed by an isotropic random field $\mu(\mathbf{x})$, satisfy a physically meaningful set of sufficient conditions so that the integral $\int_0^s ds' u_i(s')$ tends asymptotically to a Gaussian distribution. This theorem can be expressed in the useful tensor form

$$\begin{aligned}E\left\{\exp\left(ik_j \int_0^s ds' u_j(s')\right)\right\}&\\ \sim \exp\Bigg(-\tfrac{1}{2}k_i k_j E\Bigg\{\int_0^s ds' \int_0^s ds''[u_i(s') - E\{u_i(s')\}]&\\ \times [u_j(s'') - E\{u_j(s'')\}]\Bigg\}\Bigg)&\\ \times \exp\left(ik_j E\left\{\int_0^s ds' u_j(s')\right\}\right).& \qquad (15)\end{aligned}$$

This relation proves to be of considerable practical interest since it can frequently be employed to reduce the problem of finding both $B(x_i - \xi_i \mid s)$ and $\tilde{E}\{F[\mathbf{X}(s, \xi)]\}$ to the problem of determining the asymptotic behavior of certain stochastic Lagrangian integrals (see Chap. V of Neubert[5]). For the purposes of this paper, it will be assumed that all the integrals exist in at least the sense of generalized functions (see, for example, Lumley[4] or Gel'fand and Shilov[6]). It should be noted that the problem of choosing a useful set of sufficient conditions for this central limit theorem is intimately related to the problem of the existence of the integral scales which result from the asymptotic evaluation of the integrals in Eq. (15). From the material in Lumley,[4] it appears likely that a sufficient condition for the validity of Eq. (15) can be devised by requiring that the integral scales converge rapidly enough.

III. A METHOD FOR DETERMINING $B(x_i - \xi_i \mid s)$

The Lagrangian measure function $B(x_i - \xi_i \mid s)$ [which is actually, in this paper, a probability density for the case of a continuous, statistically isotropic field $\mu(\mathbf{x})$] can be determined from its characteristic function (see Sec. 3 of Monin and Yaglom[3] and Sec. 2.5 of Lumley[4]), which in Cartesian coordinates is given by

$$\varphi(\mathbf{x}, \xi \mid s) \equiv E\{\exp(ik_j[X_j(s, \xi) - \xi_j])\} \qquad (16)$$

$$= E\left\{\exp\left(ik_j \int_0^s ds' \frac{dX_j}{ds}(s', \xi)\right)\right\} \qquad (16')$$

$$\begin{aligned}\sim \exp\Bigg[-\tfrac{1}{2}k_i k_j \int_0^s ds' \int_0^s ds'' E\Bigg\{\left(\frac{dX_i}{ds}(s') - E\left\{\frac{dX_i}{ds}(s')\right\}\right)&\\ \times \left(\frac{dX_j}{ds}(s'') - E\left\{\frac{dX_j}{ds}(s'')\right\}\right)\Bigg\}\Bigg]&\\ \times \exp\left(ik_j \int_0^s ds' E\left\{\frac{dX_j}{ds}(s')\right\}\right)& \qquad (17)\end{aligned}$$

$$\begin{aligned}= \exp(-\tfrac{1}{2}k_i k_j E\{[X_i(s, \xi) - E\{X_i(s, \xi)\}]&\\ \times [X_j(s, \xi) - E\{X_j(s, \xi)\}]\})&\\ \times \exp[ik_j E\{X_j(s, \xi)\}]& \qquad (18)\end{aligned}$$

$$= \exp[-\tfrac{1}{2}k_i k_j U_{ij}(s, \xi) + ik_j m_j(s, \xi)], \qquad (19)$$

where

$$m_j = m_j(s, \boldsymbol{\xi}) \equiv E\{X_j(s, \boldsymbol{\xi})\} \tag{19'}$$

and

$$\begin{aligned} U_{ij} &= U_{ij}(s, \boldsymbol{\xi}) \\ &\equiv E\{[X_i(s, \boldsymbol{\xi}) - E\{X_i(s, \boldsymbol{\xi})\}] \\ &\quad \times [X_j(s, \boldsymbol{\xi}) - E\{X_j(s, \boldsymbol{\xi})\}]\} \qquad (20) \\ &= \int_0^s ds' \int_0^s ds'' E\left\{\left(\frac{dX_i}{ds}(s', \boldsymbol{\xi}) - E\left\{\frac{dX_i}{ds}(s', \boldsymbol{\xi})\right\}\right)\right. \\ &\quad \left.\times \left(\frac{dX_j}{ds}(s'', \boldsymbol{\xi}) - E\left\{\frac{dX_j}{ds}(s'', \boldsymbol{\xi})\right\}\right)\right\}. \qquad (20') \end{aligned}$$

Here $k_i, k_j = k_1, k_2, k_3$ are dummy variables and only stochastic quantities, such as $\int_0^s ds'(dX_j/ds)(s', \boldsymbol{\xi})$, have pertinence in $\varphi(\mathbf{k}, \boldsymbol{\xi} \mid s)$ since deterministic factors are cancelled out in Eq. (24). Note that the $X_j(s, \boldsymbol{\xi})$ are permitted to take on all their possible values in Eq. (16) and later, in Eq. (24), $B(x_i - \xi_i \mid s)$ picks the desired value for each. Therefore, the Lagrangian ensemble, rather than subensemble, expectation must be applied in Eq. (16) since $B(x_i - \xi_i \mid s)$ must be permitted to treat all possible values of $X_j(s, \boldsymbol{\xi})$ in order to be capable of determining the measure per unit s per unit $\boldsymbol{\xi}$ of the Lagrangian subensemble under consideration in Eq. (24) for each $\boldsymbol{\xi}$, s, $\mathbf{x}$. When the stochastic Lagrangian integrals $\int_0^s ds'(dX_j/ds)(s', \boldsymbol{\xi})$ go asymptotically to a Gaussian distribution, Eq. (15) renders Eq. (17). Since the components along the diagonal of the matrix $[U_{ij}]$ each represents an appropriate measure of the Lagrangian path spreading in its respective direction, $[U_{ij}]$ will be called the *Lagrangian spreading matrix*. Thus, via Eq. (24), it is seen that the Lagrangian measure function $B(x_i - \xi_i \mid s)$ is a direct consequence of the Lagrangian path spreading. Equation (20') has a fluid mechanics analog in Richardson's "nearest neighbor diffusion" [compare Eq. (113) on p. 48 of Corrsin[7] with Eq. (4.6-28) on p. 133 of Neubert[5] (this was brought to the author's attention by J. L. Lumley)]; see also Eq. (9.27) of Monin and Yaglom[3] and Eq. (7.1.3) of Tennekes and Lumley.[8] If α represents an appropriate measure of the inhomogeneity of the medium, such as the rms variation of $\mu(\mathbf{x})$ relative to its mean, then the nonstochastic limit of Eq. (20) can be given by

$$\lim_{\alpha \to 0} U_{ij} = 0. \tag{21}$$

The Lagrangian spreading matrix $[U_{ij}]$ can be expressed in terms of its principal axes by assuming the medium is statistically isotropic and choosing a suitable initial orientation for the paths. This reduces the determinant $\|U_{ij}\|$ to

$$\|U_{ij}\| = \prod_{i=1}^{3} U_{ii} = U_{11}U_{22}U_{33}, \tag{22}$$

and the components of the inverse matrix of $[U_{ij}]$ are then given by

$$U_{ii}^{-1} = 1/U_{ii} \tag{23a}$$

and

$$U_{ij}^{-1} = 0 \quad \text{for} \quad i \neq j \tag{23b}$$

(no summation on **i**). Therefore, the Lagrangian measure function for a statistically isotropic medium can be written as

$$\begin{aligned} &B(x_i - \xi_i \mid s) \\ &= \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} d\mathbf{k} \varphi(\mathbf{k}, \boldsymbol{\xi} \mid s) e^{-ik_j(x_j - \xi_j)} \qquad (24) \\ &\sim \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} d\mathbf{k} \exp\left[-\tfrac{1}{2}k_i k_j U_{ij} - ik_j(x_j - \xi_j - m_j)\right] \qquad (25) \\ &= \frac{\exp\left[-\frac{1}{2}U_{ij}^{-1}(x_i - \xi_i - m_i)(x_j - \xi_j - m_j)\right]}{2\pi(2\pi \|U_{ij}\|)^{\frac{1}{2}}} \qquad (26) \\ &= \prod_{i=1}^{3} \frac{\exp\left[-\frac{1}{2}(x_i - \xi_i - m_i)^2/U_{ii}\right]}{(2\pi U_{ii})^{\frac{1}{2}}}. \qquad (27) \end{aligned}$$

Equation (25) arises from Eq. (19), and the transformation to Eq. (26) follows when $[U_{ij}]$ is a positive-definite matrix. Equations (22)–(23b) permit Eq. (27).

Goodman[9] shows that inside an integral including $\sigma = 0$ (and under suitable conditions)

$$\lim_{b \to \infty} be^{-b^2\pi\sigma^2} = \delta(\sigma), \tag{28}$$

i.e., this limit behaves like a delta function. Thence,

$$\lim_{\alpha \to 0} \gamma(x_i - \xi_i - m_i \mid s) = \delta\left(x_i - \xi_i - \lim_{\alpha \to 0} m_i\right), \tag{29}$$

provided m_i and U_{ii} are well behaved for the integral in question, where

$$\gamma(x_i - \xi_i - m_i \mid s) \equiv \frac{\exp\left[-\frac{1}{2}(x_i - \xi_i - m_i)^2/U_{ii}(s)\right]}{[2\pi U_{ii}(s)]^{\frac{1}{2}}}. \tag{30}$$

In short,

$$\begin{aligned} \lim_{\alpha \to 0} B(x_i - \xi_i \mid s) &\sim \lim_{\alpha \to 0} \prod_{i=1}^{3} \gamma(x_i - \xi_i - m_i \mid s) \qquad (27') \\ &\equiv \lim_{\alpha \to 0} \gamma(\mathbf{x}, \boldsymbol{\xi} \mid s) \qquad (31) \\ &= \delta(\mathbf{x} - \boldsymbol{\xi} - \mathbf{m} \mid s). \qquad (29') \end{aligned}$$

For example, assume that the initial surface S_0 is a rectangle of dimensions $2\eta_0$ by $2\zeta_0$ centered at $\boldsymbol{\xi} = \mathbf{0}$

with all initial paths perpendicular to S_0 so that

$$m_2 = m_2(s) = 0, \tag{32a}$$

$$m_3 = m_3(s) = 0, \tag{32b}$$

and

$$\lim_{\alpha \to 0} m_1(s) = s \tag{32c}$$

in a statistically isotropic medium. Let

$$[\xi_i] = [0, \eta, \zeta] \tag{33}$$

and

$$[x_i] = [x, y, z]. \tag{34}$$

Then

$$\lim_{\alpha \to 0} \int_{-\eta_0}^{\eta_0} d\eta \gamma(y - \eta \mid s) = \int_{-\eta_0}^{\eta_0} d\eta \delta(y - \eta) = \begin{cases} 1, & |y| \leq \eta_0 \\ 0, & |y| > \eta_0 \end{cases}, \tag{35a}$$

$$\lim_{\alpha \to 0} \int_{-\zeta_0}^{\zeta_0} d\zeta \gamma(z - \zeta \mid s) = \int_{-\zeta_0}^{\zeta_0} d\zeta \delta(z - \zeta) = \begin{cases} 1, & |z| \leq \zeta_0 \\ 0, & |z| > \zeta_0 \end{cases}, \tag{35b}$$

and Eq. (14) reduces to

$$\begin{aligned} &\lim_{\alpha \to 0} \langle F(\mathbf{x}) \rangle \\ &= \lim_{\alpha \to 0} \int_0^\infty ds \int_{S_0} d\boldsymbol{\xi} B(x_i - \xi_i \mid s) \tilde{E}\{F[\mathbf{X}(s, \boldsymbol{\xi})]\} \\ &\sim \int_0^\infty ds \int_{-\eta_0}^{\eta_0} d\eta \int_{-\zeta_0}^{\zeta_0} d\zeta \delta(x - s) \delta(y - \eta) \delta(z - \zeta) F(s, \eta, \zeta) \\ &= \begin{cases} F(x, \eta, \zeta), & -\eta_0 \leq \eta \leq \eta_0, \quad -\zeta_0 \leq \zeta \leq \zeta_0 \\ 0, & \text{otherwise} \end{cases}, \end{aligned} \tag{36}$$

which is the necessary nonstochastic limit. If $F[\mathbf{X}(s, \boldsymbol{\xi})]$ is the Lagrangian pressure wave $p[\mathbf{X}(s, \boldsymbol{\xi})]$ of Ref. 2, Eq. (36) renders

$$\lim_{\alpha \to 0} \langle p(\mathbf{x}) \rangle \sim p_0(\eta, \zeta) e^{ik_0 x}, \tag{37}$$

which is the proper result for a uniform, nondissipative medium.

IV. CONCLUSION

In Ref. 2, it was shown that the solution of the stochastic Helmholtz equation for continuous media

$$\nabla^2 p + k_0^2 \mu^2 p = 0, \tag{38}$$

where k_0 is the free-space wavenumber, should be treated by a two-variable expansion (see Chap. 3 of Cole[10]). The resulting sound pressure wave $p[\mathbf{X}(s, \boldsymbol{\xi})]$, Eq. (39) of Neubert,[2] is intrinsically Lagrangian so that the determination of $\langle p(\mathbf{x}) \rangle$ from $p[\mathbf{X}(s, \boldsymbol{\xi})]$ invokes the stochastic Eulerian-Lagrangian problem. Thus, Eq. (14) relates $\langle p(\mathbf{x}) \rangle$ and $\tilde{E}\{p[\mathbf{X}(x, \boldsymbol{\xi})]$ via $B(x_i - \xi' \mid s)$, which is given by Eq. (27), for a statistically isotropic medium, in terms of the diagonalized Lagrangian spreading matrix $[U_{ij}]$. For the case of Fermat paths,

$$\mu_{,i} = \frac{d}{ds}\left(\mu \frac{dX_i}{ds}\right), \tag{39}$$

so that $[U_{ij}]$ can be expressed explicitly in terms of $\mu[\mathbf{X}(s, \boldsymbol{\xi})]$ and the initial conditions on S_0. However, in order to make further progress with $[U_{ij}]$ and $E\{p[\mathbf{X}(s, \boldsymbol{\xi})]\}$, it is necessary to determine the asymptotic behavior of the stochastic Lagrangian integrals involved. The same is true of the general problem represented by Eq. (14). Thus, it is seen that Eqs. (14) and (15) yield a suitable resolution of the stochastic Eulerian–Lagrangian problem for a large class of sound propagation problems if the resultant stochastic Lagrangian integrals can be evaluated asymptotically.

ACKNOWLEDGMENT

The author wishes to express his gratitude to Professor J. L. Lumley for his comments regarding this study.

* Taken from the doctoral dissertation of J. A. Neubert, "Sound Propagation in Continuous Stochastic Media," under the Interdisciplinary Program in Engineering Acoustics, The Pennsylvania State University, University Park, Pa., 1970.

[1] J. L. Lumley, J. Math. Phys. **3**, 309 (1962).

[2] J. A. Neubert, J. Acoust. Soc. Am. **48**, 1203 (1970).

[3] A. S. Monin and A. M. Yaglom, *Statistical Fluid Mechanics* (MIT Press, Cambridge, 1970).

[4] J. L. Lumley, *Stochastic Tools in Turbulence* (Academic, New York, 1971).

[5] J. A. Neubert, Ph.D. thesis, The Pennsylvania State University at University Park, 1970.

[6] I. M. Gel'fand and G. E. Shilov, *Generalized Functions, Volume I: Properties and Operations* (Academic, New York, 1964).

[7] S. Corrsin, *Mécanique de la turbulence*, Colloques Internationaux du Centre National de la Recherche Scientifique, No. 108 (C.N.R.S., Paris, 1962), pp. 27–52.

[8] H. Tennekes and J. L. Lumley, *A First Course in Turbulence* (MIT Press, Cambridge, 1971).

[9] J. W. Goodman, *An Introduction to Fourier Optics* (McGraw-Hill, New York, 1968).

[10] J. D. Cole, *Perturbation Methods in Applied Mathematics* (Blaisdell, Waltham, Mass., 1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Evaluation Techniques for Quantum Mechanical Integrals Involving Orthogonal Polynomials

TIEN CHI CHEN
IBM Research Laboratory, San Jose, California 95114

(Received 8 December 1970)

Quantum mechanical matrix elements often involve orthogonal polynomials, whose properties can be exploited for evaluation. Nullity conditions are considered for hydrogenic and harmonic oscillator radial integrals, and Gaunt's triangular condition for integrals over triple products of associated Legendre functions is generalized so that the superscripts form a triangle of even perimeter. Where the integral is nonzero, Gaussian quadratures are exact for polynomial integrands and converge under very general conditions. When the integrand is a product of three functions, one of which is a linear polynomial, the quadrature is equivalent to first developing summation-orthogonal expansions of the component functions and then integrating the product exactly.

1. INTRODUCTION

In the evaluation of quantum mechanical matrix elements, simple yet powerful integration techniques for orthogonal polynomials can often be exploited. This is the subject of the present paper.

Frequently, the interest lies in the possibility that an entire class of these integrals vanishes identically. On this nullity depend selection rules in spectral transitions, the termination of orthogonal expansions, and simplification of approximate solutions of the Schrödinger problem. If the integral is not zero, it is desirable to compute numerical values by simple and efficient algorithms.

In the following, nullity conditions will be derived for radial function integrals including the hydrogenic case studied by Pasternack and Sternheimer[1] and by Swamy, Kulkarni, and Biedenharn.[2] The investigation then leads to a generalization of Gaunt's integral[3] over products of associated Legendre functions. Evaluations by Gaussian quadrature and by expansion of the integrand are then discussed.

2. ORTHOGONAL POLYNOMIALS[4]

A simple wavefunction ψ_L often is proportional to $[w(x)]^{\frac{1}{2}}\phi_k(x)$, where $\phi_k(x)$ is an orthogonal polynomial with respect to $w(x) \geq 0$ in the interval $[a, b]$:

$$\int_a^b w(x)\phi_k(x)\phi_l(x)\,dx = A_k\delta_{kl}. \tag{2.1}$$

The subscript can be used to designate the maximum degree of the polynomial; this convenience will be assumed in this paper.[5] The quantities $w(x)$, k, and perhaps x itself, depend on the collective index L. A matrix element may then involve the integral

$$I \equiv \int_a^b w(x)\phi_n(x)Q(x)\,dx \equiv I(\phi_n Q). \tag{2.2}$$

Usually $Q(x)$ is a product of functions containing, among other things, polynomials orthogonal possibly for a different weight function, and there may be several equivalent ways to express the same integral in the form $I(\phi_n Q)$.

If $Q(x)$ were a polynomial $P_{n-1}(x)$ of degree $(n-1)$ or lower, it would possess no component along $\phi_n(x)$, and the integral would vanish:

$$I(\phi_n P_{n-1}) = \sum_{m=0}^{n-1} a_m I(\phi_n\phi_m) = 0. \tag{2.3}$$

This special property of orthogonal polynomials can be exploited for the evaluation of quantum mechanical integrals. An interesting special case presents itself when $Q(x) = \phi_m(x)\phi_l(x)$:

Theorem: $I(\phi_n\phi_m\phi_l)$ vanishes unless l, m, and n can form a triangle, that is, unless $l + m \geq n$, $m + n \geq l$, and $n + l \geq m$.

An extension to associated Legendre functions will be seen in Sec. 5.

Equation (2.3) is clearly useful for the study of nullity of integrals; it is also the basis for the Gaussian quadrature formula,[6] indicated by $J(f)$ below:

$$I(f) \equiv \int_a^b w(x)f(x)\,dx \sim \sum_{i=1}^{N} W_i f(x_i) \equiv J(f), \tag{2.4}$$

where $\{x_i\}$ are the zeros of $\phi_N(x)$. All of these zeros can be proven to lie in $[a, b]$ and to be distinct. The formula is exact if $f(x)$ is a polynomial of degree $\leq (2N-1)$. A classical proof (Ref. 5, pp. 160–61) proceeds by noting that if $g(x)$ is the $(N-1)$th-degree Lagrangian interpolation polynomial agreeing with $f(x)$ at $\{x_i\}$, then $f(x) - g(x)$ has as factors $\phi_N(x)$ and some other polynomial $P_{N-1}(x)$ of degree $(N-1)$ or lower. Therefore,

$$I(f) - I(g) = I(f-g) = I(\phi_N P_{N-1}) = 0,$$

and $I(f) = I(g)$. The latter becomes $J(f)$ after integrating out the x dependency.

An interesting summation orthogonal property follows.[7,8] $\{W_i\}$ are all positive definite and define a set of N polynomials orthogonal in the sense of weighted summation. This set turns out to be $\{\phi_k(x)\}$ itself, because

$$J(\phi_k\phi_l) = \sum_{i=1}^{N} W_i\phi_k(x_i)\phi_l(x_i) = A_k\delta_{kl}, \quad k, l \leq N-1. \tag{2.5}$$

Therefore, an arbitrary function can be expanded by purely algebraic processes in this finite set:

$$f(x) \sim \sum_{j=0}^{P} b_j\phi_j(x), \quad b_j = \frac{J(\phi_j f)}{J(\phi_j\phi_j)}, \quad P \leq N-1; \tag{2.6}$$

the result will satisfy a least-squares property with respect to weighted summation, and will closely approximate the more difficult, integration-orthogonal expansion.

The maximum allowable summation-orthogonal expansion with $P = N-1$ turns out to equal $g(x)$, the Lagrangian interpolation at $\{x_i\}$.[8] This is seen by noting that $f(x)$ and $g(x)$ have equal expansion coefficients.

3. NULLITY OF A HYDROGENIC RADIAL FUNCTION INTEGRAL

Pasternack and Sternheimer[1] discovered and proved that the hydrogenic radial function integral important in multipole transitions,

$$I_1 \equiv \int_0^\infty R_{nl}R_{nl'}r^{2-s}\,dr, \quad n > l > l', \quad s = \text{integer}, \tag{3.1}$$

vanishes for $1 < s \leq l - l' + 1$. Recently a physical explanation has been given in terms of the symmetry properties of the Coulomb field.[2]

The original proof rests on properties of the generating function for the Laguerre polynomials. We shall now show this nullity to be the explicit consequence of polynomial orthogonality. Since

$$R_{nl} = N_{nl}e^{-x/2}x^l L_{n-l-1}^{2l+1}(x), \quad \text{with} \quad x = 2ZMe^2r/n\hbar^2, \tag{3.2}$$

we have

$$I_1 \propto \int_0^\infty e^{-x}x^{2l'+1}L_{n-l'-1}^{2l'+1}(x)Q_1(x)\,dx, \tag{3.3a}$$

where

$$Q_1 = x^{l-l'+1-s}L_{n-l-1}^{2l+1}(x) \tag{3.3b}$$

is a polynomial if, and only if, $s \leq l - l' + 1$. The degree of this polynomial is $(n - l' - s)$, which is strictly less than $(n - l' - 1)$ for $1 < s$. This is precisely the Pasternack–Sternheimer condition.

The condition is extendable to the case when the two Laguerre polynomials have different n's, in which case the integral vanishes for $n - n' + 1 < s \leq l - l' + 1$. Though useful for quantum calculations, this generalized result does not apply to Eq. (3.1) with differing n's. This is because, in the radial eigenfunction for the Coulomb field, a change of the principal quantum number n necessarily alters x and, in Eq. (3.3a), the weight function will be affected.[9]

With the Coulomb context, however, a different generalization is possible. The integral

$$I_2 \equiv \int_0^\infty r^{2-s}\prod_{k=1}^{P} R_{n_k l_k}(r)\,dr \tag{3.4}$$

vanishes if there is a term $k = p$ such that $2/n_p = \sum(1/n_k)$, and $(3 - P - 2n_p + \sum n_k) < s \leq (1 - 2l_p + \sum l_k)$. As an example, if $P = 3 = p$, $n_1 = n_2 = 2n_3$, $l_1 = l_2 = n_1 - 1$, and $l_3 = 0$, then $I_2 = 0$ for $3n_3 < s \leq 4n_3 - 1$.

4. AN ISOTROPIC HARMONIC OSCILLATOR RADIAL INTEGRAL

The radial function is

$$F_{nl}(r) = C_{nl}e^{-x/2}x^{l/2}L_{(n-l-1)/2}^{l+1/2}(x),$$

with $x = M\omega r^2/\hbar$. The integral

$$I_3 \equiv \int_0^\infty F_{nl}(r)F_{n'l'}(r)r^{2-s}\,dr, \quad l \geq l', \tag{4.1}$$

vanishes for $n - n' < s \leq l - l'$, if $s + l + l'$ is an even integer.

5. A GENERALIZATION OF GAUNT'S TRIANGULAR CONDITION

Gaunt[3] studied the important integral involving a triple product of associated Legendre functions

$$I_4 \equiv \int_{-1}^{+1} P_l^m(x)P_{l'}^{m'}(x)P_{l''}^{m''}(x)\,dx, \quad m \leq l, \quad m' \leq l', \quad m'' \leq l'', \tag{5.1}$$

for the case $m'' = m + m'$, and formulated the triangular condition, namely that I_4 with $m'' = m + m'$ vanishes unless l, l', and l'' form a triangle of even perimeter. The theorem was proved using repeated integration by parts starting from the Rodriguez formula for $P_l^m(x)$.

We shall present a simpler proof of a more general theorem. With no loss of generality, we assume m, m', and m'' to be positive.

Theorem: If m, m', and m'' form a triangle of even perimeter, then I_4 will vanish unless l, l', and l'' also form a triangle of even perimeter.

Proof: Expressing $P_l^m(x)$ in Gegenbauer polynomials $\{c_n^\lambda(x)\}$ orthogonal in $[-1, +1]$ with respect to the weight function $(1 - x^2)^{\lambda-1/2}$,

$$P_l^m(x) = K_{ml}(1 - x^2)^{m/2}C_{l-m}^{m+1/2}(x), \qquad (5.2)$$

we have

$$I_4 \propto \int_{-1}^{+1}(1 - x^2)^m C_{l-m}^{m+1/2}(x)Q_4(x)\,dx, \qquad (5.3a)$$

with

$$Q_4(x) = (1 - x^2)^{(m'+m''-m)/2}C_{l'-m'}^{m'+1/2}(x)C_{l''-m''}^{m''+1/2}(x). \qquad (5.3b)$$

The triangular condition on the m's shows that the integrand is even if and only if $(l + l' + l'')$ is even. Further, $Q_4(x)$ is guaranteed to be a polynomial of degree $(l' + l'' - m)$. For $I_4 \neq 0$ the latter quantity must be no less than $l - m$, that is, $l' + l'' \geq l$. Similarly, we obtain $l'' + l \geq l'$ and $l + l' \geq l''$.

QED

An extension follows.

Theorem: The integral

$$I_5 \equiv \int_{-1}^{+1} x^n \prod_{k=1}^{P} P_{l_k}^{m_k}(x)\,dx, \quad n = \text{integer}, \qquad (5.4)$$

vanishes if either of the following conditions holds:

(a) $n + \sum l_k - \sum m_k$ is an odd integer;

(b) $\sum m_k =$ even; further, there is a term with $k = p$ such that $\sum m_k \geq 2m_p$ and $n + \sum l_k < 2l_p$ with $2n + \sum_{k\neq p}(1 - (-1)^{l_k - m_k}) \geq 0$ (i.e., a negative n will be compensated for by a sufficient number of Gegenbauer polynomials of odd order).

6. INTEGRATION BY GAUSSIAN QUADRATURES

The integral I in Eq. (2.2) may need evaluation if the test for nullity fails. Here the use of Gaussian quadratures [Eq. (2.4)] is straightforward and effective, requiring only the tabular values $\{W_i\}$ and $\{x_i\}$ and the ability to evaluate the integrand. If $Q(x)$ is a polynomial of degree m, then the quadrature with $N > (n + m)/2$ will be exact, barring round-off errors. Further, the formula need not be based on $w(x)$; other weight functions can be employed to gain exactitude, efficiency, or simplicity.

Analytical formulas for some of these integrals exist; for instance, Miller[10] has treated the case when the integrands are products of associated Legendre functions or products of generalized Laguerre functions. These same integrals are easily calculable by Gaussian quadratures by using a computer and no less accurately, particularly if the number P of components in the product is large. For example, if the evaluation cost per component is K, the total cost for an N-point quadrature is measured by KNP. Miller's analytical formulas lead usually to an P-fold coupled sum; if the evaluation effort per term is $\tilde{K}$ and the number of terms per sum is $\tilde{N}$, the total cost would be measured by $\tilde{K}\tilde{N}^P$.

For repeated evaluations, the components of the product integrand can be pretabulated at the Gaussian abscissas, replacing the evaluation cost by table look-ups. The writer has evaluated I_4, Eq. (5.1) for $\sum m =$ even, by a modification of the Legendre–Gauss quadrature[11]:

$$I_4 = \sum_{i=1}^{N} R_l^m(x_i)R_{l'}^{m'}(x_i)R_{l''}^{m''}(x_i), \quad N > (l + l' + l'')/2, \qquad (6.1)$$

with $R_l^m(x_i) = W_i^{\frac{1}{2}}P_l^m(x_i)$ stored in the computer as table entries.[12]

Further, the choice of the weight function other than $w(x)$ can transform the integrand [$\phi_n Q$ in Eq. (2.2)] into a polynomial. I_4 and $I_5, n > 0$, are exact with $\sum m =$ even, using Legendre–Gauss quadratures; Chebyshev–Gauss quadratures will be exact for $\sum m =$ odd. Standard Laguerre–Gauss quadratures are perfect for I_1 and I_2 when $(s - 2)$ is not greater than the sum of orbital quantum numbers; a generalization with $w(x) = e^{-x}x^\alpha$, $\alpha = 0$ or $-\frac{1}{2}$, will evaluate exactly I_3 for $s \leq l + l' + 1$.

7. SUMMATION-ORTHOGONAL EXPANSION OF THE PRODUCT INTEGRANDS

Where $Q(x)$ cannot be transformed into a polynomial, Gaussian quadratures can still be highly accurate, especially if the integrand shows few singularities.[6] For finite integration limits, the Gaussian quadrature with $N \to \infty$ converges to the integral if the integrand $f(x)$ in Eq. (2.4) is continuous, or even discontinuous but bounded and Riemann-integrable. For the Hermite–Gauss or generalized Laguerre–Gauss quadratures, convergence occurs if, for large $|x|$, $|w(x)f(x)|$ tends to zero faster than $1/|x|$.

Alternatively, one could adopt the costlier strategy of first approximating the integrand and then integrating the result. However, it must be noted that an accurate yet inexact integrand approximation does not necessarily imply accuracy of the integral. Conversely, an accurate integration technique may result from averaging out sizable deviations in an approximate integrand; indeed, the accuracy of Gaussian quadratures is achieved by minimizing the error of the

integral itself, while de-emphasizing integrand accuracy. One also notes that the approximation of $f(x)$ may not be adequate if $w(x)$ is nontrivial.

However, one could exploit the summation-orthogonal property associated with quadratures,[8] and be protected by the least-squares property. The weight function is not a concern as long as the proper abscissas and weights are available. Now the most powerful summation-orthogonal expansion using $\{\phi_k(x)\}$ is achieved by Gaussian quadratures, and the maximal expansion for a given N is, as noted in Sec. 2, the Lagrangian interpolation at zeros of $\phi_N(x)$. The definite integral of the interpolation is, of course, the quadrature itself.

Therefore, Gaussian quadratures actually offer a short cut to the "expansion then integrate" techniques. To ascertain the accuracy of the expansion, one could examine the magnitudes of the last few Fourier coefficients through Eq. (2.7), and change N if these were not small enough; or coefficients from two or more quadratures can be compared for consistency.[13] It is not desirable to "check the convergence" by integrating various projections of the same expansion; the act of integration is fundamentally a projection along $\phi_0(x)$ only.

The quantum-mechanical integrals not only have a nontrivial $w(x)$, the integrand further is a product of functions. Each of these functions could be individually approximated, before integration; this may be very desirable because many of these are already polynomials, with known properties. Even then, however, the result may be no better than direct quadratures.

Theorem: Let $\{g_k(x)\}$ be the $(N-1)$-degree Lagrangian interpolation polynomials agreeing with $\{f_k(x)\}$ at the zeros $\{x_i\}$ of $\phi_N(x)$. If $g_1(x)g_2(x)\cdots g_P(x)$ is a polynomial of degree $2N-1$, then $I(g_1g_2\cdots g_P) = J(f_1f_2\cdots f_P)$.

Proof: $I(g_1\cdots g_P)$ is exactly represented by the Gaussian quadrature $J(g_1\cdots g_P)$, which due to the interpolation nature of $\{g_k(x)\}$ is equal to $J(f_1\cdots f_P)$. QED

Corollary 1: $I(g_1P_N) = J(f_1P_N)$, if $P_N(x)$ is any Nth-degree polynomial. This is proved by factoring P_N conceptually into two Lagrangian polynomials.

Corollary 2: $I(g_1g_2P_1) = J(f_1f_2P_1)$, if $P_1(x)$ is any linear polynomial.

In Eq. (2.2) it might appear attractive to expand $Q(x)$ and take the projection along $\phi_n(x)$; this would yield exact results if the expansion were based on integration orthogonality. However, by Corollary 1 the Gaussian quadrature already behaves as if $Q(x)$ has been replaced by the maximal summation-orthogonal expansion.

ACKNOWLEDGMENTS

The writer acknowledges the stimulation received from Dr. R. A. Willoughby, IBM Watson Research Center, and Dr. T.-S. Shao, Bell Telephone Laboratories. He is grateful for the hospitality extended by Professor P.-O. Löwdin during the writer's stay at the Quantum Chemistry Institute, University of Uppsala, Sweden, where views in this work became crystallized through discussion with colleagues. Professor P. F. Zweifel, Virginia Polytechnic Institute, offered valuable comments on a draft of the manuscript.

[1] S. Pasternack and R. M. Sternheimer, J. Math. Phys. **3**, 1280 (1962).

[2] N. V. V. J. Swamy, R. G. Kulkarni, and L. C. Biedenharn, J. Math. Phys. **11**, 1165 (1970).

[3] J. A. Gaunt, Phil. Trans. Roy. Soc. (London) **A228**, 151 (1929), esp. pp. 192–96.

[4] See, for example, G. Szegö, *Orthogonal Polynomials*, Am. Math. Colloquium Publications 23 (American Mathematical Society, Providence, R.I., 1959), rev. ed.

[5] Our notations for orthogonal polynomials follow those in A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 2, Chap. X.

[6] See Refs. 4, 5, and also A. H. Stroud and D. Secrest, *Gaussian Integration Formulas* (Prentice-Hall, Englewood Cliffs, N.J., 1966), and P. J. Davis and P. Rabinowitz, *Numerical Integration* (Ginn-Blaisdell, Waltham, Mass., 1967).

[7] C. Lanczos, *Applied Analysis* (Prentice-Hall, Englewood Cliffs, N.J., 1956), pp. 376–78.

[8] T. C. Chen, Rev. Mod. Phys. **35**, 569 (1963).

[9] If (r/x) were independent of n, then the set of discrete radial functions $R_{nl}(r)$ for $n > l$ would become complete; on the other hand, the Coulomb field demands that $R_{nl}(r)$ be incomplete unless augmented by the continuum Coulomb functions. See H. Shull and P.-O. Löwdin, J. Chem. Phys. **23**, 1555 (1955).

[10] J. Miller, Math. Computations **17**, 84 (1963).

[11] In numerical work it is common to prefix the quadrature name by the name of the orthogonal polynomials employed. Thus, Legendre–Gauss quadratures are Gaussian quadratures based on Legendre polynomials.

[12] T. C. Chen, Ph.D. thesis, Duke University, 1957, Chap. V.

[13] T. C. Chen, (unpublished).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Gravitation as a Consequence of the Self-Interaction of the Γ Fields*

MARIO NOVELLO
Centro Brasileiro de Pesquisas Fisicas
and
Institut de Physique Théorique, Université de Genève, Genève, Switzerland

(Received 11 June 1970)

We investigate an algebraic-geometric approach to field theory. We try to prove that it is possible to understand Einstein's gravitational equation as a consequence of the self-interaction of the fundamental Γ fields of the assumed Clifford algebra. Furthermore, the antisymmetric object of the algebra obeys a set of equations that has the same structure as Maxwell's equations.

INTRODUCTION

Since the early days of the paper of Infeld–van der Waerden on spinor analysis, many authors have investigated a theory that makes some element of a Clifford algebra a more fundamental object than the metric tensor.[1] Even if we do not consider this kind of hierarchy, we can use a generalized space–time spinor algebra in the analysis of the interaction of a fermion with the gravitational field, for instance.[2] A somewhat different use of this algebra has also been made[3] (Penrose and Pirani) in studying some properties of the gravitational field and even in some models of a unified field theory.[4] In this paper we will consider the Clifford algebra as fundamental in a sense that will soon be clear.

1. THE FUNDAMENTAL OBJECT

Let us consider a set of objects e^a that can generate a universal Clifford algebra (C-algebra) over a four-dimensional differentiable manifold V_4. From the well-known property of the C-algebra,

$$\{e^a, e^b\} \tag{1}$$

is a multiple of the identity of the algebra, where, as usual,

$$\{M, N\} = MN + NM. \tag{2}$$

It is well-known[5] that the dimension of this algebra is 2^4. We will represent the elements of the basis as the set

$$\begin{aligned} &\mathbb{1}, \\ &\Gamma_\alpha, \\ &\Sigma_{\alpha\beta} = \tfrac{1}{2}(\Gamma_\alpha\Gamma_\beta - \Gamma_\beta\Gamma_\alpha), \\ &\Gamma^5 = \varphi(x)\epsilon^{\alpha\beta\lambda\mu}\Gamma_\alpha\Gamma_\beta\Gamma_\lambda\Gamma_\mu, \qquad \text{with } \Gamma^5\Gamma^5 = \mathbb{1}, \\ &\Gamma_\alpha\Gamma^5 = \varphi(x)\epsilon^{\mu\nu\rho\sigma}\Gamma_\alpha\Gamma_\mu\Gamma_\nu\Gamma_\rho\Gamma_\sigma, \end{aligned} \tag{3}$$

where the index α, β has a tensorial character. This means that if we make a transformation of coordinates

$$x^\alpha \to x'^\alpha = f^\alpha(x^\beta), \tag{4}$$

the Γ_α behave as a vector, the Γ^5 as a scalar, and $\Sigma_{\alpha\beta}$ as an antisymmetrical tensor of second order. In our choice of the representation of the algebra, the elements of the basis have a two-index property. Let us write

$$\Gamma_\alpha^{AB}, \tag{5}$$

where A, B may assume the values 1, 2, 3, or 4. We assume that there is a group of internal transformations with a space–time dependence—that is, the Γ's may suffer a transformation like

$$\Gamma_\alpha(x) \to \Gamma'^{AB}_\alpha(x) = M^A_C\Gamma^{CD}_\alpha(x)M^{-1B}_D, \tag{6}$$

where M^A_B is not a constant.

The fundamental property of the C-algebra

$$\{\Gamma_\mu(x), \Gamma_\nu(x)\} = 2g_{\mu\nu}(x)\mathbb{1} \tag{7}$$

defines the symmetric metric tensor $g_{\mu\nu}(x)$. $\mathbb{1}$ is the identity element of the algebra. We have furthermore

$$\{\Gamma_\mu(x), \Gamma^5(x)\} = 0. \tag{8}$$

2. THE SELF-INTERACTION

We will assume here[6] that the Γ's satisfy an equation of the type

$$\Gamma_{\alpha\|\beta}(x) = [U_\beta(x), \Gamma_\alpha(x)], \tag{9}$$

where the brackets, as usual, mean the commutator and the symbol $\|$ means the covariant derivative defined by

$$\Gamma_{\alpha\|\beta} = \Gamma_{\alpha|\beta} - \begin{Bmatrix} \epsilon \\ \alpha\beta \end{Bmatrix}\Gamma_\epsilon + [\tau_\beta, \Gamma_\alpha], \tag{10}$$

$$\Gamma_{\alpha|\beta} \quad \text{means} \quad \frac{\partial\Gamma_\alpha}{\partial x^\beta},$$

$$\begin{Bmatrix} \epsilon \\ \alpha\beta \end{Bmatrix} \quad \text{is the Christoffel symbol,}$$

$$\tau_\beta \quad \text{is the internal affinity.}$$

The origin of the expression (9) rests on the assumption that the correlation, at separate points, between

the Γ's, even in the existence of the group of transformation (6), does not introduce any new field. This is equivalent to assuming that the internal affinity may be expressed as a function of the objects of the C-algebra only.

The object $U_\nu^{AB}(x)$ as obtained in Ref. 6 has the form

$$U_\nu^{AB}(x) = \{\Gamma_\nu(x)(\mathbb{1} + \Gamma^5(x))\}^{AB}. \tag{11}$$

A straightforward calculation shows that the covariant derivative is not commutative and that we may write

$$\Gamma_{\alpha\|\beta\|\lambda} - \Gamma_{\alpha\|\lambda\|\beta} = R_{\alpha\epsilon\beta\lambda}\Gamma^\epsilon + [\mathbb{R}_{\beta\lambda}, \Gamma_\alpha], \tag{12}$$

where $R_{\alpha\epsilon\beta\lambda}$ is the Riemann (curvature) tensor, $\mathbb{R}_{\alpha\beta}$ is the internal curvature.

3. THE GRAVITATIONAL FIELD

From Eqs. (9) and (11) we obtain

$$(\Gamma_{\alpha\|\beta\|\lambda} - \Gamma_{\alpha\|\lambda\|\beta})g^{\alpha\beta} = 0. \tag{13}$$

So, we have

$$R_{\alpha\epsilon\beta\lambda}g^{\alpha\beta}\Gamma^\epsilon + [\mathbb{R}_{\beta\lambda}, \Gamma_\alpha]g^{\alpha\beta} = 0 \tag{14}$$

or

$$R_{\epsilon\lambda}\Gamma^\epsilon + [\mathbb{R}_{\epsilon\lambda}, \Gamma^\epsilon] = 0. \tag{15}$$

What can we say about the form of the internal curvature? It is an easy matter to prove that

$$\Gamma^5_{\|\alpha\|\beta} = 0. \tag{16}$$

From this and from the consideration that

$$\Gamma^5_{\|\alpha\|\beta} - \Gamma^5_{\|\beta\|\alpha} = [\mathbb{R}_{\alpha\beta}, \Gamma^5], \tag{17}$$

we obtain

$$[\mathbb{R}_{\alpha\beta}, \Gamma^5] = 0. \tag{18}$$

The most general expression of the internal curvature obtained as an element of the Clifford algebra and satisfying Eq. (18) has the form

$$\mathbb{R}_{\alpha\beta} = S_{\alpha\beta}\mathbb{1} + P_{\alpha\beta}\Gamma^5 + B_{\alpha\epsilon}\Gamma^\epsilon\Gamma_\beta - B_{\beta\epsilon}\Gamma^\epsilon\Gamma_\alpha, \tag{19}$$

where $S_{\alpha\beta}$, $P_{\alpha\beta}$, and $B_{\alpha\beta}$ are pure tensors that satisfy the symmetry conditions

$$S_{\alpha\beta} + S_{\beta\alpha} = 0, \tag{20}$$

$$P_{\alpha\beta} + P_{\beta\alpha} = 0. \tag{21}$$

If we put expression (19) into (15) we obtain two separate equations

$$P_{\alpha\beta} = 0, \tag{22}$$

$$R_{\alpha\beta} - 2Bg_{\alpha\beta} - 4B_{\alpha\beta} = 0, \tag{23}$$

where

$$B = B_{\alpha\beta}g^{\alpha\beta}. \tag{24}$$

From these considerations we see that the expression (9) induces a relation between the contracted Riemannian curvature (Ricci tensor) and a tensorial field. If we assume that the divergence of the tensor field $B_{\alpha\beta}(x)$ is null, then we arrive at a contradiction. So, we cannot identify $B_{\alpha\beta}$ directly as a conserved energy-momentum tensor. If we assume otherwise that

$$B_{\alpha\beta} = Mg_{\alpha\beta} + NT_{\alpha\beta}, \tag{25}$$

where

$$T^\alpha_{\beta\|\alpha} = 0, \tag{26}$$

then expression (23) assumes the form

$$R_{\alpha\beta} - \tfrac{1}{2}Rg_{\alpha\beta} = -kT_{\alpha\beta} \tag{27}$$

for a particular choice of the functions M and N.

We see that, in this case, relation (25) (where $T_{\alpha\beta}$ is the momentum tensor) implies that Eq. (23) is just Einstein's equation of gravitational theory.

4. THE ANTISYMMETRICAL OBJECT

Let us see what are the relations that Eq. (9) gives for the antisymmetrical product

$$\Sigma_{\alpha\beta} = \tfrac{1}{2}(\Gamma_\alpha\Gamma_\beta - \Gamma_\beta\Gamma_\alpha). \tag{28}$$

A straightforward calculation can show that

$$\Sigma_{\alpha\beta\|\lambda} = 2\{g_{\alpha\lambda}\Gamma_\beta - g_{\beta\lambda}\Gamma_\alpha\}(\mathbb{1} + \Gamma^5). \tag{29}$$

If we define

$$\Sigma_{\{\alpha\beta\|\lambda\}} = \frac{1}{3!}\{\Sigma_{\alpha\beta\|\lambda} - \Sigma_{\alpha\lambda\|\beta} + \Sigma_{\beta\lambda\|\alpha} - \Sigma_{\beta\alpha\|\lambda} + \Sigma_{\lambda\alpha\|\beta} - \Sigma_{\lambda\beta\|\alpha}\}, \tag{30}$$

then (29) gives

$$\Sigma_{\{\alpha\beta\|\lambda\}} = 0. \tag{31}$$

Let us now evaluate the divergence of $\Sigma^{AB}_{\alpha\beta}$. We obtain

$$\Sigma^\alpha_{\beta\|\alpha} = 6\Gamma_\beta(\mathbb{1} + \Gamma^5). \tag{32}$$

If we define the current J^{AB}_β as

$$J_\beta = 6\Gamma_\beta(\mathbb{1} + \Gamma^5), \tag{33}$$

we obtain the continuity equation

$$J^\beta_{\|\beta} = 0. \tag{34}$$

Equation (31) shows that we may introduce a potential ϑ^{AB}_α such that

$$\Sigma^{AB}_{\alpha\beta} = \vartheta^{AB}_{\alpha\|\beta} - \vartheta^{AB}_{\beta\|\alpha}. \tag{35}$$

A choice for this potential may be

$$\vartheta_\alpha = (\text{const})\Gamma_\alpha. \tag{36}$$

We see that these relations are just Maxwell's equations applied to an object that has an internal structure besides the tensor character of the usual electromagnetic field.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor J. M. Jauch for many very interesting conversations.

* This work has been supported by the Coordenação do Aperfeiçoamento de Pessoal de Nivel Superior (CAPES) and by the Centro Brasileiro de Pesquisas Fisicas.

[1] For a general review of the literature see W. L. Bade and H. Jehle, Rev. Mod. Phys. **25**, 714 (1953).
[2] D. R. Brill and J. A. Wheeler, Rev. Mod. Phys. **29**, 465 (1957).
[3] See references in A. Trautmann, F. A. E. Pirani, and A. Bondi, Brandeis Lectures on General Relativity, 1964.
[4] M. Novello, Nuovo Cimento **64**, 954 (1969).
[5] I. R. Porteous, *Topological Geometry* (Van Nostrand-Reinhold, London, 1969).
[6] M. Novello, *Vers une théorie unifiée des interactions faibles et gravitationnelles* (unpublished).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

One-Parameter Subgroups of Unitary Groups with Indefinite Metric and in Particular of the Conformal Group

J. G. F. BELINFANTE
Department of Mathematics and Department of Physics, Carnegie-Mellon University Pittsburgh, Pennsylvania 15213

AND

P. WINTERNITZ*
Department of Physics, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

(Received 14 October 1970)

Motivated by a problem concerning two-variable expansions of covariant scattering amplitudes, and by recent theories involving indefinite metrics and the conformal group, we study subgroups of unitary groups with indefinite metric. The one-parameter subgroup case reduces to finding canonical matrices for pseudo-Hermitian operators with respect to orthonormal bases. By decomposing the space on which such operators act as far as possible as an orthogonal direct sum of invariant subspaces, one obtains invariant subspaces having indecomposable primary components. The general results, summarized in tables of canonical forms valid for any finite dimension, are supplemented by more detailed tables for low dimensions, including the case of the conformal group of space–time.

I. INTRODUCTION

A. Harmonic Expansions for Noncompact Groups

In many applications of group theory in physics it is important to know the lattice of subgroups of a given group. The tools used by physicists are usually group representations realized in a definite vector space with a specified basis. For Lie groups, the simplest choices of bases can be obtained from chains of those subgroups which have Casimir invariants of some order, these invariants being used to obtain a complete set of commuting operators in the representation space. Two subgroup chains are physically equivalent if the subgroups in the one chain are conjugate to those in the other, so that we need only study the conjugacy classes of subgroups of a given group. The physical nonequivalence of various bases of group representation spaces has recently received considerable attention for the Lorentz and Poincaré groups.[1–4] Bases obtained from subgroup chains have been systematically studied for the representations of the little groups of the Poincaré group.[5,6]

We study here the classification of the conjugacy classes of all one-parameter subgroups of the unitary groups $U(p, q)$ and the special unitary groups $SU(p, q)$. In particular, we examine the one-parameter subgroups of the conformal group $SU(2, 2)$, thus preparing the way for a more general study of all the connected analytic subgroups of the conformal group. For this more general program, the classification of low-dimensional real Lie algebras given by Mubarakzjanov should prove useful.[7] The conformal group $SU(2, 2)$ is the smallest simple Lie group containing a subgroup isomorphic to the Poincaré group. Hence a study of the subgroup lattice of the conformal group should, as a by-product, shed some light on the subgroup lattice of the Poincaré group.

The choice of bases of group representation spaces by the use of subgroup chains also has applications to harmonic analysis on Lie groups and homogeneous manifolds.[8] Such harmonic analysis constitutes the mathematical foundation for recent generalizations of the partial wave analysis of scattering amplitudes.

We see that these relations are just Maxwell's equations applied to an object that has an internal structure besides the tensor character of the usual electromagnetic field.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor J. M. Jauch for many very interesting conversations.

* This work has been supported by the Coordenação do Aperfeiçoamento de Pessoal de Nivel Superior (CAPES) and by the Centro Brasileiro de Pesquisas Fisicas.

[1] For a general review of the literature see W. L. Bade and H. Jehle, Rev. Mod. Phys. **25**, 714 (1953).
[2] D. R. Brill and J. A. Wheeler, Rev. Mod. Phys. **29**, 465 (1957).
[3] See references in A. Trautmann, F. A. E. Pirani, and A. Bondi, Brandeis Lectures on General Relativity, 1964.
[4] M. Novello, Nuovo Cimento **64**, 954 (1969).
[5] I. R. Porteous, *Topological Geometry* (Van Nostrand-Reinhold, London, 1969).
[6] M. Novello, *Vers une théorie unifiée des interactions faibles et gravitationnelles* (unpublished).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

One-Parameter Subgroups of Unitary Groups with Indefinite Metric and in Particular of the Conformal Group

J. G. F. BELINFANTE
Department of Mathematics and Department of Physics, Carnegie-Mellon University Pittsburgh, Pennsylvania 15213

AND

P. WINTERNITZ*
Department of Physics, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

(Received 14 October 1970)

Motivated by a problem concerning two-variable expansions of covariant scattering amplitudes, and by recent theories involving indefinite metrics and the conformal group, we study subgroups of unitary groups with indefinite metric. The one-parameter subgroup case reduces to finding canonical matrices for pseudo-Hermitian operators with respect to orthonormal bases. By decomposing the space on which such operators act as far as possible as an orthogonal direct sum of invariant subspaces, one obtains invariant subspaces having indecomposable primary components. The general results, summarized in tables of canonical forms valid for any finite dimension, are supplemented by more detailed tables for low dimensions, including the case of the conformal group of space–time.

I. INTRODUCTION

A. Harmonic Expansions for Noncompact Groups

In many applications of group theory in physics it is important to know the lattice of subgroups of a given group. The tools used by physicists are usually group representations realized in a definite vector space with a specified basis. For Lie groups, the simplest choices of bases can be obtained from chains of those subgroups which have Casimir invariants of some order, these invariants being used to obtain a complete set of commuting operators in the representation space. Two subgroup chains are physically equivalent if the subgroups in the one chain are conjugate to those in the other, so that we need only study the conjugacy classes of subgroups of a given group. The physical nonequivalence of various bases of group representation spaces has recently received considerable attention for the Lorentz and Poincaré groups.[1–4] Bases obtained from subgroup chains have been systematically studied for the representations of the little groups of the Poincaré group.[5,6]

We study here the classification of the conjugacy classes of all one-parameter subgroups of the unitary groups $U(p, q)$ and the special unitary groups $SU(p, q)$. In particular, we examine the one-parameter subgroups of the conformal group $SU(2, 2)$, thus preparing the way for a more general study of all the connected analytic subgroups of the conformal group. For this more general program, the classification of low-dimensional real Lie algebras given by Mubarakzjanov should prove useful.[7] The conformal group $SU(2, 2)$ is the smallest simple Lie group containing a subgroup isomorphic to the Poincaré group. Hence a study of the subgroup lattice of the conformal group should, as a by-product, shed some light on the subgroup lattice of the Poincaré group.

The choice of bases of group representation spaces by the use of subgroup chains also has applications to harmonic analysis on Lie groups and homogeneous manifolds.[8] Such harmonic analysis constitutes the mathematical foundation for recent generalizations of the partial wave analysis of scattering amplitudes.

Here each subgroup chain leads to a different harmonic expansion.[9]

The standard partial wave expansion of a nonrelativistic two-body scattering amplitude for a fixed energy is an expansion in terms of the spherical functions of the rotation group $SO(3, \mathbf{R})$. Such partial wave expansions can be generalized to the scattering of relativistic particles, and also further developed in two directions. The first type of development leads to little group expansions for fixed total energy–momentum squared s, or for fixed momentum-transfer squared t. These are expansions in terms of the basis functions or finite transformation matrices of the $SO(3, \mathbf{R})$, $SO(2, 1; \mathbf{R})$, $E(2, \mathbf{R})$, and $SO(3, 1; \mathbf{R})$ little groups of the Poincaré group.[2,10–12] The second approach[1,5,12,13] leads to two-variable expansions of relativistic two-body scattering amplitudes $f(s, t)$ as a function of the two invariant Mandelstam variables s and t. The scattering amplitude is regarded as a function on a hyperboloid in relativistic momentum or velocity space, and is expanded in terms of harmonic functions of a Lorentz group acting on this hyperboloid. The subgroup chains $SO(3, 1; \mathbf{R}) \supset G \supset SO(2, \mathbf{R})$, where G is either $SO(3, \mathbf{R})$, $SO(2, 1; \mathbf{R})$, or $E(2, \mathbf{R})$, lead to three different harmonic expansions. The subgroup G in each chain here furnishes a single-variable expansion coinciding with the corresponding little group expansion. The group $SO(3, 1; \mathbf{R})$ then supplements this little group expansion with a further expansion of the relevant partial-wave amplitude in terms of functions of the other kinematical variable. Thus the $SO(3, 1; \mathbf{R})$ two-variable expansions, by incorporating the whole little group formalism, make full use of the Lorentz invariance of the scattering amplitude. The hyperboloid formalism is mainly a mathematical convenience, and one could let the $SO(3, 1; \mathbf{R})$ group act directly on the Mandelstam plane, albeit then with this induced action being different for each subgroup chain. The group $SO(3, 1; \mathbf{R})$, acting here as a group of motions on a homogeneous space of kinematical variables, is not the physical Lorentz group and could be replaced by some other group such as $SU(3)$. In the case of scattering with all four external masses equal, two-variable expansions based on an $SU(3)$ group of motions have been written, having the advantage of convenient crossing properties.[14] The scattering amplitude is expanded in terms of eigenfunctions of a differential operator $L_s^2 + L_t^2 + L_u^2$ which is symmetric in s, t, u. This operator commutes with the angular momenta in all three physical channels of a two-body process and can be interpreted as the second-order Casimir invariant for an $SU(3)$ group. The harmonic expansions in all three channels will be in terms of eigenfunctions of this one operator, making the appropriate crossing matrix block diagonal. The compact group $SU(3)$ only furnishes expansions in finite regions such as the unphysical Euclidean triangle in the Mandelstam plane, but a generalization to physical regions may be possible using the noncompact group $SU(2, 1)$.

The specific form of the harmonic expansion depends not only on the group chosen, but also on the chosen subgroup chain. It should be pointed out, however, that there are also some interesting expansions which are not based on subgroup chains.[15] All subgroups of $SU(2, 1)$ have recently been classified.[16] Since it was found that the group $SU(2, 1)$ does not contain an $E(2, \mathbf{R})$ subgroup, expansions based on it cannot fully incorporate the Poincaré little group formalism. On the other hand, expansions based on $SO(3, 1; \mathbf{R})$ do not have simple crossing properties, and one may therefore want to find another group which combines the best features of $SU(2, 1)$ and $SO(3, 1; \mathbf{R})$. The conformal group $SU(2, 2)$, locally isomorphic to $SO(4, 2; \mathbf{R})$, is a bigger group, which may have harmonic expansions fully incorporating the little group formalism and which have simple crossing properties.[17] Furthermore, the group $SU(2, 2)$ can be used for reactions involving mass-zero particles and for five-point production amplitudes. Harmonic analysis on $SU(2, 2)$ may also be of interest for another reason if this group becomes a symmetry group at very high energy. In this case one would like to study whether conformal invariance is compatible with Regge asymptotic behavior and, if so, to write a conformal invariant analog of the Regge pole formula.[18] Still more generally, one could consider harmonic analysis on any of the groups $SU(p, q)$, for which the Plancherel measure has recently been computed.[19]

B. Indefinite Metrics

The classification of one-parameter subgroups of unitary groups, with the related classification of canonical forms for Hermitian linear operators with respect to an indefinite metric, has many other applications. In quantum electrodynamics or in the theory of a mass-zero vector meson, an indefinite metric in the state vector space may help to satisfy simultaneous requirements of positive-energy density, relativistic covariance, and the subsidiary conditions on the fields.[20] (The Lorentz condition of classical electrodynamics is replaced in the Gupta theory of quantum electrodynamics by a subsidiary condition which requires the physical Hilbert subspace of the complete state vector space to be the kernel of a

certain linear operator, namely, the divergence of the positive-frequency part of the electromagnetic potential 4-vector.)

Heisenberg has also studied the use of an indefinite metric in the state vector space as a method for implementing the Pauli–Villars regularization method for obtaining finite renormalization constants in a nonlinear spinor theory.[21] The nonsingular Hermitian quantum state vector space in the Heisenberg theory is an orthogonal direct sum of an anisotropic Hilbert subspace of normal physical states and a hyperbolic subspace of dipole ghost states, both of these subspaces being invariant under the Hamiltonian operator. More recently, Kastrup has proposed a method whereby such an indefinite metric may have a connection with the conformal group.[22] A specific indefinite-metric field theory was recently suggested by Lee and Wick, in which unphysical states with complex energies are introduced. In such theories it is important to know how many real and how many complex eigenvalues can exist for a given dimension.[23]

A classification of Hermitian linear operators in Hermitian spaces with an indefinite metric is of interest for these theories, but, in this article, a complete solution of this problem is given only for finite-dimensional spaces. For the case of quantum field theories with indefinite metric in an infinite-dimensional state vector space, the present results are directly applicable only to invariant subspaces with finite dimension, a special case which is still of considerable interest, nevertheless. For example, in such theories, the Hamiltonian is assumed to be Hermitian with respect to the indefinite metric. By quantizing in a box of finite volume, the spectrum of the Hamiltonian becomes discrete, and it is reasonable to assume that most of the energy levels have only finite degeneracy. Thus the subspaces of states corresponding to a fixed finite interval of energies may be finite-dimensional vector spaces to which our results will apply. A similar application of indefinite metrics is to the theory of degenerate Bethe–Salpeter equations. Hermitian operators on finite-dimensional vector spaces with indefinite metric arise here in the study of the residues of the multiple poles of the Bethe–Salpeter Green's functions.[24] The deeper question of studying infinite-dimensional topological and pseudotopological vector spaces with an indefinite metric will not be pursued here.[25] We may finally also mention that some results on the spectra of Hermitian operators can be obtained for infinite-dimensional spaces when they are irreducible unitary representation spaces.[26]

We now turn our attention to a different problem. While we have studied unitary groups with an indefinite metric, in some of the problems of interest, we could instead have studied orthogonal groups with indefinite metric. For example, since the conformal group is locally isomorphic both to $SU(2, 2)$ and to $SO(4, 2; \mathbf{R})$, the classification of its analytic subgroups may also be obtained by studying subgroup lattices of real orthogonal groups with indefinite metric. The conjugacy classes of subgroups of real orthogonal groups with indefinite metric have been studied previously for some special cases. The most celebrated case[27] is that of the Lorentz group $SO(3, 1; \mathbf{R})$, which is also locally isomorphic to the complex special linear group $SL(2, \mathbf{C})$. The subgroups of the ordinary Lorentz group have been listed by various authors.[5,28] Another well-studied case is that of the plane Lorentz group $SO(2, 1; \mathbf{R})$, which is locally isomorphic also to $SU(1, 1)$ and to $SL(2, \mathbf{R})$. The structure of the group $SO(2, 2; \mathbf{R})$ has been extensively studied in a series of papers by Zassenhaus.[29] The conjugacy classes of the two de Sitter groups $SO(4, 1; \mathbf{R})$ and $SO(3, 2; \mathbf{R})$ were recently investigated by Philips and Wigner.[30]

Another problem related to the structure of orthogonal groups occurs in the general theory of relativity in connection with the Petrov classification of Einstein spaces.[31] Here the problem is to study the algebraic structure of the Riemann curvature tensor, which satisfies

$$R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta} = R_{\gamma\delta\alpha\beta}$$

and

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\gamma\delta\beta} + R_{\alpha\delta\beta\gamma} = 0.$$

Since the curvature tensor is symmetric under the interchange of the first two indices with the last two indices, we may regard it as a symmetric linear operator on a suitable vector space. Since the curvature tensor is also antisymmetric in the first two indices, the vector space in question can be taken as the space of antisymmetric two-index tensors. Thus, if V is the tangent space at a point of a Riemann space, then the curvature tensor may be regarded as a symmetric linear operator on the Grassmann product space $V \wedge V$. If V has metric signature (p, q), then $V \wedge V$ has an induced metric signature (p', q'), where $p' = \frac{1}{2}[p(p-1) + q(q-1)]$ and $q' = pq$. In the case of a four-dimensional Riemann space with metric signature $(+\,+\,+\,+)$, $(+\,+\,+\,-)$, or $(+\,+\,-\,-)$, the induced metric signature is $(+\,+\,+\,+\,+\,+)$, $(+\,+\,+\,-\,-\,-)$, or $(+\,+\,-\,-\,-\,-)$, respectively. For the Minkowski signature, the group relevant for studying the curvature tensor is $SO(3, 3; \mathbf{R})$, which is also locally isomorphic to the real special linear group

$SL(4, \mathbf{R})$. The subgroups of the corresponding group $SL(3, \mathbf{R})$ in three dimensions were enumerated by Nôno in connection with the classification of isotropy groups of smooth, materially uniform simple bodies.[32]

The conformal group of space–time is of interest in physics since all relativistically invariant free-field wave equations for zero-mass particles are also invariant under the conformal group.[33,34] Of current interest is the widely investigated conjecture that exact or broken conformal symmetry is relevant for scattering at very high energies and momentum transfers.[35] For a large class of relativistic interactions, conformal invariance would be implied by invariance under the one-parameter scale-transformation subgroup of the conformal group. The preliminary experimental evidence for scaling laws in deep-inelastic electroproduction has stimulated interest in the conformal group.[36] Finally, we note that the conformal group has been proposed as a dynamical group for hadrons,[37] in analogy with the use of this same group for the description of the hydrogen atom. It is also in connection with these recent developments that a detailed study of the structure of the conformal group is of interest.

II. ONE-PARAMETER SUBGROUPS

A. Structure of Pseudo-Hermitian Operators

In this section we study unitary groups $U(p, q)$ and special unitary groups $SU(p, q)$, classifying their one-parameter subgroups by using canonical forms for Hermitian linear operators on finite-dimensional vector spaces with indefinite metric. Our general approach may be described as an application of a combination of the Jordan canonical form theory with geometric algebra.[38] We make free use of the standard mathematical terminology concerning direct sums, orthogonal direct sums, invariant subspaces, and so forth.[39] For a positive-definite metric, we know from elementary quantum mechanics that the ordinary eigenvector theory suffices, and we can set up a parallel theory in the case of an indefinite metric by using generalized eigenvectors.[40] A nonzero vector ψ in a complex vector space V is said to be a generalized eigenvector of a linear operator γ on V iff $(\gamma - c1)^p\psi = 0$ for some complex number c and some positive integer p. The Dirac spinor space, which we may characterize as a four-dimensional complex vector space $\mathbf{C}^4$ equipped with the metric $(+\,+\,-\,-)$, is a familiar example of a Hermitian space with indefinite metric. For convenience we shall use the Dirac notation $\bar{\phi}\psi$ for Hermitian forms in general. Unitary and special unitary groups with indefinite metric arise naturally in the study of the geometry of Hermitian spaces.[41] In particular, the conformal group $SU(2, 2)$ is the special unitary group of Dirac spinor space.[42]

By a Hermitian linear operator γ on a Hermitian space V we mean any linear operator satisfying $\overline{(\gamma\phi)}\psi = \bar{\phi}(\gamma\psi)$ for all ϕ, ψ in V. In the case of an indefinite metric, such operators are sometimes also called "pseudo-Hermitian" in the physics literature. If S is an invariant subspace of a Hermitian linear operator γ on a Hermitian space V, then the orthogonal complement $S^\perp$ is also an invariant subspace. The primary component V_γ^c is the subspace of V consisting of zero and all generalized eigenvectors of γ corresponding to the eigenvalue c, if any. A lot is known about the primary components of a Hermitian operator when the Hermitian space V is nonsingular, that is, when $V^\perp = 0$. If γ is a Hermitian linear operator on a nonsingular Hermitian space V and if c is real, then the primary component V_γ^c is nonsingular. We recall that a subspace S of a Hermitian space V is nonsingular iff its radical $S \cap S^\perp$ is zero. If c_1 and c_2 are not complex conjugates of each other, then the primary components $V_\gamma^{c_1}$ and $V_\gamma^{c_2}$ are orthogonal to each other. Thus, when c is not real, the primary component V_γ^c is orthogonal to itself and hence totally isotropic in the sense that $\bar{\psi}\psi = 0$ for all ψ in V_γ^c. On the other hand, if c_1 and c_2 are complex conjugate to each other and if V is nonsingular, then $V_\gamma^{c_1}$ and $V_\gamma^{c_2}$ have the same dimension, and their direct sum is nonsingular. The proofs of these assertions, as well as of other theorems to be quoted below, may be found in the Appendix.[43]

We now introduce the important new concept of an elementary invariant subspace, bearing the same relation to orthogonal direct sums that the usual concept of indecomposable invariant subspace bears to ordinary direct sums. An elementary invariant subspace of an operator is an invariant subspace which cannot be written as an orthogonal direct sum of two nonzero invariant subspaces. If V is finite dimensional, it is the orthogonal direct sum of elementary invariant subspaces, $V = S_1 \perp \cdots \perp S_n$. This raises the problem of studying elementary Hermitian operators, that is, Hermitian linear operators γ on nonsingular Hermitian spaces V for which V is itself elementary. If γ is an elementary Hermitian operator on a finite-dimensional nonsingular Hermitian space V, then either $V = V_\gamma^r$, where r is real, or else $V = V_\gamma^{r+is} \oplus V_\gamma^{r-is}$, where r is real and $s > 0$. Hence either $(\gamma - r1)^n = 0$ for some integer n, or else $[(\gamma - r1)^2 + s^2 1]^p = 0$ for some integer p. If a Hermitian operator γ on a Hermitian space V is nilpotent with $\gamma^n = 0$, then the cyclic

invariant subspace S generated by a vector ψ in V is a nonsingular n-dimensional subspace iff $\bar{\psi}\gamma^{n-1}\psi \neq 0$. For an elementary Hermitian linear operator γ with $V = V_\gamma^r$, it then follows that V is cyclic as well as primary and hence indecomposable. Similarly, for an elementary Hermitian operator γ with $V = V_\gamma^{r+is} \oplus V_\gamma^{r-is}$, where $s > 0$, the two primary components are again indecomposable. Thus an elementary subspace of a Hermitian operator on a nonsingular Hermitian space is either a nonsingular indecomposable invariant subspace or else the ordinary direct sum of a pair of totally isotropic invariant subspaces.

All maximal totally isotropic subspaces of a Hermitian space have the same dimension, called the index of the Hermitian space. The index of a nonsingular Hermitian space of dimension n cannot exceed $[n/2]$ and, if the index is equal to $[n/2]$, then we say that the space has maximal index. More simply stated, a nonsingular Hermitian space has maximal index iff the numbers of $+$'s and $-$'s in its metric signature are either the same or differ by one. Elementary invariant subspaces of a Hermitian linear operator on a finite-dimensional nonsingular Hermitian space have maximal index. The classification of Hermitian operators up to conjugacy then reduces via the Witt theorem to two problems.[44] The first problem is to find the possible ways of decomposing a given nonsingular Hermitian space as an orthogonal direct sum of maximal index nonsingular subspaces. These maximal index decompositions are readily found in each case by inspection. The second problem is to find the conjugacy classes of elementary Hermitian operators on a given maximal index nonsingular space. Our solution of the second problem is to make use of explicit canonical forms. A set of vectors $\psi_1, \cdots, \psi_n$ is said to be an orthonormal basis for the nonsingular Hermitian space V with signature (p, q) iff $\bar{\psi}_i\psi_j$ is 0 when $i \neq j$, $+1$ when $i = j \leq p$, and -1 when $i = j > p$. In the case $V = V_\gamma^r$, there exists a cyclic vector ϕ in V such that $\bar{\phi}(\gamma - r1)^k\phi$ is equal to $\epsilon = \pm 1$ when $k = n - 1$ and to zero otherwise. We can find an orthonormal basis consisting of various polynomials in γ acting on this cyclic vector ϕ, yielding a canonical matrix for γ which is slightly different for even and odd dimensional spaces. In the odd-dimensional case, the quantity ϵ just determines whether the metric signature has one more $+$ sign or one more $-$ sign. In the even-dimensional case, there is only one possible maximal index metric signature, and the quantity $\epsilon = \pm 1$ serves to distinguish two nonconjugate classes of Hermitian operators having identical Jordan canonical forms.[45] The case of an elementary Hermitian operator γ with $V = V_\gamma^{r+is} \oplus V_\gamma^{r-is}$ can occur only for even dimension $n = \dim V = 2p$. In this case we can find a pair of cyclic vectors $\phi_\pm$ in $V_\gamma^{r\pm is}$, respectively, such that $\bar{\phi}_-[\gamma - (r + is)1]^k\phi_+$ is equal to 1 for $k = p - 1$ and to 0 otherwise. Having a pair of vectors gives more freedom for adjusting normalizations, so that in this case there is no analog of the quantity ϵ. Here again a canonical form is obtained by constructing an orthonormal basis consisting of linear combinations of polynomials in γ acting on ϕ_+ and ϕ_-.

In all there are five cases to consider, listed below.

Case	Metric	Eigenvalues	Other parameters
I	$(p+1, p)$	r	$\epsilon = +1$
II	$(p, p+1)$	r	$\epsilon = -1$
III	(p, p)	r	$\epsilon = +1$
IV	(p, p)	r	$\epsilon = -1$
V	(p, p)	$r \pm is$	—

These cases are all nonconjugate; but, in the fifth case, the canonical forms differing only by the sign of s are conjugate, and we must therefore restrict s to be positive (say). In the Tables I–III listing the canonical matrices, we have combined cases differing only by the value of the parameter ϵ.

B. Low-Dimensional Cases

We now study the one-parameter subgroups of two-, three-, and four-dimensional unitary and special unitary groups with indefinite metric in some detail. We apply the general classification of Hermitian linear operators given above to the low-dimensional Hermitian spaces $\mathbf{C}^2(1, 1)$, $\mathbf{C}^3(2, 1)$, $\mathbf{C}^4(2, 2)$, and $\mathbf{C}^4(3, 1)$. Historically, we first obtained these canonical forms in the low-dimensional cases by direct calculation and then generalized the results to the case of any finite dimension, but it is more elegant to do it the other way around.

The simplest illustration of the application of the general theory is to the case of Hermitian operators on the two-dimensional Hermitian space $\mathbf{C}^2(1, 1)$. In this case, the most general Hermitian matrix is of the form

$$\begin{bmatrix} A + D & B + iC \\ -B + iC & A - D \end{bmatrix},$$

where A, B, C, and D are real numbers. The eigenvalues of this matrix are given by

$$A \pm (D^2 - B^2 - C^2)^{\frac{1}{2}}.$$

There are thus three spectral cases, depending on whether D^2 is greater than, less than, or equal to $B^2 + C^2$, corresponding respectively to a pair of distinct real eigenvalues, a complex conjugate pair,

TABLE I. Canonical elementary Hermitian matrix with a single real eigenvalue r in the even-dimensional case.

$$
\begin{bmatrix}
r & 1 & 0 & & & & & & & & & 0 & -1 & 0 \\
1 & r & 1 & & & & & & & & & -1 & 0 & 1 \\
0 & 1 & r & \cdot & & & & & & & \cdot & 0 & 1 & 0 \\
 & & \cdot & \cdot & \cdot & & & & & \cdot & \cdot & \cdot & & \\
 & & & \cdot & r & 1 & 0 & 0 & -1 & 0 & \cdot & & & \\
 & & & & 1 & r & 1 & -1 & 0 & 1 & & & & \\
 & & & & 0 & 1 & r+\epsilon & +\epsilon & 1 & 0 & & & & \\
 & & & & 0 & 1 & -\epsilon & r-\epsilon & 1 & 0 & & & & \\
 & & & & 1 & 0 & -1 & 1 & r & 1 & & & & \\
 & & & \cdot & 0 & -1 & 0 & 0 & 1 & r & \cdot & & & \\
 & & \cdot & \cdot & \cdot & & & & & \cdot & \cdot & \cdot & & \\
0 & 1 & 0 & \cdot & & & & & & & \cdot & r & 1 & 0 \\
1 & 0 & -1 & & & & & & & & & 1 & r & 1 \\
0 & -1 & 0 & & & & & & & & & 0 & 1 & r
\end{bmatrix}
$$

TABLE II. Canonical elementary Hermitian matrix with a single real eigenvalue r in the odd-dimensional case.

$$
\begin{bmatrix}
r & 1 & 0 & & & & & & & & 0 & -1 & 0 \\
1 & r & 1 & & & & & & & & -1 & 0 & 1 \\
0 & 1 & r & \cdot & & & & & & \cdot & 0 & 1 & 0 \\
 & & \cdot & \cdot & \cdot & & & & \cdot & \cdot & \cdot & & \\
 & & & \cdot & r & 1 & 0 & -1 & 0 & \cdot & & & \\
 & & & & 1 & r & \epsilon & 0 & 1 & & & & \\
 & & & & 0 & 1 & r & 1 & 0 & & & & \\
 & & & & 1 & 0 & -\epsilon & r & 1 & & & & \\
 & & & \cdot & 0 & -1 & 0 & 1 & r & \cdot & & & \\
 & & \cdot & \cdot & \cdot & & & & \cdot & \cdot & \cdot & & \\
0 & 1 & 0 & \cdot & & & & & & \cdot & r & 1 & 0 \\
1 & 0 & -1 & & & & & & & & 1 & r & 1 \\
0 & -1 & 0 & & & & & & & & 0 & 1 & r
\end{bmatrix}
$$

TABLE III. Canonical elementary Hermitian matrix with a complex conjugate pair of nonreal eigenvalues $r \pm is$.

$$
\begin{bmatrix}
r & 1 & 0 & & & & & & & & 0 & -1 & is \\
1 & r & 1 & & & & & & & & -1 & is & 1 \\
0 & 1 & r & \cdot & & & & & & \cdot & is & 1 & 0 \\
 & & \cdot & \cdot & \cdot & & & & \cdot & \cdot & \cdot & & \\
 & & & \cdot & r & 1 & 0 & 0 & -1 & is & \cdot & & \\
 & & & & 1 & r & 1 & -1 & is & 1 & & & \\
 & & & & 0 & 1 & r & is & 1 & 0 & & & \\
 & & & & 0 & 1 & is & r & 1 & 0 & & & \\
 & & & & 1 & is & -1 & 1 & r & 1 & & & \\
 & & & \cdot & is & -1 & 0 & 0 & 1 & r & \cdot & & \\
 & & \cdot & \cdot & \cdot & & & & & \cdot & \cdot & \cdot & \\
0 & 1 & is & \cdot & & & & & & & \cdot & r & 1 & 0 \\
1 & is & -1 & & & & & & & & & 1 & r & 1 \\
is & -1 & 0 & & & & & & & & & 0 & 1 & r
\end{bmatrix}
$$

and a single real eigenvalue, possibly degenerate. There are only two possible breakups of $\mathbf{C}^2(1, 1)$ as an orthogonal direct sum of maximal index subspaces, namely $(+-)$ and $(+)(-)$. The correspondence between spectral cases and maximal index breakups is clearly as follows. The spectral case of a real pair of distinct eigenvalues can only correspond to the breakup $(+)(-)$, while the spectral case of a conjugate pair of nonreal eigenvalues can only correspond to the breakup $(+-)$. On the other hand, the degenerate spectral case of a single real eigenvalue could correspond to either the breakup $(+-)$ or the breakup $(+)(-)$. In terms of B, C, and D, the degenerate spectral case occurs whenever $B^2 + C^2 = D^2$; but this leads to a breakup of the type $(+)(-)$ only in the very special case $B = C = D = 0$ and of the type $(+-)$ otherwise. The canonical forms of matrices in each case, as well as the generators of representatives of each equivalence class of one-parameter subgroups of $U(1, 1)$ and $SU(1, 1)$, are given in Table IV. In determining the conjugacy classes of one-parameter subgroups of $U(1, 1)$, listed in column four, we have to remember that two matrices differing by a real nonzero factor, or by an inner automorphism of the group, generate conjugate subgroups. For the group $SU(1, 1)$, only traceless generators are permitted, giving the three conjugacy classes of one-parameter subgroups listed in column five. It is easy to see that the $SU(1, 1)$ subgroup in row one generates rotations, the one in row two generates pure Lorentz boosts along one of the space axes, and that in row three generates Euclidean translations. The limitations placed on the parameters in the matrices are necessary to take account of the equivalences otherwise existing within the individual classes. For example, since we have

$$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} r & is \\ is & r \end{pmatrix}\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = \begin{pmatrix} r & -is \\ -is & r \end{pmatrix},$$

in this case we restrict ourselves to the case $s > 0$.

In three dimensions we need only study the Hermitian space $\mathbf{C}^3(2, 1)$. Although a classification of the subgroups of $SU(2, 1)$ has already been given,[16] we repeat it in Table V for completeness as well as to conform to our present slightly different choice of canonical basis.

The possible maximal index decompositions of the three-dimensional Hermitian space $\mathbf{C}^3(2, 1)$ are $(+)(+)(-)$, $(+)(+-)$, and $(++-)$. The canonical matrices, generators, eigenvectors, and eigenvalues are summarized in Table V.

In four dimensions we have to consider both the Hermitian spaces $\mathbf{C}^4(2, 2)$ and $\mathbf{C}^4(3, 1)$. The Dirac spinor space $\mathbf{C}^4(2, 2)$ is the case of interest for the conformal group. The possible maximal index decompositions of the space $\mathbf{C}^4(2, 2)$ are $(+)(+)(-)(-)$, $(+)(+-)(-)$, $(+-)(+-)$, $(+)(+--)$, $(++-)(-)$, and $(++--)$. Again we treat each of these cases separately, as summarized in Table VI.

The results for the other four-dimensional Hermitian space $\mathbf{C}^4(3, 1)$, related to the groups $U(3, 1)$ and $SU(3, 1)$, are summarized in Table VII. These groups, which may be considered as possible complex extensions of the Lorentz group, have been studied in connection with relativistic hadron theory.[46] The maximal index decompositions of $\mathbf{C}^4(3, 1)$ are

TABLE IV. The space $\mathbf{C}^2(+-)$.

Case	Maximal index decomposition, eigenvalues, and eigenvectors	Canonical form of the matrix	Generator of $U(1, 1)$ subgroup	Generator of $SU(1, 1)$ subgroup
1.	$(+)(-)$ $r - s \sim \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $r + s \sim \begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} r - s & 0 \\ 0 & r + s \end{bmatrix}$ $-\infty < r, s < \infty$	(a) $s = 1, -\infty < r < \infty$ (b) $s = 0, r = 1$	$s = 1, r = 0$
2.	$(+-)$ $r + is \sim \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $r - is \sim \begin{bmatrix} 1 \\ -1 \end{bmatrix}$	$\begin{bmatrix} r & is \\ is & r \end{bmatrix}$ $-\infty < r < \infty, 0 < s < \infty$	$s = 1, 0 \leq r < \infty$	$s = 1, r = 0$
3.	$(+-)$ $r \sim \begin{bmatrix} 1 \\ -1 \end{bmatrix}$	$\begin{bmatrix} r + \epsilon & \epsilon \\ -\epsilon & r - \epsilon \end{bmatrix}$ $-\infty < r < \infty, \epsilon = \pm 1$	(a) $r = 1, \epsilon = \pm 1$ (b) $r = 0, \epsilon = 1$	$r = 0, \epsilon = 1$

$(+)(+)(+)(-)$, $(+)(+)(+-)$, and $(+)(++-)$. Note that the space $(+++-)$ itself is not a maximal index space and hence cannot be an elementary invariant subspace. A related consequence is that every Hermitian operator acting on $\mathbf{C}^4(3, 1)$ has at least one eigenvector with positive norm.

ACKNOWLEDGMENTS

The authors are indebted to Professor R. E. Cutkosky, Professor R. L. Ingraham, and Professor W. Noll for helpful discussions. One of us (P. W.) also thanks Professor J. Ashkin for the kind hospitality extended to him during his stay at the Physics Department of Carnegie–Mellon University.

APPENDIX

As a convenience to the reader, we sketch here the proofs of various assertions made about the structure of Hermitian operators with respect to an indefinite metric. Our first lemma generalizes the familiar arguments concerning orthogonality of eigenvectors to the case of an indefinite metric, where generalized eigenvectors have to be considered.

Lemma: If γ is a Hermitian linear operator on a finite-dimensional Hermitian space V and if c_1 and c_2 are two complex numbers which are not complex conjugates of each other, then the primary components $V_\gamma^{c_1}$ and $V_\gamma^{c_2}$ are orthogonal.

Proof: The kernels of the various powers $(\gamma - c_1 1)^p$ for increasing p form an ascending chain of subspaces which eventually terminates, so that for sufficiently large p we obtain the primary component $V_\gamma^{c_1}$. We shall show by induction on p that the kernel of $(\gamma - c_1 1)^p$ is orthogonal to the primary component $V_\gamma^{c_2}$. For $p = 1$, the argument is simple because if $(\gamma - c_1 1)\psi = 0$ and $(\gamma - c_2 1)^q \phi = 0$, then $0 = \bar{\psi}(\gamma - c_2 1)^q \phi = (c_1^* - c_2)^q \bar{\psi}\phi$, implying that $\bar{\psi}\phi = 0$. Suppose now that the kernel of $(\gamma - c_1 1)^p$ is orthogonal to $V_\gamma^{c_2}$. If ψ belongs to the kernel of $(\gamma - c_1 1)^{p+1}$, then $(\gamma - c_1 1)\psi$ belongs to the kernel of $(\gamma - c_1 1)^p$ and is therefore orthogonal to $V_\gamma^{c_2}$. For any vector $\phi \in V_\gamma^{c_2}$ there exists a smallest integer q such that $\bar{\psi}(\gamma - c_2 1)^q \phi = 0$. If $q > 0$, then

$$\overline{(\gamma - c_1 1)\psi}\,(\gamma - c_2 1)^{q-1}\phi = 0,$$

and it follows that $(c_1^* - c_2) \cdot \bar{\psi}(\gamma - c_2 1)^{q-1}\phi = 0$, so that $\bar{\psi}(\gamma - c_2 1)^{q-1}\phi = 0$, in contradiction with the assumed minimal property of q. Hence $q = 0$ and $\bar{\psi}\phi = 0$, showing that the kernel of $(\gamma - c_1 1)^{p+1}$ is orthogonal to $V_\gamma^{c_2}$ and thereby completing the inductive argument. QED

In the study of elementary invariant subspaces we used some results about cyclic vectors of nilpotent Hermitian operators. Since we can alter the trace of a Hermitian operator by adding a suitable multiple of the identity operator, it is sufficient to study the structure of traceless elementary Hermitian operators. If γ is a traceless elementary Hermitian operator on a finite-dimensional nonsingular space V, then either $V = V_\gamma^0$ or else $V = V_\gamma^{is} \oplus V_\gamma^{-is}$ for some real $s \neq 0$. Hence such operators are either nilpotent, that is, $\gamma^n = 0$ for some positive integer n, or else they satisfy the equation $(\gamma^2 + s^2 1)^p = 0$ for some $p > 0$ and $s \neq 0$. We prove the required result here for the nilpotent case more generally without assuming the operator to be elementary.

Lemma: If a Hermitian operator γ on a Hermitian space V is nilpotent with $\gamma^m = 0$, then the cyclic invariant subspace S generated by a vector ψ in V is a nonsingular m-dimensional subspace iff $\bar{\psi}\gamma^{m-1}\psi \neq 0$.

Proof: Since S is spanned by $\psi, \gamma\psi, \cdots, \gamma^{m-1}\psi$, any vector $\phi \in S \cap S^\perp$ can be written as a linear combination

$$\phi = (c_0 1 + c_1\gamma + \cdots + c_{m-1}\gamma^{m-1})\psi,$$

and $\bar{\psi}\gamma^k\phi = 0$ for all k. Since $\gamma^m = 0$, we have $0 = \bar{\psi}\gamma^{m-1}\phi = c_0\bar{\psi}\gamma^{m-1}\psi$, and, if $\bar{\psi}\gamma^{m-1}\psi \neq 0$, then we have $c_0 = 0$. By a similar argument, also $c_1 = \cdots = c_{m-1} = 0$, so that $\phi = 0$ and S is nonsingular. Moreover, this same argument also shows that the vectors $\psi, \gamma\psi, \cdots, \gamma^{m-1}\psi$ are linearly independent, and hence S has dimension m. Conversely, if $\bar{\psi}\gamma^{m-1}\psi = 0$, then any vector

$$\phi = (c_0 1 + c_1\gamma + \cdots + c_{m-1}\gamma^{m-1})\psi$$

in S satisfies $\bar{\psi}\gamma^{m-1}\phi = c_0\bar{\psi}\gamma^{m-1}\psi = 0$, and hence $\gamma^{m-1}\psi \in S \cap S^\perp$. If S were nonsingular, this would imply that $\gamma^{m-1}\psi = 0$, so that the dimension of S would be less than m. QED

The usual cyclic decomposition theorem in the theory of linear operators can now be strengthened as follows for nilpotent Hermitian operators.

Theorem: If γ is a nilpotent Hermitian operator on a finite-dimensional nonsingular Hermitian space V, then V is the orthogonal direct sum of a finite set of nonsingular cyclic invariant subspaces,

$$V = S_1 \perp \cdots \perp S_n.$$

Proof: Since γ is nilpotent, there exists a smallest integer m such that $\gamma^m = 0$. If $m = 1$, then $\gamma = 0$,

TABLE V. The space $C^3(++-)$.

Case	Maximal index decomposition, eigenvalues, and eigenvectors	Canonical form of the matrix	Generator of $U(2,1)$ subgroup	Generator of $SU(2,1)$ subgroup
1.	$(+)(+)(-)$ $a+b-c \sim \begin{bmatrix}1\\0\\0\end{bmatrix}, a+b+c \sim \begin{bmatrix}0\\1\\0\end{bmatrix},$ $a-2b \sim \begin{bmatrix}0\\0\\1\end{bmatrix}$	$\begin{bmatrix}a+b-c & 0 & 0\\ 0 & a+b+c & 0\\ 0 & 0 & a-2b\end{bmatrix}$ $-\infty < a, b < \infty, 0 \le c < \infty$	(a) $c=1, -\infty < a, b < \infty$ (b) $c=0, b=1, -\infty < a < \infty$ (c) $b=c=0, a=1$	(a) $c=1, -\infty < b < \infty, a=0$ (b) $a=c=0, b=1$
2.	$(+)(+-)$ $a \sim \begin{bmatrix}1\\0\\0\end{bmatrix}, r \pm is \sim \begin{bmatrix}0\\1\\\pm 1\end{bmatrix}$	$\begin{bmatrix}a & 0 & 0\\ 0 & r & is\\ 0 & is & r\end{bmatrix}$ $-\infty < a, r < \infty, 0 < s < \infty$	$s=1, 0 \le a < \infty, -\infty < r < \infty$	$s=1, 0 \le a = -2r < \infty$
3.	$(+)(+-)$ $a-2b \sim \begin{bmatrix}1\\0\\0\end{bmatrix}, a+b \sim \begin{bmatrix}0\\1\\-1\end{bmatrix}$	$\begin{bmatrix}a-2b & 0 & 0\\ 0 & a+b+\epsilon & \epsilon\\ 0 & -\epsilon & a+b-\epsilon\end{bmatrix}$ $-\infty < a, b < \infty, \epsilon = \pm 1$	(a) $-\infty < a < \infty, b=1, \epsilon = \pm 1$ (b) $a=1, b=0, \epsilon = \pm 1$ (c) $a=b=0, \epsilon = 1$	(a) $a=0, b=1, \epsilon = \pm 1$ (b) $a=0, b=0, \epsilon = \pm 1$
4.	$(++-)$ $r \sim \begin{bmatrix}1\\0\\-1\end{bmatrix}$	$\begin{bmatrix}r & 1 & 0\\ 1 & r & 1\\ 0 & -1 & r\end{bmatrix}$ $-\infty < r < \infty$	(a) $r=1$ (b) $r=0$	$r=0$

TABLE VI. The space $C^4(++--)$.

Case	Maximal index decomposition, eigenvalues, and eigenvectors	Canonical form of the matrix	Generator of $U(2,2)$ subgroup	Generator of $SU(2,2)$ subgroup
1.	$(+)(+)(-)(-)$ $a+b-c \sim \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, a+b+c \sim \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, a-b-d \sim \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, a-b+d \sim \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$	$\begin{bmatrix}a+b-c & 0 & 0 & 0\\ 0 & a+b+c & 0 & 0\\ 0 & 0 & a-b-d & 0\\ 0 & 0 & 0 & a-b+d\end{bmatrix}$ $-\infty < a, b < \infty, 0 \le c, d < \infty$	(a) $-\infty < a, b < \infty, c = 1, 0 \le d < \infty$ (b) $-\infty < -b \le a < \infty, c = 0, d = 1$ (c) $-\infty < a < \infty, b = 1, c = d = 0$ (d) $a = 1, b = c = d = 0$	(a) $a = 0, 0 \le b < \infty, c = 1, 0 \le d < \infty$ (b) $a = 0, 0 \le b < \infty, c = 0, d = 1$ (c) $a = 0, b = 1, c = d = 0$
2.	$(+)(+-)(-)$ $a-b \sim \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, a+b \sim \begin{bmatrix}0\\0\\0\\1\end{bmatrix}, r \pm is \sim \begin{bmatrix}0\\1\\\pm 1\\0\end{bmatrix}$	$\begin{bmatrix}a-b & 0 & 0 & 0\\ 0 & r & is & 0\\ 0 & is & r & 0\\ 0 & 0 & 0 & a+b\end{bmatrix}$ $-\infty < a, b, r < \infty, 0 < s < \infty$	$-\infty < b < a < \infty$, $-\infty < r < \infty, s = 1$	$-\infty < b < a = -r < \infty, s = 1$
3.	$(+)(+-)(-)$ $a-b \sim \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, a+b \sim \begin{bmatrix}0\\0\\0\\1\end{bmatrix}, r \sim \begin{bmatrix}0\\1\\-1\\0\end{bmatrix}$	$\begin{bmatrix}a-b & 0 & 0 & 0\\ 0 & r+\epsilon & \epsilon & 0\\ 0 & -\epsilon & r-\epsilon & 0\\ 0 & 0 & 0 & a+b\end{bmatrix}$ $-\infty < a, b, r < \infty, \epsilon = \pm 1$	(a) $-\infty < b \le a < \infty$, $r = 1, \epsilon = \pm 1$ (b) $b = 1, 1 \le a < \infty$, $r = 0, \epsilon = \pm 1$ (c) $a = 1, b = 0, r = 0$, $\epsilon = \pm 1$ (d) $a = b = r = 0, \epsilon = +1$	(a) $r = -a = 1$, $-\infty < b \le -1, \epsilon = \pm 1$ (b) $r = a = 0, b = -1$, $\epsilon = \pm 1$ (c) $a = b = r = 0, \epsilon = 1$
4.	$(+-)(+-)$ $r \pm is \sim \begin{bmatrix}1\\0\\0\\\pm 1\end{bmatrix}, t \pm iu \sim \begin{bmatrix}0\\1\\\pm 1\\0\end{bmatrix}$	$\begin{bmatrix}r & 0 & 0 & is\\ 0 & t & iu & 0\\ 0 & iu & t & 0\\ is & 0 & 0 & r\end{bmatrix}$ $-\infty < r, t < \infty, 0 < s \le u < \infty$	$s = 1, 1 \le u < \infty$, $0 \le r < \infty, -\infty < t < \infty$	$s = 1, 1 \le u < \infty$, $0 \le r = -t < \infty$

5.	$(+-)(+-)$ $r \pm is \sim \begin{bmatrix} 1 \\ 0 \\ 0 \\ \pm 1 \end{bmatrix}, t \sim \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} r & 0 & 0 & is \\ 0 & t+\epsilon & \epsilon & 0 \\ 0 & -\epsilon & t-\epsilon & 0 \\ is & 0 & 0 & r \end{bmatrix}$ $-\infty < r, t < \infty, 0 < s < \infty, \epsilon = \pm 1$	$s = 1, 0 \le r < \infty,$ $-\infty < t < \infty, \epsilon = \pm 1$	$s = 1, \epsilon = \pm 1,$ $0 \le r = -t < \infty$
6.	$(+-)(+-)$ $r \sim \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}, t \sim \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} r+\epsilon_1 & 0 & 0 & \epsilon_1 \\ 0 & t+\epsilon_2 & \epsilon_2 & 0 \\ 0 & -\epsilon_2 & t-\epsilon_2 & 0 \\ -\epsilon_1 & 0 & 0 & r-\epsilon_1 \end{bmatrix}$ $-\infty < r \le t < \infty, \epsilon_1, \epsilon_2 = \pm 1$	(a) $r = 1, 1 \le \lvert t \rvert < \infty,$ $\epsilon_1, \epsilon_2 = \pm 1$ (b) $r = 0, t = 1, \epsilon_1, \epsilon_2 = \pm 1$ (c) $r = 0, t = 0, \epsilon_1 = 1$ $\epsilon_2 = \pm 1$	(a) $r = -t = 1, \epsilon_1, \epsilon_2 = \pm 1$ (b) $r = t = 0, \epsilon_1 = 1, \epsilon_2 = \pm 1$
7.	$(+)(+--)$ $a + 3b \sim \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, a - b \sim \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$	$\begin{bmatrix} a+3b & 0 & 0 & 0 \\ 0 & a-b & -1 & 0 \\ 0 & 1 & a-b & 1 \\ 0 & 0 & 1 & a-b \end{bmatrix}$ $-\infty < a, b < \infty$	(a) $b = 1, -\infty < a < \infty$ (b) $b = 0, a = 1$ (c) $b = 0, a = 0$	(a) $a = 0, b = 1,$ (b) $a = b = 0$
8.	$(++-)(-)$ $a - b \sim \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, a + 3b \sim \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} a-b & 1 & 0 & 0 \\ 1 & a-b & 1 & 0 \\ 0 & -1 & a-b & 0 \\ 0 & 0 & 0 & a+3b \end{bmatrix}$ $-\infty < a, b < \infty$	(a) $b = 1, -\infty < a < \infty$ (b) $b = 0, a = 1$ (c) $b = a = 0$	(a) $a = 0, b = 1$ (b) $a = 0, b = 0$
9.	$(++--)$ $r + is \sim \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, r - is \sim \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$	$\begin{bmatrix} r & 1 & -1 & is \\ 1 & r & is & 1 \\ 1 & is & r & 1 \\ is & -1 & 1 & r \end{bmatrix}$ $-\infty < r < \infty, 0 < s < \infty$	$s = 1, 0 \le r < \infty$	$r = 0, s = 1$
10.	$(++--)$ $r \sim \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$	$\begin{bmatrix} r & 1 & -1 & 0 \\ 1 & r+\epsilon & \epsilon & 1 \\ 1 & -\epsilon & r-\epsilon & 1 \\ 0 & -1 & 1 & r \end{bmatrix}$ $-\infty < r < \infty, \epsilon = \pm 1$	(a) $r = 1, \epsilon = \pm 1$ (b) $r = 0, \epsilon = 1$	$r = 0, \epsilon = 1$

TABLE VII. The space $C^4(+++-)$.

Case	Maximal index decomposition, eigenvalues, and eigenvectors	Canonical form of the matrix	Generator of $U(3,1)$ subgroup	Generator of $SU(3,1)$ subgroup
1.	$(+)(+)(+)(-)$ $a+b\sim\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, a+b+c\sim\begin{bmatrix}0\\1\\0\\0\end{bmatrix},$ $a+b+c+d\sim\begin{bmatrix}0\\0\\1\\0\end{bmatrix}, a-3b-2c-d\sim\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$	$\begin{bmatrix}a+b & 0 & 0 & 0\\ 0 & a+b+c & 0 & 0\\ 0 & 0 & a+b+c+d & 0\\ 0 & 0 & 0 & a-3b-2c-d\end{bmatrix}$ $-\infty<a,b<\infty, 0\le c,d<\infty$	(a) $-\infty<a<\infty, b=1, 0\le c,d<\infty$ (b) $-\infty<a<\infty, b=0, c=1, 0\le d<\infty$ (c) $-\infty<a<\infty, b=c=0, d=1$ (d) $a=1, b=c=d=0$	(a) $a=0, b=1, 0\le c,d<\infty$ (b) $a=b=0, c=1, 0\le d<\infty$ (c) $a=b=c=0, d=1$
2.	$(+)(+)(+-)$ $a-b\sim\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, a+b\sim\begin{bmatrix}0\\1\\0\\0\end{bmatrix},$ $r\pm is\sim\begin{bmatrix}0\\0\\1\\\pm1\end{bmatrix}$	$\begin{bmatrix}a-b & 0 & 0 & 0\\ 0 & a+b & 0 & 0\\ 0 & 0 & r & is\\ 0 & 0 & is & r\end{bmatrix}$ $-\infty<a,r<\infty, 0\le b<\infty, 0<s<\infty$	$s=1, -\infty<a,r<\infty, 0\le b<\infty$	$s=1, 0\le b<\infty, -\infty<a=-r<\infty$
3.	$(+)(+)(+-)$ $a-b-c\sim\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, a-b+c\sim\begin{bmatrix}0\\1\\0\\0\end{bmatrix},$ $a+b\sim\begin{bmatrix}0\\0\\1\\-1\end{bmatrix}$	$\begin{bmatrix}a-b-c & 0 & 0 & 0\\ 0 & a-b+c & 0 & 0\\ 0 & 0 & a+b+\epsilon & \epsilon\\ 0 & 0 & -\epsilon & a+b-\epsilon\end{bmatrix}$ $-\infty<a,b<\infty, 0\le c<\infty, \epsilon=\pm1$	(a) $-\infty<a,b<\infty, c=1, \epsilon=\pm1$ (b) $-\infty<a<\infty, b=1, c=0, \epsilon=\pm1$ (c) $a=1, b=c=0, \epsilon=\pm1$ (d) $a=b=c=0, \epsilon=1$	(a) $a=0, -\infty<b<\infty, c=1, \epsilon=\pm1$ (b) $a=0, b=1, c=0, \epsilon=\pm1$ (c) $a=b=c=0, \epsilon=1$
4.	$(+)(++-)$ $a-3r\sim\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, r\sim\begin{bmatrix}0\\1\\0\\-1\end{bmatrix}$	$\begin{bmatrix}a-3r & 0 & 0 & 0\\ 0 & r & 1 & 0\\ 0 & 1 & r & 1\\ 0 & 0 & -1 & r\end{bmatrix}$ $-\infty<a,r<\infty$	(a) $r=1, -\infty<a<\infty$ (b) $r=0, a=1$ (c) $a=r=0$	(a) $a=0, r=1$ (b) $a=0, r=0$

and every subspace of V is invariant; any orthogonal direct sum decomposition of V into nonsingular lines does the trick. If $m > 1$, then, by the assumed minimal property of m, we have $\gamma^{m-1} \neq 0$. If $\bar{\psi}\gamma^{m-1}\psi = 0$ for all ψ in V, then, by polarization, we obtain $\bar{\psi}\gamma^{m-1}\phi = 0$ for all ψ, ϕ in V, and, since V is nonsingular, this would imply that $\gamma^{m-1} = 0$, a contradiction. Hence there is a vector ψ in V with $\bar{\psi}\gamma^{m-1}\psi \neq 0$, and, by the lemma, the cyclic invariant subspace S generated by ψ is a nonzero nonsingular subspace. Then we may write $V = S \perp S^{\perp}$, and the argument may be repeated with $S^{\perp}$ replacing V. QED

As a corollary of this theorem, we may conclude that elementary nilpotent Hermitian operators are indecomposable. We also prove an analogous theorem for the case of a Hermitian operator γ when $\gamma^2 + s^2 1$ is nilpotent.

Theorem: If γ is a Hermitian operator on a finite-dimensional nonsingular Hermitian space and if $\gamma^2 + s^2 1$ is nilpotent for some real $s \neq 0$, then there exist nonzero totally isotropic indecomposable subspaces $S_{\pm}$ such that $S_+ \oplus S_-$ is a nonsingular invariant subspace.

Proof: Since the nonsingular space V is the direct sum of the totally isotropic primary components $V_{\gamma}^{\pm is}$, it follows that, for every nonzero vector in V_{γ}^{is}, there exists another vector in V_{γ}^{-is} such that these two vectors are not orthogonal. If there is a vector ϕ_+ in V_{γ}^{is} such that $(\gamma - is1)^m\phi_+ \neq 0$, then there is a vector ϕ_- in V_{γ}^{-is} such that $\bar{\phi}_-(\gamma - is1)^m\phi_+ \neq 0$, and hence $(\gamma + is1)^m\phi_- \neq 0$. Let $p > 0$ be the smallest integer such that $(\gamma^2 + s^2 1)^p = 0$. If $(\gamma - is1)^{p-1}$ is zero on V_{γ}^{is}, then $(\gamma + is1)^{p-1}$ is zero on V_{γ}^{-is}, and hence $(\gamma^2 + s^2 1)^{p-1} = 0$, a contradiction. Hence there exist vectors $\phi_{\pm}$ in $V_{\gamma}^{\pm is}$ such that $\bar{\phi}_-(\gamma - is1)^{p-1}\phi_+ \neq 0$. The cyclic subspaces $S_{\pm}$ generated by $\phi_{\pm}$ are totally isotropic and primary, since $S_{\pm} \subset V_{\gamma}^{\pm is}$, and hence are indecomposable. Finally, one may verify that $\bar{\phi}_-(\gamma - is1)^{p-1}\phi_+ \neq 0$ implies that the radical of $S_+ \oplus S_-$ is zero. QED

Combining all these results, our main conclusion is that the primary components of Hermitian operators in nonsingular elementary Hermitian spaces are indecomposable.

We next study the metric signatures of the elementary subspaces.

Theorem: Elementary invariant subspaces of a Hermitian operator on a nonsingular Hermitian space are maximal-index nonsingular subspaces.

Proof: If γ is an elementary Hermitian operator on a nonsingular space V, then V is either itself indecomposable or the direct sum of two totally isotropic indecomposable subspaces. In the latter case it follows from the well-known hyperbolic enlargement theorem[38,44] that V is an even-dimensional maximal-index space. In the former case, we may subtract a real multiple of the identity from γ to obtain a nilpotent operator. If γ is nilpotent and if ϕ is a cyclic vector for V, then every other cyclic vector is of the form $p(\gamma)\phi$, where $p(\gamma)$ is a polynomial in γ whose constant term is nonzero. By constructing a suitable polynomial, it is possible to show that there exists a cyclic vector ϕ in V such that $\bar{\phi}\gamma^k\phi = 0$ for all k except for $k = n - 1$, where $n = \dim V$. We could multiply ϕ by a suitable complex factor to make $\bar{\phi}\gamma^{n-1}\phi = \pm 1$. If $m = [n/2]$ is the largest integer not exceeding $n/2$, then the vectors $\phi, \gamma\phi, \cdots, \gamma^{m-1}\phi$ span a totally isotropic subspace of V, and it follows that V has maximal index. QED

It is now a simple matter to construct orthonormal bases giving the exhibited canonical forms for elementary Hermitian operators. The choice of a canonical orthonormal basis is, of course, not unique, and questions of elegance influence the decision. Let γ be an elementary Hermitian operator acting on an n-dimensional nonsingular Hermitian space V. If $V = V_{\gamma}^{r}$, then $(\gamma - r1)^n = 0$, and there is a cyclic vector ϕ such that $\bar{\phi}(\gamma - r1)^k\phi = \epsilon\delta_{k,n-1}$, where $\epsilon = \pm 1$. In the even-dimensional case $n = 2m$, we obtain an orthonormal basis $\psi_{\pm 1}, \cdots, \psi_{\pm m}$ by writing

$$\psi_{\pm k} = [2^{m-k}(\gamma - r1)^{k-1} \pm 2^{-m+k-1}\epsilon(\gamma - r1)^{2m-k}]\phi.$$

In the odd-dimensional case $n = 2m + 1$, we put $\psi_0 = (\gamma - r1)^m\phi$, and

$$\psi_{\pm k} = [2^{m-k}(\gamma - r1)^{k-1} \pm 2^{-m+k-1}\epsilon(\gamma - r1)^{2m+1-k}]\phi.$$

If $V = V_{\gamma}^{r+is} \oplus V_{\gamma}^{r-is}$, there are vectors $\phi_{\pm}$ which are cyclic vectors for $V_{\gamma}^{r\pm is}$, respectively, such that $\bar{\phi}_-[\gamma - (r + is)1]^{k-1}\phi_+ = \delta_{k,m-1}$, where $\dim V = 2m$. In this case an orthonormal basis is given by

$$\psi_{\pm k} = 2^{m-k}[\gamma - (r + is)1]^{k-1}\phi_+ + 2^{k-m-1}[\gamma - (r - is)1]^{m-k}\phi_-.$$

* On leave of absence from the Nuclear Research Institute, Czechoslovak Academy of Sciences, Řež near Prague, Czechoslovakia. Present address: Department of Physics, University of Pittsburgh, Pittsburgh, Pa. 15213.

[1] J. A. Smorodinsky and N. J. Vilenkin, Zh. Eksp. Teor. Fiz. **46**, 1820 (1964) [Sov. Phys. JETP **19**, 1209 (1964)].

[2] M. Toller, Nuovo Cimento **53A**, 671; **54A**, 295 (1968) (contains an extensive bibliography).

[3] A. Sciarrino and M. Toller, J. Math. Phys. **8**, 1252 (1967).
[4] J. S. Lomont and H. E. Moses, J. Math. Phys. **5**, 294 (1964) [Erratum: *ibid.* **6**, 670 (1965)].
[5] I. Friš and P. Winternitz, Yad. Fiz. **1**, 889 (1965) [Sov. J. Nucl. Phys. **1**, 636 (1965)].
[6] I. Lukač, J. A. Smorodinsky, and P. Winternitz, Yad. Fiz. **7**, 192 (1968) [Sov. J. Nucl. Phys. **7**, 139 (1968)].
[7] G. M. Mubarakzjanov, Izv. Vysšhikh Učhebn. Zavedenii Mat. **1963**, No. 1 (32), 114; No. 3 (34), 99; No. 4 (35), 104 (1963) (Russian).
[8] I. M. Gel'fand, M. I. Graev, and N. J. Vilenkin, *Generalized Functions, Vol. 5: Integral Geometry and Representation Theory* (Academic, New York, 1966).
[9] N. J. Vilenkin, *Special Functions and the Theory of Group Representations*, Amer. Math. Soc. Math. Monograph Transl. No. 22 (American Mathematical Society, Providence, R.I., 1968) (Russian ed.: "Nauka," Moscow, 1965).
[10] H. Joos, Fortschr. Physik **10**, 65 (1962).
[11] J. F. Boyce, R. Delbourgo, A. Salam, and J. Strathdee, Trieste, Preprint IC/67/9, 1967.
[12] P. Winternitz, Rutherford Laboratory, Preprint RPP/T/3, 1969.
[13] M. Sheftel, J. A. Smorodinsky, and P. Winternitz, Yad. Fiz. **7**, 1325; **8**, 833 (1968) [Sov. J. Nucl. Phys. **7**, 785; **8**, 485 (1968)]; P. Pajas and P. Winternitz, Phys. Rev. D **1**, 1105 (1970).
[14] A. P. Balachandran and J. Nuyts, Phys. Rev. **172**, 1821 (1968); A. P. Balachandran, W. J. Meggs, J. Nuyts, and P. Ramond, *ibid.* **176**, 1700 (1968); J. M. Charap and B. M. Minton, J. Math. Phys. **10**, 1823 (1969).
[15] N. W. Macfadyen and P. Winternitz, J. Math. Phys. **12**, 281 (1971).
[16] W. Montgomery, L. O'Raifeartaigh, and P. Winternitz, Nucl. Phys. **B11**, 39 (1969).
[17] P. Winternitz, in *Lectures in Theoretical Physics*, edited by A. O. Barut (Gordon and Breach, New York, 1971), Vol. XIII.
[18] H. D. I. Abarbanel, M. L. Goldberger, and S. B. Treiman, Phys. Rev. Letters **22**, 500 (1969).
[19] T. Hirai, J. Math. Soc. Japan **22**, 134 (1970). The theory of unitary representations of unitary and special unitary groups with indefinite metric has been discussed in the recent mathematical physics literature. See M. I. Graev, Tr. Mosk. Mat. Obshchestva **7**, 335 (1958) [English translation: Amer. Math. Soc. Transl., Ser. 2, **66**, 1 (1968)]; P. Hillion, Nuovo Cimento **51A**, 674 (1967); J. Fischer and R. Rączka, Commun. Math. Phys. **3**, 233 (1966); **4**, 8 (1967). For the special case of $SU(2, 2)$, see T. Yao, J. Math. Phys. **8**, 1931 (1967); **9**, 1615 (1968).
[20] P. A. M. Dirac, Proc. Roy. Soc. London **A180**, 1 (1942); Commun. Dublin Inst. Advan. Studies **A1** (1943); W. Pauli, Jr., Rev. Mod. Phys. **15**, 175 (1943); S. N. Gupta, Proc. Phys. Soc. London **A63**, 681 (1950); **A64**, 850 (1951); Can. J. Phys. **35**, 961 (1957); K. Bleuler, Helv. Phys. Acta **23**, 567 (1950); K. Bleuler and W. Heitler, Progr. Theoret. Phys. (Kyoto) **5**, 600 (1950); F. J. Belinfante, Phys. Rev. **96**, 780 (1954). For further references, see also K. L. Nagy, *State Vector Spaces with Indefinite Metric in Quantum Field Theory* (Noordhoff, Groningen, 1969).
[21] W. Pauli, Jr., and F. Villars, Rev. Mod. Phys. **21**, 434 (1949); W. Heisenberg, *ibid.* **29**, 269 (1957); Nucl. Phys. **4**, 532 (1957).
[22] H. A. Kastrup, Phys. Rev. **140**, B183 (1965); **142**, 1060 (1966).
[23] T. D. Lee and G.-C. Wick, Nucl. Phys. **B9**, 209; **B10**, 1 (1969); R. E. Cutkosky, Phys. Rev. D **2**, 1386 (1970).
[24] N. Nakanishi, J. Math. Phys. **11**, 2970 (1970).
[25] It is known from examples that the Witt theorem, for example, may break down for infinite dimensions. See O. Hamara, Portugal Math. **27**, 15 (1968). The possibility of introducing a locally convex vector space topology in a general infinite-dimensional Hermitian space, subject to the requirement that the indefinite metric remain continuous, has been studied recently. See J. Jelínek and J. Virsik, Časopis Pěst. Mat. **91**, 18 (1966).
[26] R. L. Ingraham, Commun. Math. Phys. **6**, 17 (1967).
[27] E. P. Wigner, Ann. Math. **40**, 149 (1939).
[28] D. Finkelstein, Phys. Rev. **100**, 924 (1955); J. Segercrantz, Ann. Acad. Sci. Fennicae, Ser A VI, **233**, 1 (1967); R. Shaw, Quart. J. Math. Oxford **21**, 101 (1970).
[29] H. Zassenhaus, Can. Math. Bull. **1**, 31, 101, 183 (1958); G. E. Wall, J. Australian Math. Soc. **3**, 1 (1963).
[30] T. O. Philips and E. P. Wigner, in *Group Theory and Its Applications*, edited by E. M. Loebl (Academic, New York, 1968), pp. 631–76.
[31] A. Z. Petrov, *Einstein Spaces* (Pergamon, New York, 1969).
[32] T. Nôno, J. Math. Anal. Appl. **24**, 110 (1968).
[33] L. Gårding, Proc. Cambridge Phil. Soc. **41**, 49 (1945).
[34] L. Gross, J. Math. Phys. **5**, 687 (1964).
[35] H. A. Kastrup, Nucl. Phys. **58**, 561 (1964) (German); G. Mack and A. Salam, Ann. Phys. **53**, 174 (1969).
[36] H. A. Kastrup and G. Mack, Phys. Letters **22**, 331 (1966); K. Wilson, Phys. Rev. **179**, 1499 (1969); J. D. Bjorken and E. A. Paschos, *ibid.* D **1**, 3151 (1970); C. G. Callan, Jr., S. Coleman, and R. Jackiw, Ann. Phys. **59**, 42 (1970); D. G. Boulware, L. S. Brown, and R. D. Peccei, Phys. Rev. D **2**, 293 (1970), and many others.
[37] See, for example, A. O. Barut, in *Lectures in Theoretical Physics*, edited by A. O. Barut and W. E. Britten (Gordon and Breach, New York, 1968), Vol. XB, p. 377 (further references are given here).
[38] E. Artin, *Geometric Algebra* (Interscience, New York, 1957).
[39] L. Brickman and P. A. Fillmore, Can. J. Math. **19**, 810 (1967); S. Warner, *Modern Algebra* (Prentice-Hall, Englewood Cliffs, N.J., 1965), Vol. 2.
[40] See, for example, the Introduction to E. B. Dynkin, Usp. Mat. Nauk (N.S.) **2**, 59 (1947) [AMS Transl., Ser. I, **9**, 328 (1962)].
[41] N. Jacobson, *Lectures in Abstract Algebra* (Van Nostrand, Princeton, N.J., 1953); J. Dieudonné, *The Geometry of the Classical Groups* (Springer-Verlag, Berlin, 1963) (French).
[42] O. Veblen, Proc. Natl. Acad. Sci. (U.S.) **19**, 503 (1933).
[43] See also J. G. F. Belinfante, in Ref. 17.
[44] E. Witt, J. Reine Angew. Math. (Crelle) **176**, 31 (1937) [German] [further references are given by N. Jacobson, Ref. 41, 162].
[45] H. W. Turnbull, Proc. London Math. Soc. **39**, 232 (1935); J. Williamson, Duke Math. J. **3**, 715 (1937).
[46] B. Kurşunoğlu, Phys. Rev. D **1**, 1115 (1970); A. O. Barut, J. Math. Phys. **5**, 1652 (1964).

Remarks on a Class of Representations of Current Density Algebras

A. CHAKRABARTI
Centre de Physique Théorique de l'Ecole Polytechnique 17, rue Descartes, 75, Paris V, France

(Received 23 June 1970)

A class of representations of the current density algebras is considered by introducing a factorized form involving algebras containing the one satisfied by the integrated charges as a subalgebra.

Certain simple types of representations of the algebra

$$[V_0^i(\mathbf{x}), V_0^j(\mathbf{y})] = i\epsilon_{ijk}V_0^k(\mathbf{x})\delta(\mathbf{x}-\mathbf{y}), \quad x^0 = y^0 \tag{1}$$

of the isospin charge densities have been considered in Refs. 1 and 2. Our results will be closely related to those of Ref. 1.

The smeared operators

$$V^i(\varphi) = \int V_0^i(\mathbf{x})\varphi(\mathbf{x})d^3x \tag{2}$$

satisfy

$$[V^i(\varphi_1), V^j(\varphi_2)] = i\epsilon_{ijk}V^k(\varphi_1\varphi_2). \tag{3}$$

For $\varphi_1 = \varphi_2 = 1$, we have the usual isospin algebra $[\mathbf{V}(1) = \mathbf{I}]$. More generally, we see that $\mathbf{V}(\varphi)$ transforms as a vector operator under the action of $\mathbf{V}(1)$.

The factorized forms of the solutions for $\mathbf{V}(\varphi)$ found in Ref. 1 [considering the matrix elements of $\mathbf{V}(\varphi)$ acting on the states $|\mathrm{II}_3\rangle$] contain terms corresponding to the matrix elements of the generators of $O(4)$ (or E_3, related to O_4 through contraction), according to the domains of variation of some other parameters which appear.[3]

This fact suggests the following point of view, permitting considerable simplification and generalization (and also giving a clearer insight into the structure of the solutions).

Let us postulate the form

$$\mathbf{V}(\varphi) = a(\varphi)\mathbf{I} + b(\varphi)\mathbf{K}, \tag{4}$$

where

$$[I^i, I^j] = i\epsilon_{ijk}I^k, \quad [I^i, K^j] = i\epsilon_{ijk}K^k,$$
$$[K^i, K^j] = \zeta i\epsilon_{ijk}I^k. \tag{5}$$

For $\zeta = +1, -1, 0$, we obtain the algebras $O(4)$, $O(3, 1)$, and $E(3)$, respectively.

Substituting (4) in (3), we obtain, very simply and directly, the constraints

$$[a(\varphi_1), a(\varphi_2)] = [b(\varphi_1), b(\varphi_2)] = [a(\varphi_1), b(\varphi_2)] = 0, \tag{6}$$

$$a(\varphi_1)b(\varphi_2) + a(\varphi_2)b(\varphi_1) = b(\varphi_1\varphi_2), \tag{7}$$

$$a(\varphi_1)a(\varphi_2) + \zeta b(\varphi_1)b(\varphi_2) = a(\varphi_1\varphi_2), \tag{8}$$

with

$$a(1) = 1, \quad b(1) = 0. \tag{9}$$

The conditions (7), (8), and (9) correspond, respectively, to Eqs. (26), (27), and (28) of Ref. 1, where some explicit solutions are considered, along with possible physical interpretations.

Let us note that, in deriving the above relations, we did not even have to use the (known) matrix elements of the algebras $(\mathbf{I}, \mathbf{K})_\zeta$. *This is indicative of a much more general result.*

Let us consider some charge symmetry algebra,[4] say SU_2, SU_3, or something more complicated with the structure constants $c_{\alpha\beta\gamma}$, which we need not specify for the moment.

Let

$$V_\alpha(\varphi) = a(\varphi)A_\alpha + \sum_i b_i(\varphi)T_\alpha^i, \quad i = 1, 2, \cdots, n, \text{ say,} \tag{10}$$

where

$$[A_\alpha, A_\beta] = i\sum_\gamma c_{\alpha\beta\gamma}A_\gamma,$$
$$[A_\alpha, T_\beta^i] = i\sum_\gamma c_{\alpha\beta\gamma}\mu_i T^i, \tag{11}$$
$$[T_\alpha^i, T_\beta^j] = i\sum_\gamma c_{\alpha\beta\gamma}\zeta_{ij}A_\gamma.$$

(The parameters μ_i and ζ_{ij} are supposed to satisfy the usual hermiticity and other possible consistency restrictions.)

Substituting in

$$[V_\alpha(\varphi_1), V_\beta(\varphi_2)] = i\sum_\gamma c_{\alpha\beta\gamma}V_\gamma(\varphi_1\varphi_2), \tag{12}$$

we obtain (*independently of the* $c_{\alpha\beta\gamma}$ and as simply as before) the constraints

$$[a(\varphi_1), a(\varphi_2)] = [a(\varphi_1), b_i(\varphi_2)] = [b_i(\varphi_1), b_j(\varphi_2)] = 0 \tag{13}$$

and

$$a(\varphi_1)a(\varphi_2) + \sum_{ij}\zeta_{ij}b_i(\varphi_1)b_j(\varphi_2) = a(\varphi_1\varphi_2), \tag{14}$$

$$\mu_i(a(\varphi_1)b_i(\varphi_2) + a(\varphi_2)b_i(\varphi_1)) = b_i(\varphi_1\varphi_2), \tag{15}$$

with

$$a(1) = 1, \quad b_i(1) = 0. \tag{16}$$

(We have μ_i to emphasize that $\mu_i = 0$ eliminates T^i even for $\zeta_{ij} \neq 0$.)

If we consider the algebra satisfied by the Fourier transforms

$$V_\alpha(\mathbf{K}) = \int e^{-i\mathbf{K}\cdot\mathbf{x}} V_0(\mathbf{x}) d^3x, \tag{17}$$

namely,

$$[V_\alpha(\mathbf{K}_1), V_\beta(\mathbf{K}_2)] = i \sum_\gamma c_{\alpha\beta\gamma} V_\gamma(\mathbf{K}_1 + \mathbf{K}_2), \tag{18}$$

then by postulating

$$V_\alpha(\mathbf{K}) = a(\mathbf{K}) A_\alpha + \sum_i b_i(\mathbf{K}) T^i_\alpha, \quad i = 1, 2, \cdots, n, \tag{19}$$

the corresponding modifications of (13)–(16) are evident.

In this article we will make no attempt to obtain exhaustive solutions or to examine their possible significances. We will, however, try to illustrate, by taking a particularly simple case of (19), how on increasing n one can continue to find solutions for the a's and the b_i through simple modifications. The interest of increasing n (say from 1 to 2) is that the irreducible spaces of the states on which the $V(K)$'s act will now be characterized by the supplementary parameters corresponding to the larger algebra, giving more possibilities for eventual interpretations.

Let us consider the algebra

$$[I^i, I^j] = i\epsilon_{ijk} I^k, \quad [I^i, T^j_{(1,2)}] = i c_{ijk} T^k_{(1,2)},$$
$$[T^i_{(a)}, T^i_{(a)}] = \zeta_a i\epsilon_{ijk} I^k, \quad \zeta_{12} = \zeta_{21} = 0. \tag{20}$$

For $\zeta_1 = \zeta_2 = 0$, we will call this algebra $E_3(2)$. We have calculated the corresponding matrix elements in a rather different context,[5] where its connection with rigid rotators have been noted.

Let

$$\mathbf{V}(\mathbf{K}) = a(\mathbf{K})\mathbf{I} + b_1(\mathbf{K})\mathbf{T}_{(1)} + b_2(\mathbf{K})\mathbf{T}_{(2)}. \tag{21}$$

The required constraints are (apart from the usual commutativity)

$$a(\mathbf{K}_1)a(\mathbf{K}_2) + \sum_{i=1,2} \zeta_i b_i(\mathbf{K}_1) b_i(\mathbf{K}_2) = a(\mathbf{K}_1 + \mathbf{K}_2), \tag{22}$$

$$a(\mathbf{K}_1) b_i(\mathbf{K}_2) + a(\mathbf{K}_2) b_i(\mathbf{K}_1) = b_i(\mathbf{K}_1 + \mathbf{K}_2), \tag{23}$$

with

$$a(0) = 1, \quad b_i(0) = 0, \quad i = 1, 2.$$

The solutions corresponding to the restricted cases ($b_2 = 0$ or, corresponding to nonmixing of I spin, $b_1 = b_2 = 0$) need hardly be written separately. So let us note only the six following simple solutions:

$$\zeta_1 = \zeta_2 = 0: \quad a(\mathbf{K}) = e^{\lambda\cdot\mathbf{K}}, \quad b_i(\mathbf{K}) = \mu_i \cdot \mathbf{K} e^{\lambda\cdot\mathbf{K}}; \tag{24}$$

$$\zeta_1 = 1, \quad \zeta_2 = 0:$$
$$a(\mathbf{K}) = \cosh \boldsymbol{\lambda}\cdot\mathbf{K}, \quad b_1(\mathbf{K}) = \pm \sinh \boldsymbol{\lambda}\cdot\mathbf{K},$$
$$b_2(\mathbf{K}) = c_2 \sinh \boldsymbol{\lambda}\cdot\mathbf{K}; \tag{25}$$

$$\zeta_1 = -1, \quad \zeta_2 = 0:$$
$$a(\mathbf{K}) = \cos \boldsymbol{\lambda}\cdot\mathbf{K}, \quad b_1(\mathbf{K}) = \pm \sin \boldsymbol{\lambda}\cdot\mathbf{K},$$
$$b_2(\mathbf{K}) = c_2 \sin \boldsymbol{\lambda}\cdot\mathbf{K}; \tag{26}$$

$$\zeta_1 = \pm 1, \quad \zeta_2 = \mp 1:$$
$$a(\mathbf{K}) = e^{\lambda\cdot\mathbf{K}}, \quad b_1(\mathbf{K}) = \pm b_2(\mathbf{K}) = \mu\cdot\mathbf{K} e^{\lambda\cdot\mathbf{K}}; \tag{27}$$

$$\zeta_1 = +1, \quad \zeta_2 = +1:$$
$$a(\mathbf{K}) = \cosh \boldsymbol{\lambda}\cdot\mathbf{K},$$
$$b_1(\mathbf{K}) = \pm b_2(\mathbf{K}) = \pm \tfrac{1}{2} \sinh \boldsymbol{\lambda}\cdot\mathbf{K}; \tag{28}$$

$$\zeta = -1, \quad \zeta_2 = -1:$$
$$a(\mathbf{K}) = \cos \boldsymbol{\lambda}\cdot\mathbf{K},$$
$$b_1(\mathbf{K}) = \pm b_2(\mathbf{K}) = \pm \tfrac{1}{2} \sin \boldsymbol{\lambda}\cdot\mathbf{K}. \tag{29}$$

Such simple constructions as above lead to multiplicity-free representations as in Ref. 1. By introducing suitable tensor operators (and not only the generators) in the representations of the V's, one can hope to introduce a multiplicity structure if that is considered to be desirable.

We would like to make one final remark. Since, as emphasized, the conditions (6)–(9) or (22)–(23) are independent of the structure constants ($c_{\alpha\beta\gamma}$) of the charge algebra, the internal symmetry is apparently decoupled from space–time in our simple model. But it is also to be noted that, due to the presence of the a's and the b's, the states on which the current densities are supposed to act can be parametrized by using the larger algebras appearing in (10) and (11). This parametrization, in turn, depends essentially on the charge algebra we use as our starting point. In this sense the "local" nature of the density algebra may be said to interact nontrivially with the internal symmetry, even in our factorized model.

It would be of interest to examine what possible higher symmetries can be introduced in a physically significant fashion for the "intrinsic" factors of the charge densities, starting from a given symmetry group for the charges.

ACKNOWLEDGMENTS

It is a pleasure to thank C. Darzens for several interesting comments. The author has learned from Professor Y. Ne'eman (private communication) that conclusions similar to ours have since been reached, which are to be published in subsequent papers.[6,7]

[1] R. V. Mendes and Y. Ne'eman, J. Math. Phys. **11**, 3371 (1970).
[2] A. U. Klimyk, Inst. Theor. Phys. Kiev, preprint, "Representation of Current Algebra with Structural Constants of Rotation Group," 1969.
[3] These are well known. See, for example, A. Chakrabarti, J. Math. Phys. **9**, 2087 (1968), and A. Chakrabarti, M. Lévy-Nahas, and R. Sénéor, J. Math. Phys. **9**, 1274 (1968).
[4] Axial currents can be included, for example, by using the combinations $(V \pm A)$ or directly. If suitable smearing can be defined for the space components, the algebras involving them can be treated by introducing extra indices. The essential aspects of the technique stressed here will not change.
[5] A. Chakrabarti, "Remarks on Lightlike Continuous Spin and Spacelike Representations of the Poincaré Group," J. Math. Phys. (to be published). The matrix elements corresponding to the other values of ζ_1 and ζ_2 considered in (20)–(29) can be calculated using similar techniques.
[6] R. V. Mendes and Y. Ne'eman, Center for Particle Theory, Austin, Texas, Preprint CPT-49, 1970.
[7] A. Joseph, Tel-Aviv Report TAVP, 129–70, 1970.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Erratum: Half-Range Expansion Theorems in Studies of Polarized Light

[J. Math. Phys. **11**, 3416 (1970)]

E. E. BURNISTON AND C. E. SIEWERT
Departments of Mathematics and Nuclear Engineering, North Carolina State University
Raleigh, North Carolina 27607

(Received 4 February 1971)

This is a minor error (inconsequential to our proof of half-range completeness) regarding the form at infinity of the canonical solution $\mathbf{\Phi}_0(z)$. In general, Eq. (51) should read

$$\lim_{z\to\infty} \mathbf{\Phi}_0(z)\begin{bmatrix} z^{\kappa_1} & 0 \\ 0 & z^{\kappa_2} \end{bmatrix} = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix},$$

and thus Eq. (53) should read

$$\mathbf{X}^{-1}(z) \sim z\begin{bmatrix} \frac{1}{z} + \cdots & -\frac{a}{z} + \cdots \\ -\frac{b}{z} + \cdots & 1 + \cdots \end{bmatrix},$$

where a and b are constants.

In addition, a printing format change subsequent to the galley proofs has made Eq. (42) incorrect; it should read

$$\frac{|g(\mu) - 1|}{|\mu - 1|^{\alpha}} = \frac{\omega\mu\pi}{|\mu - 1|^{\alpha}\,|1 + \frac{1}{2}\omega\mu \ln [(1 - \mu)/(1 + \mu)] - \frac{1}{2}\omega\mu\pi i|}.$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Erratum: Bounds for Effective Bulk Modulus of Heterogeneous Materials

[J. Math. Phys. **10**, 2005 (1969)]

MELVIN N. MILLER
Geometric Data Corporation, Wayne, Pennsylvania 19087

(Received 11 February 1971)

Equations (2.23) and (2.24) should read as follows:

$$\frac{K^*}{(K_1K_2)^{\frac{1}{2}}} \leq \frac{1}{\alpha^{\frac{1}{2}}}\left(1 + \varphi(\alpha - 1) - \frac{\varphi(1 - \varphi)(\alpha - 1)^2}{\alpha - (\alpha - 1)\varphi + 2\gamma\{1 - \frac{4}{3}\varphi + \frac{1}{3}\beta(4\varphi - 1) + 3(\beta - 1)[G_1(1 - \varphi)^2 - G_2\varphi]^2\}}\right),$$

$$\frac{K^*}{(K_1K_2)^{\frac{1}{2}}} \geq \alpha^{\frac{1}{2}}\left(\varphi(1 - \alpha) + \alpha - \frac{\varphi(1 - \varphi)(\alpha - 1)^2}{1 + \varphi(\alpha - 1) + (3\alpha/8\beta\gamma)(3(\beta - 1)\{3[G_2\varphi^2 - G_1(1 - \varphi)^2] - \frac{4}{3}\varphi\} + 3(\beta - 1))}\right)^{-1}.$$

All figures and calculations based on these expressions in the article are correct as shown.

[1] R. V. Mendes and Y. Ne'eman, J. Math. Phys. **11**, 3371 (1970).
[2] A. U. Klimyk, Inst. Theor. Phys. Kiev, preprint, "Representation of Current Algebra with Structural Constants of Rotation Group," 1969.
[3] These are well known. See, for example, A. Chakrabarti, J. Math. Phys. **9**, 2087 (1968), and A. Chakrabarti, M. Lévy-Nahas, and R. Sénéor, J. Math. Phys. **9**, 1274 (1968).
[4] Axial currents can be included, for example, by using the combinations $(V \pm A)$ or directly. If suitable smearing can be defined for the space components, the algebras involving them can be treated by introducing extra indices. The essential aspects of the technique stressed here will not change.
[5] A. Chakrabarti, "Remarks on Lightlike Continuous Spin and Spacelike Representations of the Poincaré Group," J. Math. Phys. (to be published). The matrix elements corresponding to the other values of ζ_1 and ζ_2 considered in (20)–(29) can be calculated using similar techniques.
[6] R. V. Mendes and Y. Ne'eman, Center for Particle Theory, Austin, Texas, Preprint CPT-49, 1970.
[7] A. Joseph, Tel-Aviv Report TAVP, 129–70, 1970.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Erratum: Half-Range Expansion Theorems in Studies of Polarized Light

[J. Math. Phys. **11**, 3416 (1970)]

E. E. BURNISTON AND C. E. SIEWERT
Departments of Mathematics and Nuclear Engineering, North Carolina State University
Raleigh, North Carolina 27607

(Received 4 February 1971)

This is a minor error (inconsequential to our proof of half-range completeness) regarding the form at infinity of the canonical solution $\mathbf{\Phi}_0(z)$. In general, Eq. (51) should read

$$\lim_{z\to\infty} \mathbf{\Phi}_0(z)\begin{bmatrix} z^{\kappa_1} & 0 \\ 0 & z^{\kappa_2} \end{bmatrix} = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix},$$

and thus Eq. (53) should read

$$\mathbf{X}^{-1}(z) \sim z\begin{bmatrix} \frac{1}{z} + \cdots & -\frac{a}{z} + \cdots \\ -\frac{b}{z} + \cdots & 1 + \cdots \end{bmatrix},$$

where a and b are constants.

In addition, a printing format change subsequent to the galley proofs has made Eq. (42) incorrect; it should read

$$\frac{|g(\mu) - 1|}{|\mu - 1|^{\alpha}} = \frac{\omega\mu\pi}{|\mu - 1|^{\alpha}\,|1 + \frac{1}{2}\omega\mu \ln [(1 - \mu)/(1 + \mu)] - \frac{1}{2}\omega\mu\pi i|}.$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 6 JUNE 1971

Erratum: Bounds for Effective Bulk Modulus of Heterogeneous Materials

[J. Math. Phys. **10**, 2005 (1969)]

MELVIN N. MILLER
Geometric Data Corporation, Wayne, Pennsylvania 19087

(Received 11 February 1971)

Equations (2.23) and (2.24) should read as follows:

$$\frac{K^*}{(K_1K_2)^{\frac{1}{2}}} \leq \frac{1}{\alpha^{\frac{1}{2}}}\left(1 + \varphi(\alpha - 1) - \frac{\varphi(1 - \varphi)(\alpha - 1)^2}{\alpha - (\alpha - 1)\varphi + 2\gamma\{1 - \frac{4}{3}\varphi + \frac{1}{3}\beta(4\varphi - 1) + 3(\beta - 1)[G_1(1 - \varphi)^2 - G_2\varphi]^2\}}\right),$$

$$\frac{K^*}{(K_1K_2)^{\frac{1}{2}}} \geq \alpha^{\frac{1}{2}}\left(\varphi(1 - \alpha) + \alpha - \frac{\varphi(1 - \varphi)(\alpha - 1)^2}{1 + \varphi(\alpha - 1) + (3\alpha/8\beta\gamma)(3(\beta - 1)\{3[G_2\varphi^2 - G_1(1 - \varphi)^2] - \frac{4}{3}\varphi\} + 3(\beta - 1))}\right)^{-1}.$$

All figures and calculations based on these expressions in the article are correct as shown.