Asymptotic Theory of Wave Propagation in Spatial and Temporal Dispersive Inhomogeneous Media

KENNETH C. CHEN
New York University, Courant Institute of Mathematical Sciences
New York, New York 10012

(Received 8 June 1970)

The recently developed asymptotic theory of wave propagation is extended to slightly inhomogeneous and slowly varying anisotropic media which exhibit both spatial and temporal dispersion. A particular form of the constitutive relation is first introduced. Asymptotic solutions are then obtained by assuming a series solution "ansatz" into Maxwell's equations and the constitutive relation. The eikonal equation and the transport equation are obtained, by a procedure similar to that of Lewis, for lossless media, in which a Hermitian operator is involved. The modified transport equations obtained from other forms of the constitutive relation are given. They are interpreted as generalized Poynting theorems for appropriate physical situations. Finally, as a by-product of this work the space–time conductivity tensor of anisotropic plasmas $\boldsymbol{\sigma}(\mathbf{r}, t)$, which is the 4-dimensional Fourier–Laplace transform of $\tilde{\boldsymbol{\sigma}}(\mathbf{k}, \omega)$, is found to be very simple.

1. INTRODUCTION

Recently the asymptotic method has been developed for a large class of problems of wave propagation.[1] In particular, Lewis applied the method to obtain the time-dependent solutions of dispersive hyperbolic partial differential equations[2] and an integro-differential equation describing electromagnetic wave propagation in temporal-dispersive media.[3] He introduced formally a large parameter into the asymptotic solution, and expanded the asymptotic solution in terms of inverse power of this large parameter. He and his co-workers then applied this method to various problems of wave propagation involving time-dependent solutions.[4–7]

In this paper, the recently developed asymptotic theory is extended to the problem of wave propagation in spatial- and temporal-dispersive slightly inhomogeneous and slowly varying media. Temporal dispersion occurs when the characteristic frequencies of the medium, e.g., the resonance absorption frequency of molecules, the plasma frequency and the cyclotron frequency of plasmas, lie within the frequency range of the exciting sources. Different frequency components of the signal are absorbed and reradiated in different amounts, with the result that different frequency components propagate with different velocities. The waveform of the signal is distorted. On the other hand, if the characteristic length of the medium, e.g., molecular dimensions, lattice constants, Debye radius, is comparable to the wavelength of the wave we are studying, then spatial dispersion occurs. Weak spatial dispersion happens, for example, in birefringence, longitudinal waves, and Cerenkov radiation.[8,9] In some more complicated media, e.g., plasmas under certain conditions, strong spatial dispersion occurs.

In Sec. 2 an appropriate form of the constitutive relation for slightly inhomogeneous and slowly varying anisotropic media with both spatial and temporal dispersion is first introduced. The particular form was obtained through the study of the problem of electromagnetic wave propagation in slightly inhomogeneous and slowly varying anisotropic warm plasmas. In such a medium, the constitutive relation is obtained in Appendix A by solving the linearized Boltzmann equation with the aid of the method of characteristics. As a by-product of such a study, it is found that, for homogeneous, time-invariant anisotropic warm plasmas, the space–time-conductivity tensor $\tilde{\boldsymbol{\sigma}}(\mathbf{r}, t)$, which is the 4-dimensional Fourier–Laplace transform of $\tilde{\boldsymbol{\sigma}}(\mathbf{k}, \omega)$, is very simple.

The asymptotic parameter is then introduced into the Maxwell's equations and the constitutive relation, to obtain hyperbolic equation of convolution type with a large parameter. Then, by inserting an appropriate expression involving a "phase function" and an infinite series of "amplitude functions" into this hyperbolic equation, we derive two asymptotic governing equations: One relates the phase function and the first term of the series of amplitude functions. The other involves terms up to the second. The third-order term is discussed in Appendix B to study the limits of applicability of the asymptotic solution. As an example, we show that these general asymptotic governing equations reduce to the result obtained by Bloomberg[10] for longitudinal waves in warm plasmas.

By restricting ourselves to lossless media where the operator is Hermitian in Sec. 3, we proceed as follows: From the first of the asymptotic governing equations, we derive a first-order partial differential equation for the phase function, the "dispersion relation," which is solved by the method of characteristics. The characteristic curves are determined by integrating the characteristic equations or ray equations. The phase function is readily obtained from the eikonal equation. The transport equation for amplitude function is derived from the second asymptotic governing equation. Finally the initial value problem for the phase function and the amplitude function are formulated. The results obtained here are an extension of Lewis's treatment of wave propagation in temporal-dispersive media[3] to the more complicated spatial- and temporal-dispersive media. Let us also note that Weinberg[11] considered the case of wave propagation describable by a general differential operator. The study of hyperbolic equations of convolution type can be considered as an extension of his result.

In Sec. 4, transport equations derived from two other forms of the constitutive relation and Maxwell's equations are obtained. One form of the constitutive relation gives rise to a transport equation which asymptotically satisfies energy conservation of the wave. The transport equations obtained are interpreted as generalized Poynting theorems for appropriate physical situations.

Finally, in order to use the result here, initial conditions for the ray equation, eikonal equation, and transport equation are required. Some conditions can be obtained directly from the initial data. Others require the asymptotic solution of a "canonical problem," which is the asymptotic evaluation of the exact solution for the corresponding homogeneous media. Such solutions are ample in literature. We here refer to Lewis's work.[3]

2. ASYMPTOTIC GOVERNING EQUATIONS

In this section we derive asymptotic governing equations for electromagnetic wave propagation in spatial- and temporal-dispersive inhomogeneous and slowly varying media, which are described by the equations

$$\left(\delta_{ij}\frac{\partial^2}{\partial x_k \partial x_k} - \frac{\partial^2}{\partial x_i \partial x_j}\right)E_j - \frac{1}{\epsilon_0 c^2}\frac{\partial^2 D_i}{\partial t^2} = 0. \quad (2.1)$$

Summation notation is used throughout this work. The summation over repeated indices is from 1 to 3.

The constitutive relation for inhomogeneous and slowly varying spatial- and temporal-dispersive media is proposed to be

$$D_i(\mathbf{r}, t) = \frac{1}{(2\pi)^4}\int_0^t dt' \int_{-\infty}^{\infty} d^3\mathbf{r}' \int_{-j\sigma-\infty}^{-j\sigma+\infty} d\omega \times \int_{-\infty}^{\infty} d^3\mathbf{k}\tilde{\epsilon}_{ij}(\mathbf{r}-\mathbf{r}', t-t'; \mathbf{k}, \omega) \times \exp\left[j(\omega t' - \mathbf{k}\cdot\mathbf{r}')\right]E_j(\mathbf{r}-\mathbf{r}', t-t'). \quad (2.2)$$

Here we have assumed $E(\mathbf{r}, t) = 0$ for $t < 0$. For homogeneous and time-invariant media the dielectric tensor $\tilde{\epsilon}_{ij}(\mathbf{r}, t; \mathbf{k}, \omega)$ reduces to $\tilde{\epsilon}_{ij}(\mathbf{k}, \omega)$, which is the conventionally used time-harmonic dielectric tensor. The modification for inhomogeneous and slowly varying media is obtained from the study of electromagnetic wave propagation in anisotropic, inhomogeneous warm plasmas. In that particular problem, the constitutive relation is obtained in Appendix A by solving the linearized Boltzmann equation with the help of the method of characteristics. Other forms of the constitutive relation will be discussed later in Sec. 4.

We proceed to derive asymptotic governing equations formally. The limits of the applicability of the asymptotic solution is discussed in Appendix B. In order to derive asymptotic governing equations for (2.1) and (2.2), let us note that the asymptotic solution for large $|\mathbf{r}|$ or t of the corresponding homogeneous and time-invariant media can be obtained by the asymptotic evaluation of the exact solution in the form of an inverse Fourier–Laplace integral.[12] The asymptotic solution is of the form

$$\mathbf{E}(\mathbf{r}, t) = \exp\left[-jS(\mathbf{r}, t)\right]\sum \frac{\mathbf{A}^n(\mathbf{r}, t)}{(j\lambda)^n}, \quad (2.3)$$

with

$$\lambda = O[\max(|\mathbf{r}|, ct)].$$

Equation (2.3) will be used as the "ansatz" of asymptotic solutions for inhomogeneous and slowly varying media. Following Lewis,[3] we make the following transformation in (2.1), (2.2), and (2.3) by $\mathbf{r}\to\lambda\mathbf{r}$ and $t\to\lambda t$. The effect of this is to change the units. Equation (2.1) remains the same, while (2.2) becomes

$$D_i(\lambda\mathbf{r}, \lambda t) = \frac{1}{(2\pi)^4}\int_0^{\lambda t} dt' \int_{-\infty}^{\infty} d^3\mathbf{r}' \int_{-j\sigma-\infty}^{-j\sigma+\infty} d\omega \times \int_{-\infty}^{\infty} d^3\mathbf{k}\tilde{\epsilon}_{ij}(\lambda\mathbf{r}-\mathbf{r}', \lambda t-t'; \mathbf{k}, \omega) \times \exp\left[j(\omega t' - \mathbf{k}\cdot\mathbf{r}')\right]E_j(\lambda\mathbf{r}-\mathbf{r}', \lambda t-t'). \quad (2.4)$$

The ansatz (2.3) is now of the form

$$\mathbf{E}(\lambda\mathbf{r}, \lambda t) = \exp\left[-jS(\lambda\mathbf{r}, \lambda t)\right]\sum \frac{\mathbf{A}^n(\lambda\mathbf{r}, \lambda t)}{(j\lambda)^n}. \quad (2.5)$$

Equation (2.4) can be written as

$$D_i(\lambda\mathbf{r}, \lambda t) = \int_0^{\lambda t} dt' \int_{-\infty}^{\infty} d^3\mathbf{r}'\epsilon_{ij}(\lambda\mathbf{r}-\mathbf{r}', \lambda t - t'; \mathbf{r}', t') \times E_j(\lambda\mathbf{r}-\mathbf{r}', \lambda t - t') \quad (2.6)$$

and

$$\epsilon_{ij}(\lambda\mathbf{r}, \lambda t; \mathbf{r}', t') = \frac{1}{(2\pi)^4}\int_{-j\sigma-\infty}^{-j\sigma+\infty} d\omega \int_{-\infty}^{\infty} d^3\mathbf{k}\tilde{\epsilon}_{ij}(\lambda\mathbf{r}, \lambda t; \mathbf{k}, \omega) \times \exp[j(\omega t' - \mathbf{k}\cdot\mathbf{r}')], \quad t' > 0,$$
$$= 0, \quad t' < 0. \tag{2.7}$$

Also, the inverse transform for (2.7) is given by

$$\tilde{\epsilon}_{ij}(\lambda\mathbf{r}, \lambda t; \mathbf{k}, \omega) = \int_0^{\infty} dt' \int_{-\infty}^{\infty} d^3\mathbf{r}'\epsilon_{ij}(\lambda\mathbf{r}, \lambda t; \mathbf{r}', t') \times \exp[-j(\omega t' - \mathbf{k}\cdot\mathbf{r}')]. \tag{2.8}$$

Substitution of (2.5) into (2.6) and the expansion in λ yield

$$\begin{aligned} D_i(\lambda\mathbf{r}, \lambda t) &= \int_0^{\lambda t} dt' \int_{-\infty}^{\infty} d^3\mathbf{r}'\{\epsilon_{ij}(\lambda\mathbf{r}, \lambda t; \mathbf{r}', t') \\ &\quad - \lambda^{-1}[t'(\epsilon_{ij})_t + x_\nu'(\epsilon_{ij})_{x_\nu}] + \cdots\} \\ &\quad \times \exp\{-j[S(\lambda\mathbf{r}, \lambda t) - \lambda^{-1}S_t t' - \lambda^{-1}S_{x_\nu}x_\nu']\} \\ &\quad \times [1 + (2j\lambda)^{-2}(S_{tt}t'^2 + 2S_{tx_\nu}t'x_\nu' + S_{x_\mu x_\nu}x_\mu' x_\nu') + \cdots] \\ &\quad \times \Big(\sum_{n=0}^{\infty}\frac{1}{(j\lambda)^n}\{A_j^n(\lambda\mathbf{r}, \lambda t) - \lambda^{-1}[t'(A_j^n)_t + x_\nu'(A_j^n)_{x_\nu}] \\ &\quad + (2\lambda)^{-2}[t'^2(A_j^n)_{tt} + 2t'x_\nu'(A_j^n)_{tx_\nu} \\ &\quad + x_\mu' x_\nu'(A_j^n)_{x_\mu x_\nu}] + \cdots\}\Big). \end{aligned} \tag{2.9}$$

Introduce the local wave vector $\mathbf{k}$ and frequency ω

$$k_\nu = \lambda^{-1}S_{x_\nu}(\lambda\mathbf{r}, \lambda t), \quad \omega = -\lambda^{-1}S_t(\lambda\mathbf{r}, \lambda t), \tag{2.10}$$

and the relationship

$$\frac{\partial\tilde{\epsilon}_{ij}}{\partial k_\nu}(\lambda\mathbf{r}, \lambda t; \mathbf{k}, \omega) = \int_0^{\infty} dt' \int_{-\infty}^{\infty} d^3\mathbf{r}'(jx_\nu')\epsilon_{ij}(\lambda\mathbf{r}, \lambda t; \mathbf{r}', t') \times \exp[-j(\omega t' - \mathbf{k}\cdot\mathbf{r}')], \tag{2.11}$$

$$\frac{\partial\tilde{\epsilon}_{ij}}{\partial \omega}(\lambda\mathbf{r}, \lambda t; \mathbf{k}, \omega) = \int_0^{\infty} dt' \int_{-\infty}^{\infty} d^3\mathbf{r}'(-jt')\epsilon_{ij}(\lambda\mathbf{r}, \lambda t; \mathbf{r}', t') \times \exp[-j(\omega t' - \mathbf{k}\cdot\mathbf{r}')]. \tag{2.12}$$

With the aid of (2.8), (2.10), (2.11), and (2.12), (2.9) becomes

$$\begin{aligned} D_i(\lambda\mathbf{r}, \lambda t) &= \exp[-jS(\lambda\mathbf{r}, \lambda t)] \\ &\quad \times \Big\{\tilde{\epsilon}_{ij}(\lambda\mathbf{r}, \lambda t; \mathbf{k}, \omega)A_j^0 + \frac{1}{j\lambda}\Big[\tfrac{1}{2}\Big(-\omega_t(\tilde{\epsilon}_{ij})_{\omega\omega} \\ &\quad + 2(k_\nu)_t(\tilde{\epsilon}_{ij})_{\omega k_\nu} + 2\sum_{\mu>\nu}(k_\mu)_{x_\nu}(\tilde{\epsilon}_{ij})_{k_\mu k_\nu} \\ &\quad + (k_\nu)_{x_\nu}(\tilde{\epsilon}_{ij})_{k_\nu k_\nu}\Big)A_j^0 \\ &\quad + (\tilde{\epsilon}_{ij})_\omega(A_j^0)_t - (\tilde{\epsilon}_{ij})_{k_\nu}(A_j^0)_{x_\nu} \\ &\quad + \tilde{\epsilon}_{ij}A_j^1 + [(\tilde{\epsilon}_{ij})_{t\omega} - (\tilde{\epsilon}_{ij})_{x_\nu k_\nu}]A_j^0\Big] + \cdots\Big\}. \end{aligned} \tag{2.13}$$

Inserting (2.5) and (2.13) into (2.1) and equating the coefficients of $(j\lambda)^{-n}$, one arrives at the following:

$$[\delta_{ij}k_\nu^2 - k_ik_j - \omega^2c^{-2}\tilde{\epsilon}_{ij}(\lambda\mathbf{r}, \lambda t; \mathbf{k}, \omega)][A_j^0(\lambda\mathbf{r}, \lambda t)] = 0, \tag{2.14}$$

$$\begin{aligned} &[\delta_{ij}k_\nu^2 - k_ik_j - \omega^2c^{-2}\tilde{\epsilon}_{ij}(\lambda\mathbf{r}, \lambda t; \mathbf{k}, \omega)][A_j^1(\lambda\mathbf{r}, \lambda t)] \\ &+ \left(\delta_{ij}\frac{\partial k_\nu}{\partial x_\nu} - \frac{\partial k_i}{\partial x_j}\right)(A_j^0) - c^{-2}[\omega_t\tilde{\epsilon}_{ij}A_j^0 + 2\omega(\tilde{\epsilon}_{ij})_tA_j^0 \\ &+ 2\omega\tilde{\epsilon}_{ij}(A_j^0)_t + 2\omega(k_\nu)_t(\tilde{\epsilon}_{ij})_{k_\nu}A_j^0 + 2\omega\omega_t(\tilde{\epsilon}_{ij})_\omega A_j^0] \\ &- \tfrac{1}{2}\omega^2c^{-2}\Big(\omega_t(\tilde{\epsilon}_{ij})_{\omega\omega} + 2(k_\nu)_t(\tilde{\epsilon}_{ij})_{\omega k_\nu} + 2\sum_{\substack{\mu,\nu \\ \mu>\nu}}(k_\mu)_{x_\nu}(\tilde{\epsilon}_{ij})_{k_\mu k_\nu} \\ &+ (k_\nu)_{x_\nu}(\tilde{\epsilon}_{ij})_{k_\nu k_\nu}\Big)A_j^0 - \omega^2c^{-2}[(\tilde{\epsilon}_{ij})_\omega(A_j^0)_t - (\tilde{\epsilon}_{ij})_{k_\nu}(A_j^0)_{x_\nu}] \\ &- \omega^2c^{-2}[(\tilde{\epsilon}_{ij})_{t\omega} - (\tilde{\epsilon}_{ij})_{k_\nu x_\nu}]A_j^0 = 0. \end{aligned} \tag{2.15}$$

Equations (2.14) and (2.15) are then written in a compact form as follows:

$$[L_{ij}][A_j^0(\lambda\mathbf{r}, \lambda t)] = 0, \tag{2.16}$$

$$\begin{aligned} &[L_{ij}][A_j^1(\lambda\mathbf{r}, \lambda t)] + \frac{1}{2}\Big[\frac{\delta}{\delta t}\Big(\frac{\partial L_{ij}}{\partial\omega}\Big) - \frac{\delta}{\delta x_\nu}\Big(\frac{\partial L_{ij}}{\partial k_\nu}\Big) \\ &+ \frac{\partial}{\partial t}\Big(\frac{\partial L_{ij}}{\partial\omega}\Big) - \frac{\partial}{\partial x_\nu}\Big(\frac{\partial L_{ij}}{\partial k_\nu}\Big)\Big][A_j^0(\lambda\mathbf{r}, \lambda t)] \\ &+ \frac{\partial L_{ij}}{\partial\omega}\frac{\partial A_j^0}{\partial t} - \frac{\partial L_{ij}}{\partial k_\nu}\frac{\partial A_j^0}{\partial x_\nu} = 0, \end{aligned} \tag{2.17}$$

where

$$L_{ij} = \delta_{ij}k_\mu^2 - k_ik_j - \omega^2c^{-2}\tilde{\epsilon}_{ij}(\lambda\mathbf{r}, \lambda t; \mathbf{k}, \omega), \tag{2.18}$$

$$\frac{\delta}{\delta t} = \frac{\partial}{\partial t} + \omega_t\frac{\partial}{\partial\omega} + (k_\mu)_t\frac{\partial}{\partial k_\mu}, \tag{2.19}$$

$$\frac{\delta}{\delta x_\nu} = \frac{\partial}{\partial x_\nu} + \omega_{x_\nu}\frac{\partial}{\partial\omega} + (k_\mu)_{x_\nu}\frac{\partial}{\partial k_\mu}, \tag{2.20}$$

$$k_\mu = \lambda^{-1}S_{x_\mu}(\lambda\mathbf{r}, \lambda t), \quad \omega = -\lambda^{-1}S_t(\lambda\mathbf{r}, \lambda t). \tag{2.10}$$

After having obtained asymptotic governing equations (2.16)–(2.20), let us remove the large parameter λ from these governing equations as follows: Introduce $\lambda\mathbf{r} \to \mathbf{r}$ and $\lambda t \to t$, so that $\mathbf{r}$ and t are measured with the original unit. The asymptotic governing equations are then given by

$$[L_{ij}][A_j^0(\mathbf{r}, t)] = 0, \tag{2.21}$$

$$\begin{aligned} &\lambda^{-1}[L_{ij}][A_j^1(\mathbf{r}, t)] + \frac{1}{2}\Big[\frac{\delta}{\delta t}\Big(\frac{\partial L_{ij}}{\partial\omega}\Big) - \frac{\delta}{\delta x_\nu}\Big(\frac{\partial L_{ij}}{\partial k_\nu}\Big) \\ &+ \frac{\partial}{\partial t}\Big(\frac{\partial L_{ij}}{\partial\omega}\Big) - \frac{\partial}{\partial x_\nu}\Big(\frac{\partial L_{ij}}{\partial k_\nu}\Big)\Big][A_j^0(\mathbf{r}, t)] \\ &+ \frac{\partial L_{ij}}{\partial\omega}\frac{\partial A_j^0}{\partial t} - \frac{\partial L_{ij}}{\partial k_\nu}\frac{\partial A_j^0}{\partial x_\nu} = 0, \end{aligned} \tag{2.22}$$

with

$$L_{ij} = \delta_{ij}k_\mu^2 - k_i k_j - \omega^2 c^{-2}\tilde{\epsilon}_{ij}(\mathbf{r}, t; \mathbf{k}, \omega), \quad (2.23)$$

$$k_\mu = S_{x_\mu}(\mathbf{r}, t), \quad \omega = -S_t(\mathbf{r}, t). \quad (2.24)$$

Equations (2.19) and (2.20) remain unchanged.

We have derived asymptotic governing equations for the specific problem of electromagnetic wave propagation in spatial- and temporal-dispersive, slowly varying, and inhomogeneous media. However, the asymptotic governing equations obtained are quite general. In particular, let us show that (2.21) and (2.22) are valid for the problem of transient longitudinal wave propagation in inhomogeneous, slowly varying warm plasmas. This problem has been treated by Bloomberg[10] based on solving the linearized Vlasov equation with the aid of the method of characteristics. To obtain asymptotic governing equations for this simplified example from (2.21) and (2.22), let us note that the problem of longitudinal wave propagation is a scalar problem governed by the approximate dispersion equation

$$L(x, t) = 1 - \frac{\omega_p^2(x, t)}{\omega^2} - \frac{3k_x^2}{\omega^2}\langle V(x, t)^2\rangle, \quad (2.25)$$

with ω_p the plasma frequency

$$\omega_p(x, t)^2 = N(x, t)e^2/m\epsilon_0 \quad (2.26)$$

and $\langle V(x, t)^2\rangle$ the mean square velocity of electrons

$$\langle V(x, t)^2\rangle = \kappa T_e(x, t)/4m, \quad (2.27)$$

where $N(x, t)$ is the density of electrons, κ is Boltzmann's constant, $T_e(x, t)$ is the electronic temperature of plasmas, and m is the mass of the electron.

By using (2.25), (2.21), and (2.22), one obtains the following equations:

$$L(x, t)A^0(x, t) = 0, \quad (2.28)$$

$$\lambda^{-1}\frac{\omega}{2}L(x, t)A^1(x, t) + \frac{\partial A^0}{\partial t} + U\frac{\partial A^0}{\partial x} + \left[\frac{2\omega_p}{\omega^2}\frac{\partial \omega_p}{\partial t} - \frac{3}{2}\frac{1}{\omega}\frac{\partial \omega}{\partial t} - \frac{2U}{\omega}\frac{\partial \omega}{\partial x} + \frac{1}{2}\frac{U}{k_x}\frac{\partial k_x}{\partial x} + \frac{U}{\langle V^2\rangle}\frac{\partial\langle V^2\rangle}{\partial x} + \frac{1}{3}\left(\frac{U}{\langle V^2\rangle}\right)^2\frac{\partial\langle V^2\rangle}{\partial t}\right]A^0 = 0, \quad (2.29)$$

where

$$U = \frac{\partial \omega}{\partial k_x} = \frac{3k_x}{\omega}\langle V^2\rangle. \quad (2.30)$$

For $\omega \simeq \omega_p$, (2.29) reduces to that obtained by Bloomberg except that there are two extra terms

$$\frac{1}{3}\left(\frac{U}{\langle V^2\rangle}\right)^2\frac{\partial\langle V^2\rangle}{\partial t}A^0 \quad \text{and} \quad \lambda^{-1}\frac{\omega}{2}L(x, t)A^1.$$

The first term was ignored in Bloomberg, and the second term cannot be obtained unless a series form solution is assumed. Bloomberg did not solve (2.28) and (2.29), while the solution in Sec. 3 for the general case can be easily applied to this special case.

3. ASYMPTOTIC SOLUTIONS OF A GENERAL HERMITIAN SYSTEM OF EQUATIONS

In order to have a nontrivial solution $\mathbf{A}^0$ for (2.21), it is necessary that

$$L(\mathbf{r}, t; \mathbf{k}, \omega) = \det L_{ij} = 0, \quad k_\nu = \frac{\partial S}{\partial x_\nu}, \quad \omega = -\frac{\partial S}{\partial t}. \quad (3.1)$$

Let us assume that L_{ij} is Hermitian, i.e.,

$$L_{ij}^*(\mathbf{r}, t; \mathbf{k}, \omega) = L_{ji}(\mathbf{r}, t; \mathbf{k}, \omega). \quad (3.2)$$

It immediately follows that, for some $\mathbf{k}$, ω, $\mathbf{r}$, and t, (3.1) defines a functional relation, which is called the dispersion relation. Assuming nondegeneracy for matrix L_{ij}, one writes $\bar{\mathbf{R}}(\mathbf{r}, t; \mathbf{k}, \omega)$ as the null eigenvector:

$$L_{ij}\bar{R}_j = 0.$$

For brevity of formulas we choose the normalization constant such that

$$\bar{R}_i^*\frac{\partial L}{\partial \omega}\bar{R}_j = 1, \quad (3.3)$$

The eigenvector $\bar{R}$ characterizes the polarization of the wave.

The dispersion relation (3.1) is a first-order nonlinear partial differential equation for the phase function $S(\mathbf{r}, t)$ and can be solved by the Hamilton–Jacobi theory. One thus introduces the characteristic equations or ray equations

$$\frac{dx_\nu}{d\tau} = \frac{\partial L}{\partial k_\nu}, \quad \frac{dt}{d\tau} = -\frac{\partial L}{\partial \omega}, \quad \frac{dk_\nu}{d\tau} = -\frac{\partial L}{\partial x_\nu}, \quad \frac{d\omega}{d\tau} = \frac{\partial L}{\partial t}, \quad (3.4)$$

where τ is the parameter along the ray. On using the parameter t to replace τ, one obtains

$$\frac{dx_\nu}{dt} = -\frac{\partial L}{\partial k_\nu}\Big/\frac{\partial L}{\partial \omega} = \frac{\partial \omega}{\partial k_\nu} = V_{g_\nu},$$
$$\frac{dk_\nu}{dt} = \frac{\partial L}{\partial x_\nu}\Big/\frac{\partial L}{\partial \omega} = -\frac{\partial \omega}{\partial x_\nu}, \quad (3.5)$$
$$\frac{d\omega}{dt} = -\frac{\partial L}{\partial t}\Big/\frac{\partial L}{\partial \omega} = \frac{\partial \omega}{\partial t},$$

where $\mathbf{V}_g$ is the group velocity. The phase function satisfies the eikonal equation

$$\dot{S} = \frac{dS}{dt} = \frac{\partial S}{\partial t} + \frac{\partial S}{\partial x_\nu}\dot{x}_\nu = -\omega + k_\nu V_{g\nu} = -\omega - k_\nu \frac{\partial L}{\partial k_\nu}\Big/\frac{\partial L}{\partial \omega}. \quad (3.6)$$

Note that (3.5) and (3.6) can be solved numerically. They have already been given by Weinberg[11] for the case of a general differential operator $L(\mathbf{r}, -j\nabla)$ for the time-harmonic problem. The spatial- and temporal-dispersive, inhomogeneous, and slowly varying medium treated here is an extension of his. Without spatial dispersion, (3.6) and (3.5) reduce to the ray equation given by Lewis.[3]

Before proceeding to derive the transport equation, we give an alternative formula for the group velocity. Taking the scalar product of $\bar{R}_i^*$ and $L_{ij}\bar{R}_j$, one obtains

$$\bar{R}_i^* L_{ij}\bar{R}_j = 0. \quad (3.7)$$

Differentiation of (3.20) leads to the following:

$$\left(\bar{R}_i^*\frac{\partial L_{ij}}{\partial\omega}\bar{R}_j\right)\frac{\partial\omega}{\partial k_\nu} + \bar{R}_i^*\frac{\partial L_{ij}}{\partial k_\nu}\bar{R}_j = 0, \quad (3.8)$$

$$V_{g\nu} = \frac{\partial\omega}{\partial k_\nu} = -\bar{R}_i^*\frac{\partial L_{ij}}{\partial k_\nu}\bar{R}_j\Big/\bar{R}_i^*\frac{\partial L_{ij}}{\partial\omega}\bar{R}_j = -\bar{R}_i^*\frac{\partial L_{ij}}{\partial k_\nu}\bar{R}_j. \quad (3.9)$$

With the aid of (3.9), we are ready to derive the transport equation.

Again, assuming no degeneracy, we let

$$A^0 = \sigma\bar{R}. \quad (3.10)$$

First, taking the scalar product of $\bar{\mathbf{R}}$ and (2.22) leads to

$$\frac{\bar{R}_i^*}{2}\left[\frac{\delta}{\delta t}\left(\frac{\partial L_{ij}}{\partial\omega}\right) - \frac{\delta}{\delta x_\nu}\left(\frac{\partial L_{ij}}{\partial k_\nu}\right)\right]A_j^0 + \frac{\bar{R}_i^*}{2}\left(\frac{\partial^2 L_{ij}}{\partial t\partial\omega} - \frac{\partial^2 L_{ij}}{\partial x_\nu\partial k_\nu}\right)A_j^0 + \bar{R}_i^*\left(\frac{\partial L_{ij}}{\partial\omega}\frac{\partial A_j^0}{\partial t} - \frac{\partial L_{ij}}{\partial k_\nu}\frac{\partial A_j^0}{\partial x_\nu}\right) = 0, \quad (3.11)$$

where use has been made of the fact that

$$\bar{R}_i^* L_{ij} A_j^1 = (A_j^{1*} L_{ji}\bar{R}_i)^* = 0. \quad (3.12)$$

Inserting (3.10) into (3.11), one obtains

$$\left\{\frac{\bar{R}_i^*}{2}\left[\frac{\delta}{\delta t}\left(\frac{\partial L_{ij}}{\partial\omega}\right) - \frac{\delta}{\delta x_\nu}\left(\frac{\partial L_{ij}}{\partial k_\nu}\right)\right]\bar{R}_j + \frac{\bar{R}_i^*}{2}\left(\frac{\partial^2 L_{ij}}{\partial t\partial\omega}\right)\bar{R}_j\right\}\sigma + \bar{R}_i^*\frac{\partial L_{ij}}{\partial\omega}\bar{R}_j\frac{\partial\sigma}{\partial t} - \bar{R}_i^*\frac{\partial L_{ij}}{\partial k_\nu}\bar{R}_j\frac{\partial\sigma}{\partial x_\nu} + \left(\bar{R}_i^*\frac{\partial L_{ij}}{\partial\omega}\frac{\delta\bar{R}_j}{\delta t} - \bar{R}_i^*\frac{\partial L_{ij}}{\partial k_\nu}\frac{\delta\bar{R}_j}{\delta x_\nu}\right)\sigma = 0. \quad (3.13)$$

On using (3.3) and (3.9), we arrive at

$$\frac{d\sigma}{dt} + \alpha\sigma = 0, \quad \frac{d\sigma}{dt} = \frac{\partial\sigma}{\partial t} + V_{g\nu}\frac{\partial\sigma}{\partial x_\nu}, \quad (3.14)$$

where V_{g_ν} is given by (3.9) and α is defined as

$$\alpha = \frac{\bar{R}_i^*}{2}\left[\frac{\delta}{\delta t}\left(\frac{\partial L_{ij}}{\partial\omega}\right) - \frac{\delta}{\delta k_\nu}\left(\frac{\partial L_{ij}}{\partial k_\nu}\right)\right]\bar{R}_j + \frac{\bar{R}_i^*}{2}\left(\frac{\partial^2 L_{ij}}{\partial t\partial\omega} - \frac{\partial^2 L_{ij}}{\partial x_\nu\partial k_\nu}\right)\bar{R}_j^* + \bar{R}_i^*\frac{\partial L_{ij}}{\partial\omega}\frac{\delta\bar{R}_j}{\delta t} - \bar{R}_i^*\frac{\partial L_{ij}}{\partial k_\nu}\frac{\delta\bar{R}_j}{\delta x_\nu}. \quad (3.15)$$

Attempting now to solve the transport equation, we start by deriving a differential equation for $|\sigma|$ as follows:

$$|\sigma|^2 = \sigma^*\sigma, \quad (3.16)$$

$$\frac{d\,|\sigma|^2}{dt} = \frac{d\sigma^*}{dt}\sigma + \sigma^*\frac{d\sigma}{dt} = -(\alpha + \alpha^*)\,|\sigma|^2, \quad (3.17)$$

$$\frac{d\,|\sigma|}{dt} + \tfrac{1}{2}(\alpha + \alpha^*)\,|\sigma| = 0, \quad (3.18)$$

with

$$\alpha + \alpha^* = \bar{R}_i^*\left[\frac{\delta}{\delta t}\left(\frac{\partial L_{ij}}{\partial\omega}\right) - \frac{\delta}{\delta x_\nu}\left(\frac{\partial L_{ij}}{\partial k_\nu}\right)\right]\bar{R}_j + \bar{R}_i^*\left(\frac{\partial^2 L_{ij}}{\partial t\partial\omega} - \frac{\partial^2 L_{ij}}{\partial x_\nu\partial k_\nu}\right)\bar{R}_j + \bar{R}_i^*\frac{\partial L_{ij}}{\partial\omega}\frac{\delta\bar{R}_j}{\delta t} + \frac{\delta\bar{R}_i^*}{\delta t}\frac{\partial L_{ij}}{\partial\omega}\bar{R}_j - \bar{R}_i^*\frac{\partial L_{ij}}{\partial k_\nu}\frac{\delta\bar{R}_j}{\delta x_\nu} - \frac{\delta\bar{R}_i^*}{\delta x_\nu}\frac{\partial L_{ij}}{\partial k_\nu}\bar{R}_j. \quad (3.19)$$

On observing the symmetry of $\alpha + \alpha^*$, one writes (3.18) in the following suggestive form:

$$\frac{d}{dt}|\sigma| + \frac{1}{2}\left(\frac{\partial V_{g\nu}}{\partial x_\nu} - \frac{\partial k_\nu}{\partial k_\nu} - \frac{\partial\dot{\omega}}{\partial\omega}\right)|\sigma| = 0. \quad (3.20)$$

Here we have redefined

$$\bar{R}[\mathbf{r}, t; \mathbf{k}(\mathbf{r}, t), \omega(\mathbf{r}, t)] = \mathbf{R}(\mathbf{r}, t);$$

as a result,

$$\frac{\delta\tilde{\mathbf{R}}}{\delta t}[\mathbf{r}, t; \mathbf{k}(\mathbf{r}, t), \omega(\mathbf{r}, t)] = \frac{\partial \mathbf{R}}{\partial t}(\mathbf{r}, t),$$

$$\frac{\delta\tilde{\mathbf{R}}}{\delta x_\nu}[\mathbf{r}, t; \mathbf{k}(\mathbf{r}, t), \omega(\mathbf{r}, t)] = \frac{\partial \mathbf{R}}{\partial x_\nu}(\mathbf{r}, t).$$

The divergence of $\mathbf{V}_g$ is given by

$$\frac{\partial V_{g_\nu}}{\partial x_\nu} = -\frac{\partial}{\partial x_\nu}\left(R_i^* \frac{\partial L_{ij}}{\partial k_\nu} R_j\right) = -\left[\frac{\partial R_i^*}{\partial x_\nu}\frac{\partial L_{ij}}{\partial k_\nu} R_j + R_i^* \frac{\delta}{\delta x_\nu}\left(\frac{\partial L_{ij}}{\partial k_\nu}\right) R_j + R_i^* \frac{\partial L_{ij}}{\partial k_\nu}\frac{\partial R_j}{\partial x_\nu}\right]. \quad (3.21)$$

Also,

$$\dot{k}_\nu = R_i^* \frac{\partial L_{ij}}{\partial x_\nu} R_j, \quad (3.22)$$

$$\dot{\omega} = -R_i^* \frac{\partial L_{ij}}{\partial t} R_j. \quad (3.23)$$

Equations (3.22) and (3.23) are derived in the same way as (3.9).

To integrate the transport equation, let us invoke the following lemma.

Lemma[3,12]: If the differential equations $dx_\nu/dt = V_{g_\nu}$ admit a solution $x_\nu = x_\nu(t, \mathbf{\Gamma})$, $\nu = 1, 2, 3$, depending on the three parameters $\mathbf{\Gamma} = (\gamma_1, \gamma_2, \gamma_3)$, then the Jacobian $j(t, \mathbf{\Gamma})$ satisfies

$$\frac{d}{dt}\ln j(t, \mathbf{\Gamma}) = \frac{\partial V_{g_\nu}}{\partial x_\nu}, \quad j(t, \mathbf{\Gamma}) = \det\left(\frac{\partial x_i}{\partial \gamma_j}\right). \quad (3.24)$$

Applying this lemma to (3.20), one obtains

$$\frac{d}{dt}\ln\{|\sigma|\,[j(t, \mathbf{\Gamma})]^{\frac{1}{2}}\} - \frac{1}{2}\left(\frac{\partial \dot{k}_\nu}{\partial k_\nu} + \frac{\partial \dot{\omega}}{\partial \omega}\right) = 0. \quad (3.25)$$

This is integrated to give

$$|\sigma(t)| = |\sigma(t_0)|\left(\frac{j(t_0, \mathbf{\Gamma})}{j(t, \mathbf{\Gamma})}\right)^{\frac{1}{2}} \exp\int_{t_0}^{t} \frac{1}{2}\left(\frac{\partial \dot{k}_\nu}{\partial k_\nu} + \frac{\partial \dot{\omega}}{\partial \omega}\right) d\tau. \quad (3.26)$$

To derive the phase angle of σ, we set

$$\sigma = |\sigma| \exp(-j\theta). \quad (3.27)$$

Substitution of (3.27) into (3.14) gives

$$\frac{d|\sigma|}{dt} - j|\sigma|\frac{d\theta}{dt} + \alpha|\sigma| = 0. \quad (3.28)$$

The imaginary part of (3.28) is

$$\frac{d\theta}{dt} - (\operatorname{Im}\alpha)\theta = 0, \quad (3.29)$$

which is integrated to yield

$$\theta = \theta_0 + \int_{t_0}^{t} \operatorname{Im}\alpha \, d\tau. \quad (3.30)$$

Since $\mathbf{A}^0 = \sigma\mathbf{R}$ and $\sigma = R_i^*(\partial L_{ij}/\partial\omega)A_j^0$, (3.26) and (3.30) lead to

$$\mathbf{A}^0(t) = \left(\frac{j(t_0, \mathbf{\Gamma})}{j(t, \mathbf{\Gamma})}\right)^{\frac{1}{2}} \exp\int_{t_0}^{t}\left[\frac{1}{2}\left(\frac{\partial \dot{k}_\nu}{\partial k_\nu} + \frac{\partial \dot{\omega}}{\partial \omega}\right) - j\operatorname{Im}\alpha\right] d\tau \times \left[R_i^* \frac{\partial L_{ij}}{\partial \omega} A_j^0\right]_{t_0} \mathbf{R}(t). \quad (3.31)$$

Equation (3.31) gives the amplitude function $\mathbf{A}^0(t_0)$ at any time for any ray in terms of the initial amplitude function $\mathbf{A}^0(t_0)$. The exponential term in (3.31) provides the term due to the inhomogeneity, $\frac{1}{2}[(\partial\dot{k}_\nu/\partial k_\nu) + (\partial\dot{\omega}/\partial\omega)]$, and the phase shift, $\operatorname{Im}\alpha$, along the ray. Since the Jacobian $j(t, \mathbf{\Gamma})$ measures ray density, it appears in (3.31).

4. THE LAW OF ENERGY CONSERVATION AND OTHER FORMS OF THE CONSTITUTIVE RELATION

In the previous sections, asymptotic solutions of wave propagation in spatial- and temporal-dispersive, slightly inhomogeneous, and slowly varying media with a particular form of the constitutive relation has been obtained. In this section, we focus our attention on the transport equation and give its physical interpretation. Competitive transport equations based on other forms of the constitutive relation are discussed also. We begin by deriving a transport equation for slightly inhomogeneous and time-invariant media where energy of the propagating wave is conserved.

A. A Transport Equation for Wave Propagation in Inhomogeneous and Time-Invariant Media Based on Energy Conservation

Let us write the conservation law for the total energy of waves in its differential form in the phase space:

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x_\nu}(V_{g_\nu}u) + \frac{\partial}{\partial k_\nu}(\dot{k}_\nu u) = 0, \quad (4.1)$$

with u as energy density. Here we have assumed that energy of the wave is conserved.

In the physical space there is only one value of $\mathbf{k}(\mathbf{r}, t)$ at the point $\mathbf{r}$, t. Therefore, u is given by

$$u(\mathbf{r}, t; \mathbf{k}, \omega) = \bar{U}(\mathbf{r}, t; \mathbf{k}, \omega)\delta[\mathbf{k} - \mathbf{k}(\mathbf{r}, t)]. \quad (4.2)$$

On carrying out the integration of (4.1) with respect to $\mathbf{k}$, one has

$$\frac{\partial \bar{U}}{\partial t} + \frac{\partial}{\partial x_\nu}(V_{g\nu}\bar{U}) + \frac{\partial}{\partial k_\nu}(\dot{k}_\nu \bar{U}) = 0. \quad (4.3)$$

Note that the differentiation is carried out with respect to x_ν with $\mathbf{k}$ as constant. Functions which appear in (4.3) are

$$\bar{U} = \bar{U}[\mathbf{r}, t; \mathbf{k}(\mathbf{r}, t), \omega], \quad (4.4)$$

$$\frac{dx_\nu}{dt} = -\bar{R}_i^* \frac{\partial L_{ij}}{\partial k_\nu} \bar{R}_j = V_{g\nu}, \quad (4.5)$$

$$\frac{dk_\nu}{dt} = \bar{R}_i^* \frac{\partial L_{ij}}{\partial x_\nu} \bar{R}_j = \dot{k}_\nu, \quad (4.6)$$

$$\bar{\mathbf{R}} = \bar{\mathbf{R}}[\mathbf{r}, t; \mathbf{k}(\mathbf{r}, t), \omega], \quad (4.7)$$

$$L_{ij} = L_{ij}[\mathbf{r}, t; \mathbf{k}(\mathbf{r}, t), \omega]. \quad (4.8)$$

Redefining

$$\bar{U}[\mathbf{r}, t; \mathbf{k}(\mathbf{r}, t), \omega] = U(\mathbf{r}, t),$$
$$\bar{\mathbf{R}}[\mathbf{r}, t; \mathbf{k}(\mathbf{r}, t), \omega] = \mathbf{R}(\mathbf{r}, t), \quad (4.9)$$

one obtains

$$\frac{\partial U}{\partial t} + V_{g\nu}\frac{\partial U}{\partial x_\nu} = \frac{\partial \bar{U}}{\partial t} + V_{g\nu}\frac{\partial \bar{U}}{\partial x_\nu} + \dot{k}_\nu \frac{\partial \bar{U}}{\partial k_\nu}. \quad (4.10)$$

Equations (4.3) and (4.10) lead to

$$\frac{\partial U}{\partial t} + V_{g\nu}\frac{\partial U}{\partial x_\nu} + \frac{\partial V_{g\nu}}{\partial x_\nu} U + \left(\frac{\partial \dot{k}_\nu}{\partial k_\nu}\right) U = 0. \quad (4.11)$$

This is then the transport equation for energy-conservative propagating waves.

Let us write (4.11) in the form of the Poynting theorem. Let $\sigma\mathbf{R} = \mathbf{E}$ and $\sigma^*\mathbf{R}^* = \mathbf{E}^*$ as an approximation, and let $c^2 L_{ij}/\omega$ replace L_{ij}, since they are equivalent; we then have

$$\frac{\partial U}{\partial t} + \operatorname{div} \mathbf{S} = -\left(\frac{\partial \dot{k}_\nu}{\partial k_\nu}\right) U, \quad (4.12)$$

where

$$U = \frac{\partial}{\partial \omega}\left(E_i^* \omega \tilde{\epsilon}_{ij} E_j - \frac{c^2 k^2}{\omega^2} E_i^* E_j + \frac{c^2}{\omega^2}(\mathbf{k} \times \mathbf{E}^*)(\mathbf{k} \times \mathbf{E})\right) \quad (4.13)$$

and

$$S_\nu = -\frac{\partial}{\partial k_\nu}\left(E_i^* \omega \tilde{\epsilon}_{ij} E_j - \frac{c^2 k^2}{\omega^2} E_i^* E_j + \frac{c^2}{\omega^2}(\mathbf{k} \times \mathbf{E}^*)(\mathbf{k} \times \mathbf{E})\right). \quad (4.14)$$

In homogeneous media, the right-hand side of (4.12) vanishes. Therefore, for wave propagation in inhomogeneous media, the Poynting theorem is different from that for homogeneous media in the case when the energy of the propagating is conserved.

In the problem of wave propagation in inhomogeneous temporal-dispersive media which Lewis studied,[3] the transport equation can easily be shown to satisfy energy conservation. To do this, we consider the temporal-dispersive operator

$$L_{ij} = k_\nu A_{ij}^\nu - \omega \tilde{\epsilon}_{ij}(\omega, x_\nu). \quad (4.15)$$

The group velocity formula is

$$V_{g\nu} = R_i^* A^\nu R_j. \quad (4.16)$$

Also,

$$\dot{k}_\nu = -R_i^* \frac{\partial}{\partial x_\nu}[k_\nu A^\nu - \omega\tilde{\epsilon}(\omega, x_\nu)]R_j. \quad (4.17)$$

Equation (4.11) yields

$$\frac{dU}{dt} + \frac{\partial V_{g\nu}}{\partial x_\nu} U - R_i^* \frac{\partial A^\nu}{\partial x_\nu} R_j U = 0. \quad (4.18)$$

For the lossless case, (L.3.7.6) of Lewis is equivalent to (4.18).

B. Transport Equation Derived from the Constitutive Relation (2.2)

Let

$$U = |\sigma|^2. \quad (4.19)$$

Then (3.33) can be written as

$$\frac{dU}{dt} + \frac{\partial V_{g\nu}}{\partial x_\nu} U = \left(-\frac{\partial \dot{k}_\nu}{\partial k_\nu} + 2\frac{\partial \dot{k}_\nu}{\partial k_\nu} + \frac{\partial \dot{\omega}}{\partial \omega}\right) U. \quad (4.20)$$

Now, the first term on the right-hand side, $-(\partial \dot{k}_\nu/\partial k_\nu)U$, is due to the change of the wavenumber $\mathbf{k}$ of the wave, while $[2(\partial \dot{k}_\nu/\partial k_\nu) + (\partial\dot{\omega}/\partial\omega)]U$ is the energy supplied by the medium to ensure energy conservation of the whole system, the wave and the medium. Therefore, the propagating wave does not form a closed system.

C. Transport Equations Based on Other Forms of the Constitutive Relation

For the constitutive relation (2.2) assumed in Sec. 2, the characteristic constants are evaluated at the fixed point, i.e., $\tilde{\epsilon}_{ij}(\mathbf{r} - \mathbf{r}', t - t'; \mathbf{k}, \omega)$ has been assumed for the inhomogeneous dielectric tensor. There are other ways of incorporating the inhomogeneity. For example, we can evaluate the characteristic constants at the position of the displacement vector, i.e., $\tilde{\epsilon}_{ij}(\mathbf{r}, t; \mathbf{k}, \omega)$, or the mean value of the position of the electric field and that of displacement, i.e., $\tilde{\epsilon}_{ij}(\mathbf{r} - r'/2, t - t'/2; \mathbf{k}, \omega)$.[13]

These modifications of the constitutive relation do not change the eikonal equation. However, the transport equation is different in each case. We shall give the transport equation for each modified form of the constitutive relation. Let us begin with the constitutive relation where the dielectric tensor is evaluated at the position of the displacement vector:

$$D_i(\mathbf{r}, t) = \frac{1}{(2\pi)^4} \int_0^t dt' \int_{-\infty}^{\infty} d^3\mathbf{r}' \int_{-j\sigma-\infty}^{-j\sigma+\infty} d\omega \int_{-\infty}^{\infty} d^3\mathbf{k}\tilde{\epsilon}_{ij}(\mathbf{r}, t; \mathbf{k}, \omega) \times \exp[j(\omega t' - \mathbf{k}\cdot\mathbf{r}')]E_j(\mathbf{r}-\mathbf{r}', t-t'; \mathbf{k}, \omega). \tag{4.21}$$

Employing a procedure as given in Sec. 2, we obtain

$$[L_{ij}][A_j^0] = 0, \tag{4.22}$$

$$[L_{ij}][A_j^1] + \frac{1}{2}\left[\frac{\delta}{\delta t}\left(\frac{\partial L_{ij}}{\partial\omega}\right) - \frac{\delta}{\delta x_\nu}\left(\frac{\partial L_{ij}}{\partial k_\nu}\right) - \frac{\partial}{\partial t}\left(\frac{\partial L_{ij}}{\partial\omega}\right) + \frac{\partial}{\partial x_\nu}\left(\frac{\partial L_{ij}}{\partial k_\nu}\right)\right](A_j^0) + \frac{\partial L_{ij}}{\partial\omega}\frac{\partial A_j^0}{\partial t} - \frac{\partial L_{ij}}{\partial k_\nu}\frac{\partial A_j^0}{\partial x_\nu} = 0. \tag{4.23}$$

The transport equation derived from (4.23) can be easily shown to be

$$\frac{d}{dt}|\sigma| + \frac{1}{2}\left(\frac{\partial V_{g\nu}}{\partial x_\nu} + \frac{\partial \dot{k}_\nu}{\partial k_\nu} + \frac{\partial\dot{\omega}}{\partial\omega}\right)|\sigma| = 0. \tag{4.24}$$

For time-invariant media $\dot{\omega} = 0$, (4.24) satisfies energy conservation, since it agrees with (4.11).

On the other hand, taking the constitutive relation with the dielectric tensor evaluated at the mean value of the position of **D** and that of **E**,

$$D_i(\mathbf{r}, t) = \frac{1}{(2\pi)^4} \int_0^t dt' \int_{-\infty}^{\infty} d^3\mathbf{r}' \int_{-j\sigma-\infty}^{-j\sigma+\infty} d\omega \times \int_{-\infty}^{\infty} d^3\mathbf{k}\tilde{\epsilon}_{ij}(\mathbf{r}-\tfrac{1}{2}\mathbf{r}', t-\tfrac{1}{2}t'; \mathbf{k}, \omega) \times \exp[j(\omega t' - \mathbf{k}\cdot\mathbf{r}')]E_j(\mathbf{r}-\mathbf{r}', t-t'), \tag{4.25}$$

one arrives at the transport equation

$$\frac{d}{dt}|\sigma| + \frac{1}{2}\frac{\partial V_{g\nu}}{\partial x_\nu}|\sigma| = 0. \tag{4.26}$$

Therefore, a generalized Poynting theorem for inhomogeneous media retains the same form as that for homogeneous media:

$$\frac{\partial U}{dt} + \operatorname{div}\mathbf{S} = 0. \tag{4.27}$$

U and **S** are given by (4.12) and (4.13), respectively.

In concluding, three different forms of the constitutive relation for incorporating inhomogeneity and slow varyingness in spatial- and temporal-dispersive media have been discussed. Maxwell's equations, together with one form of the constitutive relation (4.21), satisfies energy conservation asymptotically. However, since energy of the wave is not conserved, other forms of the constitutive relation are preferred for situations where the wave itself does not form a closed system. In particular, the constitutive relation (2.2) is most suitable for describing inhomogeneous plasmas in quasi-equilibrium, for it is derived through the microscopic consideration. This study somehow limits the accuracy of the conventional derivation of the equation of radiative transfer based on energy conservation.

ACKNOWLEDGMENTS

The author would like to acknowledge Professor J. L. Yen, of the University of Toronto, Professor D. Ludwig of the Courant Institute, and Dr. R. F. Millar of the National Research Council of Canada for their discussions and assistance.

APPENDIX A: THE CONSTITUTIVE RELATION FOR ELECTROMAGNETIC WAVES IN ANISOTROPIC, INHOMOGENEOUS WARM PLASMAS

In this appendix we show that the constitutive relation for electromagnetic waves in anisotropic, inhomogeneous warm plasmas has the form assumed for general media.

The phenomenon of electromagnetic wave propagation in plasmas is determined by Maxwell's equations and constitutive relations. The constitutive relations are determined by the composition of plasmas, and can be obtained for different models: The simpler cold model is based on the electronic orbital motion. The more complicated warm model can be arrived at either in a crude approximation by introducing the pressure term into the electronic orbital motion used in the cold model or in an elaborate manner based on the kinetic theory by taking into account of the velocity spread of electrons. We derive the constitutive relation using the latter approach, but ignore the motion of ions for brevity of formulas. Briefly, the well-known method of characteristics[14] is used to solve the linearized Boltzmann equation. Then the use of a simple transformation leads to the space–time-conductivity tensor, which has not been obtained before.

Consider the collisionless Boltzmann equation for the electronic distribution function f:

$$\frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla f - \frac{e}{m}(\mathbf{E} + \mathbf{v}\times\mathbf{B})\cdot\frac{\partial f}{\partial\mathbf{v}} = 0, \tag{A1}$$

where $\mathbf{B}$ includes the externally applied magnetic field $\mathbf{B}_0 = B_0(\mathbf{r}, t)\mathbf{l}_z$ and the magnetic field of electromagnetic waves and $\mathbf{E}$ is the electric field of electromagnetic waves. Assuming $f = f_0 + f_1$, with f_1 as a small perturbation to the equilibrium Maxwellian distribution, we reduce (A1) to

$$\frac{\partial f_0}{\partial t} + \mathbf{v}\cdot\frac{\partial f_0}{\partial \mathbf{r}} - \frac{e}{m}(\mathbf{v}\times\mathbf{B}_0)\cdot\frac{\partial f_0}{\partial \mathbf{v}} = 0 \quad \text{(A2)}$$

and

$$\frac{\partial f_1}{\partial t} + \mathbf{v}\cdot\frac{\partial f_1}{\partial \mathbf{r}} - \frac{e}{m}(\mathbf{v}\times\mathbf{B}_0)\cdot\frac{\partial f_1}{\partial \mathbf{v}} = \frac{e}{m}\mathbf{E}\cdot\frac{\partial f_0}{\partial \mathbf{v}}. \quad \text{(A3)}$$

Obviously, (A2) is satisfied by a Maxwellian distribution f_0:

$$f_0 = N\frac{\langle V^2\rangle^{\frac{3}{2}}}{(2\pi)^{\frac{3}{2}}}\frac{\exp[-2(V_{0_x}^2 + V_{0_y}^2 + V_{0_z}^2)]}{\langle V^2\rangle}, \quad \text{(A4)}$$

where, for inhomogeneous and quasi-equilibrium plasmas, the mean square velocity is

$$\langle V(\mathbf{r}, t)^2\rangle = \kappa T_e/4m. \quad \text{(A5)}$$

Here

$N(\mathbf{r}, t)$ is the density of electrons,
κ is Boltzmann's constant,
T_e is the electronic temperature of plasmas.

The solution of (A3) is given by

$$f_1(\mathbf{r}, \mathbf{v}, t) = \int_0^t dt_0 \int_{-\infty}^{\infty} d^3\mathbf{r}_0 d^3\mathbf{v}_0 g(\mathbf{r}_0, \mathbf{v}_0, t_0) \times \delta(\mathbf{r} - \mathbf{R})\delta(\mathbf{v} - \mathbf{V}), \quad \text{(A6)}$$

with

$$g(\mathbf{r}_0, \mathbf{v}_0, t_0) = \frac{e}{m}\mathbf{E}\cdot\frac{\partial f_0}{\partial \mathbf{v}_0}. \quad \text{(A7)}$$

Under the influence of the external magnetic field, $B_0(\mathbf{r}, t)\mathbf{l}_z$, $\mathbf{R}(\mathbf{r}_0, \mathbf{v}_0, t_0)$, and $\mathbf{V}(\mathbf{r}_0, \mathbf{v}_0, t_0)$ are the classical asymptotic position and velocity of electrons in terms of their initial conditions $\mathbf{r}_0$ and $\mathbf{v}_0$. Let us assume that the plasma is not very hot so that electrons with large gyrating radius r_g are negligible, i.e., for most electrons $r_g = |\mathbf{v}_0|/\omega_H = |m\mathbf{v}_0/eB_0|$ is very small. Therefore, the assumption of weak inhomogeneity of the magnetic field leads to

$$\omega_H(\mathbf{r}_0 + m\mathbf{v}_0 \times \mathbf{l}_z/eB_0, t) \sim \omega_H(\mathbf{r}_0, t). \quad \text{(A8)}$$

Then, $\mathbf{R}$ and $\mathbf{V}$ are given by

$$\begin{aligned}\mathbf{R} &= \mathbf{r}_0 + \mathbf{v}_{0_\rho}\frac{\sin\omega_H\tau}{\omega_H} + \mathbf{v}_{0_\rho}\times\mathbf{l}_z\frac{1-\cos\omega_H\tau}{\omega_H} + v_{0_z}\tau\mathbf{l}_z \\ &= \mathbf{r}_0 + \mathbf{R}',\end{aligned}$$

$$\begin{aligned}\mathbf{V} &= v_{0_\rho}\cos\omega_H\tau + \mathbf{v}_{0_\rho}\times\mathbf{l}_z\sin\omega_H\tau + v_{0_z}\tau\mathbf{l}_z, \\ \mathbf{v}_{0_\rho} &= \mathbf{v}_0 - v_{0_z}\mathbf{l}_z, \\ \omega_H &\sim \omega_H(\mathbf{r}_0, t) = eB_0(\mathbf{r}_0, t)m^{-1}, \\ \tau &= t - t_0.\end{aligned} \quad \text{(A9)}$$

To see that (A6) is the solution of (A3), one simply differentiates (A6) with respect to t as follows:

$$\begin{aligned}\frac{\partial f_1}{\partial t}(\mathbf{r}, \mathbf{v}, t) &= \int_{-\infty}^{\infty} d^3\mathbf{r}_0 d^3\mathbf{v}_0 g(\mathbf{r}_0, \mathbf{v}_0, t)\delta(\mathbf{r} - \mathbf{r}_0)\delta(\mathbf{v} - \mathbf{v}_0) \\ &\quad + \int_0^t dt_0 \int_{-\infty}^{\infty} d^3\mathbf{r}_0 d^3\mathbf{v}_0 g(\mathbf{r}_0, \mathbf{v}_0, t_0) \times \frac{\partial}{\partial t}[\delta(\mathbf{r} - \mathbf{R})\delta(\mathbf{v} - \mathbf{V})] \\ &= g(\mathbf{r}, \mathbf{v}, t) + \int_0^t dt_0 \int_{-\infty}^{\infty} d^3\mathbf{r}_0 d^3\mathbf{v}_0 \times \left(\frac{\partial \mathbf{R}}{\partial t}\cdot\frac{\partial}{\partial \mathbf{R}}[\delta(\mathbf{r} - \mathbf{R})]\delta(\mathbf{v} - \mathbf{V}) + \frac{\partial \mathbf{V}}{\partial t}\cdot\frac{\partial}{\partial \mathbf{V}}[\delta(\mathbf{v} - \mathbf{V})]\delta(\mathbf{r} - \mathbf{R})\right) \\ &= g(\mathbf{r}, \mathbf{v}, t) - \int_0^t dt_0 \int_{-\infty}^{\infty} d^3\mathbf{r}_0 d^3\mathbf{v}_0 \times \left(\mathbf{V}\cdot\frac{\partial}{\partial \mathbf{r}}[\delta(\mathbf{r} - \mathbf{R})]\delta(\mathbf{v} - \mathbf{V}) - \frac{e}{m}\mathbf{V}\times\mathbf{B}_0\cdot\frac{\partial}{\partial \mathbf{v}}[\delta(\mathbf{v} - \mathbf{V})]\delta(\mathbf{r} - \mathbf{R})\right) \\ &= g(\mathbf{r}, \mathbf{v}, t) - \left(\mathbf{v}\cdot\frac{\partial}{\partial \mathbf{r}} - \frac{e}{m}\mathbf{v}\times\mathbf{B}_0\cdot\frac{\partial}{\partial \mathbf{v}}\right)f_1(\mathbf{r}, \mathbf{v}, t).\end{aligned} \quad \text{(A10)}$$

Therefore, the current induced can be expressed as

$$J_I(\mathbf{r}, t) = -e\int_{-\infty}^{\infty} \mathbf{v} f_1(\mathbf{r}, \mathbf{v}, t) d^3\mathbf{v} \quad \text{(A11)}$$

$$= -\frac{e^2}{m}\int_0^t dt_0 \int_{-\infty}^{\infty} d^3\mathbf{r}_0\, d^3\mathbf{v}_0 \mathbf{V}(\mathbf{r}_0, \mathbf{v}_0, t_0)\frac{\partial f_0}{\partial \mathbf{v}_0} \cdot \mathbf{E}(\mathbf{r}_0, t_0)\delta(\mathbf{r} - \mathbf{R}). \quad \text{(A12)}$$

The last step follows from the substitution of (A6) into (A11) and a subsequent integration with respect to $\mathbf{v}$. Let us now change the variables of integration from $dt_0 d^3\mathbf{v}_0$ to $dt_0 d^3\mathbf{R}'$. Writing $\mathbf{R}'$ in terms of $\mathbf{v}_0$ explicitly, and vice versa,

$$\begin{aligned}x' &= v_{0_x}\frac{\sin\omega_H\tau}{\omega_H} - v_{0_y}\frac{1-\cos\omega_H\tau}{\omega_H}, \\ y' &= v_{0_x}\frac{1-\cos\omega_H\tau}{\omega_H} + v_{0_y}\frac{\sin\omega_H\tau}{\omega_H}, \\ z' &= v_{0_z}\tau\end{aligned} \quad \text{(A13)}$$

and

$$v_{0_x} = \frac{\omega_H^2}{2(1-\cos\omega_H\tau)}\left(\frac{\sin\omega_H\tau}{\omega_H}x' + \frac{1-\cos\omega_H\tau}{\omega_H}y'\right),$$

$$v_{0_y} = \frac{\omega_H^2}{2(1-\cos\omega_H\tau)}\left(-\frac{1-\cos\omega_H\tau}{\omega_H}x' + \frac{\sin\omega_H\tau}{\omega_H}y'\right), \quad (A14)$$

$$v_{0_z} = \tau^{-1}z',$$

one easily obtains the Jacobian of the transformation

$$\frac{\partial(v_{0_x}, v_{0_y}, v_{0_z})}{\partial(x', y', z')} = \frac{\omega_H^2}{2\tau(1-\cos\omega_H\tau)}. \quad (A15)$$

Then the integration of (A12) over $\mathbf{r}_0$ leads to

$$\begin{aligned} J_I(\mathbf{r}, t) &= -\frac{e^2}{m}\int_0^t dt'\int_{-\infty}^{\infty} d^3\mathbf{R}'\mathbf{V}[\mathbf{r}-\mathbf{R}', \mathbf{v}_0(\mathbf{R}', t-t', \omega_H), t'] \\ &\times \frac{\partial f_0}{\partial \mathbf{v}_0}[\mathbf{r}-\mathbf{R}', \mathbf{v}_0(\mathbf{R}', t-t', \omega_H), t'] \\ &\times \frac{\omega_H^2}{2(t-t')[1-\cos\omega_H(t-t')]}\cdot\mathbf{E}(\mathbf{r}-\mathbf{R}', t'), \end{aligned} \quad (A16)$$

with

$$\omega_H = \omega_H(\mathbf{r}-\mathbf{R}', t'). \quad (A17)$$

It appears that (A16) has a singularity at $t' = 0$. However, a closer examination with the aid of (A4) and (A14) gives a contrary conclusion. Let us now generalize the well-known convolution integral to the weakly inhomogeneous medium by introducing additional variables into the kernel:

$$\begin{aligned} \mathbf{J}_I(\mathbf{r}, t) &= \int_0^t dt'\int_{-\infty}^{\infty} d^3\mathbf{r}'\boldsymbol{\sigma}(\mathbf{r}', t'; \mathbf{r}-\mathbf{r}', t-t')\cdot\mathbf{E}(\mathbf{r}', t') \\ &= \int_0^t dt'\int_{-\infty}^{\infty} d^3\mathbf{r}'\boldsymbol{\sigma}(\mathbf{r}-\mathbf{r}', t-t'; \mathbf{r}', t') \\ &\quad\cdot\mathbf{E}(\mathbf{r}-\mathbf{r}', t-t'). \end{aligned} \quad (A18)$$

It immediately follows that

$$\begin{aligned} \boldsymbol{\sigma}(\mathbf{r}-\mathbf{r}', t-t'; \mathbf{r}', t') &= -e^2m^{-1}\mathbf{V}[\mathbf{r}-\mathbf{r}', \mathbf{v}_0(\mathbf{r}', t', \omega_H), t-t'] \\ &\times \frac{\partial f_0}{\partial \mathbf{v}_0}[\mathbf{r}-\mathbf{r}', \mathbf{v}_0(\mathbf{r}', t', \omega_H), t-t'] \\ &\times \frac{\omega_H^2}{2t'(1-\cos\omega_H t')}, \end{aligned} \quad (A19)$$

where f_0 is given by (A4) and the characteristic constants are

$$\omega_H = \omega_H(\mathbf{r}-\mathbf{r}', t-t'), \quad N = N(\mathbf{r}-\mathbf{r}', t-t'),$$
$$\langle V^2\rangle = \langle V(\mathbf{r}-\mathbf{r}', t-t')^2\rangle. \quad (A20)$$

Thus, through the simple transformation (A14), we have arrived at the conductivity tensor in $\mathbf{r}-t$ space, which has not been given before. The simplicity of the form is rather striking. Even for weakly inhomogeneous plasmas, (A19) does not involve any integration. To show that the space–time-conductivity tensor (A19) derived for homogeneous media is equivalent to the conventional time-harmonic conductivity tensor, let us apply the four-dimensional Fourier–Laplace transformation to (A19):

$$\begin{aligned} \tilde{\boldsymbol{\sigma}}(\mathbf{k}, \omega) &= \int_0^\infty dt'\int_{-\infty}^{\infty} d^3\mathbf{r}'\exp[-j(\omega t'-\mathbf{k}\cdot\mathbf{r}')]\tilde{\boldsymbol{\sigma}}(\mathbf{r}', t') \\ &= \int_0^\infty dt'\int_{-\infty}^{\infty} d^3\mathbf{r}'\exp[-j(\omega t'-\mathbf{k}\cdot\mathbf{r}')]\left(-\frac{e^2}{m}\right) \\ &\times\left(\mathbf{V}[\mathbf{v}_0(\mathbf{r}', t')]\frac{\partial f_0}{\partial \mathbf{v}_0}[\mathbf{v}_0(\mathbf{r}', t')]\frac{\omega_H^2}{2t'(1-\cos\omega_H t')}\right). \end{aligned} \quad (A21)$$

Transformation of variables of integration from $\mathbf{r}'$ into $\mathbf{v}_0$, with the transformation given by (A13), and the replacement of τ by t' lead to

$$\tilde{\boldsymbol{\sigma}}(\mathbf{k}, \omega) = \int_0^\infty dt'\int_{-\infty}^{\infty} d^3\mathbf{v}_0\exp\{-j[\omega t'-\mathbf{k}\cdot\mathbf{r}'(\mathbf{v}_0)]\} \times\left(-\frac{e^2}{m}\right)\left(\mathbf{V}(\mathbf{v}_0)\frac{\partial f_0}{\partial \mathbf{v}_0}(\mathbf{v}_0)\right). \quad (A22)$$

Also, let

$$v_x = v_\perp\cos\alpha, \quad v_y = v_\perp\sin\alpha, \quad v_z = v_\parallel. \quad (A23)$$

Then

$$v_{0_x} = v_\perp\cos(\alpha+\omega_H\tau), \quad v_{0_y} = v_\perp\sin(\alpha+\omega_H\tau). \quad (A24)$$

Transformation of variables of integration from $\mathbf{v}_0$ to $\mathbf{v}$ and use of (A24) give

$$\begin{aligned} \tilde{\boldsymbol{\sigma}}(\mathbf{k}, \omega) &= \int_0^\infty dt'\int_{-\infty}^{\infty} d^3\mathbf{v} \\ &\times\exp\left(-j\omega t' + j\frac{k_\perp v_\perp}{\omega_H}\times[\sin(\alpha+\omega_H t')-\sin\alpha] + jk_\parallel v_\parallel\omega_H t'\right) \\ &\times[v_\perp\cos(\alpha)\mathbf{l}_x + v_\perp\sin(\alpha)\mathbf{l}_y + v_\parallel\mathbf{l}_z] \\ &\times\left(\cos(\alpha+\omega_H t)\frac{\partial f_0}{\partial v_\perp}\mathbf{l}_x + \sin(\alpha+\omega_H t)\frac{\partial f_0}{\partial v_\perp}\mathbf{l}_y + \frac{\partial f_0}{\partial v_\parallel}\mathbf{l}_z\right). \end{aligned} \quad (A25)$$

This is equivalent to (10·39) of Montgomery *et al.*[15]

Finally, let us note that the inhomogeneous and slowly varying characteristic constants ω_H, N, and $\langle V^2 \rangle$ are evaluated at $\mathbf{r} - \mathbf{r}'$ and $t - t'$. This agrees with the general form (2.2) for the constitutive relation in inhomogeneous, and slowly varying, spatial- and temporal-dispersive media.

APPENDIX B: LIMITS OF APPLICABILITY OF ASYMPTOTIC SOLUTIONS

Previously, in arriving at (2.14) and (2.15), we have expanded the governing equations (2.1) and (2.4) up to the second-order term. In case one proceeds to the third-order term, the equation obtained is

$$L_{ij}A_j^2 + \frac{1}{2}\left[\frac{\delta}{\delta t}\left(\frac{\partial L_{ij}}{\partial \omega}\right) - \frac{\delta}{\delta x_\nu}\left(\frac{\partial L_{ij}}{\partial k_\nu}\right) + \frac{\partial^2 L_{ij}}{\partial t \partial \omega} - \frac{\partial^2 L_{ij}}{\partial x_\nu \partial k_\nu}\right]A_j^1 + \frac{\partial L_{ij}}{\partial \omega}\frac{\partial A_j^1}{\partial t} - \frac{\partial L_{ij}}{\partial k_\nu}\frac{\partial A_j^1}{\partial x_\nu} = -\mathfrak{L}_{ij}A_j^0, \quad \text{(B1)}$$

where $\mathfrak{L}_{ij}$ involves many second-order terms.

The solution of (B1) for A_j^1 is

$$A_j^1 = A_j^0 \int_0^{\lambda t} \sigma^{-1} R_i^* \mathfrak{L}_{ij}(\sigma R_j)\, d\tau, \quad A_j^0 = \sigma R_j, \quad \text{(B2)}$$

where we have taken only the force term in the solution, since the solution of the homogeneous equation for A_j^1 has the same form as that of the equation for A_j^0.

To satisfy $A_j^1 \ll \lambda A_j^1$, it is necessary that

$$\int_0^{\lambda t} \sigma^{-1} R_i^* \mathfrak{L}_{ij}(\sigma R_j)\, d\tau \ll \lambda. \quad \text{(B3)}$$

Equation (B3) reduces to the following two conditions: First, at caustics the Jacobian $j(t, \mathbf{\Gamma})$ is zero. Then integration by parts leads to the appearance of $j(t, \mathbf{\Gamma})$ in the denominator of the integrand. As a result, (B3) is violated. Therefore,

$$j(t, \mathbf{\Gamma}) \neq 0. \quad \text{(B4)}$$

One can reduce the inequality (B3) to some simplified inequalities. However, such a calculation is very tedious. We shall give a simple estimate. Let the medium be slowly varying:

$$R_i^* \frac{\partial \tilde{\epsilon}_{ij}}{\partial x_\nu} R_j, \quad R_i^* \frac{\partial \tilde{\epsilon}_{ij}}{\partial t} R_j \ll 1. \quad \text{(B5)}$$

Then each term of (B3) can easily be shown to satisfy the inequality. Equation (B5) is then the second condition.

Finally, in contrast to geometrical optics in nondispersive media, these conditions are not sufficient. In order that ansatz (2.3) is an assumed solution for dispersive media, it is imperative that

$$\lambda_{\text{char}}, \quad c/\omega_{\text{char}} \ll \lambda, \quad \text{(B6)}$$

where

λ_{char} is the characteristic length of spatial dispersion,

ω_{char} is the characteristic frequency of temporal dispersion.

Their significance has already been discussed in the Introduction. Equation (B6) is required so that the first term of (2.3) gives dominant contribution for homogeneous and time-invariant dispersive media.

[1] R. M. Lewis and J. B. Keller, Courant Institute, New York University, Research Report No. Em-194, 1964.
[2] R. M. Lewis, Proceedings of the Symposium on Asymptotic Methods for Differential Equations at University of Wisconsin, 1964.
[3] R. M. Lewis, Arch. Ratl. Mech. Anal. **20**, 191 (1965).
[4] N. Bleinstein and R. M. Lewis, "Asymptotic Expansions of Solutions of Initial-Boundary Value Problems for a Dispersive Hyperbolic Equation,"in *Electromagnetic Wave Theory*, J. Brown, Ed. (Pergamon, New York, 1967).
[5] R. Handelsman and R. M. Lewis, J. Math. Phys. **9**, 260 (1968).
[6] B. Granoff and R. M. Lewis, Phil. Trans. Roy. Soc., Ser. A, **262**, 387 (1967).
[7] R. M. Lewis and J. K. Cohen, J. Math. Phys. **11**, 296 (1970).
[8] A. A. Rukhadze and V. P. Silin, Usp. Fiz. Nauk **74**, 223 (1961) [Sov. Phys. Usp. **4**, 459 (1961)].
[9] V. L. Ginzburg, Zh. Eksp. Teor. Fiz. **34**, 1593 (1958) [Sov. Phys. JETP **7**, 1096 (1958)].
[10] H. Bloomberg, Phys. Fluids **11**, 1023 (1968).
[11] S. Weinberg, Phys. Rev. **126**, 1899 (1962).
[12] R. Aris, *Vectors, Tensors, and the Basic Equations of Fluid Mechanics* (Prentice-Hall, Englewood Cliffs, N.J., 1962).
[13] B. B. Kadomtsev, *Plasma Turbulence* (Academic, New York, 1965).
[14] P. M. Platzman and S. J. Buchbaum, Phys. Fluids **4**, 1288 (1961).
[15] D. C. Montogomery and D. A. Tidman, *Plasma Kinetic Theory* (McGraw-Hill, New York, 1964).

Two-Dimensional Quarter Space Problems in One-Speed Transport Theory*

A. LEONARD
Nuclear Engineering Division, Stanford University, Stanford, California 94305

(Received 20 July 1970)

Methods derived from the theory of several complex variables are used as a means of analyzing a class of two-dimensional transport problems in a scattering and absorbing quarter space ($0 \le x_1, 0 \le x_2, -\infty \le x_3 \le \infty$) described by a linear, one-speed Boltzmann equation. Using Fourier transformation and the Bochner decomposition, the multivariable analog of the Wiener–Hopf factorization, we find the Green's function in transform space, which solves all source problems having a solution bounded at infinity. The transform of the density asymptotically far from the corner ($x_1 = x_2 = 0$) is determined explicitly, while the remainder is given in terms of the solution to a pair of Fredholm equations.

1. INTRODUCTION

For the past forty years, the Wiener–Hopf technique[1] has proven to be a powerful tool in the analysis of integral equations over the half-line with a difference kernel. For that reason, one of its many applications has been to one-speed, linear transport in a half-space.[2] The method is based on Fourier transformation and relies heavily upon the theory of functions of a complex variable.

In this paper, we use a similar approach generalized to two complex variables to study two-dimensional transport in a quarter space. Here the basic integral equation is over a quarter plane, with the kernel depending upon distance in the plane. In Sec. 2, double Fourier transformation of the transport equation yields a two-variable Wiener–Hopf problem for four unknown transforms corresponding to the densities in each quarter space. A similar mathematical problem arises in the theory of electromagnetic wave diffraction from a right angle dielectric wedge. Although an exact solution to the diffraction problem is not yet available,[3] the Bochner decomposition,[4] the multivariable analog of the Wiener–Hopf factorization, was used in one of the analyses[5] and is found to be a useful tool for our analysis as discussed in Sec. 3. In Sec. 4, the asymptotic contribution to the transforms is found explicitly by one-dimensional Wiener–Hopf analyses and is then subtracted, yielding an equation for a new set of four unknown functions representing the transforms of "transient" densities which are nonnegligible only near the corner. The properties of this new equation allow, by subsequent manipulations in Sec. 5, the solution to be expressed in terms of the solution to a pair of Fredholm equations derived in Sec. 6. Analogous to the one-dimensional problem for a finite slab,[6] these Fredholm equations appear to represent the interaction of the "transient" densities in the two quarter spaces adjacent to the scattering and absorbing quarter space. It is shown that this pair of equations may be solved by iteration.

2. FOURIER TRANSFORMATION OF THE TRANSPORT EQUATION

We consider one-speed neutron transport in a quarter space (Q), $0 \le x_1$, $0 \le x_2$, $-\infty \le x_3 \le \infty$, with isotropic scattering and a given source distribution $S(\mathbf{r}) = S(x_1, x_2)$. The integral transport equation for the neutron density $\rho(\mathbf{r})$ is

$$\rho(\mathbf{r}) = c\int_Q \frac{e^{-|\mathbf{r}-\mathbf{r}'|}}{4\pi\,|\mathbf{r}-\mathbf{r}'|^2}\left[\rho(\mathbf{r}') + \frac{S(\mathbf{r}')}{c}\right] d\mathbf{r}', \quad (2.1)$$

where distances are in units of a mean free path, $\mathbf{r}$ and $\mathbf{r}'$ are three-dimensional vectors, and c is the mean number of neutrons emitted per collision. Letting $\rho(\mathbf{r}) = \rho(x_1, x_2)$, $S(\mathbf{r}) = S(x_1, x_2)$, and performing the x_3 integration, the transport equation (2.1) becomes

$$\rho(x_1, x_2) = c\int_0^\infty\int_0^\infty dx_1'\,dx_2' K(|\mathbf{x}-\mathbf{x}'|)\left(\rho(x_1', x_2') + \frac{S(x_1', x_2')}{c}\right), \quad (2.2)$$

where $|\mathbf{x}-\mathbf{x}'|^2 = (x_1 - x_1')^2 + (x_2 - x_2')^2$ and the kernel K has the integral representation

$$K(s) = \frac{1}{2\pi s}\int_1^\infty \frac{e^{-ts}\,dt}{t(t^2-1)^{\frac{1}{2}}}. \quad (2.3)$$

For convenience, we now consider the integral equation for $\varphi(x_1, x_2)$, $-\infty \le x_1,\ x_2 < \infty$, with a specific inhomogeneous term as follows:

$$\varphi(x_1, x_2) = c\int_0^\infty\int_0^\infty dx_1'\,dx_2' K(|\mathbf{x}-\mathbf{x}'|)\varphi(x_1', x_2') + \begin{cases}\exp(-a_1x_1 - a_2x_2), & x_1, x_2 \ge 0\\ 0, & \text{otherwise}\end{cases}. \quad (2.4)$$

If φ is determined for arbitrary a_1 and a_2 in a strip in the complex plane, then φ plays the role of a Green's function in transform space. Specifically, if in (2.2) we let $V(x_1, x_2)$ denote the inhomogeneous term

$$V(x_1, x_2) = \int_0^\infty \int_0^\infty dx_1' \, dx_2' K(|\mathbf{x} - \mathbf{x}'|) S(x_1', x_2')$$

and $\tilde{V}(a_1, a_2)$ be the double Laplace transform of $V(x_1, x_2)$,

$$\tilde{V}(a_1, a_2) = \int_0^\infty \int_0^\infty dx_1 \, dx_2 \exp(-a_1 x_1 - a_2 x_2) V(x_1, x_2),$$

then we assert without proof that the solution to (2.2) may be given in terms of $\tilde{V}$ and φ as follows:

$$\rho(x_1, x_2) = \left(\frac{1}{2\pi i}\right)^2 \int_{\Gamma_1} \int_{\Gamma_2} \varphi(x_1, x_2; a_1, a_2) \times \tilde{V}(-a_1, -a_2) \, da_1 \, da_2,$$

where Γ_1 and Γ_2 are vertical contours to the left of all the singularities of $\tilde{V}(-a_1, -a_2)$ and to the right of all singularities of $\varphi(x_1, x_2; a_1, a_2) \equiv \varphi(x_1, x_2)$.

We now use a simple device to transform the integral over the quarter plane to an integral over the whole (x_1, x_2) plane so that we may make use of the convolution theorem of Fourier transforms. Let

$$\varphi_i(x_1, x_2) = \varphi(x_1, x_2) \chi_i(x_1, x_2), \tag{2.5}$$

where χ_i is the characteristic function of the ith quadrant:

$$\chi_1(x_1, x_2) = \begin{cases} 1, & 0 \le x_1 \le \infty, \quad 0 \le x_2 \le \infty \\ 0, & \text{elsewhere} \end{cases};$$

$$\chi_2(x_1, x_2) = \begin{cases} 1, & -\infty \le x_1 < 0, \quad 0 \le x_2 \le \infty \\ 0, & \text{elsewhere} \end{cases};$$

$$\chi_3(x_1, x_2) = \begin{cases} 1, & -\infty \le x_1 < 0, \quad -\infty \le x_2 < 0 \\ 0, & \text{elsewhere} \end{cases};$$

$$\chi_4(x_1, x_2) = \begin{cases} 1, & 0 \le x_1 \le \infty, \quad -\infty \le x_2 < 0 \\ 0, & \text{elsewhere} \end{cases}.$$

In terms of the φ_i, Eq. (2.2) may be rewritten as

$$\varphi_1 + \varphi_2 + \varphi_3 + \varphi_4 = c \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} dx_1' \, dx_2' K(|\mathbf{x} - \mathbf{x}'|) \varphi_1(x_1', x_2') + \chi_1(x_1, x_2) \exp(-a_1 x_1 - a_2 x_2). \tag{2.6}$$

Taking the double Fourier transform of the above equation, we find

$$\{1 - c\hat{K}[(k_1^2 + k_2^2)^{\frac{1}{2}}]\} \Phi_1(k_1, k_2) = -\Phi_2(k_1, k_2) - \Phi_3(k_1, k_2) - \Phi_4(k_1, k_2) + 1/(a_1 - ik_1)(a_2 - ik_2), \tag{2.7}$$

where the Φ_i are the double transforms of the φ_i:

$$\Phi_i(k_1, k_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} dx_1 \, dx_2 \times \exp(ik_1 x_1 + ik_2 x_2) \varphi_i(x_1, x_2), \quad i = 1, \cdots, 4, \tag{2.8}$$

and

$$\hat{K}(k) = (\tan^{-1} k)/k. \tag{2.9}$$

With $c < 1$ and Re $(a_1, a_2) \ge 0$, $\varphi_1(x_1, x_2)$ will be a bounded function of x_1 and x_2. From (2.3) and (2.4), the Φ_i will be analytic in the following sets of half-planes:

Φ_1: $\{\text{Im}\, k_1 > 0, \text{Im}\, k_2 > 0\}$,

Φ_2: $\{\text{Im}\, k_1 < 1, \text{Im}\, k_2 > 0\}$,

Φ_3: $\{\text{Im}\, k_1 < 1, \text{Im}\, k_2 < 0\} \cup \{\text{Im}\, k_1 < 0, \text{Im}\, k_2 < 1\} \cup \{(\text{Im}\, k_1)^2 + (\text{Im}\, k_2)^2 < 1\}$,

Φ_4: $\{\text{Im}\, k_1 > 0, \text{Im}\, k_2 < 1\}$.

This is illustrated in Fig. 1. Note that all four of the Φ_i have a common tube of analyticity:

$$T_\Phi = \{\text{Im}\, k_1 > 0, \text{Im}\, k_2 > 0\} \cap \{(\text{Im}\, k_1)^2 + (\text{Im}\, k_2)^2 < 1\}. \tag{2.10}$$

The goal is to determine $\Phi_1(k_1, k_2)$, which by inverse Fourier transformation gives the neutron density in the quarter space. What we have, then, is

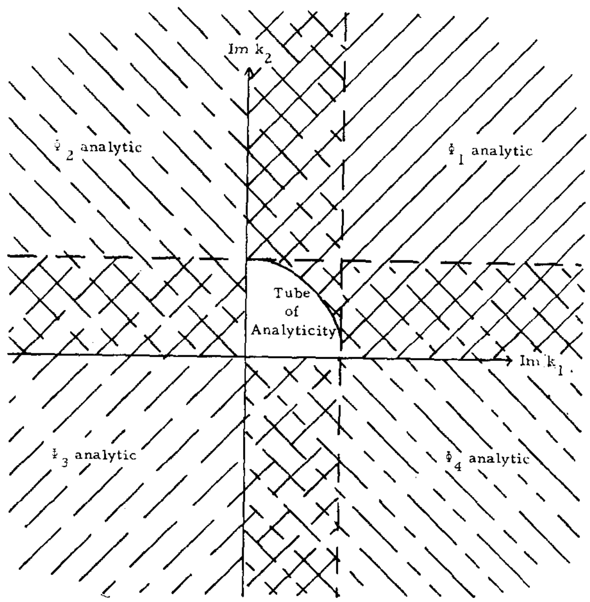


FIG. 1. Domains of analyticity of the $\Phi_i(k_1, k_2)$ with Re a_1, Re $a_2 \ge 0$.

the analogous Wiener–Hopf problem for two complex variables. Instead of two unknown functions which are analytic in a common strip and each analytic in opposite half-planes, we have four functions $\{\Phi_i, i = 1, 2, 3, 4\}$, which are analytic in the common tube T_Φ, each being analytic in a respective set of half-planes. Unfortunately, the factorization of the function $1 - c\hat{K}$, although possible and very useful, as will be shown, does not seem to yield a closed form solution, as it does in the one-dimensional problem.

3. BOCHNER DECOMPOSITION OF $1 - c\hat{K}$

In the analogous one-dimensional problem, one factors the function

$$\Lambda(k) = 1 - c\hat{K}(k)$$

into a product of two functions $H^+(k)$ and $H^-(k)$, respectively analytic in the upper and lower k plane with a common strip of analyticity corresponding to the domains of analyticity of the unknowns Φ^+ and Φ^-. This is the Wiener–Hopf factorization, which, of course, is the key to the Wiener–Hopf technique. Subsequent manipulations and application of the Liouville theorem yield the closed-form solution to the one-dimensional problem.

In the present two-dimensional problem, we intend to make use of the factorization of

$$\Lambda[(k_1^2 + k_2^2)^{\frac{1}{2}}] \equiv 1 - cK((k_1^2 + k_2^2)^{\frac{1}{2}}) \tag{3.1}$$

into a product of four functions:

$$\Lambda = H_1 H_2 H_3 H_4, \tag{3.2}$$

where, as we shall see, the H_i have the following regions of analyticity:

$$\begin{aligned} &H_1: \{\mathrm{Im}\, k_1 \geq 0, \mathrm{Im}\, k_2 \geq 0\} \cup T_\Lambda, \\ &H_2: \{\mathrm{Im}\, k_1 \leq 0, \mathrm{Im}\, k_2 \geq 0\} \cup T_\Lambda, \\ &H_3: \{\mathrm{Im}\, k_1 \leq 0, \mathrm{Im}\, k_2 \leq 0\} \cup T_\Lambda, \\ &H_4: \{\mathrm{Im}\, k_1 \geq 0, \mathrm{Im}\, k_2 \leq 0\} \cup T_\Lambda, \end{aligned} \tag{3.3a}$$

where T_Λ is the tube

$$T_\Lambda = \{(\mathrm{Im}\, k_1)^2 + (\mathrm{Im}\, k_2)^2 < \varkappa_0^2\}, \tag{3.3b}$$

and $\varkappa_0$ satisfies

$$1 - (c/\varkappa_0) \tanh^{-1} \varkappa_0 = 0. \tag{3.4}$$

The conditions for the existence of such a factorization and the method of calculation are given in a theorem of Bochner.[4] Let $f(k_1, k_2)$ be analytic and of bounded L_2 norm in a tube T, $\beta_j \leq \mathrm{Im}\, k_j \leq \alpha_j$. The L_2 norm of f is defined by

$$\|f\|_2 = \left(\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |f(\eta_1 + i\xi_1, \eta_2 + i\xi_2)|^2 \, d\eta_1 \, d\eta_2\right)^{\frac{1}{2}}, \tag{3.5}$$

where the integration is confined to the tube T. According to Bochner, this function is uniquely decomposable (up to additive constants) into a sum of four functions, $f = f_1 + f_2 + f_3 + f_4$, each of which is analytic and bounded in respective radial tubular domains:

$$\begin{aligned} &f_1: \{\mathrm{Im}\, k_1 > \beta_1, \mathrm{Im}\, k_2 > \beta_2\}, \\ &f_2: \{\mathrm{Im}\, k_1 < \alpha_1, \mathrm{Im}\, k_2 > \beta_2\}, \\ &f_3: \{\mathrm{Im}\, k_1 < \alpha_1, \mathrm{Im}\, k_2 < \alpha_2\}, \\ &f_4: \{\mathrm{Im}\, k_1 > \beta_1, \mathrm{Im}\, k_2 < \alpha_2\}. \end{aligned}$$

The f_i may be given in terms of Cauchy integrals. Letting $[f]_{\sigma_i}\pm$ denote the following integrals of f,

$$[f]_{\sigma_1}\pm = \frac{1}{2\pi i}\int_{\Gamma_1\pm} \frac{f(z_1, k_2)\, dz_1}{z_1 - k_1}, \tag{3.6}$$

$$[f]_{\sigma_2}\pm = \frac{1}{2\pi i}\int_{\Gamma_2\pm} \frac{f(k_1, z_2)\, dz_2}{z_2 - k_2}, \tag{3.7}$$

where the contours $\Gamma_j^\pm$ are depicted in Fig. 2, we find that the f_j are given by

$$f_1 = [f]_{\sigma_1^+\sigma_2^+}, \tag{3.8a}$$

$$f_2 = [f]_{\sigma_1^-\sigma_2^+}, \tag{3.8b}$$

$$f_3 = [f]_{\sigma_1^-\sigma_2^-}, \tag{3.8c}$$

$$f_4 = [f]_{\sigma_1^+\sigma_2^-}. \tag{3.8d}$$

To obtain the product decomposition of $\Lambda = 1 - c\hat{K}$, one must first take the logarithm and determine the additive decomposition of $\ln(\Lambda)$. The desired result is then obtained by exponentiation. For convenience, however, we shall reduce the decomposition problem for Λ to one that has already been considered by Kraut in his analysis of an elastic wave propagation problem.[7]

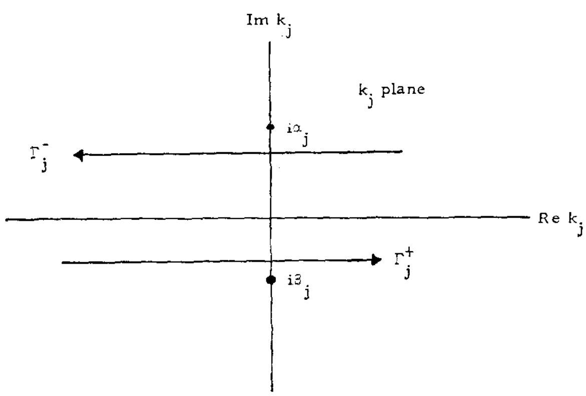


FIG. 2. Integration paths in the tubular domain T.

We have the following integral representation of $\Lambda(k)$:

$$\Lambda(k) = \frac{1-c}{\varkappa_0^2}\frac{k^2+\varkappa_0^2}{k^2+1}\exp\left[\frac{2k^2}{\pi}\int_1^\infty \frac{\theta(1/t)\,dt}{(t^2+k^2)t}\right], \quad k^2 = k_1^2 + k_2^2, \quad (3.9)$$

where θ is given by

$$\theta(\sigma) = \tan^{-1}(c\pi\sigma/2\lambda(\sigma)), \quad \theta(0) = 0, \quad (3.10)$$

$$\lambda(\sigma) = 1 - c\sigma\tanh^{-1}\sigma. \quad (3.11)$$

This is a simple variation of a representation well known in one-dimensional transport theory.[8] A more convenient expression is given by

$$\Lambda(k) = \exp\left(-\frac{2}{\pi}\int_1^\infty \frac{\theta(1/t)t\,dt}{(t^2+k^2)} - 2\int_{\varkappa_0}^1 \frac{t\,dt}{t^2+k^2}\right). \quad (3.12)$$

Thus, to achieve the product decomposition of Λ, we see by the above integral representation that we need only additively decompose the function

$$V(k_1, k_2; t) = 1(/t^2 + k_1^2 + k_2^2). \quad (3.13)$$

Since t is real and $\geq \varkappa_0$, V is analytic in the tube T_Λ [Eq. (3.3b)]. It is an easy matter to verify that V also has bounded L_2 norm in T_Λ. By a simple modification of a calculation by Kraut,[7] we obtain the following decomposition:

$$1/(t^2 + k_1^2 + k_2^2) = V_1(k_1, k_2; t) + V_2(k_1, k_2; t) + V_3(k_1, k_2; t) + V_4(k_1, k_2; t), \quad (3.14)$$

with

$$V_1(k_1, k_2; t) = \frac{1}{t^2 + k_1^2 + k_2^2} \times \left[\frac{1}{4} - \frac{k_1}{2\pi(k_2^2+t^2)^{\frac{1}{2}}}\ln\left(\frac{k_2 + (k_2^2+t^2)^{\frac{1}{2}}}{it}\right) - \frac{k_2}{2\pi(k_1^2+t^2)^{\frac{1}{2}}}\ln\left(\frac{k_1+(k_1^2+t^2)^{\frac{1}{2}}}{it}\right)\right]; \quad (3.15)$$

$$V_2(k_1, k_2; t) = V_1(-k_1, k_2; t), \quad (3.16)$$

$$V_3(k_1, k_2; t) = V_1(-k_1, -k_2; t), \quad (3.17)$$

$$V_4(k_1, k_2; t) = V_1(k_1, -k_2; t). \quad (3.18)$$

In the above, the principle branch of the logarithm is to be taken, and we will arbitrarily choose the branches of the radicals so that $(k_i^2)^{\frac{1}{2}} = +k_i$. Thus, the H_i, which were defined in the factorization of $\Lambda(k)$ [Eq. (3.2)], are given explicitly by

$$H_i(k_1, k_2) = \exp\left(-\frac{2}{\pi}\int_1^\infty \theta\left(\frac{1}{t}\right)V_i(k_1, k_2; t)t\,dt - 2\int_{\varkappa_0}^1 V_i(k_1, k_2; t)t\,dt\right). \quad (3.19)$$

One can now verify without much difficulty that the H_i, as given above, are analytic in their respective domains as per (3.3).

4. SUBTRACTION OF THE ASYMPTOTIC SOLUTION

With the definition (3.1), Eq. (2.7) becomes

$$\Lambda(k_1, k_2)\Phi_1(k_1, k_2) = -\Phi_2(k_1, k_2) - \Phi_3(k_1, k_2) - \Phi_4(k_1, k_2) + 1/(a_1 - ik_1)(a_2 - ik_2), \quad (4.1)$$

where, here and in the following, $\Lambda(k_1, k_2) = \Lambda[(k_1^2 + k_2^2)^{\frac{1}{2}}]$. Using a Bochner decomposition on the term $\Lambda\Phi_1$,[9] we can derive a set of integral equations relating Φ_1, Φ_2, Φ_3, and Φ_4. The decomposition yields

$$\Lambda\Phi_1 = [\Lambda\Phi_1]_1 + [\Lambda\Phi_1]_2 + [\Lambda\Phi_1]_3 + [\Lambda\Phi_1]_4, \quad (4.2)$$

where the operations $[\ ]_i$, $i = 1, \cdots, 4$, are defined by Eqs. (3.6)–(3.8), with the corresponding contours in Fig. 2 confined to the tube T_Φ. The uniqueness of this decomposition allows us to equate terms on the rhs of (4.2) to corresponding terms on the rhs of (4.1) as follows:

$$[\Lambda\Phi_1]_1 = 1/(a_1 - ik_1)(a_2 - ik_2), \quad (4.3)$$

$$[\Lambda\Phi_1]_j = -\Phi_j, \quad j = 2, 3, 4. \quad (4.4)$$

If in (4.3) [with Re $(a_1, a_2) > 0$] the integration contours are taken to be the real axes and k_1 and k_2 approach these contours (from above), we obtain a singular integral equation for $\Phi_1(k_1, k_2)$ on the real axes:

$$\tfrac{1}{4}\Lambda(k_1, k_2)\Phi_1(k_1, k_2) + \frac{1}{4\pi i}\int_{-\infty}^{+\infty}\frac{\Lambda(z_1, k_2)\Phi_1(z_1, k_2)\,dz_1}{z_1 - k_1} + \frac{1}{4\pi i}\int_{-\infty}^{+\infty}\frac{\Lambda(k_1, z_2)\Phi_1(k_1, z_2)\,dz_2}{z_2 - k_2} - \frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{\Lambda(z_1, z_2)\Phi_1(z_1, z_2)\,dz_1\,dz_2}{(z_1 - k_1)(z_2 - k_2)} = \frac{1}{(a_1 - ik_1)(a_2 - ik_2)}, \quad (4.5)$$

where the integrals are computed as principal values. If one could solve the above equation for Φ_1, then Φ_2, Φ_3, and Φ_4 would follow from (4.4).

In a study of diffraction of electromagnetic waves from a quarter space, Kraut and Lehman[3] encounter similar mathematical problems. They derive an equation analogous to (4.5) and prove that the solution may be obtained by iteration if a certain parameter is less than unity. If this parameter is close to 1, the calculated convergence rate is slow. We could do the same here for $c < 1$. Instead, we will derive an iterative scheme which takes advantage of the analyticity of the various functions and which yields a convergence rate which is relatively fast regardless of the value of c. The transform of the asymptotic flux distribution (away from the corner) is the zeroth-order term in this scheme. Higher-order terms produce a significant correction only very near the corner.

To begin the analysis of (4.1), we redefine the problem in terms of a new set of unknowns $\hat{\Phi}_i$, $i = 1, \cdots, 4$, which represent differences between the Φ_i and the transforms of the asymptotic distributions. The asymptotic distributions are derived in Appendix A by assuming, for example, that if x_2 is large and positive, then, under certain conditions on (a_1, a_2), the purely absorbing quarter space $(x_2 < 0,\ x_1 > 0)$ can be replaced by a medium which has the same properties as quarter space (Q) occupying the positive quadrant. The result is a one-dimensional Wiener–Hopf equation which is easily solved in both cases (large x_1, large x_2).

Referring to (A7) and (A12), we see that Φ_1, given by

$$\begin{aligned}\Phi_1(k_1, k_2) = 1/[(a_1 - ik_1)(a_2 - ik_2)H_1(k_1, k_2)\\ \times H_2(-ia_1, k_2)H_3(-ia_1, -ia_2)\\ \times H_4(k_1, -ia_2)] + \hat{\Phi}_1(k_1, k_2), \quad (4.6)\end{aligned}$$

will produce the desired asymptotic behavior for large x_1 or x_2 with $\hat{\Phi}_1$ analytic in the upper (k_1, k_2) planes and yielding the correction near the corner. Similarly, to get correct asymptotic behavior in the quarter planes adjacent to the position quadrant, we choose to define $\hat{\Phi}_2$ and $\hat{\Phi}_4$ as follows:

$$\Phi_2(k_1, k_2) = \frac{1}{(a_1 - ik_1)}\frac{1}{(a_2 - ik_2)} \times \left(1 - \frac{H_2(k_1, k_2)H_3(k_1, -ia_2)}{H_2(-ia_1, k_2)H_3(-ia_1, -ia_2)}\right) + \hat{\Phi}_2(k_1, k_2), \quad (4.7)$$

$$\Phi_4(k_1, k_2) = \frac{1}{(a_1 - ik_1)}\frac{1}{(a_2 - ik_2)} \times \left(1 - \frac{H_4(k_1, k_2)H_3(-ia_1, k_2)}{H_4(k_1, -ia_2)H_3(-ia_1, -ia_2)}\right) + \hat{\Phi}_4(k_1, k_2). \quad (4.8)$$

Note that in (4.7) we have substituted the ratio $H_2(k_1, k_2)/H_2(-ia_1, k_2)$ in preference to the choice indicated by (A13), namely, $H_2(k_1, -ia_2)/H_2(-ia_1, -ia_2)$, and have made a similar substitution in (4.8). The reason for these changes as well as the definition of Φ_3,

$$\Phi_3(k_1, k_2) = -\frac{1}{(a_1 - ik_1)}\frac{1}{(a_2 - ik_2)} \times \left(1 - \frac{H_3(-ia_1, k_2)H_3(k_1, -ia_2)}{H_3(k_1, k_2)H_3(-ia_1, -ia_2)}\right) + \hat{\Phi}_3(k_1, k_2), \quad (4.9)$$

is given below.

The problem, now defined in terms of hatted variables, is

$$\begin{aligned}\Lambda(k_1, k_2)\hat{\Phi}_1(k_1, k_2) = -\hat{\Phi}_2(k_1, k_2) - \hat{\Phi}_3(k_1, k_2)\\ - \hat{\Phi}_4(k_1, k_2) + \hat{S}(k_1, k_2), \quad (4.10)\end{aligned}$$

where (4.6)–(4.9) were used in (4.1) and the source term $\hat{S}$ is given by

$$\begin{aligned}\hat{S}(k_1, k_2) = \frac{1}{(a_1 - ik_1)(a_2 - ik_2)H_3(-ia_1, -ia_2)}\\ \times \Bigg(-\frac{H_2(k_1, k_2)H_3(k_1, k_2)H_4(k_1, k_2)}{H_2(-ia_1, k_2)H_4(k_1, -ia_2)}\\ + \frac{H_2(k_1, k_2)H_3(k_1, -ia_2)}{H_2(-ia_1, k_2)}\\ - \frac{H_3(-ia_1, k_2)H_3(k_1, -ia_2)}{H_3(k_1, k_2)}\\ + \frac{H_4(k_1, k_2)H_3(-ia_1, k_2)}{H_4(k_1, -ia_2)}\Bigg). \quad (4.11)\end{aligned}$$

Our motivation for making the choice of $\hat{\Phi}_2$, $\hat{\Phi}_3$, and $\hat{\Phi}_4$, as defined by (4.7)–(4.9), becomes clearer by noting that the residues of $\hat{S}$ vanish identically at $k_1 = -ia_1$ and at $k_2 = -ia_2$. Also, because

$$\lim_{k_j \to \infty} H_i(k_1, k_2) = 1 + O(\ln k_j/k_j), \quad j = 1, 2,\ i = 1, \cdots, 4, \quad (4.12)$$

we find that

$$\lim_{k_j \to \infty} \hat{S}(k_1, k_2) = O(\ln k_j/k_j^2), \quad j = 1, 2. \quad (4.13)$$

This latter fact will be useful in our derivation of a convergent iterative solution.

5. SOLUTION FOR $\hat{\Phi}_1$ IN TERMS OF $\hat{\Phi}_2$ AND $\hat{\Phi}_4$

We now demonstrate that assuming that $\hat{\Phi}_2$ and $\hat{\Phi}_4$ are known leads directly to the solution for $\hat{\Phi}_1$. Later we shall determine $\hat{\Phi}_2$ and $\hat{\Phi}_4$. For this purpose and many of the remaining calculations we require the following factorization of Λ in the variable k_2:

$$\Lambda(k_1, k_2) = (\varkappa_0^2 + k_1^2 + k_2^2)E(k_2, k_1)E(-k_2, k_1), \quad (5.1)$$

where

$$E(k_2, k_1) = \exp\left(-\frac{1}{\pi}\int_1^\infty \frac{\theta(1/t)t\,dt}{(t^2 + k_1^2)^{\frac{1}{2}}[(t^2 + k_1^2)^{\frac{1}{2}} - ik_2]}\right)\Big/(1 + k_1^2)^{\frac{1}{2}} - ik_2. \tag{5.2}$$

The above result follows most simply from (3.12). The function $E(k_2, k_1)$ is analytic in the upper k_2 plane while $E(-k_2, k_1)$ is analytic in the lower k_2 plane.

With the substitution (5.1), Eq. (4.10) reads

$$(\varkappa_0^2 + k_1^2 + k_2^2)E(k_2, k_1)E(-k_2, k_1)\hat{\Phi}_1 = -\hat{\Phi}_2 - \hat{\Phi}_3 - \hat{\Phi}_4 + \hat{S}. \tag{5.3}$$

Before proceeding further, we need to derive some useful relations between the $\hat{\Phi}_i$ at infinity. Multiplying (5.3) by k_1 and letting $k_1 \to \infty$, we find that

$$\nu_1(k_2) = -\nu_2(k_2) - \nu_3(k_2) - \nu_4(k_2), \tag{5.4}$$

where

$$\nu_j(k_2) = \lim_{k_1 \to \infty} ik_1\hat{\Phi}_j(k_1, k_2), \quad j = 1, 2, 3, 4, \tag{5.5}$$

and where we have used (4.13) and the fact that $\Lambda \to 1$ when either k_1 or $k_2 \to \infty$. Because ν_1 and ν_2 are analytic in the upper half-plane, because ν_3 and ν_4 are analytic in the lower half-plane, and because all the ν_j vanish at infinity, we must have that

$$\nu_1(k_2) + \nu_2(k_2) = 0, \tag{5.6}$$

$$\nu_3(k_2) + \nu_4(k_2) = 0. \tag{5.7}$$

Similarly, if we define the limits

$$\mu_j(k_1) = \lim_{k_2 \to \infty} ik_2\hat{\Phi}_j(k_1, k_2), \quad j = 1, 2, 3, 4, \tag{5.8}$$

then we find

$$\mu_1(k_1) + \mu_4(k_1) = 0, \tag{5.9}$$

$$\mu_2(k_1) + \mu_3(k_1) = 0. \tag{5.10}$$

Now, dividing Eq. (5.3) by $E(-k_2, k_1)$ and performing the operation $[\ ]_{\sigma_2+}$ [see Eqs. (3.6)–(3.8)] on the result yields

$$(\varkappa_0^2 + k_1^2 + k_2^2)E(k_2, k_1)\hat{\Phi}_1 + \mu_4(k_1) = \left[\frac{\hat{S} - \hat{\Phi}_2}{E(-k_2, k_1)}\right]_{\tilde{\sigma}_2+}, \tag{5.11}$$

where the operations $[\ ]_{\tilde{\sigma}_i+}$ correspond to the contours $\tilde{\Gamma}_i^{\pm}$, as shown in Fig. 3 and where we have used the facts that (1) the lhs of (5.11) is analytic in the upper k_2 plane and is square-integrable in T_Φ [because of (5.9)] and (2) the same is true of $(\hat{\Phi}_3 + \hat{\Phi}_4)/E(-k_2, k_1)$ except that it is analytic in the lower k_2 plane.

Now dividing (5.11) by $(\varkappa_0^2 + k_1^2 + k_2^2)E(k_2, k_1)$ gives, for $\hat{\Phi}_1$,

$$\hat{\Phi}_1(k_1, k_2) = \frac{1}{(\varkappa_0^2 + k_1^2 + k_2^2)E(k_2, k_1)} \times \left\{\left[\frac{\hat{S} - \hat{\Phi}_2}{E(-k_2, k_1)}\right]_{\sigma_2+} - \mu_4(k_1)\right\}. \tag{5.12}$$

Thus $\hat{\Phi}_1$ is expressed in terms of known functions and the unknown functions $\hat{\Phi}_2$ and μ_4.

A similar expression may be developed in terms of $\hat{\Phi}_4$ and ν_2 by using the factorization of Λ in k_1:

$$\Lambda(k_1, k_2) = (\varkappa_0^2 + k_1^2 + k_2^2)E(k_1, k_2)E(-k_1, k_2). \tag{5.13}$$

We find, analogous to (5.12), that $\hat{\Phi}_1$ may be represented as

$$\hat{\Phi}_1(k_1, k_2) = \frac{1}{(\varkappa_0^2 + k_1^2 + k_2^2)E(-k_1, k_2)} \times \left\{\left[\frac{\hat{S} - \hat{\Phi}_4}{E(-k_1, k_2)}\right]_{\sigma_1} - \nu_2(k_2)\right\}. \tag{5.14}$$

Of course, at this point both $\hat{\Phi}_2$ and $\hat{\Phi}_4$ are unknown. The above expressions for $\hat{\Phi}_1$ were simply derived by a modified one-dimensional Wiener–Hopf analysis, the modification consisting of factoring out the zeros of Λ and using the Wiener–Hopf factorization of the remainder. In the next section, however, a coupled pair of Fredholm equations for $\hat{\Phi}_2$ and $\hat{\Phi}_4$ will be derived which will be shown to be solvable by iteration. Once this is done, we will return to (5.12) or (5.14) to obtain $\hat{\Phi}_1$.

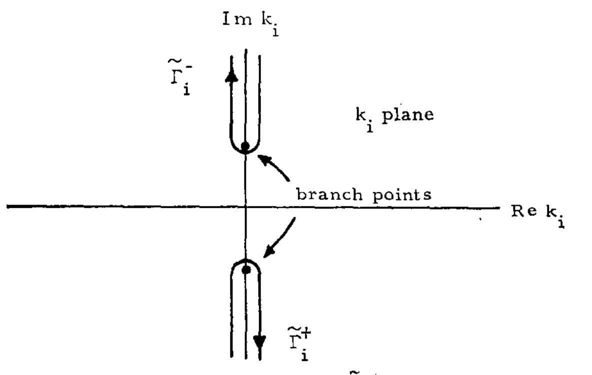


FIG. 3. The contours $\tilde{\Gamma}_i^{\pm}$.

6. COUPLED FREDHOLM EQUATIONS FOR $\hat{\Phi}_2$ AND $\hat{\Phi}_4$

We return to (5.13) and divide by $E(-k_1, k_2)$. This time, however, we perform the operation $[\ ]_{\sigma_1^-}$ and find

$$0 = \left[\frac{\hat{S} - \hat{\Phi}_4}{E(-k_1, k_2)}\right]_{\tilde{\sigma}_1^-} - \frac{\hat{\Phi}_2 + \hat{\Phi}_3}{E(-k_1, k_2)} + \nu_2 - \nu_4. \quad (6.1)$$

Now we multiply the above equation by $E(-k_1, k_2)$ and perform the operation $[\ ]_{\sigma_2^+}$. This filters out the term $\hat{\Phi}_3$, which is analytic in the lower k_2 plane, giving

$$\hat{\Phi}_2(k_1, k_2) = -\left[E(-k_1, k_2)\left(\left[\frac{\hat{\Phi}_4 - \hat{S}}{E(-k_1, k_2)}\right]_{\tilde{\sigma}_1^-} + \nu_2 - \nu_4\right)\right]_{\sigma_2^+}. \quad (6.2)$$

Another equation relating $\hat{\Phi}_2$ and $\hat{\Phi}_4$ may be derived by dividing by $E(-k_2, k_1)$ in (5.3) and performing the operation $[\ ]_{\sigma_2^-}$. We find

$$0 = \left[\frac{\hat{S} - \hat{\Phi}_2}{E(-k_2, k_1)}\right]_{\tilde{\sigma}_2^-} - \frac{\hat{\Phi}_3 + \hat{\Phi}_4}{E(-k_2, k_1)} - \mu_2 + \mu_4. \quad (6.3)$$

Multiplying (6.3) by $E(-k_2, k_1)$ and performing $[\ ]_{\sigma_1^+}$ gives the second desired equation:

$$\hat{\Phi}_4(k_1, k_2) = -\left[E(-k_2, k_1)\left(\left[\frac{\hat{\Phi}_2 - \hat{S}}{E(-k_2, k_1)}\right]_{\tilde{\sigma}_2^-} - \mu_2 + \mu_4\right)\right]_{\sigma_1^+}. \quad (6.4)$$

The contours Γ_i^+ (corresponding to the operations $[\ ]_{\sigma_i^+}$) are as shown in Fig. 2 and, for the moment, are confined to the tube T_Λ. This is sufficient to ensure analyticity of $S(z_1, z_2)$ if $(ia_1, ia_2) \in T_\Lambda \cap \{\mathrm{Im}\,(ia_1), \mathrm{Im}\,(ia_2) \geq 0\}$.

If (6.2) is evaluated for $k_1 \in \Gamma_1^+$ and $k_2 \in \tilde{\Gamma}_2^-$ to obtain $\hat{\Phi}_2(k_1, k_2)$ and (6.4) is evaluated for $k_1 \in \tilde{\Gamma}_1^-$ and $k_2 \in \Gamma_2^+$ to obtain $\hat{\Phi}_4(k_1, k_2)$, then the two equations represent a pair of coupled Fredholm equations for $\hat{\Phi}_2$ and $\hat{\Phi}_4$. The quantities $1/(z_i - k_i)$ are clearly bounded. In contemplating a Neumann series solution to (6.2) and (6.4), it would be advantageous to make $1/(z_i - k_i)$ as small as possible by lowering the Γ_i^+ contours as far as possible into the lower half-planes. Guided by our experience in one-dimensional problems, we might hope to achieve deformation of the contours Γ_i^+ to $\tilde{\Gamma}_i^+$ shown in Fig. 3, which possibly would have the additional advantage of producing a real-valued kernel. However, this deformation is not possible for the following reason. In (6.2), for example, if the integration variable z_2 approaches the line $(-i\infty, -i)$, then the branch point of $E(-z_1, z_2)$ in the z_1 plane approaches the real z_1 axis at a point depending on the value of z_2 [refer to the representation of $E(-z_1, z_2)$ given by (5.2)]. This is shown in Fig. 4. Thus it would not be possible to deform Γ_1^+ and Γ_2^+, as shown in Fig. 3, because the contours $\tilde{\Gamma}_1^-$ and $\tilde{\Gamma}_2^-$ could not be preserved.

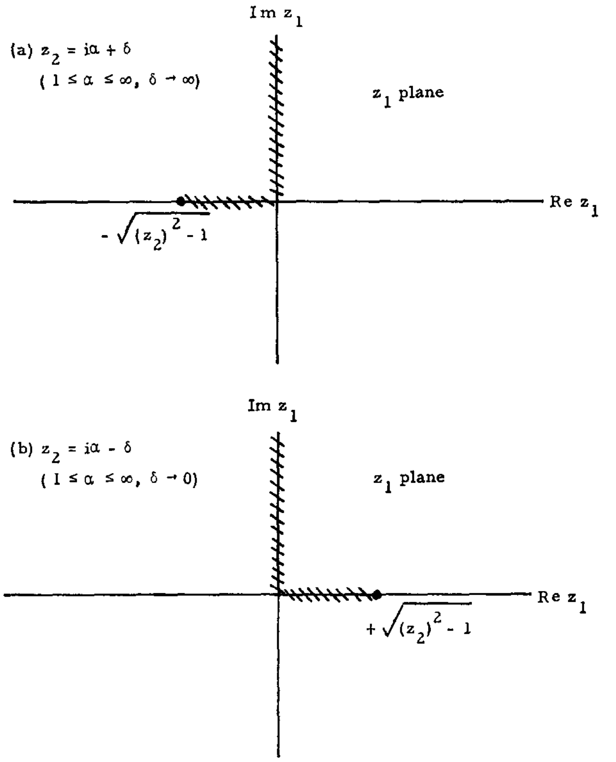


FIG. 4. Possible choice of branch cuts of $E(-z_1, z_2)$.

An iterative scheme which avoids this difficulty and leads to real-valued kernels will now be derived. First, we note that because

$$\left[\frac{\hat{\Phi}_4}{E(-k_1, k_2)}\right]_{\tilde{\sigma}_1^-} = \frac{\hat{\Phi}_4}{E(-k_1, k_2)} - \mu_4 - \left[\frac{\hat{\Phi}_4}{E(-k_1, k_2)}\right]_{\tilde{\sigma}_1^+}, \quad (6.5)$$

it follows that (6.2) may be rewritten as

$$\hat{\Phi}_2(k_1, k_2) = \left[E(-k_1, k_2)\left[\frac{\hat{\Phi}_4}{E(-k_1, k_2)}\right]_{\tilde{\sigma}_1^+}\right]_{\sigma_2^+} + R_2(k_1, k_2), \quad (6.6)$$

and, similarly, (6.4) may be rewritten as

$$\hat{\Phi}_4(k_1, k_2) = \left[E(-k_2, k_1)\left[\frac{\hat{\Phi}_2}{E(-k_2, k_1}\right]_{\tilde{\sigma}_2{}^+}\right]_{\sigma_1{}^+} + R_4(k_1, k_2), \tag{6.7}$$

where

$$R_j(k_1, k_2) = T_j(k_1, k_2) + U_j(k_1, k_2), \quad j = 2, 4, \tag{6.8}$$

and

$$T_2(k_1, k_2) = \left(\frac{1}{2\pi i}\right)^2 \int_{\Gamma_2{}^+} \frac{dz_2 E(-k_1, z_2)}{(z_2 - k_2)} \int_{\tilde{\Gamma}_1{}^-} \frac{dz_1 \hat{S}(z_1, z_2)}{E(-z_1, z_2)(z_1 - k_1)}, \tag{6.9}$$

$$T_4(k_1, k_2) = \left(\frac{1}{2\pi i}\right)^2 \int_{\Gamma_1{}^+} \frac{dz_1 E(-k_2, z_1)}{(z_1 - k_1)} \int_{\tilde{\Gamma}_2{}^-} \frac{dz_2 \hat{S}(z_1, z_2)}{E(-z_2, z_1)(z_2 - k_2)}, \tag{6.10}$$

$$U_2(k_1, k_2) = \frac{1}{2\pi i} \int_{\Gamma_2{}^+} \frac{dz_2 E(-k_1, z_2) \nu_2(z_2)}{z_2 - k_2}, \tag{6.11}$$

$$U_4(k_1, k_2) = \frac{1}{2\pi i} \int_{\Gamma_1{}^+} \frac{dz_1 E(-k_2, z_1) \mu_4(z_1)}{z_1 - k_1}. \tag{6.12}$$

The contours in (6.6) and (6.7) may now be deformed into the lower half-planes. From a careful inspection of the functions R_2, R_4, $E(-k_2, k_1)$, and $E(-k_1, k_2)$, we find that $\hat{\Phi}_2$ and $\hat{\Phi}_4$ have branch cuts as shown in Fig. 5. Accordingly, it is useful to define discontinuities of various functions as follows:

$k_1, z_1 \in (-i\infty, -i0)$:

$$\hat{\Phi}_{2+}(z_1, z_2) - \hat{\Phi}_{2-}(z_1, z_2) \equiv \begin{cases} \psi_2(z_1, z_2), & z_2 \in (-i, -i\varkappa_0) \\ \chi_2(z_1, z_2), & z_2 \in (-i\infty, -i) \end{cases}, \tag{6.13}$$

$$\frac{E_+(-k_1, z_2)}{E_+(-z_1, z_2)} - \frac{E_-(-k_1, z_2)}{E_-(-z_1, z_2)} \equiv G(k_1, z_1, z_2), \quad z_2 \in (-i\infty, -i), \tag{6.14}$$

$k_2, z_2 \in (-i\infty, -i0)$:

$$\hat{\Phi}_4^+(z_1, z_2) - \hat{\Phi}_4^-(z_1, z_2) \equiv \begin{cases} \psi_4(z_1, z_2), & z_1 \in (-i, -i\varkappa_0) \\ \chi_4(z_1, z_2), & z_1 \in (-i\infty, -i) \end{cases}, \tag{6.15}$$

$$\frac{E^+(-k_2, z_1)}{E^+(-z_2, z_1)} - \frac{E^-(-k_2, z_1)}{E^-(-z_2, z_1)} \equiv G(k_2, z_2, z_1), \quad z_1 \in (-i\infty, -i), \tag{6.16}$$

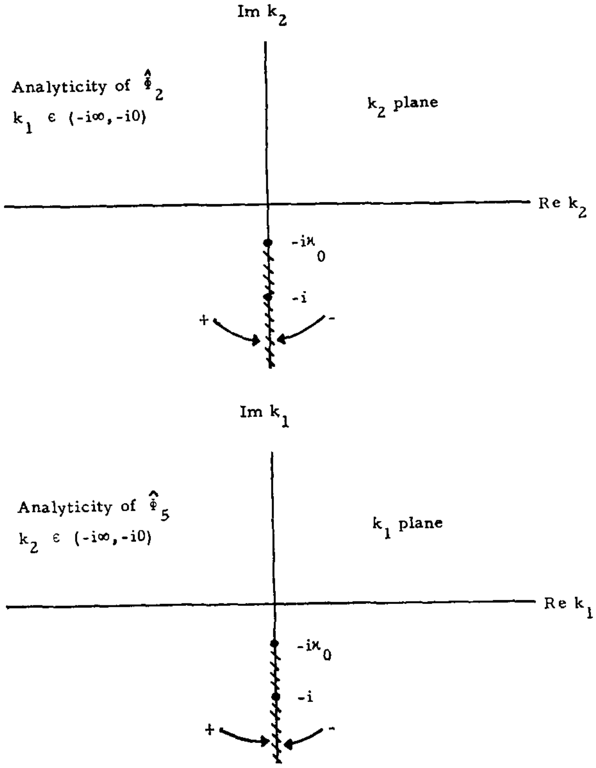


FIG. 5. Branch cuts of $\hat{\Phi}_2$ and $\hat{\Phi}_4$.

where the superscript ($\pm$) denotes a limit in the k_1 plane and the subscript ($\pm$) denotes a limit in the k_2 plane (see Fig. 5).

Collapsing the Γ_i^+ about the negative imaginary axes,[10] we can rewrite (6.6) and (6.7) as

$$\hat{\Phi}_2(k_1, k_2) = \left(\frac{1}{2\pi i}\right)^2 \int_{-i\infty}^{-i} \frac{dz_2}{z_2 - k_2} \times \left(\int_{-i\infty}^{-i} \frac{dz_1 G(k_1, z_1, z_2) \chi_4(z_1, z_2)}{z_1 - k_1} + \int_{-i}^{-i\varkappa_0} \frac{dz_1 G(k_1, z_1, z_2) \psi_4(z_1, z_2)}{z_1 - k_1}\right) + R_2(k_1, k_2), \tag{6.17}$$

$$\hat{\Phi}_4(k_1, k_2) = \left(\frac{1}{2\pi i}\right)^2 \int_{-i\infty}^{-i} \frac{dz_1}{z_1 - k_1} \times \left(\int_{-i\infty}^{-i} \frac{dz_2 G(k_2, z_2, z_1) \chi_2(z_1, z_2)}{z_2 - k_2} + \int_{-i}^{-i\varkappa_0} \frac{dz_2 G(k_2, z_2, z_1) \psi_2(z_1, z_2)}{z_2 - k_2}\right) + R_4(k_1, k_2). \tag{6.18}$$

In Appendix B, integral representations for T_2 and T_4, the known parts of R_2 and R_4, are derived by using formulas analogous to (6.5) in (6.9) and (6.10) with contour integration.

Thus we need only determine ψ_2, χ_2, ν_2, ψ_4, χ_4, and μ_4 to obtain $\hat{\Phi}_2$ and $\hat{\Phi}_4$ from (6.17) and (6.18), respectively. The discontinuities of $\hat{\Phi}_2$ and $\hat{\Phi}_4$ as calculated by just these equations provide four of the equations for the above functions. Computing the discontinuity of (6.17) across the cut $(-i, -i\varkappa_0)$ in the k_2 plane, we find that $\psi_2(k_1, k_2)$ is simply given as the discontinuity of R_2:

$$\psi_2(k_1, k_2) = R_{2+}(k_1, k_2) - R_{2-}(k_1, k_2),$$
$$k_1 \in (-i\infty, -i0), \quad k_2 \in (-i, -i\varkappa_0). \quad (6.19)$$

Similarly, from (6.18) we compute $\psi_4(k_1, k_2)$ to be given by

$$\psi_4(k_1, k_2) = R_4^+(k_1, k_2) - R_4^-(k_1, k_2),$$
$$k_1 \in (-i, -i\varkappa_0), \quad k_2 \in (-i\infty, -i0). \quad (6.20)$$

Next, computing the discontinuities of (6.17) and (6.18) across $(-i\infty, -i)$ in the k_1 and k_2 planes, respectively, yields two coupled Fredholm equations for χ_2 and χ_4:

$$\chi_2(k_1, k_2) = \frac{1}{2\pi i}\int_{-i\infty}^{-i} \frac{dz_1 G(k_1, z_1, k_2)\chi_4(z_1, k_2)}{z_1 - k_1} + [R_{2+}(k_1, k_2) - R_{2-}(k_1, k_2)], \quad (6.21)$$

$$\chi_4(k_1, k_2) = \frac{1}{2\pi i}\int_{-i\infty}^{-i} \frac{dz_2 G(k_2, z_2, k_1)\chi_2(k_1, z_2)}{z_2 - k_2} + [R_4^+(k_1, k_2) - R_4^-(k_1, k_2)]. \quad (6.22)$$

The kernels are real and are continuous because $G(k_i, z_i, k_j)$ is analytic in k_i and vanishes at $k_i = z_i$.

We define the linear operators L_2 and L_4,

$$L_2(\gamma)(k_1, k_2) = \frac{1}{2\pi i}\int_{-i\infty}^{-i} \frac{dz_1 G(k_1, z_1, k_2)\gamma(z_1, k_2)}{z_1 - k_1}, \quad (6.23)$$

$$L_4(\gamma)(k_1, k_2) = \frac{1}{2\pi i}\int_{-i\infty}^{-i} \frac{dz_2 G(k_2, z_2, k_1)\gamma(k_1, z_2)}{z_2 - k_2}, \quad (6.24)$$

and write an iterative scheme for solving (6.21) and (6.22) as follows:

$$\chi_2^{(n)} = L_2(\chi_4^{(n)}) + [R_{2+} - R_{2-}], \quad (6.25)$$

$$\chi_4^{(n)} = L_4(\chi_4^{(n-1)}) + [R_4^+ - R_4^-],$$
$$\chi_4^{(0)} \equiv 0, \quad n = 1, 2, \cdots. \quad (6.26)$$

To justify this procedure, we must demonstrate that the norm of L_4L_2 satisfies $\|L_4L_2\| < 1$ in some Banach space. In Appendix C, we show indeed that this is true for all $0 < c \gtrsim 2$ and that the convergence is uniform.

As a final step, we relate the unknown functions ν_2 and μ_4 to the functions ψ_2, χ_2 and ψ_4, χ_4 as follows. As in the development of (6.2), but using the factorization $\Lambda = (H_1H_4)(H_2H_3)$, we find that

$$\hat{\Phi}_2(k_1, k_2) = \left[H_2H_3\left[\frac{\hat{\Phi}_4}{H_2H_3}\right]_{\tilde{\sigma}_1+}\right]_{\tilde{\sigma}_2+} + \left[H_2H_3\left[\frac{\hat{S}}{H_2H_3}\right]_{\tilde{\sigma}_1-}\right]_{\sigma_2+}. \quad (6.27)$$

Multiplying (6.27) by ik_2 and letting $k_2 \to \infty$, we have

$$\nu_2(k_2) = \frac{i}{4\pi^2}\int_{-i\infty}^{-i\varkappa_0} \frac{dz_2}{z_2 - k_2} \times \int_{-i\infty}^{-i\varkappa_0} \frac{dz_1[\hat{\Phi}_4^+(z_1, z_2) - \hat{\Phi}_4^-(z_1, z_2)]}{H_3(z_1, z_2)} \times \left(\frac{1}{H_{2+}(z_1, z_2)} - \frac{1}{H_{2-}(z_1, z_2)}\right), \quad (6.28)$$

where $\hat{\Phi}_4^+ - \hat{\Phi}_4^-$ is related to ψ_4 and χ_4 by (6.15). Similarly, we find that $\mu_4(k_1)$ is given by

$$\mu_4(k_1) = \frac{i}{4\pi^2}\int_{-i\infty}^{-i\varkappa_0} \frac{dz_1}{z_1 - k_1} \times \int_{-i\infty}^{-i\varkappa_0} \frac{dz_2[\hat{\Phi}_{2+}(z_1, z_2) - \hat{\Phi}_{2-}(z_1, z_2)]}{H_3(z_1, z_2)} \times \left(\frac{1}{H_4^+(z_1, z_2)} - \frac{1}{H_4^-(z_1, z_2)}\right), \quad (6.29)$$

where $\hat{\Phi}_{2+} - \hat{\Phi}_{2-}$ is related to ψ_2 and χ_2 by (6.13).

Equations (6.19)–(6.22) and (6.28) and (6.29) comprise the required set of six equations for the six unknowns ψ_2, χ_2, ν_2, ψ_4, χ_4, and μ_4.

7. CONCLUSION

Assuming an inhomogeneous term $\exp(-a_1x_1 - a_2x_2)$ in the transport equation (2.2), we find that the double Fourier transform of the flux in the quarter space, $\Phi_1(k_1, k_2)$, is given by (4.6) in terms of the transform of the flux away from the corner whose properties are discussed in Appendix A, and a correction $\hat{\Phi}_1(k_1, k_2)$. The function $\hat{\Phi}_1(k_1, k_2)$ is given in terms of $\hat{\Phi}_2(k_1, k_2)$ or $\hat{\Phi}_4(k_1, k_2)$ by (5.12) or (5.14), respectively. The functions $\hat{\Phi}_2$ and $\hat{\Phi}_4$ have singularities only on the imaginary axes and are given by (6.17) and (6.18) in terms of functions $\psi_i(z_1, z_2)$, $\chi_i(z_1, z_2)$, $i = 2, 4$, $\nu_2(k_2)$, and $\mu_4(k_1)$. The ψ_i are given by (6.19) and (6.21), while the χ_i satisfy Fredholm equations (6.21) and (6.22). Uniform convergence to the solution of these equations may be obtained by iteration for values of c: $0 \leq c \lesssim 2$. The functions ν_2 and μ_4 are related to the ψ_i and χ_i by (6.28) and (6.29).

A more direct approach to the solution appears to be possible by obtaining Fredholm equations for Φ_2 and Φ_4 (instead of $\hat{\Phi}_2$ and $\hat{\Phi}_4$) starting from (4.1) and bypassing the subtraction of the asymptotic terms. There are two objections to this approach. First, the asymptotic behavior of the flux does not appear in a natural, relatively simple way. Second, the properties of the resultant Fredholm equations for Φ_2 and Φ_4 appear to be very sensitive to the value of $\varkappa_0$ (and hence the value of c) and to the location of the points ia_1 and ia_2. Convergence is an open question. In particular, residues at the poles $k_1 = -ia_1$ and $k_2 = -ia_2$ are unknown and must be carried along during the iteration.

APPENDIX A: ASYMPTOTIC SOLUTIONS

Far away from the boundary $x_1 = 0$, the spatial distribution of the density $\varphi(x_1, x_2)$ in x_1 will tend to the source distribution $\exp(-a_1 x)$ if $\operatorname{Re}(a_1)$ and $\operatorname{Re}(a_2)$ are small enough. Later in this appendix we give specific upper bounds which must be satisfied. Thus,

$$\varphi(x_1, x_2) \to \exp(-a_1 x_1)\alpha_2(x_2), \quad x_1 \to +\infty. \tag{A1}$$

To determine the distribution in x_2 for large x_1 given by α_2, we substitute the rhs above into the transport equation (2.4) and extend the integration on x_1 to $(-\infty, +\infty)$:

$$\begin{aligned}\exp(-a_1x_1)\alpha_2(x_2) &= c\int_{-\infty}^{\infty} dx_1' \int_0^{\infty} dx_2' K(|\mathbf{x}-\mathbf{x}'|)\exp(-a_1x_1')\alpha_2(x_2') \\ &\quad + \begin{cases}\exp(-a_1x_1 - a_2x_2), & x_2 \geq 0 \\ 0, & x_2 < 0\end{cases}.\end{aligned} \tag{A2}$$

Equation (A2) represents a one-dimensional Wiener-Hopf problem in x_2. Fourier transformation in x_2 of the above equation yields

$$\Lambda[(k_2^2 - a_1^2)^{\frac{1}{2}}]A_2^+(k_2) = -A_2^-(k_2) + 1/(a_2 - ik_2), \tag{A3}$$

where Λ is defined in (3.1) and $A_2^\pm$ is the transform of α_2 for $x_2 \gtrless 0$:

$$A_2^+(k_2) = \int_0^{\infty} \exp(ik_2x_2)\alpha_2(x_2)\,dx_2, \tag{A4}$$

$$A_2^-(k_2) = \int_{-\infty}^{0} \exp(ik_2x_2)\alpha_2(x_2)\,dx_2. \tag{A5}$$

The appropriate factorization of $\Lambda[(k_2^2 - a_1^2)^{\frac{1}{2}}]$ is given by

$$\Lambda[(k_2^2 - a_1^2)^{\frac{1}{2}}] = [H_1(-ia_1, k_2)H_2(-ia_1, k_2)] \times [H_3(-ia_1, k_2)H_4(-ia_1, k_2)], \tag{A6}$$

where the first factor in square brackets above is analytic and nonzero in the upper k_2 plane while the same is true for the second factor in the lower k_2 plane.

Using the factorization (A6), we easily solve (A3) to give

$$A_2^+(k_2) = [(a_2 - ik_2)H_1(-ia_1, k_2)H_2(-ia_1, k_2) \times H_3(-ia_1, -ia_2)H_4(-ia_1, -ia_2)]^{-1} \tag{A7}$$

and

$$A_2^-(k_2) = \frac{1}{a_2 - ik_2} \times \left(1 - \frac{H_3(-ia_1, k_2)H_4(-ia_1, k_2)}{H_3(-ia_1, -ia_2)H_4(-ia_1, -ia_2)}\right). \tag{A8}$$

To interpret the above results without going into any great detail, we see from (5.13) and (5.2) that

$$H_1(-ia_1, k_2)H_2(-ia_1, k_2) = \left(\frac{(\varkappa_0^2 - a_1^2)^{\frac{1}{2}} - ik_2}{(1 - a_1^2)^{\frac{1}{2}} - ik_2}\right) \times \exp\left(-\frac{1}{\pi}\int_1^{\infty} \frac{\theta(1/t)t\,dt}{(t^2 - a_1^2)^{\frac{1}{2}}[(t^2 - a_1^2)^{\frac{1}{2}} - ik_2]}\right). \tag{A9}$$

Thus H_1H_2 for $k_1 = -ia_1$ has a zero at $k_2 = -i(\varkappa_0^2 - a_1^2)^{\frac{1}{2}}$ and a branch point at $k_2 = -i(1 - a_1^2)^{\frac{1}{2}}$ with a corresponding branch cut in the lower half-plane. (The apparent pole at the branch point is canceled by the exponential term which goes to zero at that point.) Fourier inversion of A_2^+ will give by contour integration two discrete exponential terms for $\alpha_2(x_2)$, $x_2 > 0$, plus an integral over the branch cut as follows:

$$\begin{aligned}\lim_{x_1\to\infty} \varphi(x_1, x_2) &= \exp(-a_1x_1)\frac{1}{2\pi}\int_{-\infty}^{+\infty} \exp(-ik_2x_2)A_2^+(k_2)\,dk_2 \\ &= \exp(-a_1x_1)\Big(R_1\exp(-a_2x_2) \\ &\quad + R_2\exp[-(\varkappa_0^2 - a_1^2)^{\frac{1}{2}}x_2] \\ &\quad + \int_{(1-a_1^2)^{\frac{1}{2}}}^{\infty} \exp(-kx_2)D(k)\,dk\Big),\end{aligned} \tag{A10}$$

where residues R_1 and R_2 and the function $D(k)$ associated with the discontinuity across the branch cut are easily obtained by using (A7) in (A9). For $x_2 < 0$, $\alpha_2(x_2)$ will consist only of an integral term over the branch cut of $H_3(-ia_1, k_2)H_4(-ia_1, k_2)$ in the upper half k_2 plane. We have thus determined the behavior of $\varphi(x_1, x_2)$ for large positive x_1 for all x_2.

Similarly, we expect for large positive x_2 that

$$\varphi(x_1, x_2) \to \exp(-a_2 x_2)\alpha_1(x_1), \quad x_2 \to +\infty. \quad \text{(A11)}$$

If Fourier transforms of α_1 are defined as follows:

$$A_1^+(k_1) = \int_0^\infty \exp(ik_1x_1)\alpha_1(x_1)\,dx_1,$$

$$A_1^-(k_1) = \int_{-\infty}^0 \exp(ik_1x_1)\alpha_1(x_1)\,dx_1,$$

then an analysis similar to that above for α_2 gives

$$A_1^+(k_1) = [(a_1 - ik_1)H_1(k_1, -ia_2)H_4(k_1, -ia_2) \times H_2(-ia_1, -ia_2)H_3(-ia_1, -ia_2)]^{-1}, \quad \text{(A12)}$$

and

$$A_1^-(k_1) - \frac{1}{a_1 - ik_1} \times \left(1 - \frac{H_2(k_1, -ia_2)H_3(k_1, -ia_2)}{H_2(-ia_1, -ia_2)H_3(-ia_1, -ia_2)}\right). \quad \text{(A13)}$$

The large x_2 behavior of $\varphi(x_1, x_2)$ for $x_1 > 0$ will have the same form as given in (A10), with a_1, a_2, x_1, and x_2 replaced by a_2, a_1, x_2, and x_1, respectively.

An earlier statement was made that Re (a_1) and Re (a_2) must be small enough if (A1) and (A11) are to be valid expansions. From (A10) and the equivalent expression for large x_2 we see that Re (a_1) and Re (a_2) must satisfy

$$\text{Re}(a_1) < \text{Re}(\varkappa_0^2 - a_2^2)^{\frac{1}{2}}$$

and

$$\text{Re}(a_2) < \text{Re}(\varkappa_0^2 - a_1^2)^{\frac{1}{2}},$$

or, if a_1 and a_2 are real,

$$a_1^2 + a_2^2 < \varkappa_0^2.$$

APPENDIX B: THE FUNCTIONS $T_2(k_1, k_2)$ AND $T_4(k_1, k_2)$

The function T_2 is given by (6.9) or, equivalently, by

$$T_2(k_1, k_2) = \left[E(-k_1, k_2)\left[\frac{\hat{S}(k_1, k_2)}{E(-k_1, k_2)}\right]_{\sigma_1-}\right]_{\sigma_2+}, \quad \text{(B1)}$$

where $\hat{S}$ is given by (4.11). For the first two terms within the brackets in (4.11), we apply the operations shown on the rhs of (B1), and for the second pair of terms we use, for any suitably behaved $B(k_1, k_2)$,

$$\left[\frac{B(k_1, k_2)}{E(-k_1, k_2)}\right]_{\sigma_1-} = \frac{B(k_1, k_2)}{E(-k_1, k_2)} - \left[\frac{B(k_1, k_2)}{E(-k_1, k_2)}\right]_{\sigma_1+}. \quad \text{(B2)}$$

After subsequent manipulations we find

$$\begin{aligned} H_3(-ia_1, -ia_2)T_2(k_1, k_2) &= \frac{H_3(k_1, -ia_2)}{(a_1 - ik_1)(a_2 - ik_2)} \times \left(\frac{H_2(k_1, k_2)}{H_2(-ia_1, k_2)} - \frac{H_2(k_1, -ia_2)}{H_2(-ia_1, -ia_2)}\right) \\ &- \frac{1}{2\pi i(a_1 - ik_1)H_4(-ia_1, -ia_2)} \times \int_{-i\infty}^{-i} \frac{dz_2 G(k_1, -ia_1, z_2)H_4(-ia_1, z_2)}{(z_2 - k_2)(a_2 - iz_2)} \\ &+ \frac{1}{2\pi i}\int_{-i\infty}^{-i\varkappa_0} \frac{dz_2}{(z_2 - k_2)(a_2 - iz_2)} \times \left(\frac{E_+(-k_1, z_2)}{H_{2+}(-ia_1, z_2)} - \frac{E_-(-k_1, z_2)}{H_{2-}(-ia_1, z_2)}\right) \\ &- \frac{1}{4\pi^2}\int_{-i\infty}^{-i} \frac{dz_2 H_3(-ia_1, z_2)}{(z_2 - k_2)(a_2 - iz_2)} \times \int_{-i\infty}^{-i\varkappa_0} \frac{dz_1 G(k_1, z_1, z_2)}{(z_1 - k_1)(a_1 - iz_1)} \\ &\times \left(\frac{H_4^+(z_1, z_2)}{H_4^+(z_1, -ia_2)} - \frac{H_4^-(z_1, z_2)}{H_4^-(z_1, -ia_2)}\right). \end{aligned} \quad \text{(B3)}$$

Similarly, we find that $T_4(k_1, k_2)$ is given by

$$\begin{aligned} H_3(-ia_1, -ia_2)T_4(k_1, k_2) &= \frac{H_3(-ia_1, k_2)}{(a_1 - ik_1)(a_2 - ik_2)} \times \left(\frac{H_4(k_1, k_2)}{H_4(k_1, -ia_2)} - \frac{H_4(-ia_1, k_2)}{H_4(-ia_1, -ia_2)}\right) \\ &- \frac{1}{2\pi i}\frac{1}{(a_2 - ik_2)H_2(-ia_1, -ia_2)} \times \int_{-i\infty}^{-i} \frac{dz_1 G(k_2, -ia_2, z_1)H_2(z_1, -ia_2)}{(z_1 - k_1)(a_1 - iz_1)} \\ &+ \frac{1}{2\pi i}\int_{-i\infty}^{-i\varkappa_0} \frac{dz_1}{(z_1 - k_1)(a_1 - iz_1)} \times \left(\frac{E^+(-k_2, z_1)}{H_4^+(z_1, -ia_2)} - \frac{E^-(-k_2, z_1)}{H_4^-(z_1, -ia_2)}\right) \\ &- \frac{1}{4\pi^2}\int_{-i\infty}^{-i} \frac{dz_1 H_3(z_1, -ia_2)}{(z_1 - k_1)(a_1 - iz_1)} \times \int_{-i\infty}^{-i\varkappa_0} \frac{dz_2 G(k_2, z_2, z_1)}{(z_2 - k_2)(a_2 - iz_2)} \\ &\times \left(\frac{H_{2+}(z_1, z_2)}{H_{2+}(-ia_1, z_2)} - \frac{H_{2-}(z_1, z_2)}{H_{2-}(-ia_1, z_2)}\right). \end{aligned} \quad \text{(B4)}$$

APPENDIX C: CONVERGENCE OF THE ITERATION SCHEME

In this appendix we will show that the Fredholm equations derived in Sec. 6 may be solved by iteration and that the convergence is uniform. The integral operator in question is L_2L_4 given by (6.23) and (6.24). Since the Fredholm kernels belonging to L_2 and L_4 will be seen to be positive, we can define the norm of the operator L_2L_4 as

$$\|L_2L_4\| = \max_{k_1,k_2\in(-i\infty,-i)} \left| \frac{B(k_1,k_2)}{4\pi^2}\int_{-i\infty}^{-i} dz_1 \times \int_{-i\infty}^{-i} \frac{dz_2 G(k_1,z_1,k_2)G(k_2,z_2,z_1)}{(z_1-k_1)(z_2-k_2)B(z_1,z_2)} \right|, \quad \text{(C1)}$$

where $B(k_1, k_2)$ is a positive, bounded function to be chosen later. The corresponding norm on the function space is

$$\|f\| = \max_{k_1,k_2\in(-i\infty,-i)} |B(k_1,k_2)f(k_1,k_2)|. \quad \text{(C2)}$$

To prove uniform convergence, we must show that $\|L_2L_4\| < 1$. For convenience we switch to positive real variables

$$ik_j \to \eta_j, \quad iz_j \to \zeta_j, \quad j = 1, 2$$

and consider the quantity $-(1/2\pi)G(k_2, z_2, z_1)/(z_2 - k_2)$. Using (6.16) and (5.2), we find that

$$-\frac{1}{2\pi}\frac{G(k_2,z_2,z_1)}{z_2-k_2} = \frac{1}{2\pi i}\frac{G(-i\eta_2,-i\zeta_2,-i\zeta_1)}{\zeta_2-\eta_2} = \frac{D_0D_1D_2}{\pi}\frac{\sin(\psi_0-\psi_1)}{\eta_2-\zeta_2}, \quad \text{(C3)}$$

where

$$D_0 = \left(\frac{\zeta_1^2+\zeta_2^2-1}{\zeta_1^2+\eta_2^2-1}\right)^{\frac{1}{2}}, \quad \text{(C4)}$$

$$D_1 = \exp\left[-\frac{1}{\pi}\int_1^{\zeta_1}\theta(1/t)t\,dt \times \left(\frac{1}{\zeta_1^2+\eta_2^2-t^2} - \frac{1}{\zeta_1^2+\zeta_2^2-t^2}\right)\right], \quad \text{(C5)}$$

$$D_2 = \exp\left[-\frac{1}{\pi}\int_{\zeta_1}^{\infty}\frac{\theta(1/t)t\,dt}{(t^2-\zeta_1^2)^{\frac{1}{2}}} \times \left(\frac{1}{\eta_2+(t^2-\zeta_1^2)^{\frac{1}{2}}} - \frac{1}{\zeta_2+(t^2-\zeta_1^2)^{\frac{1}{2}}}\right)\right], \quad \text{(C6)}$$

$$\psi_0 = \sin^{-1}\left(\frac{(\eta_2-\zeta_2)(\zeta_1^2-1)^{\frac{1}{2}}}{(\zeta_1^2+\zeta_2^2-1)^{\frac{1}{2}}(\zeta_1^2+\eta_2^2-1)^{\frac{1}{2}}}\right), \quad \text{(C7)}$$

$$\psi_1 = \frac{1}{\pi}\int_1^{\zeta_1}\frac{\theta(1/t)t\,dt}{(\zeta_1^2-t^2)^{\frac{1}{2}}}\left(\frac{\zeta_2}{\zeta_1^2+\zeta_2^2-t^2} - \frac{\eta_2}{\zeta_1^2+\eta_2^2-t^2}\right). \quad \text{(C8)}$$

The quantities ψ_0, ψ_1, and $\eta_2 - \zeta_2$ always have the same sign. Furthermore,

$$|\psi_0| \leq \pi/2,$$

and, using $\theta(1/t) \leq \pi$, we obtain

$$|\psi_1| \leq \left|\int_1^{\zeta_1}\frac{t\,dt}{(\zeta_1^2-t^2)^{\frac{1}{2}}}\left(\frac{\zeta_2}{\zeta_1^2+\zeta_2^2-t^2} - \frac{\eta_2}{\zeta_1^2+\eta_2^2-t^2}\right)\right| = \left|\tan^{-1}\left(\frac{(\zeta_1^2-1)^{\frac{1}{2}}}{\zeta_2}\right) - \tan^{-1}\left(\frac{(\zeta_1^2-1)^{\frac{1}{2}}}{\eta_2}\right)\right| = |\psi_0|. \quad \text{(C9)}$$

Thus the left-hand side of (C3) is positive, real, and bounded by

$$-\frac{1}{2\pi i}\frac{G(-i\eta_2,-i\zeta_2,-i\zeta_1)}{\zeta_2-\eta_2} \leq \frac{D_0D_1D_2}{\pi}\frac{\sin\theta_0}{\eta_2-\zeta_2} = \frac{D_1D_2D_3}{\pi}, \quad \text{(C10)}$$

where

$$D_3 = \frac{(\zeta_1^2-1)^{\frac{1}{2}}}{\zeta_1^2+\eta_2^2-1}. \quad \text{(C11)}$$

We now choose

$$B(z_1,z_2) = B(-i\zeta_1,-i\zeta_2) = (\zeta_1^2+\zeta_2^2-1)\exp\left[\frac{1}{\pi}\int_1^{\zeta_1}\frac{\theta(1/t)t\,dt}{\zeta_1^2+\zeta_2^2-t^2} + \frac{1}{\pi}\int_{\zeta_1}^{\infty}\frac{\theta(1/t)t\,dt}{(t^2-\zeta_1^2)^{\frac{1}{2}}}\left(\frac{1}{\zeta_2+(t^2-\zeta_1^2)^{\frac{1}{2}}}\right)\right], \quad \text{(C12)}$$

consistent with the behavior of χ_2 and χ_4 at infinity. Substituting (C12) and (C10) into (C1) and using carefully selected inequalities which are too numerous to repeat here, we find that the norm of L_2L_4 is bounded by

$$\|L_2L_4\| = \max_{\eta_1,\eta_2\in(1,\infty)}\left|\frac{\exp(K)}{\pi^2}\int_1^{\infty}\frac{(\eta_2^2-1)^{\frac{1}{2}}\,d\zeta_1}{\zeta_1^2+\eta_2^2-1} \times \int_1^{\infty}\frac{(\zeta_1^2-1)^{\frac{1}{2}}\,d\zeta_2}{\zeta_1^2+\zeta_2^2-1}\right| \leq \tfrac{1}{4}\exp(K), \quad \text{(C13)}$$

where

$$K = \max_{\lambda\in(1,\infty)}\left|\frac{1}{\pi}\int_1^{\infty}\frac{\theta(1/t)t\,dt}{(t^2-1)^{\frac{1}{2}}[1+(t^2-1)^{\frac{1}{2}}]} - \left(\frac{1}{\pi}\int_1^{\infty}\frac{\theta(1/t)t\,dt}{(t^2-1)^{\frac{1}{2}}[\lambda+(t^2-1)^{\frac{1}{2}}]} - \frac{1}{\pi}\int_1^{\lambda}\frac{\theta(1/t)\,dt}{\lambda^2+1-t^2}\right)\right|. \quad \text{(C14)}$$

Finally, we obtain the following interesting result for $0 \leq c \not< 2$. In this case one can show that the term in braces above is positive, so that

$$K \leq \frac{1}{\pi}\int_1^{\infty}\frac{\theta(1/t)t\,dt}{(t^2-1)^{\frac{1}{2}}[1+(t^2-1)^{\frac{1}{2}}]} = \ln\left(\frac{1}{E(i,i)}\right), \quad \text{(C15)}$$

where $E(k_2, k_1)$ is defined in (5.2). Thus,

$$\|L_2L_4\| \leq 1/4E(i, i). \quad \text{(C16)}$$

The result (C16) should be compared with analogous results in one-dimensional slab geometry.[11]

Using the inequality $\theta(1/t)t \leq \pi$, valid also for $0 < c \lessapprox 2$, we find further that

$$\|L_2L_4\| \leq \frac{1}{4}\left(\frac{\sqrt{2}+1}{\sqrt{2}-1}\right)^{1/\sqrt{2}} \approx 0.87. \quad \text{(C17)}$$

Due to the numerous inequalities employed to achieve this result, we suspect that the actual norm of L_2L_4 is significantly lower than the above upper bound.

* Part of this research was supported by a National Science Foundation Grant.

[1] N. Wiener and E. Hopf, Sitzber. Preuss Acad. Wiss., 696 (1931).
[2] See, e.g., B. Davison, *Neutron Transport Theory* (Oxford U.P., London, 1958).
[3] E. A. Kraut and G. W. Lehman, J. Math. Phys. **10**, 1340 (1969).
[4] S. Bochner, Am. J. Math. **59**, 732 (1937).
[5] J. Radlow, Intern. J. Eng. Sci. **2**, 275 (1964).
[6] See, e.g., A. Leonard and T. W. Mullikin, J. Math. & Phys. **44**, 327 (1965).
[7] E. A. Kraut, J. Math. Phys. **9**, 1481 (1968).
[8] See, e.g., K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Mass., 1967).
[9] The density $\varphi(x_1, x_2)$ is bounded along the edge of the quarter space so that $\Phi_1 \sim 1/k_i$ for large k_i, and hence $\Lambda\Phi_1$ has a bounded L_2 norm in the tube T_Φ.
[10] The contours at infinity give no contribution. See Ref. 9.
[11] See Ref. 6, p. 331.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Exactly Solvable Cell Model with a Melting Transition

D. J. Gates
Mathematics Department, Imperial College, London S.W.7, England

(Received 19 June 1970)

A new cell model for classical particle systems is presented and analyzed. In this model the particles are confined to congruent, interconnected, cubic cells of volume ω centered on the points of a cubic lattice with lattice spacing $1/\gamma$. The particles interact via a 2-body potential of the form $q(\mathbf{r}) + \omega^{-1}K(\gamma\mathbf{r})$. The paper deals with the limiting form of this model in which the cells are very large but their separation is much larger. The free energy density is defined by

$$a(\rho, T) \equiv \lim_{\omega\to\infty} \lim_{\gamma\to 0} \tilde{a}(\rho, T, \gamma, \omega),$$

where $\tilde{a}(\rho, T, \gamma, \omega)$ is the free energy density at density ρ, temperature T, and arbitrary γ and ω. For a very general class of functions q and K, it is proved that $a(\rho, T)$ is given by a variational principle. For a certain class of functions K (including $K \leq 0$), $a(\rho, T)$ is given by the Lebowitz–Penrose generalization of the van der Waals–Maxwell theory. For a different class of functions K the system has crystalline states. When K is chosen so that only particles in nearest-neighbor cells interact and K is isotropic, it is proved that the most general crystalline state of the system has a density distribution with two values ρ_+ and ρ_- arranged in a checkerboard (sodium chloride) pattern. For the special case with K repulsive, $K(0) = 0$ and $q = 0$, the system has a second-order melting transition from a crystalline to a fluid state, with no critical temperature. Various correlation functions are defined and evaluated. In the 1-dimensional nearest-neighbor case, the results include exact versions of the Ornstein–Zernike theory for both fluid and crystalline states. Magnetic systems are also considered. Different special cases of the model yield precisely the Weiss theory of ferromagnetism and the Néel–van Vleck theory of antiferromagnetism.

1. INTRODUCTION

This paper deals with a new cell model for many-body systems. The model is of a general type in that it applies to particle systems and magnetic spin systems, and allows a wide choice of interaction potentials. It is not a realistic model for these systems, but, nevertheless, it exhibits many of their properties and has the advantage of being very amenable to exact treatment. In particular, it has crystalline (or antiferromagnetically ordered) states and a melting transition which can be studied in detail.

The explanation of the crystalline state and the phenomenon of melting from the principles of statistical mechanics is an outstanding unsolved problem of theoretical physics. Several simplified models have been studied, but even these are not very well understood. The early theories are of the mean-field type, due mainly to Kirkwood and Monroe.[1] The Lennard-Jones and Devonshire theory,[2] and the model of this paper, are related to these. Recent work[3] has shown that these theories are derivable from the statistical mechanics of a model system. It has been shown[4] that where $E(k_2, k_1)$ is defined in (5.2). Thus,

$$\|L_2L_4\| \leq 1/4E(i, i). \quad \text{(C16)}$$

The result (C16) should be compared with analogous results in one-dimensional slab geometry.[11]

Using the inequality $\theta(1/t)t \leq \pi$, valid also for $0 < c \lessapprox 2$, we find further that

$$\|L_2L_4\| \leq \frac{1}{4}\left(\frac{\sqrt{2}+1}{\sqrt{2}-1}\right)^{1/\sqrt{2}} \approx 0.87. \quad \text{(C17)}$$

Due to the numerous inequalities employed to achieve this result, we suspect that the actual norm of L_2L_4 is significantly lower than the above upper bound.

* Part of this research was supported by a National Science Foundation Grant.

[1] N. Wiener and E. Hopf, Sitzber. Preuss Acad. Wiss., 696 (1931).
[2] See, e.g., B. Davison, *Neutron Transport Theory* (Oxford U.P., London, 1958).
[3] E. A. Kraut and G. W. Lehman, J. Math. Phys. 10, 1340 (1969).
[4] S. Bochner, Am. J. Math. 59, 732 (1937).
[5] J. Radlow, Intern. J. Eng. Sci. 2, 275 (1964).
[6] See, e.g., A. Leonard and T. W. Mullikin, J. Math. & Phys. 44, 327 (1965).
[7] E. A. Kraut, J. Math. Phys. 9, 1481 (1968).
[8] See, e.g., K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Mass., 1967).
[9] The density $\varphi(x_1, x_2)$ is bounded along the edge of the quarter space so that $\Phi_1 \sim 1/k_i$ for large k_i, and hence $\Lambda\Phi_1$ has a bounded L_2 norm in the tube T_Φ.
[10] The contours at infinity give no contribution. See Ref. 9.
[11] See Ref. 6, p. 331.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Exactly Solvable Cell Model with a Melting Transition

D. J. Gates
Mathematics Department, Imperial College, London S.W.7, England

(Received 19 June 1970)

A new cell model for classical particle systems is presented and analyzed. In this model the particles are confined to congruent, interconnected, cubic cells of volume ω centered on the points of a cubic lattice with lattice spacing $1/\gamma$. The particles interact via a 2-body potential of the form $q(\mathbf{r}) + \omega^{-1}K(\gamma\mathbf{r})$. The paper deals with the limiting form of this model in which the cells are very large but their separation is much larger. The free energy density is defined by

$$a(\rho, T) \equiv \lim_{\omega\to\infty} \lim_{\gamma\to 0} \tilde{a}(\rho, T, \gamma, \omega),$$

where $\tilde{a}(\rho, T, \gamma, \omega)$ is the free energy density at density ρ, temperature T, and arbitrary γ and ω. For a very general class of functions q and K, it is proved that $a(\rho, T)$ is given by a variational principle. For a certain class of functions K (including $K \leq 0$), $a(\rho, T)$ is given by the Lebowitz–Penrose generalization of the van der Waals–Maxwell theory. For a different class of functions K the system has crystalline states. When K is chosen so that only particles in nearest-neighbor cells interact and K is isotropic, it is proved that the most general crystalline state of the system has a density distribution with two values ρ_+ and ρ_- arranged in a checkerboard (sodium chloride) pattern. For the special case with K repulsive, $K(0) = 0$ and $q = 0$, the system has a second-order melting transition from a crystalline to a fluid state, with no critical temperature. Various correlation functions are defined and evaluated. In the 1-dimensional nearest-neighbor case, the results include exact versions of the Ornstein–Zernike theory for both fluid and crystalline states. Magnetic systems are also considered. Different special cases of the model yield precisely the Weiss theory of ferromagnetism and the Néel–van Vleck theory of antiferromagnetism.

1. INTRODUCTION

This paper deals with a new cell model for many-body systems. The model is of a general type in that it applies to particle systems and magnetic spin systems, and allows a wide choice of interaction potentials. It is not a realistic model for these systems, but, nevertheless, it exhibits many of their properties and has the advantage of being very amenable to exact treatment. In particular, it has crystalline (or anti-ferromagnetically ordered) states and a melting transition which can be studied in detail.

The explanation of the crystalline state and the phenomenon of melting from the principles of statistical mechanics is an outstanding unsolved problem of theoretical physics. Several simplified models have been studied, but even these are not very well understood. The early theories are of the mean-field type, due mainly to Kirkwood and Monroe.[1] The Lennard-Jones and Devonshire theory,[2] and the model of this paper, are related to these. Recent work[3] has shown that these theories are derivable from the statistical mechanics of a model system. It has been shown[4] that this system does crystallize in a certain sense, but the form of the crystal structure and the nature of the melting transition (in particular, its order) are not known. Different approximate treatments give conflicting results.[1,5]

Another model is the gas of hard spheres or hard disks for which the computer experiments of Adler and Wainwright[6] suggest the existence of a crystal state and a melting transition of the first order. However, there is no theoretical proof of this result. A similar model is the hard-square lattice gas,[7] for which some accurate expansion techniques have been found by Gaunt and Fisher. These expansions indicate the existence of a second-order melting transition, but again this has not been proved.

The need for models which, like the model of this paper, can be treated exactly is therefore apparent. The model may be useful as a "testing ground" for general formal theories of melting, such as the theory of "group invariance of states" and "broken symmetries," which has been formulated recently.[8]

2. THE CELL MODEL

In this section we define and discuss the details of the cell model. In this model the particles are confined to congruent, similarly oriented, ν-dimensional cubes $\{\omega(\mathbf{y}/\gamma)\}$, called cells, of equal volume ω, where $\omega(\mathbf{x})$ is centered at a point $\mathbf{x}$. The vector $\mathbf{y}$ can lie only on the points of a ν-dimensional cubic lattice $\mathbf{Z}^\nu$ with unit cell of side unity ($\mathbf{Z}^\nu$ is the space of all ν-tuples of integers). Hence the cells $\{\omega(\mathbf{y}/\gamma)\}$ are centered at points of a cubic lattice with unit cell of side $1/\gamma$ (see Fig. 1). The particles can move freely from one cell to another. To make this physically possible, one can imagine very fine tubes of negligible volume connecting the cells. Each particle interacts with other particles, both in the same cell and in other cells, via the 2-body potential

$$v(\mathbf{r}, \gamma, \omega) \equiv q(\mathbf{r}) + \omega^{-1}K(\gamma\mathbf{r}), \tag{2.1}$$

where $q(\mathbf{r})$ is called the *short-range or reference potential* and $\omega^{-1}K(\gamma\mathbf{r})$ is called the *long-range potential.*

We shall consider the free energy density $\tilde{a}(\rho, T, \gamma, \omega)$ (defined below) of a system of such particles with average density ρ and temperature T, and evaluate its limit

$$a(\rho, T) \equiv \lim_{\omega\to\infty} \lim_{\gamma\to 0} \tilde{a}(\rho, T, \gamma, \omega). \tag{2.2}$$

This limiting free energy density $a(\rho, T)$ describes a system in which the distance $1/\gamma$ between the cells and the volume ω of each cell are both very large. Since the limit $\gamma \to 0$ is taken first, the separation of the cells

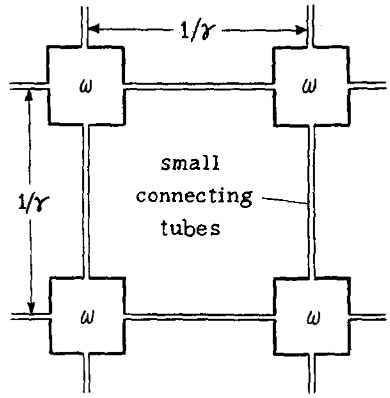


FIG. 1. Illustration of the cell model.

is much larger than their dimensions, i.e.,

$$\gamma^{-1} \gg \omega^{1/\nu}. \tag{2.3}$$

The range of the potential $\omega^{-1}K(\gamma\mathbf{r})$ also becomes infinite as $\gamma \to 0$, so that "as seen by the cells" this potential has a fixed range. For example, the interaction potential of two particles at the centers $\mathbf{y}_1/\gamma$ and $\mathbf{y}_2/\gamma$ of $\omega(\mathbf{y}_1/\gamma)$ and $\omega(\mathbf{y}_2/\gamma)$ is $\omega^{-1}K(\mathbf{y}_1 - \mathbf{y}_2)$, which is independent of γ. The need for the factor $1/\omega$ in the potential can be understood by considering the contribution from $\omega^{-1}K(\gamma\mathbf{r})$ to the potential energy of a single particle interacting with every other particle, for a state of uniform density ρ. An estimate of this contribution is

$$u(\gamma, \omega) \equiv \rho \int_{\Omega(\mathbf{Z}^\nu)} d\mathbf{r}\, \omega^{-1}K(\gamma\mathbf{r}), \tag{2.4}$$

where $\Omega(\mathbf{Z}^\nu)$ is the union of all the $\omega(\mathbf{y}/\gamma)$'s, i.e.,

$$\Omega(\mathbf{Z}^\nu) \equiv \bigcup_{\mathbf{y}\in\mathbf{Z}^\nu} \omega(\mathbf{y}/\gamma). \tag{2.5}$$

This gives

$$\begin{aligned} u &= \rho \sum_{\mathbf{y}\in\mathbf{Z}^\nu} \frac{1}{\omega} \int_{\omega(\mathbf{y}/\gamma)} d\mathbf{r}\, K(\gamma\mathbf{r}) \\ &= \rho \sum_{\mathbf{y}\in\mathbf{Z}^\nu} \frac{1}{\bar{\omega}} \int_{\bar{\omega}(\mathbf{y})} d\mathbf{r}\, K(\mathbf{r}), \end{aligned} \tag{2.6}$$

where $\bar{\omega}(\mathbf{y})$ is a cube of volume $\bar{\omega} \equiv \gamma^\nu\omega$, centered at $\mathbf{y}$. Assuming K is continuous, we obtain

$$\frac{1}{\bar{\omega}} \int_{\bar{\omega}(\mathbf{y})} d\mathbf{r}\, K(\mathbf{r}) \to K(\mathbf{y}) \quad \text{as} \quad \gamma \to 0, \tag{2.7}$$

and hence

$$\begin{aligned} \lim_{\omega\to\infty} \lim_{\gamma\to 0} u(\gamma, \omega) &= \lim_{\gamma\to 0} u(\gamma, \omega) \\ &= \rho \sum_{\mathbf{y}\in\mathbf{Z}^\nu} K(\mathbf{y}), \end{aligned} \tag{2.8}$$

which is finite if the sum converges [see condition (2.16)]. Without the factor $1/\omega$ in the long-range potential, this limit would be infinite, and the system would behave catastrophically in the limit $\omega \to \infty$.

The present model is very similar to the model considered by Lebowitz and Penrose and others.[3,4,9] They used a long-range potential of the form $\gamma^\nu K(\gamma \mathbf{r})$, without confining particles to cells. The factor γ^ν is needed for the same reason as the factor $1/\omega$ is needed in the present model. The behavior of the two models is also very similar, as is shown in Sec. 3.

We define the free energy density $\tilde{a}(\rho, T, \gamma, \omega)$ in the usual way as follows. Let D be a set of points of $\mathbf{Z}^\nu$ forming a cube, and let $\Omega(D)$ be the set comprising all the $\omega(\mathbf{y}/\gamma)$'s for which $\mathbf{y} \in D$, i.e.,

$$D \subset \mathbf{Z}^\nu, \quad \Omega(D) \equiv \bigcup_{\mathbf{y} \in D} \omega\left(\frac{\mathbf{y}}{\gamma}\right) \subset \Omega(\mathbf{Z}^\nu). \tag{2.9}$$

Thus $\Omega(D)$, which depends on γ and ω, has total volume

$$|\Omega(D)| = \omega\,|D|, \tag{2.10}$$

where $|D|$ is the number of points in D. Then we define

$$\tilde{a}(\rho, T, \gamma, \omega) \equiv -kT \lim_{\substack{N, |D| \to \infty \\ N/|D| \to \omega\rho}} \frac{1}{|\Omega(D)|} \log Z(N, D, T, \gamma, \omega), \tag{2.11}$$

where

$$Z(N, D, T, \gamma, \omega) \equiv \frac{1}{N!\,\Lambda^{\nu N}} \int_{\Omega(D)} d\mathbf{x}_1 \cdots \int_{\Omega(D)} d\mathbf{x}_N \exp\left(\frac{-U_N}{kT}\right), \tag{2.12}$$

which is the partition function for N particles in $\Omega(D)$. Also, Λ is the thermal wavelength and

$$U_N \equiv \sum_{1 \le a < b \le N} v(\mathbf{x}_a - \mathbf{x}_b, \gamma, \omega) \quad \text{for} \quad N \ge 2. \tag{2.13}$$

Note that, in (2.11), the average density of the particles $N/|\Omega(D)|$ tends to ρ, while ω and γ remain constant. The definition (2.12) expresses the fact that particles can move from cell to cell.

To complete the definition of the model, we must specify conditions on the functions $q(\mathbf{r})$ and $K(\mathbf{r})$ which ensure the existence of the limits (2.2) and (2.11). We shall assume always that

$$q(\mathbf{r}) = q(-\mathbf{r}) \quad \text{and} \quad K(\mathbf{r}) = K(-\mathbf{r}), \tag{2.14}$$

$$K(r) \text{ is continuous and bounded,} \tag{2.15}$$

$$|K(\mathbf{r})| < C\,|\mathbf{r}|^{-\nu-\epsilon}, \tag{2.16}$$

where C and ϵ are positive constants. Further conditions are needed, and these may be of two types:

Type-I systems: q satisfies[10] the stability and tempering conditions of Fisher[11] [Eqs. (3.9a)–(3.9c) and (3.11a)]. No further conditions on K. (2.17)

Type-II systems: $q \ge 0$ and q satisfies the tempering condition (3.11a) of Fisher.[11] $K(\mathbf{r})$ is also a "positive-plus-positive" potential in the sense of Fisher [Eqs. (3.6a) and (3.6b)]. (2.18)

For Type-I systems, it follows from (2.15), (2.16), and (2.17) that $v(\mathbf{r}, \gamma, \omega)$ satisfies the stability and tempering conditions (3.9a)–(3.9c) and (3.11a) of Fisher, which imply the existence of $\tilde{a}$. For Type-II systems, it follows from (2.15), (2.16), and (2.18) that $v(\mathbf{r}, \gamma, \omega)$ is a positive-plus-positive potential and satisfies the tempering condition (3.11a) of Fisher, which again implies the existence of $\tilde{a}$.

For Type-II systems, v need not have a core. For example, a possible choice is $q = 0$, which we consider further in Sec. 5.

3. VARIATIONAL PRINCIPLE AND VAN DER WAALS–MAXWELL THEORY

In this section we present the basic results for the cell model, and give an outline of their derivation. First, we need some definitions. Let $\mathcal{C}(\rho)$ be the set of all functions $n \equiv \{n(\mathbf{y}) : \mathbf{y} \in \mathbf{Z}^\nu\}$, whose values $n(\mathbf{y})$ are (a) nonnegative, (b) periodic with respect to $\mathbf{y}$ (with unspecified period), and (c) have average value ρ, i.e.,

$$\frac{1}{|\Gamma(n)|} \sum_{\mathbf{y} \in \Gamma(n)} n(\mathbf{y}) = \rho, \tag{3.1}$$

where $\Gamma(n) \subset \mathbf{Z}^\nu$ is the unit cell of n and $|\Gamma|$ is the number of points in Γ. Let $G(n, T)$ be defined for any $n \in \mathcal{C}(\rho)$ by

$$G(n, T) \equiv \frac{1}{|\Gamma(n)|} \sum_{\mathbf{y} \in \Gamma(n)} \left(a^0[n(\mathbf{y}), T] + \tfrac{1}{2} n(\mathbf{y}) \sum_{\mathbf{y}' \in \mathbf{Z}^\nu} n(\mathbf{y}') K(\mathbf{y} - \mathbf{y}') \right), \tag{3.2}$$

where $a^0(\rho, T)$ is the free energy density of a continuum system, called the *reference system*, with the 2-body potential $q(\mathbf{r})$, i.e.,

$$a^0(\rho, T) \equiv -kT \lim_{\substack{N, |V| \to \infty \\ N/|V| \to \rho}} \frac{1}{|V|} \log Z^0(N, V, T), \tag{3.3}$$

where

$$Z^0(N, V, T) \equiv \frac{1}{N!\,\Lambda^{\nu N}} \int_V d\mathbf{x}_1 \cdots \int_V d\mathbf{x}_N \exp\left(\frac{-Q_N}{kT}\right), \tag{3.4}$$

$$Q_N \equiv \sum_{1 \le a < b \le N} q(\mathbf{x}_a - \mathbf{x}_b), \tag{3.5}$$

and V is a cube (in the ν-dimensional real number space) of volume $|V|$.

The basic result about the cell model is given by the following variational principle.

Theorem 1: For systems both of Type I and of Type II, the free energy density $a(\rho, T)$, defined by (2.2), exists, is convex in ρ, and is given by

$$a(\rho, T) = \inf_{n \in \mathcal{C}(\rho)} G(n, T). \tag{3.6}$$

The result (3.6) simply states that, to find $a(\rho, T)$, one minimizes the free energy functional $G(n, T)$ over all possible local density functions $n(\mathbf{y})$. This is similar to the well-known thermodynamic principle of "minimizing the free energy." One must use an infimum in (3.6) rather than a minimum, because one may need to make $\Gamma(n)$ arbitrarily large to minimize G, which would mean that the minimum could not be attained for any $n \in \mathcal{C}(\rho)$. (This happens for 2-phase states.)

Theorem 1 should be compared with a corresponding result of Ref. 3, Part I (Theorem 2). The latter result is the same except that $n(\mathbf{y})$ is replaced by an integrable function, and the sums in the definition of G are replaced by integrals. The proofs of the two results are also very similar, and so we shall only give an outline of the proof of Theorem 1 here, and refer the reader to Ref. 3 for details.

Outline of Proof (with T omitted from the notation): From (2.9) and (2.12), we can write

$$Z(N, D, \gamma, \omega) = \sum_{\{N(\mathbf{y})\} \in \mathcal{S}(N,D)} \left(\prod_{\mathbf{y} \in D} \frac{1}{N(\mathbf{y})!\, \Lambda^{\nu N(\mathbf{y})}} \int_{\omega(\mathbf{y}/\gamma)^{N(\mathbf{y})}} \right) \times d\mathbf{x}_1 \cdots d\mathbf{x}_N \exp(-U_N/kT), \tag{3.7}$$

where $\mathcal{S}(N, D)$ is the set of functions $\{N(\mathbf{y})\}$ such that $\mathbf{y} \in D$ and

$$\sum_{\mathbf{y} \in D} N(\mathbf{y}) = N, \tag{3.8}$$

where the $N(\mathbf{y})$'s are nonnegative integers. The notation in (3.7) indicates that there are $N(\mathbf{y})$ volume integrations over each $\omega(\mathbf{y}/\gamma)$. Since $\omega^{-1}K(\gamma\mathbf{r})$ varies slowly with $\mathbf{r}$ for small γ, the contribution from this potential, to the interaction between a particle in $\omega(\mathbf{y}/\gamma)$ and a particle in $\omega(\mathbf{y}'/\gamma)$, is approximately $\omega^{-1}K(\mathbf{y} - \mathbf{y}')$. Also, the contribution from $q(\mathbf{r})$ is almost zero because the cell separation becomes infinite as $\gamma \to 0$. Hence we can write

$$U_N \simeq \sum_{\mathbf{y} \in D} U^0_{N(\mathbf{y})} + \tfrac{1}{2} \sum_{\mathbf{y},\mathbf{y}' \in D} N(\mathbf{y})N(\mathbf{y}')\omega^{-1}K(\mathbf{y} - \mathbf{y}'), \tag{3.9}$$

where $U^0_{N(\mathbf{y})}$ is the potential energy due to $q(\mathbf{r})$ for a system of $N(\mathbf{y})$ particles in $\omega(\mathbf{y}/\gamma)$. This gives

$$Z(N, D, \gamma, \omega) \simeq \sum_{\{N(\mathbf{y})\} \in \mathcal{S}(N,D)} \left(\prod_{\mathbf{y} \in D} Z^0[N(\mathbf{y}), \omega] \right) \times \exp\left(-\frac{1}{2kT\omega} \sum_{\mathbf{y},\mathbf{y}' \in D} N(\mathbf{y})N(\mathbf{y}')K(\mathbf{y} - \mathbf{y}') \right). \tag{3.10}$$

The log of this sum can be approximated by the log of the maximum term. Setting

$$n(\mathbf{y}) \equiv N(\mathbf{y})/\omega \tag{3.11}$$

in the resulting expression gives

$$-\frac{kT}{|\Omega(D)|} \log Z(N, D, \gamma, \omega) = \min_{\{\omega n(\mathbf{y})\} \in \mathcal{S}(N,D)} \frac{1}{|D|} \left(-\sum_{\mathbf{y} \in D} \frac{1}{\omega} kT \log Z^0[\omega n(\mathbf{y}), \omega] + \tfrac{1}{2} \sum_{\mathbf{y},\mathbf{y}' \in D} n(\mathbf{y})n(\mathbf{y}')K(\mathbf{y} - \mathbf{y}') \right) + \text{corrections}. \tag{3.12}$$

If ω is large, one can replace

$$-\omega^{-1}kT \log Z^0[\omega n(\mathbf{y}), \omega]$$

by its limit $a^0[n(\mathbf{y})]$, as $\omega \to \infty$, plus a small correction. Also, if ω is arbitrarily large, the numbers $n(\mathbf{y})$ can be arbitrarily close to any real number, consistent with (3.8); i.e., any function in the set $\mathcal{C}_D(N/|\Omega(D)|)$, where

$$\mathcal{C}_D(\rho) \equiv \left\{ n : n(\mathbf{y}) \geq 0 \text{ for } \mathbf{y} \in D, \text{ and } \frac{1}{|D|} \sum_{\mathbf{y} \in D} n(\mathbf{y}) = \rho \right\}, \tag{3.13}$$

can be approximated arbitrarily closely by an $n(\mathbf{y})$ of the form (3.11).

To find $a(\rho)$, we must first let $|D| \to \infty$ with $N/|\Omega(D)| \to \rho$ in (3.12), which by (2.11) gives $\tilde{a}(\rho, \gamma, \omega)$; then we must let $\gamma \to 0$ and finally $\omega \to \infty$. By the method of Ref. 3, the corrections can be shown to vanish in the above triple limit. Hence we have

$$a(\rho) = \lim_{|D| \to \infty} \min_{n \in \mathcal{C}_D(\rho)} \frac{1}{|D|} \times \left(\sum_{\mathbf{y} \in D} a^0[n(\mathbf{y})] + \tfrac{1}{2} \sum_{\mathbf{y},\mathbf{y}' \in D} n(\mathbf{y})n(\mathbf{y}')K(\mathbf{y} - \mathbf{y}') \right). \tag{3.14}$$

This is itself a variational principle for $a(\rho, T)$. One can show that it is equivalent to (3.6), by using the method of Sec. 3 in Ref. 3. The convexity of $a(\rho, T)$ follows[3] from the convexity of $\tilde{a}(\rho, T, \gamma, \omega)$.

A basic property of the model is that, like the model considered in Refs. 3, 4, and 9, it yields a generalization of the van der Waals–Maxwell theory of the gas–liquid transition for a certain class of functions K. More precisely, if we put

$$\hat{K}(\mathbf{p}) \equiv \sum_{\mathbf{s}\in \mathbf{Z}^\nu} K(\mathbf{s}) \exp(2\pi i \mathbf{p}\cdot\mathbf{s}) \tag{3.15}$$

and

$$\alpha \equiv \sum_{\mathbf{s}\in \mathbf{Z}^\nu} K(\mathbf{s}) = \hat{K}(0), \tag{3.16}$$

we then have the following.

Theorem 2: (a) If the function K is such that $\hat{K} \geq 0$, then

$$a(\rho, T) = G(\rho, T) = a^0(\rho, T) + \tfrac{1}{2}\alpha\rho^2. \tag{3.17}$$

(b) If $\hat{K}(\mathbf{p}) \geq \hat{K}(0)$ for all $\mathbf{p}$ (of which a special case is $K \leq 0$), then

$$a(\rho, T) = CE[a^0(\rho, T) + \tfrac{1}{2}\alpha\rho^2], \tag{3.18}$$

where $CEf(\rho)$, called the convex envelope of f, is defined for any f as the maximal convex function not exceeding $f(\rho)$.

This theorem, which is due essentially to Lebowitz and Penrose,[9] can be deduced from Theorem 1 by adapting the method of Ref. 4 (Sec. 5). The pressure corresponding to (3.18) is given by the Maxwell construction applied to a generalized van der Waals equation. Note that for Type-II systems $\alpha \geq 0$, so that (b) does not apply.

The other results of Refs. 3 and 4 can also be easily adapted. In particular, one can show that for some functions K the Eqs. (3.17) and (3.18) do not apply. Instead, the functional $G(n, T)$ is minimized by a nonuniform, periodic function $n(\mathbf{y})$, so that the system has a crystalline phase. It is this phenomenon which is investigated in more detail in the following sections.

4. THE CASE OF NEAREST-NEIGHBOR INTERACTIONS

To study the above-mentioned phenomenon of crystallization, we consider some special cases of the model. The simplest case is obtained by choosing

$$K(0) = \alpha, \quad K(\mathbf{s}) = 0 \quad \text{for all} \quad \mathbf{s} \neq 0, \quad \mathbf{s} \in \mathbf{Z}^\nu. \tag{4.1}$$

This gives $\hat{K}(\mathbf{p}) = \hat{K}(0)$, so that (3.17) and (3.18) apply for all α.

More interesting is the special case with

$$K(\mathbf{s}) = 0 \quad \text{for} \quad |\mathbf{s}| > 1, \quad \mathbf{s} \in \mathbf{Z}^\nu, \tag{4.2}$$

which implies that only particles in nearest-neighbor cells interact via the long-range potential. Let us put

$$K_0 \equiv K(0), \tag{4.3}$$

and let the system be isotropic so that

$$K(\mathbf{s}) = K_1 \quad \text{for} \quad |\mathbf{s}| = 1. \tag{4.4}$$

One easily shows that

$$\hat{K} \geq 0 \quad \text{if and only if} \quad K_0 \geq 2\nu\,|K_1|, \tag{4.5}$$

so that $a(\rho, T)$ is given by (3.17), where

$$\alpha = K_0 + 2\nu K_1. \tag{4.6}$$

One can also show that

$$\hat{K}(\mathbf{p}) \geq K(0) \quad \text{if and only if} \quad K_1 \leq 0, \tag{4.7}$$

so that $a(\rho, T)$ is given by (3.18). These results are indicated in Fig. 2 by the regions (1) and (2). We shall find that systems in the remaining region (3) have a crystalline phase.

We now derive an equation of state which holds for all K_0 and K_1, including the region (3). First, we define the function MEf, called the midpoint envelope of f, by

$$MEf(\rho) \equiv \inf_h \tfrac{1}{2}[f(\rho + h) + f(\rho - h)] \tag{4.8}$$

for any function $f(\rho)$. We shall prove the following.

Theorem 3: For the nearest-neighbor cell model, defined above,

$$a(\rho, T) = CE\{ME[a^0(\rho, T) + (K_0 - \tfrac{1}{2}\alpha)\rho^2] + (\alpha - K_0)\rho^2\}. \tag{4.9}$$

For values of ρ and T, where the bracket { } coincides with its convex envelope, one can also write

$$a(\rho, T) = G(n^*, T). \tag{4.10}$$

Here $n^*(\mathbf{y}, \rho, T)$ minimizes $G(n, T)$ for $n \in \mathcal{C}(\rho)$, and

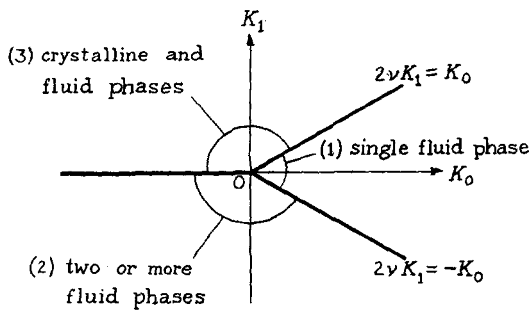


FIG. 2. The properties of the nearest-neighbor cell model as determined by the interaction constants K_0 and K_1.

has the form[12]

$$n^*(\mathbf{y}, \rho, T) = \rho_+(\rho, T) \quad \text{for } \sum y_k \text{ even},$$
$$= \rho_-(\rho, T) \quad \text{for } \sum y_k \text{ odd}, \quad (4.11)$$

where $\frac{1}{2}(\rho_+ + \rho_-) = \rho$ and $y_1 \cdots y_\nu$ are the components of $\mathbf{y}$: i.e., the system has a local density n^* with two (possibly equal) values ρ_+ and ρ_- arranged in a checkerboard (sodium chloride) pattern (see Fig. 3). When the bracket { } differs from its convex envelope, the system has two phases, both with a density of the form (4.11).

Before proving this theorem, we note some of its consequences. Using the properties[4] of the operation ME, one can show that (4.9) reduces to the van der Waals–Maxwell results (3.17) and (3.18) under conditions (4.5) and (4.7). Also, Theorems 2 and 3 of Ref. 4, together with (4.9), imply that (3.17) and (3.18) do *not* hold in region (3) of Fig. 2. More precisely, they imply the following.

Corollary 1: If $K_1 > 0$ and $2\nu K_1 > K_0$, and the function $a^0(\rho, T) + (K_0 - \frac{1}{2}\alpha)\rho^2$ is not convex in ρ (i.e., T is sufficiently low), then there are values of ρ for which

$$a(\rho, T) < CE[a^0(\rho, T) + \tfrac{1}{2}\alpha\rho^2]. \quad (4.12)$$

The set of such values of ρ includes (a) those intervals where $a^0 + \frac{1}{2}\alpha\rho^2$ differs from its convex envelope and also (b) those intervals where $a^0 + (K_0 - \frac{1}{2}\alpha)\rho^2$ differs from its midpoint envelope.

In the intervals (a) and (b) it follows from Theorem 3 that the system is in a state with (or has at least one phase with) a density of the form (4.11), where $\rho_+ \neq \rho_-$. Such states can be described as *spatially ordered* or *crystalline.*

We prove Theorem 3 by using Theorem 1 to obtain upper and lower bounds on $a(\rho, T)$. In the present case, it follows from (3.2) that

$$G(n) = \frac{1}{|\Gamma|} \sum_{\mathbf{y}\in\Gamma} \Big(a^0[n(\mathbf{y})] + \tfrac{1}{2}K_0 n(\mathbf{y})^2 + K_1 \sum_{k=1}^{\nu} n(\mathbf{y})n(\mathbf{y} + \mathbf{e}_k) \Big), \quad (4.13)$$

where $\mathbf{e}_1 \cdots \mathbf{e}_\nu$ are the unit vectors in $\mathbf{Z}^\nu$. An upper bound on $a(\rho)$ can be obtained by noting, from (3.6), that

$$a(\rho) \le G(n') \quad \text{for any} \quad n' \in \mathcal{C}(\rho). \quad (4.14)$$

Choosing n' to be given by the right side of (4.11) yields

$$G(n') = \tfrac{1}{2}[a^0(\rho_+) + a^0(\rho_-) + \tfrac{1}{2}K_0\rho_+^2 + \tfrac{1}{2}K_0\rho_-^2] + \nu K_1 \rho_+\rho_-. \quad (4.15)$$

Since $\frac{1}{2}(\rho_+ + \rho_-) = \rho$, we can put $\rho_\pm = \rho \pm h$, where h is any positive constant such that $\rho \pm h$ are in the domain of a^0. Then (4.14) and (4.15) yield

$$a(\rho) \le \tfrac{1}{2}\{a^0(\rho + h) + a^0(\rho - h) + (\tfrac{1}{2}K_0 - \nu K_1) \times [(\rho + h)^2 + (\rho - h)^2]\} + 2\nu K_1\rho^2. \quad (4.16)$$

Since h is arbitrary, we can minimize the right side with respect to h. Via (4.8), this gives

$$a(\rho) \le ME[a^0(\rho) + (\tfrac{1}{2}K_0 - \nu K_1)\rho^2] + 2\nu K_1\rho^2. \quad (4.17)$$

Using (4.6) and noting that $a(\rho)$ is convex, we obtain

$$a(\rho) \le CE\{ME[a^0(\rho) + (K_0 - \tfrac{1}{2}\alpha)\rho^2] + (\alpha - K_0)\rho^2\}. \quad (4.18)$$

To obtain a lower bound on $a(\rho)$, we first express G in a more convenient form. Let us put

$$\varphi(\rho) \equiv a^0(\rho) + \tfrac{1}{2}K_0\rho^2. \quad (4.19)$$

Since $n(\mathbf{y})$ is periodic, it follows that

$$\sum_{\mathbf{y}\in\Gamma} \varphi[n(\mathbf{y})] = \sum_{\mathbf{y}\in\Gamma} \varphi[n(\mathbf{y} + \mathbf{a})] \quad \text{for any} \quad \mathbf{a} \in \mathbf{Z}^\nu. \quad (4.20)$$

In particular, this holds if $\mathbf{a}$ is any of the $\mathbf{e}_k$. Thus (4.13) reduces to

$$G(n) = \frac{1}{|\Gamma|} \sum_{\mathbf{y}\in\Gamma} \frac{1}{\nu} \sum_{k=1}^{\nu} g_k(n, \mathbf{y}), \quad (4.21)$$

where

$$g_k(n, \mathbf{y}) \equiv \tfrac{1}{2}\varphi[n(\mathbf{y})] + \tfrac{1}{2}\varphi[n(\mathbf{y} + \mathbf{e}_k)] + \nu K_1 n(\mathbf{y})n(\mathbf{y} + \mathbf{e}_k)$$
$$= \tfrac{1}{2}\psi[n(\mathbf{y})] + \tfrac{1}{2}\psi[n(\mathbf{y} + \mathbf{e}_k)] + \tfrac{1}{2}\nu K_1[n(\mathbf{y}) + n(\mathbf{y} + \mathbf{e}_k)]^2, \quad (4.22)$$

where

$$\psi(\rho) \equiv \varphi(\rho) - \nu K_1\rho^2. \quad (4.23)$$

[$g_k(n, \mathbf{y})/\nu$ can be interpreted as the free energy of the bond between the cell at $\mathbf{y}$ and the cell at $\mathbf{y} + \mathbf{e}_k$.] From the definition of $ME\psi$ it follows that for all ρ_1 and ρ_2

$$\tfrac{1}{2}\psi(\rho_1) + \tfrac{1}{2}\psi(\rho_2) \ge ME\psi(\tfrac{1}{2}\rho_1 + \tfrac{1}{2}\rho_2), \quad (4.24)$$

which with (4.22) gives

$$g_k(n, \mathbf{y}) \ge \xi[\tfrac{1}{2}n(\mathbf{y}) + \tfrac{1}{2}n(\mathbf{y} + \mathbf{e}_k)], \quad (4.25)$$

where

$$\xi(\rho) \equiv ME\psi(\rho) + 2\nu K_1\rho^2. \quad (4.26)$$

ρ_+	ρ_-	ρ_+	ρ_-	ρ_+
ρ_-	ρ_+	ρ_-	ρ_+	ρ_-
ρ_+	ρ_-	ρ_+	ρ_-	ρ_+

FIG. 3. The crystal structure in the nearest-neighbor cell model (2-dimensional case) as given by (4.11).

But the definition of $CE\xi$ implies

$$\begin{aligned}\xi(\rho) &\geq CE\xi(\rho)\\ &= CE\{ME[a^0(\rho) + (\tfrac{1}{2}K_0 - \nu K_1)\rho^2] + 2\nu K_1\rho^2\}\\ &\equiv \zeta(\rho), \text{ say.} \end{aligned} \tag{4.27}$$

This gives

$$g_k(n, \mathbf{y}) \geq \zeta[\tfrac{1}{2}n(\mathbf{y}) + \tfrac{1}{2}n(\mathbf{y} + \mathbf{e}_k)]. \tag{4.28}$$

Substituting this in (4.21) and using the convexity of ζ gives

$$\begin{aligned}G(n) &\geq \zeta\left(\frac{1}{|\Gamma|}\sum_{y\in\Gamma}\frac{1}{\nu}\sum_{k=1}^{\nu}[\tfrac{1}{2}n(\mathbf{y}) + \tfrac{1}{2}n(\mathbf{y} + \mathbf{e}_k)]\right)\\ &= \zeta(\rho) \quad \text{for all} \quad n \in \mathfrak{C}(\rho),\end{aligned} \tag{4.29}$$

where the equality follows from (3.1). Since (4.29) holds for all n, it follows that

$$\inf_{n\in\mathfrak{C}(\rho)} G(n) \geq \zeta(\rho), \tag{4.30}$$

which together with (3.6) and (4.18) proves (4.9). The statements (4.10) and (4.11) follow from the argument leading to (4.18). This completes the proof of Theorem 3.

5. NEAREST-NEIGHBOR INTERACTIONS AND IDEAL REFERENCE SYSTEM

To analyze in detail the kind of thermodynamic behavior predicted by the equation of state (4.9), we consider in this section the special case

$$q(\mathbf{r}) = 0 \quad \text{for all } \mathbf{r}, \quad K_0 = 0. \tag{5.1}$$

Consequently, the reference system is an ideal gas. From (2.17) and (2.18) we see that (5.1) is a Type-II system, so that $K(\mathbf{r})$ must satisfy the conditions (2.18). One can easily show that these conditions are satisfied if and only if

$$K_1 \geq 0. \tag{5.2}$$

We now have

$$\alpha = 2\nu K_1 \geq 0, \tag{5.3}$$

so that (4.9) becomes

$$a(\rho, T) = CE\{ME[a^0(\rho, T) - \tfrac{1}{2}\alpha\rho^2] + \alpha\rho^2\}. \tag{5.4}$$

Also, a^0 is the free energy density of an ideal gas, i.e.,

$$a^0(\rho, T) = kT[\rho\log(\Lambda^\nu\rho) - \rho], \tag{5.5}$$

where Λ is the thermal wavelength.

Equation (5.4) can be simplified by using Lemma 5 of Ref. 4, which states that, for any function $f(\rho)$ and any constants L and M,

$$ME[f(\rho) + L\rho + M] = MEf(\rho) + L\rho + M, \tag{5.6}$$

$$CE[f(\rho) + L\rho + M] = CEf(\rho) + L\rho + M. \tag{5.7}$$

Substituting (5.5) in (5.4) and using (5.6) and (5.7) gives

$$a(\rho, T) = A\chi(B\rho) + C\rho, \tag{5.8}$$

where

$$\chi(\eta) \equiv CE\{ME[\eta\log\eta - \tfrac{1}{2}\eta^2] + \eta^2\}, \tag{5.9}$$

$$A(T) \equiv (kT)^2/\alpha, \tag{5.10}$$

$$B(T) \equiv \alpha/(kT), \tag{5.11}$$

$$C(T) \equiv kT[\log(\Lambda^\nu kT/\alpha) - 1]. \tag{5.12}$$

The numbers A and B are nonnegative, while A is an increasing and B a decreasing function of T. Hence, an immediate consequence of (5.8) is that there is *no critical temperature*. As T varies, the graph of $A\chi(B\rho)$ against ρ changes only in scale. The temperatures at which $C = 0$ are not significant because C appears only in the combination $C\rho$. The result (5.8) can be simplified further by using the following.

Lemma 1: The function

$$a_r(\eta) \equiv ME(\eta\log\eta - \tfrac{1}{2}\eta^2) + \eta^2 \tag{5.13}$$

is strictly convex.

Proof: We first prove that

$$a_r''(\eta) > 0 \quad \text{for} \quad \eta \neq 1. \tag{5.14}$$

From (4.13), one can write, for any continuous function f,

$$MEf(\eta) = \tfrac{1}{2}f[\eta + \delta(\eta)] + \tfrac{1}{2}f[\eta - \delta(\eta)], \tag{5.15}$$

where $\delta(\eta) \geq 0$ is the function which minimizes the right side. It follows that the right side is stationary with respect to variations in δ, which implies

$$f'[\eta + \delta(\eta)] = f'[\eta - \delta(\eta)]. \tag{5.16}$$

Using this, we deduce from (5.15) that

$$\frac{d}{d\eta}MEf(\eta) = f'[\eta + \delta(\eta)]. \tag{5.17}$$

The results (5.16) and (5.17) have a simple graphical interpretation. Let $MLg(\eta)$, called the *midpoint locus* of g, be defined for any g, as the locus of the midpoints of the horizontal chords of $g(\eta)$. It follows that if $MLf'(\eta)$ is single valued, then

$$\begin{aligned}\frac{d}{d\eta}MEf(\eta) &= f'(\eta), &&\text{where} \quad \delta(\eta) = 0,\\ &= MLf'(\eta), &&\text{where} \quad \delta(\eta) > 0.\end{aligned} \tag{5.18}$$

We now choose

$$f(\eta) \equiv \eta\log\eta - \tfrac{1}{2}\eta^2, \tag{5.19}$$

so that

$$f'(\eta) = \log \eta - \eta + 1. \tag{5.20}$$

By sketching f' (see Fig. 4) and its midpoint locus, we see that

$$\begin{aligned} \delta(\eta) &= 0 \quad \text{for} \quad \eta \leq 1 \\ &> 0 \quad \text{for} \quad \eta > 1. \end{aligned} \tag{5.21}$$

Hence from (5.18) we have

$$\begin{aligned} \theta(\eta) &\equiv \frac{d}{d\eta} MEf(\eta) \\ &= \log \eta - \eta + 1 \qquad \text{for} \quad \eta \leq 1 \\ &= ML[\log \eta - \eta + 1] \quad \text{for} \quad \eta \geq 1, \end{aligned} \tag{5.22}$$

and from (5.13)

$$a_r'(\eta) = \theta(\eta) + 2\eta. \tag{5.23}$$

From (5.22) we immediately have

$$\theta'(\eta) > 0 \quad \text{for} \quad \eta < 1, \tag{5.24}$$

so that

$$a_r''(\eta) > 2 > 0 \quad \text{for} \quad \eta < 1. \tag{5.25}$$

To complete the proof of (5.14), we use (5.16), (5.17), (5.20), and (5.22) to obtain, for $\eta > 1$,

$$\begin{aligned} \theta(\eta) &= \log(\eta + \delta) - (\eta + \delta) + 1 \\ &= \log(\eta - \delta) - (\eta - \delta) + 1. \end{aligned} \tag{5.26}$$

It follows from this that $\delta(\eta)$ is given implicitly by the equation

$$\delta \coth \delta = \eta. \tag{5.27}$$

[This implies that $\delta(\eta)$, and hence $\theta(\eta)$, are single valued, as required.] Differentiating (5.26) with respect to η and using (5.27) to obtain

$$\delta' = [\cosh(2\delta) - 1]/[\sinh(2\delta) - 2\delta] \tag{5.28}$$

yields, after simplification,

$$\begin{aligned} \theta'(\eta) &= -2 \\ &\quad + [\cosh(2\delta) - 1 - 2\delta^2]/\{\delta[\sinh(2\delta) - 2\delta]\}. \end{aligned} \tag{5.29}$$

By sketching graphs one finds that $\cosh x > 1 + \frac{1}{2}x^2$ and $\sinh x > x$ for $x > 0$, so that the second term in

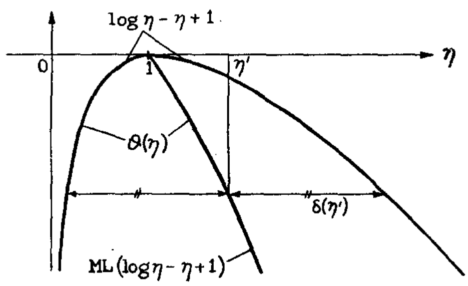


FIG. 4. Illustration of the functions $\theta(\eta)$ and $\delta(\eta)$.

(5.29) is positive for $\delta > 0$. It follows that

$$\theta'(\eta) > -2 \quad \text{for} \quad \eta > 1 \tag{5.30}$$

and from (5.23) that

$$a_r''(\eta) > 0 \quad \text{for} \quad \eta > 1. \tag{5.31}$$

This together with (5.25) proves (5.14). To complete the proof of Lemma 1, we note from (5.22) (see Fig. 4) that $\theta(\eta)$ is continuous, and is zero for $\eta = 1$, which, with (5.23), implies

$$a_r'(1+) = a_r'(1-) = 2. \tag{5.32}$$

This together with (5.14) proves Lemma 1. [The result (5.30) means that the midpoint locus shown in Fig. 4 has a gradient > -2.]

It follows from Lemma 1 that the *CE* in (5.9) can be dropped, so that (5.8) becomes

$$a(\rho, T) = Aa_r(B\rho) + C\rho. \tag{5.33}$$

Hence $a(\rho, T)$ is a strictly convex function of ρ, so that its graph has no straight line segments. This means that *the system has no first-order transitions.* Furthermore, since $a_r''(\eta)$ is discontinuous at $\eta = 1$, it follows that $\partial^2 a(\rho, T)/\partial\rho^2$ is discontinuous at $B\rho = 1$. Thus *a second-order transition occurs when*

$$\rho = kT/\alpha, \tag{5.34}$$

and this transition persists for all temperatures. The function $a_r(\eta)$, called the *reduced free energy*, determines the shape of the isotherms of $a(\rho, T)$.

The chemical potential is given by

$$\mu(\rho, T) \equiv \frac{\partial}{\partial \rho} a(\rho, T) = kT\mu_r(B\rho) + C, \tag{5.35}$$

where μ_r, called the *reduced chemical potential*, is defined by

$$\mu_r(\eta) \equiv a_r'(\eta). \tag{5.36}$$

The function μ_r, which gives the shape of the isotherms of $\mu(\rho, T)$, is sketched in Fig. 5(a). To find the gradient of $\mu_r(\eta)$ at $\eta = 1+$, we use (5.29) to obtain

$$\theta' \to -\tfrac{3}{2} \quad \text{as} \quad \delta \to 0 \tag{5.37}$$

(see also Fig. 4), which with (5.23) gives

$$\mu_r'(\eta) = a_r''(\eta) \to \tfrac{1}{2} \quad \text{as} \quad \eta \to 1+. \tag{5.38}$$

Figure 5(a) also gives the shape of the isotherms of the density $\rho(\mu, T)$ in the grand ensemble, because $\rho(\mu, T)$ is just the inverse function of $\mu(\rho, T)$ for constant T.

Note that for $\eta \geq 1$ the function $\mu_r(\eta)$ is given parametrically in terms of δ by Eqs. (5.23), (5.26), and (5.27). One can eliminate δ as follows. From these

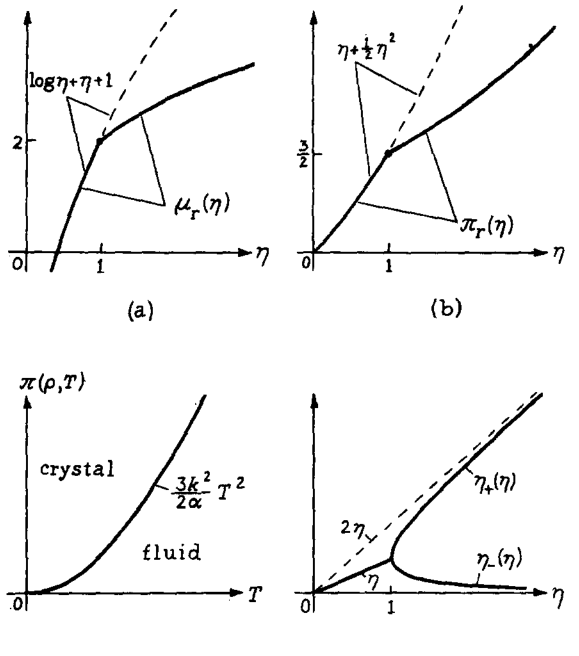


FIG. 5. (a) The reduced chemical potential μ_r which gives the true chemical potential through Eq. (5.35). (b) The reduced pressure π_r which gives the true pressure through Eq. (5.43). (c) the pressure–temperature phase diagram. (d) The functions $\eta_\pm$ which give the densities $\rho_\pm$ through Eq. (5.51).

equations we have

$$\begin{aligned}\theta &= \tfrac{1}{2}\log(\eta+\delta)(\eta-\delta) - \eta + 1\\ &= \mu_r - 2\eta,\end{aligned} \tag{5.39}$$

which gives

$$\delta = [\eta^2 - \exp(2\mu_r - 2\eta - 2)]^{\frac{1}{2}}. \tag{5.40}$$

Substituting this in (5.27) gives an implicit equation for μ_r of the form $\phi(\mu_r, \eta) = 0$. Alternatively, μ_r is given directly for $\eta > 1$ by the graphical construction

$$\mu_r(\eta) = \text{the locus of the midpoints of the chords of } (\log\eta + \eta + 1) \text{ of gradient 2.} \tag{5.41}$$

To find the canonical pressure

$$\pi(\rho, T) \equiv \rho^2 \frac{\partial}{\partial\rho}\left(\frac{a(\rho, T)}{\rho}\right), \tag{5.42}$$

we use (5.33) and obtain

$$\pi(\rho, T) = A\pi_r(B\rho), \tag{5.43}$$

where

$$\pi_r(\eta) \equiv \eta^2 \frac{\partial}{\partial\eta}\left(\frac{a_r(\eta)}{\eta}\right), \tag{5.44}$$

which we call the *reduced pressure*. The following results are easily deduced from the properties of a_r:

$$\begin{aligned}&\pi_r(\eta) = \eta + \tfrac{1}{2}\eta^2 \quad \text{for} \quad \eta \le 1,\\ &\pi_r'(\eta) = \begin{cases} 1 & \text{for} \quad \eta = 0\\ 2 & \text{for} \quad \eta = 1-\\ \tfrac{1}{2} & \text{for} \quad \eta = 1+\end{cases}\\ &\quad \to 1 \quad \text{as} \quad \eta \to \infty.\end{aligned} \tag{5.45}$$

The function π_r is sketched in Fig. 5(b). From (5.43) and (5.45) we obtain

$$\pi(\rho, T) = \rho kT + \tfrac{1}{2}\alpha\rho^2 \quad \text{for} \quad \rho \le kT/\alpha, \tag{5.46}$$

so that the transition occurs when

$$\pi = (3k^2/2\alpha)T^2 \tag{5.47}$$

or, equivalently, when

$$\pi = (3\alpha/2)\rho^2. \tag{5.48}$$

The phase diagram for π is sketched in Fig. 5(c). It would be a simple matter to plot the isotherms of π accurately by using a parametric formula for π_r in terms of δ.

In the crystalline phase the local density n^* has the form (4.11). The functions $\rho_\pm(\rho, T)$ are found by minimizing the right side of (4.15), which in the present case yields

$$kT\log(\Lambda^\nu\rho_\pm) + \alpha\rho_\mp = m, \tag{5.49}$$

where m is a Lagrange multiplier. This can be written as

$$\rho_\pm = (\Lambda^{-\nu}e^{m/kT})\exp(-\alpha\rho_\mp/kT), \tag{5.50}$$

which closely resembles the integral equation of the mean field theory of melting.[1,5] The solution is easily shown to be

$$\rho_\pm(\rho, T) = (kT/\alpha)\eta_\pm(\alpha\rho/kT), \tag{5.51}$$

where

$$\eta_\pm(\eta) \equiv \eta \pm \delta(\eta) \tag{5.52}$$

and $\delta(\eta)$ is given by (5.27). The functions $\eta_\pm$ are sketched in Fig. 5(d). The important features are that ρ_+ and ρ_- become rapidly unequal at the onset of crystallization and that at high densities or low temperatures the system approaches the density distribution $\rho_+ = 2\rho$ and $\rho_- = 0$.

6. CORRELATION FUNCTIONS

In order to understand better the structure of the different phases of the model, it is useful to consider the correlation functions. We consider first the k-particle distribution function $n_k(\mathbf{x}_1 \cdots \mathbf{x}_k, N, D, T, \gamma, \omega)$ for N particles in $\Omega(D)$ [see (2.9)], defined in the usual way[13] in terms of the 2-body potential (2.1).

Note that n_k is defined only for $\mathbf{x}_i \in \Omega(D)$. Following Fisher,[13] we define a space-averaged, infinite-volume distribution function by

$$\bar{n}_k(\mathbf{r}_1 \cdots \mathbf{r}_{k-1}, \rho, T, \gamma, \omega) \equiv \lim_{\substack{N,|D|\to\infty \\ N/|D|\to\rho\omega}} \frac{1}{|\Omega(D)|} \int_{\Omega(D)} d\mathbf{x}\, n_k(\mathbf{x}, \mathbf{x}+\mathbf{r}_1, \cdots \mathbf{x}+\mathbf{r}_{k-1}, N, D, T, \gamma, \omega). \quad (6.1)$$

To understand the structure of the system in individual cells, we consider the *short-range distribution function*

$$\bar{n}_k^{\mathrm{S}}(\mathbf{r}^{k-1}, \rho, T) \equiv \lim_{\omega\to\infty} \lim_{\gamma\to 0} \bar{n}_k(\mathbf{r}^{k-1}, \rho, T, \gamma, \omega), \quad (6.2)$$

where $\mathbf{r}^m \equiv (\mathbf{r}_1 \cdots \mathbf{r}_m)$. A similar function was defined and studied by the author in Ref. 14. Some general results for $\bar{n}_k^{\mathrm{S}}$ can be obtained from Eqs. (10)–(13) of this reference by replacing integrals with respect to $\mathbf{y}$ by summations with $\mathbf{y}$ in $\mathbf{Z}^\nu$ and by replacing n_ρ by n^*. For the particular system considered in Sec. 5, Eqs. (10) and (12) of Ref. 14 give

$$\begin{aligned}\bar{n}_2^{\mathrm{S}}(\mathbf{r}, \rho, T) &= \rho^2 && \text{for} \quad \rho \le kT/\alpha \text{ (fluid)} \\ &= \tfrac{1}{2}(\rho_+^2 + \rho_-^2) && \text{for} \quad \rho \ge kT/\alpha \text{ (crystal)}.\end{aligned} \quad (6.3)$$

The result for the crystal phase means that, in the short range (i.e., in individual cells), the system "looks like" a mixture of equal volumes of two different fluids of densities ρ_+ and ρ_-. This is expected, because there is no crystalline structure in individual cells.

To understand the structure of the general system over distances of order γ^{-1} [i.e., on the scale of the long-range potential $\omega^{-1}K(\gamma\mathbf{r})$ and of the cell separation], we consider the *long-range distribution function*

$$\bar{n}_k^{\mathrm{L}}(\mathbf{s}^{k-1}, \rho, T) \equiv \lim_{\omega\to\infty} \lim_{\gamma\to 0} \bar{n}_k\left(\frac{\mathbf{s}_1}{\gamma} \cdots \frac{\mathbf{s}_{k-1}}{\gamma}, \rho, T, \gamma, \omega\right). \quad (6.4)$$

Some general results for this function can be obtained from Eqs. (15) and (16) of Ref. 14 by again replacing integrations over $\mathbf{y}$ by sums. For the particular system considered in Sec. 5, these equations yield

$$\begin{aligned}\bar{n}_2^{\mathrm{L}}(\mathbf{s}, \rho, T) &= \rho^2 \quad \text{if} \quad \rho \le kT/\alpha \text{ (fluid)} \\ &\left.\begin{aligned} &= \tfrac{1}{2}(\rho_+^2 + \rho_-^2) && \text{for} \ \textstyle\sum s_\sigma \text{ even} \\ &= \rho_+\rho_- && \text{for} \ \textstyle\sum s_\sigma \text{ odd}\end{aligned}\right\} \\ &\qquad \text{if} \quad \rho \ge kT/\alpha \text{ (crystal)},\end{aligned} \quad (6.5)$$

where $s_1, \cdots, s_\nu$ are the components of $\mathbf{s}$. For the crystal phase, $\bar{n}_2^{\mathrm{L}}$ is therefore periodic with the same symmetry as n^*. It does not tend to ρ^2 as $|\mathbf{s}| \to \infty$, i.e., there is "long-range order."

Finally, we consider the *modified Ursell correlation function* $\hat{u}_k(\mathbf{x}_1 \cdots \mathbf{x}_k, N, D, T, \gamma, \omega)$, defined in terms of the n_k in the usual way.[15] We define their space averages $\bar{u}_k(\mathbf{r}^{k-1}, \rho, T, \gamma, \omega)$ as in (6.1). To understand the relevance of the Ornstein–Zernike theory[15] to the cell model, we consider the *weighted Ursell function*

$$\bar{u}_k^{\mathrm{W}}(\mathbf{s}^{k-1}, \rho, T) \equiv \lim_{\omega\to\infty} \lim_{\gamma\to 0} \omega^{k-1} \bar{u}_k\left(\frac{\mathbf{s}_1}{\gamma}, \cdots \frac{\mathbf{s}_{k-1}}{\gamma}, \rho, T, \gamma, \omega\right). \quad (6.6)$$

The general results for this function can be obtained from Eqs. (18)–(24) of Ref. 14 by again replacing the integrals over $\mathbf{y}$, $\mathbf{s}$, and $\mathbf{p}$ by sums in $\mathbf{Z}^\nu$.

For the system of Sec. 4, with nearest-neighbor interactions, which, in addition, is one dimensional ($\nu = 1$) and has $K_0 = 0$, we deduce from (20), (21), and (22) of Ref. 14 that for one-phase states

$$\bar{u}_2^{\mathrm{W}}(s) = \tfrac{1}{2}\mathcal{G}(0, s) + \tfrac{1}{2}\mathcal{G}(1, s+1), \quad (6.7)$$

where ρ and T dependence is omitted from the notation, and $\mathcal{G}(y, y')$ is the solution of the difference equation

$$\tfrac{1}{2}\alpha\mathcal{G}(y+1, y') + \tfrac{1}{2}\alpha\mathcal{G}(y-1, y') + a_2^0[n^*(y)]\mathcal{G}(y, y') = kT\delta_{y,y'}, \quad (6.8)$$

subject to the boundary condition

$$\mathcal{G}(y, y') \to 0 \quad \text{as} \quad |y - y'| \to \infty, \quad (6.9)$$

where

$$a_2^0(\rho) \equiv \frac{\partial^2 a^0(\rho)}{\partial \rho^2}. \quad (6.10)$$

In the case $\alpha < 0$, which yields the van der Waals–Maxwell result (3.18), we have $n^* = \rho$ for one-phase states, and (6.8) has the solution

$$\mathcal{G}(y, y') = kT[(a_2^0)^2 - \alpha^2]^{-\frac{1}{2}} \lambda^{|y-y'|}, \quad (6.11)$$

where

$$\lambda \equiv \{[(a_2^0)^2 - \alpha^2]^{\frac{1}{2}} - a_2^0\}/\alpha. \quad (6.12)$$

Since $a_2^0 > -\alpha$ for one-phase states (i.e., $a^0 + \frac{1}{2}\alpha\rho^2$ is convex), it follows that $0 < \lambda < 1$, so that (6.9) is satisfied. Combining (6.12) with (6.7) now gives

$$\bar{u}_2^{\mathrm{W}}(s, \rho, T) = kT[(a_2^0)^2 - \alpha^2]^{-\frac{1}{2}} \lambda^{|s|}. \quad (6.13)$$

This has the form of the one-dimensional Ornstein–Zernike formula.[15] In particular, $\bar{u}_2^{\mathrm{W}}$ becomes very long range ($\lambda \to 1$) as the critical point is approached ($a_2^0 + \alpha \to 0$). Unlike the Ornstein–Zernike formula, (6.13) is exact and holds for all $s \in \mathbf{Z}^1$.

In the case $\alpha > 0$ and $q = 0$, which yields the crystal states of Sec. 5, we have[12]

$$\begin{aligned} n^*(y) &= \rho \quad \text{if} \quad \rho \le kT/\alpha \\ &\left.\begin{aligned}&= \rho_+ \text{ for } y \text{ even} \\ &= \rho_- \text{ for } y \text{ odd}\end{aligned}\right\} \quad \text{if} \quad \rho \ge kT/\alpha, \end{aligned} \quad (6.14)$$

where ρ_+ and ρ_- are given by (5.53). Also, from (5.5), we have

$$a_2^0(\rho, T) = kT/\rho. \tag{6.15}$$

For $\rho < kT/\alpha$ we again obtain (6.11)–(6.13), but now with $-1 < \lambda < 0$, so that $\bar{u}_2^W$ alternates in sign as s varies. Now $\bar{u}_2^W$ becomes very long range ($\lambda \to -1$) at the freezing point. This is in contrast with the behavior of the thermodynamic functions, whose analytic forms in the fluid phase do not indicate the freezing transition [cf. (5.46)]. For $\rho > kT/\alpha$, one obtains from (6.8), after a fair amount of work,

$$\mathcal{G}(y, y') = h(y)h(y')g(y - y'), \tag{6.16}$$

where

$$h(y) \equiv \rho_+^{\frac{1}{2}} \quad \text{for } y \text{ even}$$
$$\equiv \rho_-^{\frac{1}{2}} \quad \text{for } y \text{ odd} \tag{6.17}$$

and

$$g(s) \equiv (1 - \eta_+\eta_-)^{-\frac{1}{2}}\kappa^{|s|}. \tag{6.18}$$

Here $\eta_\pm$ are given by (5.51) and (5.52), and

$$\kappa \equiv [(1 - \eta_+\eta_-)^{\frac{1}{2}} - 1]/(\eta_+\eta_-)^{\frac{1}{2}}. \tag{6.19}$$

Combining these with (6.7) yields, for $\rho > kT/\alpha$,

$$\bar{u}_2^W(s, \rho, T) = \rho g(s) \qquad \text{for } s \text{ even}$$
$$= (\rho_+\rho_-)^{\frac{1}{2}}g(s) \quad \text{for } s \text{ odd}. \tag{6.20}$$

From (5.52) and (5.27) we have

$$\eta_+\eta_- = \delta^2/\sinh^2\delta < 1 \quad \text{for} \quad \delta > 0, \tag{6.21}$$

which implies that κ is real and $-1 < \kappa < 0$, so that $g(s)$ alternates in sign as s varies. The result (6.20) again resembles the Ornstein–Zernike formula, but is modified by the crystal structure. Again $\bar{u}_2^W$ becomes very long range at the melting transition because, from (6.19) and (6.21), $\kappa \to -1$ as $\delta \to 0$. It is interesting to note that in the present case ($\alpha > 0$) the function $\bar{u}_2^W$ becomes long range at the fluid–crystal transition *for all temperatures*, while for the previous case ($\alpha < 0$) it becomes long range only at the critical temperature and density of the gas–liquid transition.

The function $\bar{u}_2^W$ is related to the pressure π [Eq. (5.42)] by the compressibility formula

$$\frac{kT\rho}{\partial\pi(\rho)/\partial\rho} = \sum_{s\in Z^1} \bar{u}_2^W(s, \rho). \tag{6.22}$$

This can be either deduced from the general results of Ref. 14, or verified directly from (6.13) and (6.20). For the gas–liquid system ($\alpha < 0$) the sum in (6.22) diverges at the critical point, and hence $\partial\pi/\partial\rho \to 0$ as expected. But, for the crystal–fluid system ($\alpha > 0$) the sum does *not* diverge at the melting or the freezing side of the transition. Even though $\bar{u}_2^W$ becomes long range, the fact that it alternates in sign results in a finite sum (note that $\sum |\bar{u}_2^W|$ does diverge[16]). Hence $\partial\pi/\partial\rho$ tends to (different) nonzero values on each side of the transition, as indicated earlier in (5.45) and Fig. 5(b). It would be interesting to know the extent to which these results can be generalized to melting transitions in more realistic models.

The results of this section can be derived by the method of functional differentiation outlined in Ref. 14. The derivation is not quite rigorous but is based, like that of Lebowitz and Penrose,[9] on the assumption that certain limits and derivatives exist.

7. MAGNETIC SYSTEMS

As pointed out in Sec. 1, the model of this paper can be applied to magnetic systems, and yields as special cases both the Weiss theory of ferromagnetism and the Néel–van Vleck theory of antiferromagnetism.[17] In this section we outline how this occurs.

We consider a system of N spins $\sigma_1, \cdots, \sigma_N$ arranged on those sites $\mathbf{x}_1, \cdots, \mathbf{x}_N$ of a ν-dimensional lattice (of arbitrary lattice constant) which lie in the set $\Omega(D)$ defined by (2.9), i.e., only the cells $\{\omega(\mathbf{y}/\gamma)\}$ contain spins. The Hamiltonian is

$$\sum_{1 \le a < b \le N} \sigma_a\sigma_b v(\mathbf{x}_a - \mathbf{x}_b, \gamma, \omega), \tag{7.1}$$

where v is given by (2.1). Using the canonical formalism, rather than the more usual grand canonical formalism, one can define the free energy per spin $\tilde{a}(\rho, T, \gamma, \omega)$, where ρ is now the magnetization per spin and

$$-1 \le \rho \le 1. \tag{7.2}$$

The free energy $a(\rho, T)$ of the model is then defined by (2.2).

One finds that the variational principle (3.6) again holds if the condition $n(\mathbf{y}) \ge 0$ in the definition of $\mathcal{C}(\rho)$ is replaced by

$$-1 \le n(\mathbf{y}) \le 1. \tag{7.3}$$

Here, the random function $n(\mathbf{y})$ represents the local magnetization in the cell at $\mathbf{y}$.

Theorem 2 also holds for the magnetic case. We show that the Weiss theory of ferromagnetism follows from Eq. (3.18) of Theorem 2 if $q = 0$. In this case[18]

$$a^0(\rho, T) = kT\{\tfrac{1}{2}(1 + \rho)\log[\tfrac{1}{2}(1 + \rho)] + \tfrac{1}{2}(1 - \rho)\log[\tfrac{1}{2}(1 - \rho)]\}. \tag{7.4}$$

In the canonical formalism, the magnetic field $H(\rho, T)$ is given as a function of magnetization ρ by

$$H(\rho, T) = \frac{\partial a(\rho, T)}{\partial \rho}. \tag{7.5}$$

Combining this with (7.4) and (3.18) gives, with $\alpha < 0$,

$$H(\rho, T) = 0, \quad \text{for} \quad |\rho| \leq (1 + kT/\alpha)^{\frac{1}{2}},$$
$$= kT \tanh^{-1} \rho + \alpha\rho, \quad \text{otherwise}, \quad (7.6)$$

which is just the main result of the Weiss theory.[18]

All the results of Sec. 4 hold for magnetic systems without modification. The crystalline states can here be interpreted as states with antiferromagnetic ordering. We now show that the special case considered in Sec. 5 yields precisely the Néel-van Vleck theory of antiferromagnetism. In this case (5.4) holds, with a^0 given by (7.4). The simplification leading to (5.8) does not apply here, so that the behavior of the system is more complicated. However, one can prove the analog of Lemma 1; viz., if $\alpha > 0$, the function

$$\psi(\rho, T) \equiv ME[a^0(\rho, T) - \tfrac{1}{2}\alpha\rho^2] + \alpha\rho^2 \quad (7.7)$$

is a strictly convex function of ρ for all $T > 0$. This means that

$$a(\rho, T) = \psi(\rho, T) \quad (7.8)$$

and that $a(\rho, T)$ is strictly convex, so that from (7.5) ρ is a continuous function of H. Hence, there are no two-phase states (consisting of two phases with different magnetization) and consequently no ferromagnetic transitions. (This question seems to have been overlooked in the original derivations.) Now, using essentially the argument of Sec. 5, we obtain from (7.5) and (7.8)

$$H(\rho, T) = kT \tanh^{-1} \rho + \alpha\rho \quad \text{for} \quad |\rho| \geq (1 - kT/\alpha)^{\frac{1}{2}} \quad (7.9)$$

and

$$H(\rho, T) = kT \tanh^{-1} (\rho \pm \delta) - \alpha(\rho \pm \delta) + 2\alpha\rho \quad \text{for} \quad |\rho| \leq (1 - kT/\alpha)^{\frac{1}{2}}, \quad (7.10)$$

where $\alpha > 0$ and $\delta(\rho, T)$ is given implicitly by the equation

$$\rho^2 = 1 + \delta^2 - 2\delta \coth (2\alpha\delta/kT). \quad (7.11)$$

[The two alternative expressions for H in (7.10) are equal by virtue of (7.11), cf. (5.26)]. This is precisely the main result of the Néel-van Vleck theory.[19] It implies that there is ordering whenever $|\rho| \leq (1 - kT/\alpha)^{\frac{1}{2}}$. Consequently, there is no ordering for any ρ if $T > T_N$, where

$$T_N \equiv \alpha/k, \quad (7.12)$$

which is called the Néel temperature. Some other results are given by Garrett.[17] Besides these, one can show that, although $H(\rho, T)$ is a continuous function of ρ, its gradient is discontinuous at the order-disorder transition. (I have obtained some other exact results about the magnetization and susceptibility curves, analogous to those in Sec. 5, which seem to be new but not worthy of publication.)

The derivation of the Néel-van Vleck theory outlined here is an advance on the original derivations in two main respects: (i) It is based on a statistical model, and (ii) the existence and the checkerboard symmetry of the antiferromagnetic states has been proved, not assumed, i.e., the introduction of "sublattices" has been avoided.

8. DISCUSSION

The main results of this paper are summarized in the abstract.

It would be interesting to study cases of the model which are more general than the nearest-neighbor case. The simplest problems arise in one-dimensional systems. For example, in the case of nearest- and next-nearest-neighbor interactions, what is the crystal structure? It seems that the periods 2 and 3 are both possible, or perhaps the period 6 occurs, or even a transition between two crystalline states of different periods. If the next-neighbor interactions are sufficiently attractive, it seems that the system could have both a gas-liquid transition (of the first-order van der Waals-Maxwell type) and a second-order liquid-crystal transition. Such a system may have a triple point, which would be interesting to study. Possibly, this also happens in the nearest-neighbor case of Sec. 5 if $K_0 < 0$.

For the case where the interactions extend to even more neighbors, it is not easy to see what the periods of the crystal states will be, especially in more than one dimension. Possibly, some sort of group theory could be used to study this problem in a general way. The foundations of such a theory have already been formulated.[8]

The model is related to the mean-field theory of melting [see Eqs. (3.6) and (5.50) and Refs. 1 and 5]. We therefore suspect that the latter theory will give a similar second-order melting transition with no critical temperature, but this has yet to be proved.

Apart from its obvious artificiality, the model does not properly duplicate the mechanism by which melting and freezing are currently believed to occur, namely,[7] by the geometrical disordering and ordering of hard (or sufficiently hard) spherical particles. The model in this paper "freezes" simply because repulsion between particles in neighboring cells favors a nonuniform density. The model is more realistic for antiferromagnets, since it roughly duplicates the forces which cause opposite alignment of neighboring spins in real antiferromagnets.

ACKNOWLEDGMENTS

I am grateful to O. E. Maly, C. Orton, E. R. Smith, and especially Professor O. Penrose for helpful advice, and to Professor C. Domb for acquainting me with some useful references. Also, I gratefully acknowledge financial support from "The Royal Commission for the Exhibition of 1851."

[1] J. G. Kirkwood and E. Monroe, J. Chem. Phys. **8**, 623 (1940).
[2] J. E. Lennard-Jones and A. F. Devonshire, Proc. Phys. Soc. (London) **A163**, 53 (1937); **A165**, 1 (1938).
[3] D. J. Gates and O. Penrose, Commun. Math. Phys. **15**, 255 (1969) (Part I); **16**, 231 (1970) (Part II).
[4] D. J. Gates and O. Penrose, Commun. Math. Phys. **17**, 194 (1970).
[5] R. Brout, Physica **29**, 1041 (1963).
[6] B. J. Adler and T. E. Wainwright, Phys. Rev. **127**, 359 (1962).
[7] D. S. Gaunt and M. E. Fisher, J. Chem. Phys. **43**, 2840 (1965). For a related model which is solvable at one special temperature, see M. E. Fisher, J. Math. Phys. **4**, 278 (1963), and for a very recent solvable model see R. J. Baxter, J. Math. Phys. **11**, 3116 (1970).
[8] D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969), Chaps. 6 and 7; G. Emch, H. J. F. Knops, and E. J. Verboven, J. Math. Phys. **11**, 1655 (1970). The cell model illustrates explicitly many of the general results in these references: e.g., violation of the "uniform clustering" condition (Emch *et al.*) is illustrated by our Eq. (6.5). A different clustering condition ($\Sigma\,|\bar{u}_2^W| < \infty$) breaks down at the melting transition [see our Eq. (6.20) et seq.]. Also, absence of a critical point for melting transitions (Emch *et al.*) is confirmed by the cell model. Furthermore, the variational principle (3.6) (or its grand canonical equivalent[8]) is closely related to that obtained by Ruelle.
[9] J. L. Lebowitz and O. Penrose, J. Math. Phys. **7**, 98 (1966); M. Kac, G. E. Uhlenbeck, and P. C. Hemmer, J. Math. Phys. **4**, 216, (1963).
[10] We do not need to give $q(\mathbf{r})$ a hard core here, because the "corridors" used in Refs. 3 and 9 are empty in the present model.
[11] M. E. Fisher, Arch. Ratl. Mech. Anal. **17**, 377 (1964).
[12] There are two choices for n^*, namely (4.11) and the function obtained by interchanging ρ_+ and ρ_- in (4.11). On the other hand, the space-averaged correlations (Sec. 6) are unique.
[13] M. E. Fisher, J. Math. Phys. **6**, 1643 (1965).
[14] D. J. Gates, J. Phys., Gen. Phys. (Proc. Phys. Soc. London) **3**, L11 (1970).
[15] J. L. Lebowitz and J. K. Percus, J. Math. Phys. **4**, 248 (1963), Eqs. (2.2) and (5.10).
[16] This corresponds to the "staggered compressibility" of Gaunt and Fisher.[7]
[17] L. Néel, Ann. Phys. **5**, 256 (1933); J. H. van Vleck, J. Chem. Phys. **9**, 85 (1941); C. G. B. Garrett, J. Chem. Phys. **19**, 1154 (1951). The Néel–van Vleck theory does not agree well with experiment in certain details. A better, but not exactly solvable, model is the antiferromagnetic Ising model: see M. F. Sykes and M. E. Fisher, Physica **28**, 919, 939 (1962). For another exactly solvable antiferromagnetic model, see M. E. Fisher, Proc. Roy. Soc. (London) **A234**, 66 (1960); **A256**, 502 (1960).
[18] This formalism is not widely used. See, however, D. ter Haar, *Elements of Statistical Mechanics* (Rinehart, New York, 1954), Eqs. (12.225) and (12.226).
[19] C. G. B. Garrett, Ref. 16, Eqs. (5)–(7).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Properties of the Residues in the Veneziano Model*

P. N. Dobson, Jr., and W. A. Simmons
Department of Physics and Astronomy, University of Hawaii
Honolulu, Hawaii

(Received 19 August 1970)

Explicit formulas for the expansion coefficients of the Veneziano amplitude are derived. We show that the problem of positivity can be reduced to a study of the zeros of certain polynomials.

I. INTRODUCTION

Since the original proposal by Veneziano[1] of a simple functional form for a scattering amplitude with crossing symmetry and duality, there has been intensive study of both the mathematical properties[2] of the model and its applications to hadron phenomenology. Despite this intense investigation, many fundamental features of even the simplest form of the Veneziano model remain unresolved. In particular, the poles which appear in the familiar "partial fraction" expansion of the model should have positive residues; otherwise, the poles would have to be interpreted as due to negative norm intermediate states (ghosts). While much work has been done on this problem,[2–5] there has yet to be a definitive determination of a region of model parameters for which all residues are positive. It is true that a solution to this problem would not necessarily eliminate the ghost problem in the Veneziano model. Models such as that put forward by Nambu[6] and co-workers, which deal with the factorizability of the residues, indicate that positive residues in the four-point function might have to be interpreted as sums of residues from sets of degenerate poles, some of which are ghosts. Nonetheless, a first step would seem to be a solution of the positivity problem for the simple model.

Our purpose in this paper is to present a number of mathematical results which are central to the study of the positivity of the residues. We derive explicit formulas for the coefficients of the partial wave expansion of the Veneziano amplitude and show that the problem of positivity can be reduced to a study of

ACKNOWLEDGMENTS

I am grateful to O. E. Maly, C. Orton, E. R. Smith, and especially Professor O. Penrose for helpful advice, and to Professor C. Domb for acquainting me with some useful references. Also, I gratefully acknowledge financial support from "The Royal Commission for the Exhibition of 1851."

[1] J. G. Kirkwood and E. Monroe, J. Chem. Phys. **8**, 623 (1940).

[2] J. E. Lennard-Jones and A. F. Devonshire, Proc. Phys. Soc. (London) **A163**, 53 (1937); **A165**, 1 (1938).

[3] D. J. Gates and O. Penrose, Commun. Math. Phys. **15**, 255 (1969) (Part I); **16**, 231 (1970) (Part II).

[4] D. J. Gates and O. Penrose, Commun. Math. Phys. **17**, 194 (1970).

[5] R. Brout, Physica **29**, 1041 (1963).

[6] B. J. Adler and T. E. Wainwright, Phys. Rev. **127**, 359 (1962).

[7] D. S. Gaunt and M. E. Fisher, J. Chem. Phys. **43**, 2840 (1965). For a related model which is solvable at one special temperature, see M. E. Fisher, J. Math. Phys. **4**, 278 (1963), and for a very recent solvable model see R. J. Baxter, J. Math. Phys. **11**, 3116 (1970).

[8] D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969), Chaps. 6 and 7; G. Emch, H. J. F. Knops, and E. J. Verboven, J. Math. Phys. **11**, 1655 (1970). The cell model illustrates explicitly many of the general results in these references: e.g., violation of the "uniform clustering" condition (Emch *et al.*) is illustrated by our Eq. (6.5). A different clustering condition ($\sum |\bar{u}_2^W| < \infty$) breaks down at the melting transition [see our Eq. (6.20) et seq.]. Also, absence of a critical point for melting transitions (Emch *et al.*) is confirmed by the cell model. Furthermore, the variational principle (3.6) (or its grand canonical equivalent[8]) is closely related to that obtained by Ruelle.

[9] J. L. Lebowitz and O. Penrose, J. Math. Phys. **7**, 98 (1966); M. Kac, G. E. Uhlenbeck, and P. C. Hemmer, J. Math. Phys. **4**, 216, (1963).

[10] We do not need to give $q(\mathbf{r})$ a hard core here, because the "corridors" used in Refs. 3 and 9 are empty in the present model.

[11] M. E. Fisher, Arch. Ratl. Mech. Anal. **17**, 377 (1964).

[12] There are two choices for n^*, namely (4.11) and the function obtained by interchanging ρ_+ and ρ_- in (4.11). On the other hand, the space-averaged correlations (Sec. 6) are unique.

[13] M. E. Fisher, J. Math. Phys. **6**, 1643 (1965).

[14] D. J. Gates, J. Phys., Gen. Phys. (Proc. Phys. Soc. London) **3**, L11 (1970).

[15] J. L. Lebowitz and J. K. Percus, J. Math. Phys. **4**, 248 (1963), Eqs. (2.2) and (5.10).

[16] This corresponds to the "staggered compressibility" of Gaunt and Fisher.[7]

[17] L. Néel, Ann. Phys. **5**, 256 (1933); J. H. van Vleck, J. Chem. Phys. **9**, 85 (1941); C. G. B. Garrett, J. Chem. Phys. **19**, 1154 (1951). The Néel–van Vleck theory does not agree well with experiment in certain details. A better, but not exactly solvable, model is the antiferromagnetic Ising model: see M. F. Sykes and M. E. Fisher, Physica **28**, 919, 939 (1962). For another exactly solvable antiferromagnetic model, see M. E. Fisher, Proc. Roy. Soc. (London) **A234**, 66 (1960); **A256**, 502 (1960).

[18] This formalism is not widely used. See, however, D. ter Haar, *Elements of Statistical Mechanics* (Rinehart, New York, 1954), Eqs. (12.225) and (12.226).

[19] C. G. B. Garrett, Ref. 16, Eqs. (5)–(7).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Properties of the Residues in the Veneziano Model*

P. N. Dobson, Jr., and W. A. Simmons
Department of Physics and Astronomy, University of Hawaii
Honolulu, Hawaii

(Received 19 August 1970)

Explicit formulas for the expansion coefficients of the Veneziano amplitude are derived. We show that the problem of positivity can be reduced to a study of the zeros of certain polynomials.

I. INTRODUCTION

Since the original proposal by Veneziano[1] of a simple functional form for a scattering amplitude with crossing symmetry and duality, there has been intensive study of both the mathematical properties[2] of the model and its applications to hadron phenomenology. Despite this intense investigation, many fundamental features of even the simplest form of the Veneziano model remain unresolved. In particular, the poles which appear in the familiar "partial fraction" expansion of the model should have positive residues; otherwise, the poles would have to be interpreted as due to negative norm intermediate states (ghosts). While much work has been done on this problem,[2–5] there has yet to be a definitive determination of a region of model parameters for which all residues are positive. It is true that a solution to this problem would not necessarily eliminate the ghost problem in the Veneziano model. Models such as that put forward by Nambu[6] and co-workers, which deal with the factorizability of the residues, indicate that positive residues in the four-point function might have to be interpreted as sums of residues from sets of degenerate poles, some of which are ghosts. Nonetheless, a first step would seem to be a solution of the positivity problem for the simple model.

Our purpose in this paper is to present a number of mathematical results which are central to the study of the positivity of the residues. We derive explicit formulas for the coefficients of the partial wave expansion of the Veneziano amplitude and show that the problem of positivity can be reduced to a study of the zeros of certain polynomials. While the results contained here have probably been derived by other workers interested in this problem, they have not been published, and their collection in one paper will prove useful to other investigators.

II. PROPERTIES OF THE RESIDUES IN THE VENEZIANO MODEL

In the body of this paper we will confine our attention to the function[1]

$$V(s, t) = \frac{\Gamma(1 - \alpha(s))\Gamma(1 - \alpha(t))}{\Gamma(1 - \alpha(s) - \alpha(t))}, \qquad (1)$$

where $\alpha(x) = \alpha_0 + \alpha' x$. Remarks on the general case will be found in the Appendix. As is well known, this function has the partial fraction expansion[2]

$$V(s, t) = \sum_{K=1}^{\infty} \frac{\Gamma(K + \alpha(t))}{\Gamma(K)\Gamma(\alpha(t))} \frac{1}{\alpha(s) - K}.$$

In general, each pole in the expansion appears in a number of partial waves. Since it is the residue corresponding to a state with definite angular momentum that is of physical interest, we must consider the partial wave expansion of the expression

$$T_K(\alpha) = \frac{\Gamma(K + \alpha(t))}{\Gamma(\alpha(t))} = \alpha(\alpha + 1)\cdots(\alpha + K - 1). \qquad (2)$$

For convenience, we will consider only the case of equal mass scattering (such as $\pi\pi$ scattering). Then we can write

$$\alpha(t) = \alpha_0 + \alpha' t = 2\alpha' k^2 z + (\alpha_0 - 2\alpha' k^2) \equiv az + b,$$

where z is the cosine of the scattering angle and k is the center of mass momentum. $T_K(\alpha)$ is thus seen to be a polynomial of degree K in z, and so we can expand it as follows:

$$T_K(\alpha) = \sum_{J=0}^{K} (2J + 1) C_J^K(a, b) P_J(z).$$

To find the coefficients C_J^K, it is convenient first to write $T_K(\alpha)$ in the form

$$T_K(\alpha) = \sum_{m=1}^{K} S_K^{(m)} \alpha^m, \qquad (3)$$

where we have the explicit expression for the Stirling number

$$S_K^{(m)} = \sum_{l=0}^{K-m} \sum_{n=0}^{l} \frac{(-1)^{K+m+n}}{l!} \times \binom{K-1+l}{K-m+l} \binom{2K-m}{K-m-l} \binom{l}{n} n^{K-m+l}.$$

For computational purposes, it is more convenient to generate the $S_K^{(m)}$ from the recursion relations:

$$S_K^{(1)} = (K-1)!, \quad S_K^{(K)} = 1,$$
$$S_K^{(m)} = S_{K-1}^{(m-1)} + (K-1) S_{K-1}^{(m)}, \quad 1 < m < K. \qquad (4)$$

Now we substitute $\alpha(t) = az + b$ in (3), use the binomial expansion, and obtain

$$T_K(\alpha) = \sum_{j=0}^{K} \sum_{m=j}^{K} S_K^{(m)} \binom{m}{j} b^{m-j} a^j z^j,$$

where, for convenience in reordering the sum, we have defined $S_K^{(0)} = 0$ for all K. Further reduction is made by using the expansions

$$z^{2k} = \sum_{r=0}^{k} (4r + 1) \gamma_{k-r}^{2r} P_{2r}(z)$$

and

$$z^{2k+1} = \sum_{r=0}^{k} (4r + 3) \gamma_{k-r}^{2r+1} P_{2r+1}(z),$$

where

$$\gamma_m^J = (J + 2m)!/2^m m!\,(2J + 2m + 1)!!.$$

This leads us to the result

$$C_J^K(a, b) = \sum_{m=0}^{[(K-J)/2]} \gamma_m^J a^{J+2m} B_{K-J-2m}^K(b),$$

where $[x]$ means the largest integer contained in x and

$$B_N^K(b) = \sum_{l=0}^{N} S_K^{(l+K-N)} \binom{l + K - N}{l} b^l. \qquad (5)$$

The functions $B_N^K(b)$ have a number of interesting properties. For example, it follows from (3) and (5) that

$$B_{K-n}^K(b) = \frac{1}{n!} \frac{d^n}{db^n} T_K(b). \qquad (6)$$

The recursion relations (4) for the $S_K^{(m)}$ imply the following recursion relations for the B_N^K:

$$B_0^K = 1, \quad B_K^K = T_K(b) = \Gamma(b + K)/\Gamma(b),$$
$$B_N^K = (b + K - 1) B_{N-1}^{K-1} + B_N^{K-1}, \quad 0 < N < K.$$

These, in turn, lead to

$$C_J^K = a\left[\left(\frac{J}{2J+1}\right) C_{J-1}^{K-1} + \left(\frac{J+1}{2J+1}\right) C_{J+1}^{K-1}\right] + (b + K - 1) C_J^{K-1}.$$

When considering residues, we are interested in $C_J^K(a, b)$ evaluated at a_K and b_K determined by the pole condition

$$\alpha(s) = K.$$

This means

$$a_K = a_1 + (K-1)/2, \quad b_K = b_1 - (K-1)/2,$$

where

$$a_1 = -\tfrac{1}{2}(\alpha_0 + 4\alpha'\mu^2 - 1),$$

$$b_1 = \tfrac{1}{2}(3\alpha_0 + 4\alpha'\mu^2 - 1),$$

and μ is the particle mass.

For our purposes in the next section, it is useful to consider some general properties of the $B_N^K(b_K)$ considered as functions of b_1. First we write

$$B_K^K(b_K) = b_K(b_K + 1) \cdots (b_K + K - 1).$$

Now consider K even; i.e., $K = 2k$. Then

$$b_{2k} = b_1 - k + \tfrac{1}{2}$$

and

$$\begin{aligned} B_{2k}^{2k}(b_{2k}) &= (b_1 - k + \tfrac{1}{2}) \cdots (b_1 - \tfrac{1}{2}) \\ &\quad \times (b_1 + \tfrac{1}{2}) \cdots (b_1 + k - \tfrac{1}{2}) \\ &= (b_1^2 - \tfrac{1}{4})(b_1^2 - \tfrac{9}{4}) \cdots [b_1^2 - (k - \tfrac{1}{2})^2]. \end{aligned}$$

The function is thus seen to depend only on even powers of b_1. Furthermore, from Eq. (6), we see that if K is even, $B_N^K(b_K)$ will depend only on even or odd powers of b_1 according to whether N is even or odd.

Now consider the case of odd K; i.e., $K = 2k + 1$. Then,

$$b_{2k+1} = b_1 - k$$

and

$$\begin{aligned} B_{2k+1}^{2k+1}(b_{2k+1}) &= (b_1 - k) \cdots (b_1 - 1) \\ &\quad \times b_1(b_1 + 1) \cdots (b_1 + k) \\ &= b_1(b_1^2 - 1)(b_1^2 - 4) \cdots (b_1^2 - k^2). \end{aligned}$$

This is a function involving only odd powers of b_1. Equation (6) again implies that $B_N^K(b_K)$ depends only on even or odd powers of b_1 according to whether N is even or odd.

III. POSITIVITY OF THE RESIDUES

We look first at the leading trajectory. Here we desire to have

$$C_K^K(a_K, b_K) = \gamma_0^K(a\quad)^K \geq 0,$$

which is easily satisfied for all K if

$$a_1 \geq 0.$$

This means we must require

$$\alpha_0 \leq 1 - 4\alpha'\mu^2.$$

The first daughter trajectory gives us

$$C_{K-1}^K(a_K, b_K) = \gamma_0^{K-1}(a_K)^{K-1} B_1^K(b_K) \geq 0.$$

But the formulas of the last section easily provide the result

$$B_1^K(b_K) = K b_1,$$

so that positivity of the first daughter residues is guaranteed by taking

$$b_1 \geq 0 \quad \text{or} \quad \alpha_0 \geq \tfrac{1}{3} - \tfrac{4}{3}\alpha'\mu^2.$$

This is, of course, the same condition derived by Oehme.[4]

Before considering additional trajectories, it is well to make some use of the properties of the B_N^K derived at the end of the last section. Basically, these properties allow us to write the residue in the form

$$\begin{aligned} C_J^K(a_K, b_K) &= (a_K)^J f_J^K(a_K^2, b_1^2), \qquad K - J \text{ even}, \\ &= (a_K)^J b_1 g_J^K(a_K^2, b_1^2), \quad K - J \text{ odd}. \end{aligned}$$

The useful feature of this is that the factors in front of the functions f_J^K and g_J^K are already positive under the restrictions developed for the leading and first daughter trajectory. Furthermore, f_J^K and g_J^K are polynomials in the variables a_K^2 and b_1^2, and they are positive for sufficiently large $|a_K|$ or $|b_1|$. This means that we will have positivity outside some region centered on $b_1 = 0$, $a_1 = -(K - 1)/2$, whose boundary is determined from $f_J^K = 0$ or $g_J^K = 0$. The main task now is to determine to what extent these regions encroach on the quadrant

$$a_1 \geq 0, \quad b_1 \geq 0. \tag{7}$$

Let us now turn to the second daughter trajectory. Here we find that

$$f_{K-2}^K = \frac{K!}{2!\,(2K-3)!!}\left(b_1^2 + \frac{1}{2K-1}\,a_K^2 - \frac{K+1}{12}\right),$$

so that we have positivity outside the ellipses

$$b_1^2 + \frac{1}{2K-1}\left(a_1 + \frac{K-1}{2}\right)^2 = \frac{K+1}{12}$$

for $K = 2, 3, 4, \cdots$. Numerical calculations show that the case $K = 2$ dominates in the quadrant of interest, so that we have positivity of residues for the first three trajectories in the quadrant (7) for points outside the ellipse

$$b_1^2 + \tfrac{1}{3}(a_1 + \tfrac{1}{2})^2 = \tfrac{1}{4}. \tag{8}$$

Note that the point $a_1 = \frac{1}{4}$, $b_1 = \frac{1}{4}$, which corresponds to $\alpha_0 = \frac{1}{2}$ and $\alpha'\mu^2 = 0$, lies on this ellipse. This is just the well-known decoupling of the 0^+ daughter of the f^0 in $\pi\pi$ scattering in the limit of zero pion mass.

For the third daughter we have

$$g_{K-3}^K = \frac{K!}{3!\,(2K-5)!!}\left(b_1^2 + \frac{3}{2K-3}\,a_K^2 - \frac{K+1}{4}\right).$$

Thus we have positivity outside the ellipses

$$b_1^2 + \frac{3}{2K-3}\left(a_1 + \frac{K-1}{2}\right)^2 = \frac{K+1}{4},$$

for $K = 3, 4, 5, \cdots$. It is not difficult to show that none of these ellipses extends into the quadrant (7), and so no new restriction arises.

The computational effort begins to become formidable when we consider daughters beyond the third. The fourth and fifth daughter trajectories have residues depending upon the functions

$$f^K_{K-4} = \frac{K!}{4!\,(2K-7)!!} \times \left[\left(b_1^2 - \frac{K+1}{4}\right)^2 + \frac{6}{2K-5}\,a_K^2\left(b_1^2 - \frac{K+1}{12}\right) + \frac{3}{(2K-5)(2K-3)}\,a_K^4 - \frac{(K+1)(5K+4)}{120}\right] \quad (9)$$

and

$$g^K_{K-5} = \frac{K!}{5!\,(2K-9)!!} \times \left[[b_1^2 - \tfrac{5}{12}(K+1)]^2 + \frac{10}{2K-7}\,a_K^2\left(b_1^2 - \frac{K+1}{4}\right) + \frac{15}{(2K-5)(2K-3)}\,a_K^4 - \frac{(K+1)(5K+2)}{72}\right]. \quad (10)$$

It is not too difficult to show that Eq. (10) does not give rise to any boundary curve extending into (8). For Eq. (9), however, the cases $K = 4, \cdots, 9$ do produce boundaries entering (8). Tedious numerical checking shows that none of these boundaries gets outside the ellipse (7) in this quadrant, so that in fact no new restrictions on the parameters are required for positivity of the fourth and fifth daughter residues.

IV. DISCUSSION

The investigation outlined in the last section makes it seem a reasonable conjecture that the simple Veneziano function has positive residues for parameters in the region described by (7) and (8). This region is shown in Fig. 1. If we also require that $\alpha'\mu^2$ be nonnegative, the region is further restricted as shown. We have, in fact, shown that this region suffices for the

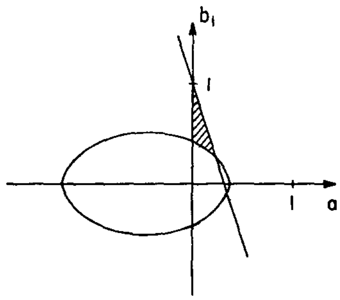


FIG. 1. Region of positive residues. The first six trajectories have positive residues for a_1, b_1 in the first quadrant outside the ellipse shown. Only in the shaded region is $\alpha'\mu^2$ positive as well. This region of positivity is shown in Fig. 2 plotted in the $(\alpha_0, \alpha'\mu^2)$ plane.

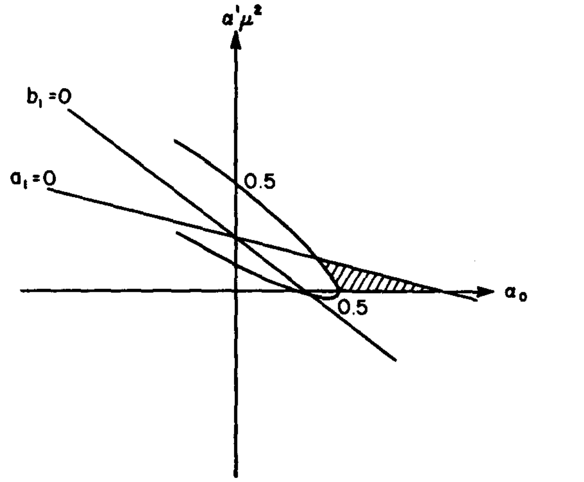


FIG. 2. Region of positive residues for the first six trajectories in the $(\alpha_0, \alpha'\mu^2)$ plane.

positivity of the residues of the first six trajectories (see Fig. 2). It is clear that the brute force approach of the last section eventually becomes unworkable and is not suited to the study of trajectories beyond the fifth daughter. Our formulation of the problem, however, reduces it to a study of the positions of zeros of certain polynomials, the f^K_J and g^K_J. There exists a considerable classical literature on this subject, which provides some hope of an eventual solution.

APPENDIX

The discussion of the text is based upon the special form (1). However, in applications the more general form

$$V_{N,P}(\alpha(S), \alpha(t)) \equiv \frac{\Gamma(N - \alpha(t))\Gamma(N - \alpha(S))}{\Gamma(N + P - \alpha(S) - \alpha(t))} \quad \text{with } P \leq N \quad (A1)$$

is often used. We shall indicate here how the partial wave expansion above can be extended to (A1).

From the properties of the gamma function, we have

$$V_{N,P}(X, Y) = \frac{\Gamma(N - Y)}{\Gamma(P - Y)} \cdot \frac{\Gamma(N - X)\Gamma(P - Y)}{\Gamma(N + P - X - Y)}$$

which, after some manipulation, gives the general partial fraction expansion

$$V_{N,P}(X, Y) = (-1)^{N-P} \sum_{K=1}^{\infty} \frac{\Gamma(Y - P + K)}{\Gamma(K)\Gamma(Y - N + 1)} \times \left(\frac{1}{N + K - 1 - X}\right).$$

According to the definition (2), the residues are proportional to

$T_{K'}(Y - N + 1)$, where $K' = K + N - P - 1$.

Since we identify $\alpha(t)$ with y, the formulas of the text are modified by changing K to K' and α_0 to $\alpha_0 - N + 1$.

Since the forms

$$\Gamma(N - X)\Gamma(n - Y)/\Gamma(L - X - Y), \quad N \neq n,$$

have been shown[7,8] to be linear combinations of terms of the type (A1), we need not consider them further. The formulas derived here yield the partial wave expansions of the most general Veneziano amplitudes.

* Work supported in part by the U.S. Atomic Energy Commission.

[1] G. Veneziano, Nuovo Cimento 57A, 190 (1968).
[2] J. Yellin, UCRL Preprints No. UCRL 18637 and UCRL 18665 (unpublished). See also Refs. in these papers.
[3] F. Wagner, Nuovo Cimento 63A, 393 (1969).
[4] R. Oehme, Nuovo Cimento Letters 1, 420 (1969).
[5] J. Shapiro, Phys. Rev. 179, 1345 (1969).
[6] Y. Nambu and P. Frampton, University of Chicago, Preprint EFI 69-56, 1969.
[7] M. L. Whippman, Phys. Rev. D 1, 701 (1970).
[8] R. E. Kreps and N. S. Milgram, University of Toronto, preprint, 1969.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Four-Body Problem: A Complete System of Angular Functions*

HAROLD WILLIAM GALBRAITH†
Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania 19104

(Received 28 May 1970)

A complete set of angular functions for the four-body problem is given. Such functions form the basis for irreducible representations of the orthogonal group O_9, reduced according to the chain $O_9 \supset O_3^1 \times O_3^2 \times O_3^3 \supset O_3 \supset O_2$. The transformation properties of the functions are given and hence a matrix representation of the permutation group on four objects is explicitly specified. The reduction of this representation yields functions suited to Bose–Einstein or Fermi–Dirac statistics.

1. INTRODUCTION

In 1965 the method of k-harmonics was used by Dragt[1] to obtain a complete set of angular functions for the three-body problem. The motivation for the new technique was a desire to treat all of the particles on the same footing. The invariance group of a free 3-particle system in the center of mass is the orthogonal group in six dimensions, O_6. Owing to the difficulty of the reduction of O_6 with respect to the actual rotation group of the three-particle system, Dragt was led to consider inner automorphisms of O_6 generated by operators of the permutation group. He found that there exists a subalgebra of the O_6 Lie algebra with the remarkable property that each of the operators of this subalgebra commutes with the operators of the alternating subgroup $A_3 \subset S_3$. This (A_3-commutative) subalgebra is just the Lie algebra of the group of three-dimensional unitary matrices SU_3. Hence, SU_3 became the "A_3-democratic" subgroup of O_6, and Dragt obtained his complete set of functions through reduction of SU_3 according to $SU_3 \supset O_3 \supset O_2$, where O_3 and O_2 are the three- and two-dimensional rotation groups of the system.

In the general n-body case, the significance of the word "k-harmonic" is understood through a group theoretical analysis of the system as carried out by Lévy-Leblond and Lurçat.[2] The n-particle phase space has a spherical structure, and, in the three-particle case, the group O_6 comprises the set of rotations connecting all points of the six-dimensional, three-body phase space sphere. O_6 is therefore "transitive" on the three-particle phase space. Further, Lévy-Leblond and Lurçat have shown that, in the n-particle case, any group transitive on phase space [$3n - 4$ sphere in $3(n - 1)$ dimensions] may be taken as the starting group of the chain, e.g., for $n = 3$, SU_3 is transitive on the 5-sphere of three-body phase space.

If one assumes a polynomial form, homogeneous and of degree k, for the basis functions, SU_3 tensor traces are equated to zero by requiring that the polynomials satisfy the six-dimensional Laplace equation, and the quantum number k describes the simultaneous localization of the three-body system.[3] This is useful since k gives information about a "global" or true "three-body" property of the state [as opposed to earlier schemes which rely on a description of the system through a (2 + 1) particle state].

In the present work we follow the "global" method of Lévy-Leblond and Lurçat to obtain a complete set of basis functions or k-harmonics for the case of a free, equal mass, four-particle system. The resulting basis functions form a complete set on the nine-dimensional spherical phase space carrying symmetric representations of the orthogonal group O_9. Since there are no transitive compact, connected Lie groups other than O_{3n-3} in the case of n-even,[2] $n > 2$, we must use O_9 here.

are modified by changing K to K' and α_0 to $\alpha_0 - N + 1$.

Since the forms

$$\Gamma(N - X)\Gamma(n - Y)/\Gamma(L - X - Y), \quad N \neq n,$$

have been shown[7,8] to be linear combinations of terms of the type (A1), we need not consider them further. The formulas derived here yield the partial wave expansions of the most general Veneziano amplitudes.

* Work supported in part by the U.S. Atomic Energy Commission.

[1] G. Veneziano, Nuovo Cimento **57A**, 190 (1968).
[2] J. Yellin, UCRL Preprints No. UCRL 18637 and UCRL 18665 (unpublished). See also Refs. in these papers.
[3] F. Wagner, Nuovo Cimento **63A**, 393 (1969).
[4] R. Oehme, Nuovo Cimento Letters **1**, 420 (1969).
[5] J. Shapiro, Phys. Rev. **179**, 1345 (1969).
[6] Y. Nambu and P. Frampton, University of Chicago, Preprint EF1 69-56, 1969.
[7] M. L. Whippman, Phys. Rev. D **1**, 701 (1970).
[8] R. E. Kreps and N. S. Milgram, University of Toronto, preprint, 1969.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Four-Body Problem: A Complete System of Angular Functions*

HAROLD WILLIAM GALBRAITH†
Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania 19104

(Received 28 May 1970)

A complete set of angular functions for the four-body problem is given. Such functions form the basis for irreducible representations of the orthogonal group O_9, reduced according to the chain $O_9 \supset O_3^1 \times O_3^2 \times O_3^3 \supset O_3 \supset O_2$. The transformation properties of the functions are given and hence a matrix representation of the permutation group on four objects is explicitly specified. The reduction of this representation yields functions suited to Bose–Einstein or Fermi–Dirac statistics.

1. INTRODUCTION

In 1965 the method of k-harmonics was used by Dragt[1] to obtain a complete set of angular functions for the three-body problem. The motivation for the new technique was a desire to treat all of the particles on the same footing. The invariance group of a free 3-particle system in the center of mass is the orthogonal group in six dimensions, O_6. Owing to the difficulty of the reduction of O_6 with respect to the actual rotation group of the three-particle system, Dragt was led to consider inner automorphisms of O_6 generated by operators of the permutation group. He found that there exists a subalgebra of the O_6 Lie algebra with the remarkable property that each of the operators of this subalgebra commutes with the operators of the alternating subgroup $A_3 \subset S_3$. This (A_3-commutative) subalgebra is just the Lie algebra of the group of three-dimensional unitary matrices SU_3. Hence, SU_3 became the "A_3-democratic" subgroup of O_6, and Dragt obtained his complete set of functions through reduction of SU_3 according to $SU_3 \supset O_3 \supset O_2$, where O_3 and O_2 are the three- and two-dimensional rotation groups of the system.

In the general n-body case, the significance of the word "k-harmonic" is understood through a group theoretical analysis of the system as carried out by Lévy-Leblond and Lurçat.[2] The n-particle phase space has a spherical structure, and, in the three-particle case, the group O_6 comprises the set of rotations connecting all points of the six-dimensional, three-body phase space sphere. O_6 is therefore "transitive" on the three-particle phase space. Further, Lévy-Leblond and Lurçat have shown that, in the n-particle case, any group transitive on phase space [$3n - 4$ sphere in $3(n - 1)$ dimensions] may be taken as the starting group of the chain, e.g., for $n = 3$, SU_3 is transitive on the 5-sphere of three-body phase space.

If one assumes a polynomial form, homogeneous and of degree k, for the basis functions, SU_3 tensor traces are equated to zero by requiring that the polynomials satisfy the six-dimensional Laplace equation, and the quantum number k describes the simultaneous localization of the three-body system.[3] This is useful since k gives information about a "global" or true "three-body" property of the state [as opposed to earlier schemes which rely on a description of the system through a (2 + 1) particle state].

In the present work we follow the "global" method of Lévy-Leblond and Lurçat to obtain a complete set of basis functions or k-harmonics for the case of a free, equal mass, four-particle system. The resulting basis functions form a complete set on the nine-dimensional spherical phase space carrying symmetric representations of the orthogonal group O_9. Since there are no transitive compact, connected Lie groups other than O_{3n-3} in the case of n-even,[2] $n > 2$, we must use O_9 here.

We note further that, even in the case of three particles, SU_3 is not "S_3-democratic," but "A_3-democratic." This idea leads to a study, in the four-body case, of the subgroups of S_4 and application of the "democracy" concept to subgroups of O_9. Hence, one arrives at the chain $O_9 \supset O_3^1 \times O_3^2 \times O_3^3 \supset O_3 \supset O_2$.[4]

In Sec. 2, center of mass variables are defined and the infinitesimal generators for the groups O_9, $O_3^1 \times O_3^2 \times O_3^3$, O_3, and O_2 are written out in terms of them. In this section we also see that the $O_3^1 \times O_3^2 \times O_3^3$ subgroup is "V_4-democratic."[5]

In Sec. 3 we construct the required basis functions, distinguished by the eight quantum numbers $\{k, s, l_1, l_2, l_3, m_1, m_2, m_3\}$. (Of course, the coupling of the l_j to get a total angular momentum label L, M is equivalent to the reduction of $O_3^1 \times O_3^2 \times O_3^3$ with respect to $O_3 \supset O_2$.) The functions are constructed as an $O_3^1 \times O_3^2 \times O_3^3$ scalar polynomial piece, homogeneous of degree k, multiplied by a product of three $Y_{l_j,m_j}(\theta_j, \phi_j)$ functions. The requirement that the whole function satisfy nine-dimensional Laplace equation is equivalent to restricting ourselves to the symmetric representations of O_9. The O_9 label is k, i.e., its value is simply related to the eigenvalue of the quadratic Casimir operator of O_9. Our eighth label, s, was discovered through the recursion formula which results from the requirement that the functions satisfy the nine-dimensional Laplace equation. Solutions denoted in this way possess simple transformation properties under the group V_4 and the (23) interchange operation. This is helpful since any of the operators of S_4 can be written as a product of operators from V_4, the (23) operator, and the (123) cyclic permutation.

In Sec. 4 we study the problem of the construction of operators which commute with the generators of the $O_3^1 \times O_3^2 \times O_3^3$ subgroup of O_9, and demonstrate that there are no "S_4-democratic," independent, eighth operators which can be constructed from the elements of the Lie algebra of O_9.

In Sec. 5, a matrix representation of S_4 is given in terms of the solutions of Sec. 3. This matrix representation may subsequently be reduced, and states obeying Bose–Einstein and Fermi–Dirac statistics constructed.

In Appendix A sample functions are given. In Appendix B we calculate the number of S_4-symmetric, angular momentum zero states.

The functions given here are useful for the determination of the binding energy and wavefunctions of a bound system of 4-particles and for concrete calculations in the case of ^{4}He.

2. CENTER OF MASS COORDINATES AND INFINITESIMAL OPERATORS

Consider a four-particle, equal mass[6] system with $\mathbf{r}_i$ the position vector of particle i in the laboratory system. The problem of finding a suitable transformation into the center of mass has been discussed by Lévy-Leblond.[4] He has shown that a useful set of c.m. variables are

$$\begin{aligned} \boldsymbol{\xi}_1 &= \tfrac{1}{2}(\mathbf{r}_1 + \mathbf{r}_4 - \mathbf{r}_2 - \mathbf{r}_3), \\ \boldsymbol{\xi}_2 &= \tfrac{1}{2}(\mathbf{r}_2 + \mathbf{r}_4 - \mathbf{r}_1 - \mathbf{r}_3), \\ \boldsymbol{\xi}_3 &= \tfrac{1}{2}(\mathbf{r}_3 + \mathbf{r}_4 - \mathbf{r}_1 - \mathbf{r}_2), \\ \boldsymbol{\xi}_4 &= \tfrac{1}{2} \textstyle\sum_1^4 \mathbf{r}_j. \end{aligned} \tag{1}$$

The free Hamiltonian operator becomes

$$H = -\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial \boldsymbol{\xi}_1^2} + \frac{\partial^2}{\partial \boldsymbol{\xi}_2^2} + \frac{\partial^2}{\partial \boldsymbol{\xi}_3^2}\right). \tag{2}$$

We now form a 9-vector of position

$$\boldsymbol{\rho} = \begin{bmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \\ \boldsymbol{\xi}_3 \end{bmatrix}. \tag{3}$$

Considering the group of 9×9 orthogonal matrices acting on $\boldsymbol{\rho}$, through matrix multiplication, we see that H is a 9-scalar and therefore has O_9 as its full symmetry group.

The Weyl infinitesimal generators for the O_9 group may be taken as

$$\Lambda_{i\alpha;j\beta} = i\left(\xi_{i\alpha}\frac{\partial}{\partial \xi_{j\beta}} - \xi_{j\beta}\frac{\partial}{\partial \xi_{i\alpha}}\right). \tag{4}$$

Here $\xi_{i\alpha}$ is the ith component of the αth relative position vector. The $\Lambda_{i\alpha;j\beta}$ satisfy the commutation relation

$$\begin{aligned} [\Lambda_{i\alpha;j\beta}, \Lambda_{i'\alpha';j'\beta'}] = i(&\delta_{\alpha\alpha'}\delta_{ii'}\Lambda_{j\beta;j'\beta'} + \delta_{\beta\beta'}\delta_{jj'}\Lambda_{i\alpha;i'\alpha'} \\ &- \delta_{\beta\alpha'}\delta_{ji'}\Lambda_{i\alpha;j'\beta'} - \delta_{\alpha\beta'}\delta_{ij'}\Lambda_{i\beta;i'\alpha'}). \end{aligned} \tag{5}$$

Elements of the direct product subgroup $O_3^1 \times O_3^2 \times O_3^3$ are simultaneous rotations in each of the three-dimensional spaces, one for each vector $\boldsymbol{\xi}_j$. Hence, we have

$$\mathcal{R}(O_3^1 \times O_3^2 \times O_3^3)\boldsymbol{\rho} \to \boldsymbol{\rho}' = \begin{bmatrix} R_1 & 0 & 0 \\ 0 & R_2 & 0 \\ 0 & 0 & R_3 \end{bmatrix}\begin{bmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \\ \boldsymbol{\xi}_3 \end{bmatrix}. \tag{6}$$

The matrices $\{R_j : j = 1, 2, 3\}$ are 3×3 orthogonal. Obviously, if $R_1 = R_2 = R_3$, we have a rotation of the total four-body system in the center of mass. The infinitesimal generators of O_3^α are

$$l_{\alpha i} = \tfrac{1}{2}\epsilon_{i\beta\gamma}\Lambda_{\beta\alpha;\gamma\alpha}, \quad i = 1, 2, 3. \tag{7}$$

The rotation group of the 4-particle system in the center of mass, O_3, is generated by the three operators

$$L_i = \sum_{\alpha=1}^{3} l_{\alpha i}, \quad i = 1, 2, 3. \tag{8}$$

The utility of the variables ξ_j can be seen by considering the transformations of the ξ_j induced by the action of the operators of the permutation group S_4.

S_4 has 4! elements, each of which may be expressed as a product of elements from the set

$$G = \{i, (23), (123), (24)(13), (14)(23), (34)(12)\}, \tag{9}$$

where $(ijk \cdots t)$ means the cyclic permutation $\binom{ijkl \cdots t}{jkl \cdots ti}$. The operators $\{i, (24)(13), (14)(23), (34)(12)\}$ are a subgroup of S_4 which we label V_4. We find, for example, that

$$[(24)(13) + (14)(23) + (34)(12)]\boldsymbol{\rho} = -\boldsymbol{\rho}. \tag{10}$$

Hence, defining

$$\Sigma \equiv [(24)(13) + (14)(23) + (34)(12)] \tag{11}$$

gives Σ as the nine-dimensional parity operator.

The product functions $Y_{l_1}^{m_1}(\hat{\xi}_1) Y_{l_2}^{m_2}(\hat{\xi}_2) Y_{l_3}^{m_3}(\hat{\xi}_3)$ are bases for the irreducible representations of the $O_3^1 \times O_3^2 \times O_3^3$ group.

Since the Weyl infinitesimal generators of $O_3^1 \times O_3^2 \times O_3^3$ all commute with the operators of V_4, basis functions for the irreducible representations of $O_3^1 \times O_3^2 \times O_3^3$ can change at most by a phase when acted upon by the operators of V_4. One easily has

$$\begin{aligned}
&(24)(13) Y_{l_1}^{m_1}(\hat{\xi}_1) Y_{l_2}^{m_2}(\hat{\xi}_2) Y_{l_3}^{m_3}(\hat{\xi}_3) \\
&\qquad = (-)^{l_1+l_3} Y_{l_1}^{m_1}(\hat{\xi}_1) Y_{l_2}^{m_2}(\hat{\xi}_2) Y_{l_3}^{m_3}(\hat{\xi}_2), \\
&(14)(23) Y_{l_1}^{m_1}(\hat{\xi}_1) Y_{l_2}^{m_2}(\hat{\xi}_2) Y_{l_3}^{m_3}(\hat{\xi}_3) \\
&\qquad = (-)^{l_2+l_3} Y_{l_1}^{m_1}(\hat{\xi}_1) Y_{l_2}^{m_2}(\hat{\xi}_2) Y_{l_3}^{m_3}(\hat{\xi}_3), \\
&(34)(12) Y_{l_1}^{m_1}(\hat{\xi}_1) Y_{l_2}^{m_2}(\hat{\xi}_2) Y_{l_3}^{m_3}(\hat{\xi}_3) \\
&\qquad = (-)^{l_1+l_2} Y_{l_1}^{m_1}(\hat{\xi}_1) Y_{l_2}^{m_2}(\hat{\xi}_2) Y_{l_3}^{m_3}(\hat{\xi}_3),
\end{aligned} \tag{12}$$

and

$$\begin{aligned}
\sum Y_{l_1}^{m_1}(\hat{\xi}_1) Y_{l_2}^{m_2}(\hat{\xi}_2) Y_{l_3}^{m_3}(\hat{\xi}_3) &= [(-)^{l_1+l_3} + (-)^{l_2+l_3} + (-)^{l_1+l_2}] \\
&\quad \times Y_{l_1}^{m_1}(\hat{\xi}_1) Y_{l_2}^{m_2}(\hat{\xi}_2) Y_{l_3}^{m_3}(\hat{\xi}_3).
\end{aligned} \tag{13}$$

3. A COMPLETE SET OF ANGULAR FUNCTIONS FOR THE FOUR-BODY PROBLEM

This section will be divided into three parts. First, starting from the free Schrödinger equation of four noninteracting particles, we show the necessity of introducing harmonic polynomials in 9-space. We write the Laplacian in a spherical coordinate system, and obtain a recursion formula relating coefficients of the homogeneous polynomials.

Next, we solve the $l_1 = l_2 = l_3 = 0$ case, introducing the needed label. We also derive the transformation properties of the solutions under the (23) interchange.

Finally, following the technique of the $l_1 = l_2 = l_3 = 0$ case, we write the complete set of angular functions for arbitrary l_j, and obtain the transformation properties under (23).

A. The Schrödinger Equation

The free wave equation in the center of mass is

$$-\frac{\hbar^2}{2\mu}(\nabla_{\xi_1}^2 + \nabla_{\xi_2}^2 + \nabla_{\xi_3}^2)\psi(\xi_1, \xi_2, \xi_3) = E\psi(\xi_1, \xi_2, \xi_3), \tag{14}$$

where $\nabla_{\xi_j}^2$ is the usual 3-Laplacian of vector ξ_j. The solutions, $\psi(\xi_1, \xi_2, \xi_3)$, may be expanded in terms of the angular functions of the 9-sphere

$$\psi(\xi_1, \xi_2, \xi_3) = \sum_k R_k(\rho) U_k\left(\frac{\xi_1}{\rho}, \frac{\xi_2}{\rho}, \frac{\xi_3}{\rho}\right), \tag{15}$$

where ρ is the length of the nine-position vector

$$\rho^2 = \textstyle\sum_1^3 \xi_j^2. \tag{16}$$

The U_k are the required angular functions. They are related to the solutions of the 9-Laplace equation by

$$U_k\left(\frac{\xi_1}{\rho}, \frac{\xi_2}{\rho}, \frac{\xi_3}{\rho}\right) = \frac{1}{\rho^k} P_k(\xi_1, \xi_2, \xi_3), \tag{17}$$

where

$$P_k(\xi_1, \xi_2, \xi_3) = \sum_{n_1, n_2, n_3} a_{n_1 n_2 n_3}(l_1, l_2, l_3) \xi_1^{n_1} \xi_2^{n_2} \xi_3^{n_3} \times Y_{l_1}^{m_1}(\hat{\xi}_1) Y_{l_2}^{m_2}(\hat{\xi}_2) Y_{l_3}^{m_3}(\hat{\xi}_3). \tag{18}$$

In (18) the sum is over each n_j such that

$$\textstyle\sum_1^3 n_j = k; \tag{19}$$

k is the degree of the $O_3^1 \times O_3^2 \times O_3^3$ scalar polynomial piece.

The $a_{n_1 n_2 n_3}(l_1, l_2, l_3)$ of (18) are related by the requirement that $P_k(\xi_1, \xi_2, \xi_3)$ be harmonic. Using this fact and Eq. (14) gives

$$\frac{1}{\rho^8}\frac{d}{d\rho}\rho^8\frac{d}{d\rho}R_k(\rho) + \left(E - \frac{k(k+7)}{\rho^2}\right)R_k(\rho) = 0. \tag{20}$$

Hence we must solve

$$\Delta_9 P_k(\xi_1, \xi_2, \xi_3) = 0. \tag{21}$$

The Laplacian now has the form

$$\Delta_9 = \textstyle\sum_1^3 \left(\frac{1}{\xi_j^2}\frac{\partial}{\partial \xi_j}\xi_j^2\frac{\partial}{\partial \xi_j} - \mathbf{l}_j^2\frac{(\theta_j, \phi_j)}{\xi_j^2}\right). \tag{22}$$

We label the k-polynomial part by

$$\mathfrak{F}^{k,l_1,l_2,l_3}(\xi_1,\xi_2,\xi_3)=\sum_{n_1n_2n_3} a_{n_1n_2n_3}(l_1,l_2,l_3)\xi_1^{n_1}\xi_2^{n_2}\xi_3^{n_3}. \tag{23}$$

Now Eq. (21) becomes a differential equation for $\mathfrak{F}^{k,l_1,l_2,l_3}(\xi_1,\xi_2,\xi_3)$:

$$\sum_{j=1}^{3}\left(\frac{1}{\xi_j^2}\frac{\partial}{\partial\xi_j}\xi_j^2\frac{\partial}{\partial\xi_j}-\frac{l_j(l_j+1)}{\xi_j^2}\right)\mathfrak{F}^{k,l_1,l_2,l_3}(\xi_1,\xi_2,\xi_3)=0. \tag{24}$$

From these results we have

$$\begin{aligned}\sum_{n_1n_2n_3} a_{n_1n_2n_3}(l_1,l_2,l_3)&\\ \times\{[n_1(n_1+1)-l_1(l_1+1)]\xi_1^{n_1-2}\xi_2^{n_2}\xi_3^{n_3}&\\ +[n_2(n_2+1)-l_2(l_2+1)]\xi_1^{n_1}\xi_2^{n_2-2}\xi_3^{n_3}&\\ +[n_3(n_3+1)-l_3(l_3+1)]\xi_1^{n_1}\xi_2^{n_2}\xi_3^{n_3-2}\}&=0. \end{aligned}\tag{25}$$

Equating to zero the coefficients of terms of like form, we get the required relation between the $a_{n_1n_2n_3}(l_1, l_2, l_3)$,

$$\begin{aligned}[n_1(n_1+1)-l_1(l_1+1)]a_{n_1n_2n_3}&\\ =-[(n_2+2)(n_2+3)-l_2(l_2+1)]a_{n_1-2,n_2+2,n_3}&\\ -[(n_3+2)(n_3+3)-l_3(l_3+1)]a_{n_1-2,n_2,n_3+2}.&\end{aligned}\tag{26}$$

B. Solution of the $l_1 = l_2 = l_3 = 0$ Case

Specializing to the case of $l_1 = l_2 = l_3 = 0$, one has the recursion relation

$$\begin{aligned}n_1(n_1+1)a_{n_1n_2n_3}=-(n_2+2)(n_2+3)a_{n_1-2,n_2+2,n_3}&\\ -(n_3+2)(n_3+3)a_{n_1-2,n_2,n_3+2}.&\end{aligned}\tag{27}$$

It follows from (27) that

$$a_{n_1n_2n_3}=0,\quad n_1,n_2,n_3=\text{odd}, \tag{28}$$

and that the a_{0,n_2,n_3} are all independent. Here $\mathfrak{F}^k(\xi_1,\xi_2,\xi_3)$ is the complete function of (18), and the only condition which we have at our disposal for determining the coefficients is Eq. (27).

The $\mathfrak{F}^k(\xi_1,\xi_2,\xi_3)$ are eigenfunctions of seven operators and the degree of degeneracy is the number of solutions to the equation

$$n_2+n_3=k, \tag{29}$$

with n_2, n_3, k being even. This is just $k/2+1$.

We may express the $a_{n_1n_2n_3}$ in terms of the independent $a_{0n_2n_3}$ by iterating (27):

$$\begin{aligned}a^{\mathrm{o}}_{n_1n_2n_3}&=-\sum_{j=0}^{n_1/4-\frac12}\frac{(n_1/2)!\,(n_2+n_1+1-2j)!\,(n_3+1+2j)!}{j!\,(n_1/2-j)!\,(n_1+1)!\,(n_2+1)!\,(n_3+1)!}\\ &\quad\times a_{0,n_2+n_1-2j,n_3+2j}\\ &\quad-\sum_{j=0}^{n_1/4-\frac12}\frac{(n_1/2)!\,(n_3+n_1+1-2j)!\,(n_2+1+2j)!}{j!\,(n_1/2-j)!\,(n_1+1)!\,(n_2+1)!\,(n_3+1)!}\\ &\quad\times a_{0,n_2+2j,n_3+n_1-2j},\end{aligned}\tag{30}$$

and

$$\begin{aligned}a^{\mathrm{e}}_{n_1n_2n_3}&=\sum_{j=0}^{n_1/4}\frac{(n_1/2)!\,(n_2+n_1+1-2j)!\,(n_3+1+2j)!}{j!\,(n_1/2-j)!\,(n_1+1)!\,(n_2+1)!\,(n_3+1)!}\\ &\quad\times a_{0,n_2+n_1-2j,n_3+2j}\\ &\quad+\sum_{j=0}^{n_1/4-1}\frac{(n_1/2)!\,(n_3+1+n_1-2j)!\,(n_2+1+2j)!}{j!\,(n_1/2-j)!\,(n_1+1)!\,(n_2+1)!\,(n_3+1)!}\\ &\quad\times a_{0,n_2+2j,n_3+n_1-2j}.\end{aligned}\tag{31}$$

In (30) and (31) the superscript (e, o) means that $n_1/2$ is (even, odd).

We could at this point construct a new operator from the elements of the O_9 Lie algebra (see Sec. 4). Then the requirement that the operator so constructed be diagonal on $\mathfrak{F}^k(\xi_1,\xi_2,\xi_3)$ would break the $k/2+1$ fold degeneracy. Instead, however, we proceed as follows: The independent coefficients may be written as $\{a_{0,k-s,s}\colon s=0,2,\cdots,k\}$. By putting $a_{0,k-s,s}$ to unity for some particular s and setting all other a_{0,n_2,n_3} to zero, we obtain a polynomial which is specified by k and s. Since s takes on a different value for each degenerate state, it may be taken as our missing label.

We now work out the transformation properties of $\mathfrak{F}^{k,s}(\xi_1,\xi_2,\xi_3)$ under the (23) operator. Since operators of V_4 change $\boldsymbol{\xi}_j$ at the most by a sign, all $\mathfrak{F}^{k,s}(\xi_1,\xi_2,\xi_3)$ are invariant under V_4. Hence, we say loosely that s is "V_4-democratic." Therefore, we need only determine the transformation properties of $\mathfrak{F}^{k,s}$ under the operators (23) and (123). (123) is worked out later. However, since the action of (23) is determined by the 2–3 interchange symmetry exhibited in Eqs. (30) and (31), we work it out here.

According to (27), (30), and (31), we have

$$a_{n_1n_2n_3}=\sum_{s'=0,2,\cdots}^{k}C_{s'}(n_1,n_2,n_3)a_{0,k-s',s'}, \tag{32a}$$

$$a_{n_1n_3n_2}=\sum_{s'=0,2,\cdots}^{k}C_{s'}(n_1,n_2,n_3)a_{0,s',k-s'}. \tag{32b}$$

Using the definition of the $\mathfrak{F}^k$ and $\mathfrak{F}^{k,s}$, one then has

$$\mathfrak{F}^k(\xi_1,\xi_2,\xi_3)=\sum_{n_1n_2n_3}\left(\sum_{s'}C_{s'}(n_1,n_2,n_3)a_{0,k-s',s'}\right)\xi_1^{n_1}\xi_2^{n_2}\xi_3^{n_3} \tag{33}$$

and

$$\mathfrak{F}^{k,s}(\xi_1,\xi_2,\xi_3)=\sum_{n_1n_2n_3}C_s(n_1,n_2,n_3)\xi_1^{n_1}\xi_2^{n_2}\xi_3^{n_3}. \tag{34}$$

But $\mathfrak{F}^k(\xi_1,\xi_2,\xi_3)$ may also be written as

$$\sum a_{n_1n_3n_2}\xi_1^{n_1}\xi_2^{n_3}\xi_3^{n_2},$$

giving

$$\mathcal{F}^k(\xi_2, \xi_3)\xi_1, = \sum_{n_1 n_2 n_3}\left(\sum_{s'=0,2,\cdots}^{k} C_{s'}(n_1, n_2, n_3) a_{0,s',k-s'}\right)\xi_1^{n_1}\xi_2^{n_3}\xi_3^{n_2}. \quad (35)$$

Hence,

$$\mathcal{F}^{k,k-s} = \sum_{n_1 n_2 n_3} C_s(n_1, n_2, n_3)\xi_1^{n_1}\xi_2^{n_3}\xi_3^{n_2}. \quad (36)$$

Comparing (36) with (34), we have the sought-for result

$$(23)\mathcal{F}^{k,s}(\xi_1, \xi_2, \xi_3) = \mathcal{F}^{k,k-s}(\xi_1, \xi_2, \xi_3). \quad (37)$$

We may obtain the $\mathcal{F}^{k,s}(\xi_1, \xi_2, \xi_3)$ explicitly by using Eqs. (30) and (31) along with the requirement that $a_{0,k-s,s}$ be unity:

$$\begin{aligned}\mathcal{F}^{k,s}(\xi_1, \xi_2, \xi_3) &= -\sum_{n_1}^{\mathrm{o}}\left(\sum_{j_1=0}^{s/2}\frac{(n_1/2)!\,(k-s+1)!\,(s+1)!\,\theta(n_1-2-4j_1)}{j_1!\,(n_1/2-j_1)!\,(n_1+1)!\,(k-n_1-s+2j_1+1)!\,(s-2j_1+1)!}\xi_1^{n_1}\xi_2^{k-n_1-s+2j_1}\xi_3^{s-2j_1}\right.\\ &\left.+\sum_{j_2=0}^{(k-s)/2}\frac{(n_1/2)!\,(s+1)!\,(k-s+1)!\,\theta[n_1-2(2j_2+1)]}{j_2!\,(n_1/2-j_2)!\,(n_1+1)!\,(k-s-2j_2+1)!\,(s+2j_2-n_1+1)!}\xi_1^{n_1}\xi_2^{k-s-2j_2}\xi_3^{s+2j_2-n_1}\right)\\ &+\sum_{n_1}^{\mathrm{e}}\left(\sum_{j_1=0}^{s/2}\frac{(n_1/2)!\,(k-s+1)!\,(s+1)!\,\theta(n_1-4j_1)}{j_1!\,(n_1/2-j_1)!\,(n_1+1)!\,(k-n_1-s+2j_1+1)!\,(s-2j_1+1)!}\xi_1^{n_1}\xi_2^{k-n_1-s+2j_1}\xi_3^{s-2j_1}\right.\\ &\left.+\sum_{j_2=0}^{(k-s)/2}\frac{(n_1/2)!\,(s+1)!\,(k-s+1)!\,\theta[n_1-4(j_2+1)]}{j_2!\,(n_1/2-j_2)!\,(n_1+1)!\,(k-s-2j_2+1)!\,(s+2j_2-n_1+1)!}\xi_1^{n_1}\xi_2^{k-s-2j_2}\xi_3^{s+2j_2-n_1}\right). \quad (38)\end{aligned}$$

In (38) (o, e) requires a sum over $n_1/2$ (odd, even). Equation (38) gives the complete solution for the case of $l_1 = l_2 = l_3 = L = M = 0$.

C. Complete Set of Angular Functions

We now obtain a complete set of four-body states for arbitrary l_j by the method of the previous section. Consider again Eq. (25):

$$\begin{aligned}\sum_{n_1 n_2 n_3} a_{n_1 n_2 n_3}\{&[n_1(n_1+1) - l_1(l_1+1)]\xi_1^{n_1-2}\xi_2^{n_2}\xi_3^{n_3} + [n_2(n_2+1) - l_2(l_2+1)]\xi_1^{n_1}\xi_2^{n_2-2}\xi_3^{n_3}\\ &+ [n_3(n_3+1) - l_3(l_3+1)]\xi_1^{n_1}\xi_2^{n_2}\xi_3^{n_3-2}\} = 0. \quad (25)\end{aligned}$$

We observe the following:

(a) The only nonzero coefficients are

$$\{a_{l_1+j_1, l_2+j_2, l_3+j_3}: j_1 \geq 0, j_2 \geq 0, j_3 \geq 0, j_1 + j_2 + j_3 = k - \textstyle\sum_1^3 l_j\};$$

(b) the coefficients $a_{l_1, l_2+j_2, l_3+j_3}$ are all independent.

Equation (26) now reads

$$\begin{aligned}[(l_1+j_1)(l_1+j_1+1) - l_1(l_1+1)]a_{l_1+j_1, l_2+j_2, l_3+j_3} &= -[(l_2+j_2+2)(l_2+j_2+3) - l_2(l_2+1)]a_{l_1+j_1-2, l_2+j_2+2, l_3+j_3}\\ &\quad - [(l_3+j_3+2)(l_3+j_3+3) - l_3(l_3+1)]a_{l_1+j_1-2, l_2+j_2, l_3+j_3+2}. \quad (39)\end{aligned}$$

Iterating (39) gives

$$\begin{aligned}a^{\mathrm{e}}_{l_1+j_1, l_2+j_2, l_3+j_3} &= \sum_{t_1=0}^{j_1/4}\frac{(j_1/2)!\prod_{p=0,2,\cdots}^{j_1-2-2t_1}[(j_2+2+p)(j_2+3+p+2l_2)]}{(j_1/2-t_1)!\,t_1!\prod_{p=0,2,\cdots}^{j_1-2}[(j_1-p)(j_1-p+2l_1+1)]}\left\{\begin{matrix}\prod_{p=2,4,\cdots}^{2t_1>0}[(j_3+p)(j_3+p+2l_2+1)]\\ 1 \quad (t_1=0)\end{matrix}\right\}\\ &\times a_{l_1, j_2+j_1+l_2-2t_1, j_3+l_3+2t_1} + \sum_{t_2=0}^{j_1/4-1}\frac{(j_1/2)!\prod_{p=0,2,\cdots}^{j_1-2-2t_2}[(j_3+2+p)(j_3+3+p+2l_3)]}{(j_1/2-t_2)!\,(t_2)!\prod_{p=0,2,\cdots}^{j_1-2}[(j_1-p)(j_1-p+1+2l_1)]}\\ &\times\left\{\begin{matrix}\prod_{p=2,4,\cdots}^{2t_2>0}[(j_2+p)(j_2+p+1+2l_3)]\\ 1 \quad (t_2=0)\end{matrix}\right\}a_{l_1, l_2+j_2+2t_2, l_3+j_3+j_1-2t_2} \quad (40)\end{aligned}$$

and

$$a^{o}_{l_1+j_1,l_2+j_2,l_3+j_3} = -\sum_{t_1=0}^{j_1/4-\frac{1}{2}} \frac{(j_1/2)! \prod_{p=0,2,\cdots}^{j_1-2-2t_1} [(j_2+2+p)(j_2+3+p+2l_2)]}{(j_1/2-t_1)!\, t_1! \prod_{p=0,2,\cdots}^{j_1-2} [(j_1-p)(j_1-p+1+2l_1)]} \times \begin{Bmatrix} \prod_{p=2,4,\cdots}^{2t_1>0} [(j_3+p)(j_3+p+1+2l_3)] \\ 1 \quad (t_1=0) \end{Bmatrix} a_{l_1,j_2+j_1+l_2-2t_1,j_3+l_3+2t_1}$$
$$- \sum_{t_2=0}^{j_1/4-\frac{1}{2}} \frac{(j_1/2)! \prod_{p=0,2,\cdots}^{j_1-2-2t_2} [(j_3+2+p)(j_3+3+p+2l_3)]}{\prod_{p=0,2,\cdots}^{j_1-2} [(j_1-p)(j_1-p+1+2l_1)](j_1/2-t_2)!\, t_2!} \times \begin{Bmatrix} \prod_{p=2,4,\cdots}^{2t_2>0} [(j_2+p)(j_2+p+1+2l_2)] \\ 1 \quad (t_2=0) \end{Bmatrix} a_{l_1,l_2+j_2+2t_2,l_3+j_3+j_1-2t_2}. \tag{41}$$

Here the superscript (o, e) refers to $j_1/2$ (odd, even). Defining

$$\kappa = k - \textstyle\sum_1^3 l_j \tag{42}$$

and fixing $a_{l_1,l_2+\kappa-s,l_3+s}$ as unity while the remaining independent coefficients are set equal to zero gives

$$P^{k,s,l_1,l_2,l_3}_{m_1m_2m_3}(\xi_1,\xi_2,\xi_3)$$
$$= \Bigg\{ -\sum_{j_1}^{o} \Bigg[\sum_{t_1=0}^{s/2} \binom{j_1/2}{t_1} \frac{\prod_{p=0,2,\cdots}^{j_1-2-2t_1} [(p+2+\kappa-j_1-s+2t_1)(p+3+\kappa-j_1-s+2t_1+2l_2)]}{\prod_{p=0,2,\cdots}^{j_1-2} [(j_1-p)(j_1-p+1+2l_1)]}$$
$$\times \begin{Vmatrix} \prod_{p=2,4,\cdots}^{2t_1>0} [(s-2t_1+p)(s-2t_1+p+1+2l_3)] \\ 1 \quad (t_1=0) \end{Vmatrix} \theta[j_1-2(2t_1+1)]\xi_1^{l_1+j_1}\xi_2^{l_2+\kappa-j_1-s+2t_1}\xi_3^{l_3+s-2t_1}$$
$$+ \sum_{t_2=0}^{(\kappa-s)/2} \binom{j_1/2}{t_2} \frac{\prod_{p=0,2,\cdots}^{j_1-2+2t_2} [(2+p+s+2t_2-j_1)(3+p+s-j_1+2t_2+2l_3)]}{\prod_{p=0,2,\cdots}^{j_1-2} [(j_1-p)(j_1-p+1+2l_1)]}$$
$$\times \begin{Vmatrix} \prod_{p=2,4,\cdots}^{2t_2>0} [(\kappa-s+2t_2+p)(\kappa-s+2t_2+p+1+2l_2)] \\ 1 \quad (t_2=0) \end{Vmatrix}$$
$$\times\, \theta[j_1-2(2t_2+1)]\xi_1^{l_1+j_1}\xi_2^{l_2+\kappa-s-2t_2}\xi_3^{l_3+s-j_1+2t_2} \Bigg]$$
$$+ \sum_{j_1(\neq 0)}^{e} \Bigg[\sum_{t_1=0}^{s/2} \binom{j_1/2}{t_1} \frac{\prod_{p=0,2,\cdots}^{j_1-2-2t_1} [(\kappa-s-j_1+2t_1+2+p)(\kappa-s-j_1+2t_1+3+p+2l_2)]}{\prod_{p=0,2,\cdots}^{j_1-2} [(j_1-p)(j_1-p+1+2l_1)]}$$
$$\times \begin{Vmatrix} \prod_{p=2,4,\cdots}^{2t_1>0} [(s-2t_1+p)(s-2t_1+p+1+2l_3)] \\ 1 \quad (t_1=0) \end{Vmatrix} \theta(j_1-4t_1)\xi_1^{l_1+j_1}\xi_2^{l_2+\kappa-j_1-s+2+2t_1}\xi_3^{l_3+s-2t_1}$$

$$
+ \sum_{t_2=0}^{(\kappa-s)/2} \binom{j_1/2}{t_2} \frac{\prod_{p=0,2,\cdots}^{j_1-2-2t_2} [(2+p+s-j_1+2t_2)(3+p-j_1+s+2t_2+2l_3)]}{\prod_{p=0,2,\cdots}^{j_1-2} [(j_1-p)(j_1-p+1+2l_1)]}
$$

$$
\times \left\| \begin{matrix} \prod_{p=2,4,\cdots}^{2t_2>0} [(\kappa-s-2t_2+p)(\kappa-s-2t_2+p+1+2l_2)] \\ 1 \quad (t_2=0) \end{matrix} \right\| \theta[j_1-4(t_2+1)]
$$

$$
\left. \times\, \xi_1^{j_1+l_1} \xi_2^{l_2+\kappa-s-2t_2} \xi_3^{l_3-j_1+s+2t_2} \right] + \left. \delta_{j_1,0} \xi_1^{l_1} \xi_2^{l_2+\kappa-s} \xi_3^{l_3+s} \right\} Y_{l_1}^{m_1}(\hat{\xi}_1) Y_{l_2}^{m_2}(\hat{\xi}_2) Y_{l_3}^{m_3}(\hat{\xi}_3). \tag{43}
$$

In Eq. (43) the superscript (o, e) on the summation indicates that j_1 is to be summed over its values such that $j_1/2$ is (odd, even). The notation $\binom{j_1/2}{t_1}$ is a binomial coefficient, and the matrix $\| \;\; \|$ symbol means, e.g., take 1 when $t_{\binom{1}{2}} = 0$ or take the upper form when $t_{\binom{1}{2}} \neq 0$.

We note the following which are easily derived from Eqs. (25) and (39):

(c) The j_k are all even. (This follows from the fact that the Laplace operator reduces the degree of a homogeneous polynomial by 2.)

(d) κ is always even.

(e) The spectrum of s is $0, 2, \cdots, \kappa$ or $\kappa/2 + 1$ values. [Obviously s here is a generalization of that of Sec. 3A, and the solutions of 3A are contained in Eq. (43).]

Due to the (23) symmetry expressed in (40) and (41), we have as before

$$
(23)\mathcal{J}^{k,s} = \mathcal{J}^{k,\kappa-s}. \tag{44}
$$

Equations (43) and (17) give the desired angular functions for the four-body problem explicitly.

4. A COMPLETE SET OF OBSERVABLES FOR THE FOUR-BODY PROBLEM

We now study the construction of operators which commute with the Casimir operators from the chain $O_9 \supset O_3^1 \times O_3^2 \times O_3^3 \supset O_3 \supset O_2$, and show that there are no missing S_4 invariant operators.

The chain $O_3 \supset O_3^1 \times O_3^2 \times O_3^3 \supset O_3 \supset O_2$ gives six Casimir operators:

$$
O_9 : \Lambda^2 = \sum_{i\alpha j\beta}{}_1^3 \Lambda_{i\alpha;j\beta}^2,
$$

$$
O_3^1 \times O_3^2 \times O_3^3 : \mathbf{l}_\alpha^2 = \sum_i{}_1^3 l_{i\alpha}^2, \quad \alpha = 1, 2, 3,
$$

$$
O_3 : \mathbf{L}^2 = \sum_\alpha{}_1^3 L_\alpha^2,
$$

$$
O_2 : L_3.
$$

The coupling of the three angular momenta requires the use of an intermediate coupling operator. This operator is defined by [7,8]

$$
T = \epsilon_{ijk} l_{i1} l_{j2} l_{k3}. \tag{45}
$$

The fact that one more operator is required to completely reduce O_9 means that the O_9 representations, when restricted to the $O_3^1 \times O_3^2 \times O_3^3$ subgroup, are not multiplicity free. Hence, an operator is needed which differentiates between equivalent $O_3^1 \times O_3^2 \times O_3^3$ representations within a given O_9 representation.

Since O_3 is a subgroup of $O_3^1 \times O_3^2 \times O_3^3$, we need only find an operator which commutes with the infinitesimal generators of $O_3^1 \times O_3^2 \times O_3^3$, and, since all infinitesimal generators of the O_9 Lie algebra commute with Λ^2, we consider only those operators which, when constructed from the $\Lambda_{i\alpha;j\beta}$, are invariant under $O_3^1 \times O_3^2 \times O_3^3$ rotations. The Weyl infinitesimal generators of O_9 form a skew-symmetric, 9×9 matrix

$$
[\Lambda_{i\alpha;j\beta}] = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} \\ \Lambda_{21} & \Lambda_{22} & \Lambda_{23} \\ \Lambda_{31} & \Lambda_{32} & \Lambda_{33} \end{bmatrix} = [\Lambda_{\sigma\tau}], \tag{46}
$$

where the submatrix $\Lambda_{\sigma\tau}$ is

$$
\Lambda_{\sigma\tau} = [\Lambda_{i\sigma;j\tau}]. \tag{47}
$$

$O_3^1 \times O_3^2 \times O_3^3$ rotations induce transformations on the elements of the O_9 algebra according to

$$
[\Lambda'_{\sigma\tau}] = \begin{bmatrix} R_1 & 0 & 0 \\ 0 & R_2 & 0 \\ 0 & 0 & R_3 \end{bmatrix} [\Lambda_{\sigma\tau}] \begin{bmatrix} \tilde{R}_1 & 0 & 0 \\ 0 & \tilde{R}_2 & 0 \\ 0 & 0 & \tilde{R}_3 \end{bmatrix}. \tag{48}
$$

Due to the orthogonality of the R_j submatrices, we are led to consider quadratic forms in the $\Lambda_{i\alpha;j\beta}$. Such forms are conveniently generated in terms of traces of products of the matrices $\Lambda_{\sigma\tau}$.

Using equation (48), one may demonstrate that $\mathrm{Tr}\,[(\Lambda_{\sigma\tau}\tilde{\Lambda}_{\sigma\tau})^n]$ are $O_3^1 \times O_3^2 \times O_3^3$ invariant operators. Using this result combined with the permutational properties of the elements $\Lambda_{i\alpha;j\beta}$, one finds to second order in the $\Lambda_{i\alpha;j\beta}$

$$\Omega^{(2)} = \mathrm{Tr}\,(\Lambda_{12}\tilde{\Lambda}_{12}) + \mathrm{Tr}\,(\tilde{\Lambda}_{12}\Lambda_{12}) + \mathrm{Tr}\,(\Lambda_{23}\tilde{\Lambda}_{23}) + \mathrm{Tr}\,(\tilde{\Lambda}_{23}\Lambda_{23}) + \mathrm{Tr}\,(\Lambda_{13}\tilde{\Lambda}_{13}) + \mathrm{Tr}\,(\tilde{\Lambda}_{13}\Lambda_{13}), \quad (49)$$

as an $O_3^1 \times O_3^2 \times O_3^3$ invariant operator which is symmetric with respect to particle permutations. However, it is easily checked that

$$\Omega^{(2)} = \Lambda^2 - 2\sum_1^3 \mathbf{l}_j^2, \quad (50)$$

and $\Omega^{(2)}$ is not independent.

One can go on to search for higher-order invariants by the same technique, but to no avail, as we now show.

The solutions for $k = 2$, $l_1 = l_2 = l_3 = 0$ are expressed by

$$\mathcal{T}^2 = a_{020}\xi_2^2 + a_{002}\xi_3^2 - (a_{002} + a_{020})\xi_1^2. \quad (51)$$

We require

$$\Omega\mathcal{T}^2 = \omega\mathcal{T}^2. \quad (52)$$

The eigenvalue equation (52) must give a nontrivial relation between the coefficients a_{020} and a_{002}, breaking the degeneracy. Since Ω is an S_4 invariant operator, we have

$$(12)\mathcal{T}^2 = \pm\mathcal{T}^2 \quad (53)$$

and

$$(123)\mathcal{T}^2 = \exp\left(\frac{2\pi ni}{3}\right)\mathcal{T}^2. \quad (54)$$

It is easily shown that (53), (54), and (51) are incompatible. Hence, we have the general result: *There are no independent, $O_3^1 \times O_3^2 \times O_3^3$ and S_4 invariant operators for the four-body problem in the chain $O_9 \supset O_3^1 \times O_3^2 \times O_3^3 \supset O_3$.*

We now ask: What is the maximum S_4 symmetry which an independent eighth operator can exhibit? To answer this, we first note the basic result that $O_3^1 \times O_3^2 \times O_3^3$ invariant operators may only be formed from $\{\boldsymbol{\xi}_j \cdot \boldsymbol{\xi}_j, \nabla_{\xi_j}^2, \boldsymbol{\xi}_j \cdot \nabla_{\xi_j}\}$, $j = 1, 2, 3$. Hence, any $O_3^1 \times O_3^2 \times O_3^3$ invariant operator is V_4-democratic. Therefore, we study an operator such that

$$(23)\Omega(23)^{-1} = \pm\Omega \quad (55)$$

and

$$(123)\Omega(123)^{-1} = \pm\Omega. \quad (56)$$

It is obvious, however, that Ω^2 is then an S_4 invariant, implying that Ω will not break the degeneracy. Therefore, one must choose *either* (55) or (56), implying that simple transformation properties under (23) necessarily imply complicated properties with respect to (123) and vice versa. Therefore, we have justification for the solutions of Sec. 3 which have simple transformation properties under (23).

5. TRANSFORMATION PROPERTIES UNDER PARTICLE PERMUTATIONS

In this section we determine the transformation properties of the solutions contained in Eq. (43), under the action of the operators G, giving explicitly the matrix representation of S_4 in the basis of Sec. 3.

We begin by again dividing the general solution into a polynomial part and spherical-harmonic part

$$P^{k,s,l_1,l_2,l_3}_{m_1m_2m_3}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3) = \mathcal{T}^{k,s}(\xi_j)\prod_{j=1}^{3} Y_{l_j}^{m_j}(\hat{\xi}_j). \quad (57)$$

First consider the polynomial piece $\mathcal{T}^{k,s}(\xi_1, \xi_2, \xi_3)$. From Sec. 3C we have

$$(23)\mathcal{T}^{k,s} = \mathcal{T}^{k,\kappa-s}. \quad (58)$$

For the (123) cyclic permutation we write formally

$$(123)\mathcal{T}^{k,s} = \sum_{j=0,2,\cdots}^{\kappa} A_{sj}\mathcal{T}^{k,j}. \quad (59)$$

One may determine the matrix elements A_{sj} in the following way: Each distinct state $\mathcal{T}^{k,s}$ was constructed by requiring that $a_{l_1,l_2+\kappa-s,l_3+s}$ be unity while the other independent coefficients were zero. We have

$$\mathcal{T}^{k,s}(\xi_1, \xi_2, \xi_3) = \sum_{\substack{j_1j_2j_3 \\ j_1\neq 0}} a_{l_1+j_1,l_2+j_2,l_3+j_3}\xi_1^{l_1+j_1}\xi_2^{l_2+j_2}\xi_3^{l_3+j_3} + \xi_1^{l_1}\xi_2^{l_2+\kappa-s}\xi_3^{l_3+s}. \quad (60)$$

Clearly, the term $\xi_1^{l_1}\xi_2^{l_2+\kappa-s}\xi_3^{l_3+s}$ is unique to the particular $\mathcal{T}^{k,s}$, so that on applying (123) to $\mathcal{T}^{k,s}$ we look for the coefficient of the term $\xi_1^{l_1}\xi_2^{l_2+\kappa-j}\xi_3^{l_3+j}$ which is A_{sj}. We get:

(a) $A_{sj} = 0$ if $(l_2 + \kappa - j - l_1)/2$ is not integer;

(b) if $(l_2 + \kappa - j - l_1)/2 = \text{odd}$,

$$A_{sj} = -\binom{[l_2 + \kappa - j - l_1]/2}{[l_3 - l_1 + s]/2} \frac{\prod_{p=0,2,\cdots}^{l_2+\kappa-j-2-l_3-s} [(p + 2 + l_3 - l_2 + j)(p + 3 + l_3 + l_2 + j)]}{\prod_{p=0,2,\cdots}^{l_2+\kappa-j-l_1-2} [(l_2 + \kappa - j - l_1 - p)(l_2 + l_1 + 1 + \kappa - j - p)]}$$

$$\times \left\| \begin{matrix} \prod_{p=2,4,\cdots}^{l_3-l_1+s} [(l_1 - l_3 + p)(l_1 + l_3 + p + 1)] \\ 1 \quad (l_1 = l_3 + s) \end{matrix} \right\| \theta[l_2 + \kappa - j + l_1 - 2(l_3 + s + 1)]\theta(l_1 - l_3);$$

(c) if $(l_2 + \kappa - j - l_1)/2 = \text{odd}$,

$$A_{sj} = -\binom{[l_2 + \kappa - j - l_1]/2}{[l_2 + \kappa - s - l_3 - j]/2} \frac{\prod_{p=0,2,\cdots}^{2(l_2+\kappa-j-1)-l_1-l_3-s} [(2 + p + l_1 - l_3)(3 + p + l_1 + l_3)]}{\prod_{p=0,2,\cdots}^{l_2+\kappa-j-l_1-2} [(l_2 + \kappa - j - l_1 - p)(l_2 + l_1 + \kappa - j - p + 1)]}$$

$$\times \left\| \begin{matrix} \prod_{p=2,4,\cdots}^{l_2+\kappa-s-l_3-j} [(l_3 - l_2 + j + p)(l_2 + l_3 + j + p + 1)] \\ 1 \quad (s + l_3 + j = l_2 + \kappa) \end{matrix} \right\|$$

$$\times\, \theta(j - l_2 - \kappa - l_1 + 2l_3 - 2 + 2s)\theta(l_3 + j - l_2);$$

(d) if $(l_2 - l_1 + \kappa - j)/2 = \text{even}$,

$$A_{sj} = \binom{[l_2 + \kappa - j - l_1]/2}{[l_3 - l_1 + s]/2} \frac{\prod_{p=0,2,\cdots}^{l_2+\kappa-j-2-l_3-s} [(p + 2 + l_3 - l_2 + j)(p + 3 + l_3 + l_2 + j)]}{\prod_{p=0,2,\cdots}^{l_2+\kappa-j-l_1-2} [(l_2 + \kappa - j - l_1 - p)(l_2 + l_1 + 1 + \kappa - j - p)]}$$

$$\times \left\| \begin{matrix} \prod_{p=2,4,\cdots}^{l_3-l_1+s} [(l_1 - l_3 + p)(l_1 + l_3 + p + 1)] \\ 1 \quad (l_1 = l_3 + s) \end{matrix} \right\|$$

$$\times\, \theta(l_1 - l_3)\theta(l_2 + l_1 + \kappa - j - 2l_3 - 2s)\theta(l_2 - l_1 + \kappa - j - 1);$$

(e) if $(l_2 - l_1 + \kappa - j)/2 = \text{even}$,

$$A_{sj} = \binom{[l_2 + \kappa - j - l_1]/2}{[l_2 + x - s - l_3 - j]/2} \frac{\prod_{p=0,2,\cdots}^{2(l_2+\kappa-j-1)-l_1-l_3-s} [(2 + p + l_1 - l_3)(3 + p + l_1 + l_3)]}{\prod_{p=0,2,\cdots}^{l_2+\kappa-j-l_1-2} [(l_2 + \kappa - j - l_1 - p)(l_2 + l_1 + \kappa - j - p + 1)]}$$

$$\times \left\| \begin{matrix} \prod_{p=2,4,\cdots}^{l_2+\kappa-s-l_3-j} [(l_3 - l_2 + j + p)(l_2 + l_3 + j + p + 1)] \\ 1 \quad (s + l_3 + j = l_2 + \kappa) \end{matrix} \right\|$$

$$\times\, \theta(l_3 + j - l_2)\theta(l_2 - l_1 + \kappa - j - 1)\theta(-l_2 - l_1 + 2l_3 - \kappa + j + 2s - 4);$$

(f) $A_{sj} = A_{l_1-l_3,\, l_2+\kappa-l_1} = 1.$

Coefficients not found here are zero.

We have in (58) and (59) (with the A_{sj} determined) completely specified the transformation properties of the polynomial piece $\mathcal{P}^{k,s}$ under S_4.

We now work out the S_4 transformation properties of the spherical harmonic part, considering the special case of $L = 0$.

We denote functions with eigenvalues $\{l_1, l_2, l_3, T, L, M\}$ by $\mathcal{X}_{TLM}^{l_1l_2l_3}(\xi_j)$, and functions which are obtained through the use of standard coupling operators, e.g., $\mathbf{l}_{12}^2$ by $\mathcal{Y}_{l_{12}LM}^{l_1l_2l_3(12)}(\xi_j)$. We may write

$$\mathcal{X}_{TLM}^{l_1l_2l_3}(\xi_j) = \sum_{m_1m_2m_3} \Lambda_{l_1l_2l_3m_1m_2m_3}(T, L, M) Y_{l_1}^{m_1}(\xi_1) Y_{l_2}^{m_2}(\xi_2) Y_{l_3}^{m_3}(\xi_3), \tag{61}$$

with the $\Lambda_{l_1l_2l_3m_1m_2m_3}(T, L, M)$ the desired transformation brackets. Noting that T, the symmetric coupling operator, may be written

$$T = \frac{i}{4} [\mathbf{l}_{12}^2, \mathbf{l}_{23}^2], \tag{62}$$

one has

$$\mathcal{X}_{000}^{l_1l_2l_3}(\xi_j) = \mathcal{Y}_{l_j00}^{l_1l_2l_3(ik)}(\xi_j), \quad i \neq j \neq k = 1, 2, 3, \tag{63}$$

and

$$\Lambda_{l_1l_2l_3m_1m_2m_3}(0, 0, 0) = \frac{(-)^{l_1+m_1}}{(2l_1 + 1)^{\frac{1}{2}}} \langle l_2l_3m_2m_3 \mid l_1 - m_1 \rangle. \tag{64}$$

It now follows directly that

$$(24)(13)\mathfrak{X}_{000}^{l_1l_2l_3} = (-)^{l_1+l_3}\mathfrak{X}_{000}^{l_1l_2l_3}, \qquad (65)$$

$$(14)(23)\mathfrak{X}_{000}^{l_1l_2l_3} = (-)^{l_2+l_3}\mathfrak{X}_{000}^{l_1l_2l_3}, \qquad (66)$$

$$(34)(12)\mathfrak{X}_{000}^{l_1l_2l_3} = (-)^{l_1+l_2}\mathfrak{X}_{000}^{l_1l_2l_3}, \qquad (67)$$

$$(23)\mathfrak{X}_{000}^{l_1l_2l_3} = (-)^{l_2+l_3-l_1}\mathfrak{X}_{000}^{l_1l_3l_2}, \qquad (68)$$

$$(123)\mathfrak{X}_{000}^{l_1l_2l_3} = \mathfrak{X}_{000}^{l_3l_1l_2}. \qquad (69)$$

We now summarize the results:

$$(24)(13)P_{000}^{ksl_1l_2l_3}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3) = (-)^{l_1+l_3}P_{000}^{ksl_1l_2l_3}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3), \qquad (70)$$

$$(14)(23)P_{000}^{ksl_1l_2l_3}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3) = (-)^{l_2+l_3}P_{000}^{ksl_1l_2l_3}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3), \qquad (71)$$

$$(34)(12)P_{000}^{ksl_1l_2l_3}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3) = (-)^{l_1+l_2}P_{000}^{ksl_1l_2l_3}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3), \qquad (72)$$

$$(23)P_{000}^{ksl_1l_2l_3}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3) = (-)^{l_2+l_3-l_1}P_{000}^{k,\kappa-s,l_1l_3l_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3), \qquad (73)$$

$$(123)P_{000}^{ksl_1l_2l_3}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3) = \sum_{j=0,2,\cdots}^{\kappa} A_{sj}[(123)]P_{000}^{kjl_3l_1l_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3). \qquad (74)$$

Here

$$P_{000}^{ksl_1l_2l_3}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3) = \mathfrak{I}^{ks}(\xi_1, \xi_2, \xi_3)\mathfrak{X}_{000}^{l_1l_2l_3}(\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3). \qquad (75)$$

Equations (70)–(74) give us the full matrix representation of the permutation group S_4 in our basis set for the important case of $L = 0$.

CONCLUSION

A complete set of basis functions for the four-particle problem has been developed. The basis was given in terms of homogeneous, harmonic polynomials, and the properties of the polynomials under the permutation group was determined.

Construction of the eighth operator, "missing" from the chain $O_9 \supset O_3^1 \times O_3^2 \times O_3^3 \supset O_3 \supset O_2$, was discussed in detail including proof that an S_4-symmetric, independent operator is impossible.

In Appendix A sample states are given, and in Appendix B we calculate the number of symmetric $L = 0$ states of a given k.

The solutions given may be used for determination of the wavefunctions and energies of a four-particle bound system, hence lending themselves to calculations of ^{4}He.[9] They may also be used directly to investigate symmetries and selection rules in the case of a single elementary particle decaying into four identical particles.

ACKNOWLEDGMENTS

My first acknowledgment goes to Dr. R. D. Amado of the University of Pennsylvania whose guidance and help has proven invaluable in this work. Secondly, I would like to thank Dr. J. D. Louck of Los Alamos Scientific Laboratory for his helpful discussions, particularly concerning Sec. 4.

APPENDIX A: DIMENSION FORMULA AND SAMPLE STATES

The number of independent homogeneous polynomial functions of degree k in nine variables is[10]

$$^9D_{[k]} = (k + 8)!/k!\,8!\,. \qquad (A1)$$

Since the Laplacian reduces the degree of such polynomials by two, it imposes $(k + 6)!/(k - 2)!\,8!$ conditions on the coefficients. Hence, the number of harmonic functions of degree k is

$$n(k) = (k + 8)!/k!\,8! - (k + 6)!/(k - 2)!\,8!\,. \qquad (A2)$$

In Sec. 3 we determined that in $\{ksl_1l_2l_3m_1m_2m_3\}$ s takes on $\kappa/2 + 1$ values with $\kappa \in \{0, 2, \cdots, \binom{k}{k-1}\}$ depending on whether k is even or odd. The l_j are all possible solutions to

$$\textstyle\sum_1^3 l_j = k - \kappa. \qquad (A3)$$

Hence, we obtain the following dimension formula for the basis of Sec. 3:

$$N(k) = \sum_{P=\binom{0}{1},\binom{2}{3},\cdots}^{k} \sum_{l_1,l_2,l_3=0}^{\infty} \left(\frac{k-P}{2} + 1\right) \times (2l_1 + 1)(2l_2 + 1)(2l_3 + 1)\delta_{P,l_1+l_2+l_3}. \qquad (A4)$$

The sum on P is over even (odd) values if k is even (odd). It is not hard to satisfy oneself that

$$N(k) = n(k). \qquad (A5)$$

Below we list all solutions of degrees 1, 2, 3, 4, and some of higher degree.

$$k = 1, \quad n(1) = 9$$

$$P_{m_100}^{10100}(\boldsymbol{\xi}_j) = \xi_1 Y_1^{m_1}(\hat{\xi}_1)$$
$$P_{0m_20}^{10010}(\boldsymbol{\xi}_j) = \xi_2 Y_1^{m_2}(\hat{\xi}_2)$$
$$P_{00m_3}^{10001}(\boldsymbol{\xi}_j) = \xi_3 Y_1^{m_3}(\hat{\xi}_3)$$

$$k = 2, \quad n(2) = 44$$

$$P_{000}^{20000}(\boldsymbol{\xi}_j) = \xi_2^2 - \xi_1^2$$
$$P_{000}^{22000}(\boldsymbol{\xi}_j) = \xi_3^2 - \xi_1^2$$
$$P_{m_1m_20}^{20110}(\boldsymbol{\xi}_j) = \xi_1\xi_2 Y_1^{m_1}(\hat{\xi}_1)Y_1^{m_2}(\hat{\xi}_2)$$
$$P_{m_10m_3}^{20101}(\boldsymbol{\xi}_j) = \xi_1\xi_3 Y_1^{m_1}(\hat{\xi}_1)Y_1^{m_3}(\hat{\xi}_3)$$
$$P_{0m_2m_3}^{20011}(\boldsymbol{\xi}_j) = \xi_2\xi_3 Y_1^{m_2}(\hat{\xi}_2)Y_1^{m_3}(\hat{\xi}_3)$$
$$P_{m_100}^{20200}(\boldsymbol{\xi}_j) = \xi_1^2 Y_2^{m_1}(\hat{\xi}_1)$$
$$P_{0m_20}^{20020}(\boldsymbol{\xi}_j) = \xi_2^2 Y_2^{m_2}(\hat{\xi}_2)$$
$$P_{00m_3}^{20002}(\boldsymbol{\xi}_j) = \xi_3^2 Y_2^{m_3}(\hat{\xi}_3)$$

$$k = 3, \quad n(3) = 156$$

$$P^{30111}_{m_1m_2m_3}(\boldsymbol{\xi}_j) = \xi_1\xi_2\xi_3 Y_1^{m_1}(\hat{\xi}_1)Y_1^{m_2}(\hat{\xi}_2)Y_1^{m_3}(\hat{\xi}_3)$$
$$P^{30120}_{m_1m_20}(\boldsymbol{\xi}_j) = \xi_1\xi_2^2 Y_1^{m_1}(\hat{\xi}_1)Y_2^{m_2}(\hat{\xi}_2)$$
$$P^{30021}_{0m_2m_3}(\boldsymbol{\xi}_j) = \xi_2^2\xi_3 Y_2^{m_2}(\hat{\xi}_2)Y_1^{m_3}(\hat{\xi}_3)$$
$$P^{30201}_{m_10m_3}(\boldsymbol{\xi}_j) = \xi_1^2\xi_3 Y_2^{m_1}(\hat{\xi}_1)Y_1^{m_3}(\hat{\xi}_3)$$
$$P^{30102}_{m_10m_3}(\boldsymbol{\xi}_j) = \xi_1\xi_3^2 Y_1^{m_1}(\hat{\xi}_1)Y_2^{m_3}(\hat{\xi}_3)$$
$$P^{30012}_{0m_2m_3}(\boldsymbol{\xi}_j) = \xi_2\xi_3^2 Y_1^{m_2}(\hat{\xi}_2)Y_2^{m_3}(\hat{\xi}_3)$$
$$P^{30210}_{m_1m_20}(\boldsymbol{\xi}_j) = \xi_1^2\xi_2 Y_2^{m_1}(\hat{\xi}_1)Y_1^{m_2}(\hat{\xi}_2)$$
$$P^{30100}_{m_100}(\boldsymbol{\xi}_j) = (\xi_1\xi_2^2 - \tfrac{6}{10}\xi_1^3)Y_1^{m_1}(\hat{\xi}_1)$$
$$P^{32100}_{m_100}(\boldsymbol{\xi}_j) = (\xi_1\xi_3^2 - \tfrac{6}{10}\xi_1^3)Y_1^{m_1}(\hat{\xi}_1)$$
$$P^{30010}_{0m_20}(\boldsymbol{\xi}_j) = (-\tfrac{10}{6}\xi_1^2\xi_2 + \xi_2^3)Y_1^{m_2}(\hat{\xi}_2)$$
$$P^{32010}_{0m_20}(\boldsymbol{\xi}_j) = (-\xi_1^2\xi_2 + \xi_2\xi_3^2)Y_1^{m_2}(\hat{\xi}_2)$$
$$P^{30001}_{00m_3}(\boldsymbol{\xi}_j) = (-\xi_1^2\xi_3 + \xi_2^2\xi_3)Y_1^{m_3}(\hat{\xi}_3)$$
$$P^{32001}_{00m_3}(\boldsymbol{\xi}_j) = (-\tfrac{10}{6}\xi_1^2\xi_3 + \xi_3^3)Y_1^{m_3}(\hat{\xi}_3)$$
$$P^{30300}_{m_100}(\boldsymbol{\xi}_j) = \xi_1^3 Y_3^{m_1}(\hat{\xi}_1)$$
$$P^{30030}_{0m_20}(\boldsymbol{\xi}_j) = \xi_2^3 Y_3^{m_2}(\hat{\xi}_2)$$
$$P^{30003}_{00m_3}(\boldsymbol{\xi}_j) = \xi_3^3 Y_3^{m_3}(\hat{\xi}_3)$$

$$k = 4, \quad n(4) = 450$$

$$P^{40000}_{000}(\boldsymbol{\xi}_j) = \xi_2^4 + \xi_1^4 - \tfrac{20}{6}\xi_1^2\xi_2^2$$
$$P^{42000}_{000}(\boldsymbol{\xi}_j) = \xi_2^2\xi_3^2 + \tfrac{6}{10}\xi_1^4 - (\xi_1^2\xi_2^2 + \xi_1^2\xi_3^2)$$
$$P^{44000}_{000}(\boldsymbol{\xi}_j) = \xi_3^4 + \xi_1^4 - \tfrac{20}{6}\xi_1^2\xi_3^2$$
$$P^{40200}_{m_100}(\boldsymbol{\xi}_j) = (\xi_1^2\xi_2^2 - \tfrac{3}{7}\xi_1^4)Y_2^{m_1}(\hat{\xi}_1)$$
$$P^{40020}_{0m_20}(\boldsymbol{\xi}_j) = (\xi_2^4 - \tfrac{7}{3}\xi_1^2\xi_2^2)Y_2^{m_2}(\hat{\xi}_2)$$
$$P^{40002}_{00m_3}(\boldsymbol{\xi}_j) = (\xi_2^2\xi_3^2 - \xi_1^2\xi_3^2)Y_2^{m_3}(\hat{\xi}_3)$$
$$P^{42200}_{m_100}(\boldsymbol{\xi}_j) = (\xi_1^2\xi_3^2 - \tfrac{3}{7}\xi_1^4)Y_2^{m_1}(\hat{\xi}_1)$$
$$P^{42020}_{0m_20}(\boldsymbol{\xi}_j) = (\xi_2^2\xi_3^2 - \xi_1^2\xi_2^2)Y_2^{m_2}(\hat{\xi}_2)$$
$$P^{42002}_{00m_3}(\boldsymbol{\xi}_j) = (\xi_3^4 - \tfrac{7}{3}\xi_1^2\xi_3^2)Y_2^{m_3}(\hat{\xi}_3)$$
$$P^{40110}_{m_1m_20}(\boldsymbol{\xi}_j) = (\xi_1\xi_2^3 - \xi_1^3\xi_2)Y_1^{m_1}(\hat{\xi}_1)Y_1^{m_2}(\hat{\xi}_2)$$
$$P^{40101}_{m_10m_3}(\boldsymbol{\xi}_j) = (\xi_1\xi_2^2\xi_3 - \tfrac{3}{5}\xi_1^3\xi_3)Y_1^{m_1}(\hat{\xi}_1)Y_1^{m_3}(\hat{\xi}_2)$$
$$P^{40011}_{0m_2m_3}(\boldsymbol{\xi}_j) = (\xi_2^3\xi_3 - \tfrac{5}{3}\xi_1^2\xi_2\xi_3)Y_1^{m_2}(\hat{\xi}_2)Y_1^{m_3}(\hat{\xi}_3)$$
$$P^{42110}_{m_1m_20}(\boldsymbol{\xi}_j) = (\xi_1\xi_3^2\xi_2 - \tfrac{3}{5}\xi_1^3\xi_2)Y_1^{m_1}(\hat{\xi}_1)Y_1^{m_2}(\hat{\xi}_2)$$
$$P^{42101}_{m_10m_3}(\boldsymbol{\xi}_j) = (\xi_1\xi_3^3 - \xi_1^3\xi_3)Y_1^{m_1}(\hat{\xi}_1)Y_1^{m_3}(\hat{\xi}_3)$$
$$P^{42011}_{0m_2m_3}(\boldsymbol{\xi}_j) = (\xi_3^3\xi_2 - \tfrac{5}{3}\xi_1^2\xi_2\xi_3)Y_1^{m_2}(\hat{\xi}_2)Y_1^{m_3}(\hat{\xi}_3)$$
$$P^{40400}_{m_100}(\boldsymbol{\xi}_j) = \xi_1^4 Y_4^{m_1}(\hat{\xi}_1)$$
$$P^{40040}_{0m_20}(\boldsymbol{\xi}_j) = \xi_2^4 Y_4^{m_2}(\hat{\xi}_2)$$
$$P^{40004}_{00m_3}(\boldsymbol{\xi}_j) = \xi_3^4 Y_4^{m_3}(\hat{\xi}_3)$$
$$P^{40310}_{m_1m_20}(\boldsymbol{\xi}_j) = \xi_1^3\xi_2 Y_3^{m_1}(\hat{\xi}_1)Y_1^{m_2}(\hat{\xi}_2)$$
$$P^{40301}_{m_10m_3}(\boldsymbol{\xi}_j) = \xi_1^3\xi_3 Y_3^{m_1}(\hat{\xi}_1)Y_1^{m_3}(\hat{\xi}_3)$$
$$P^{40130}_{m_1m_20}(\boldsymbol{\xi}_j) = \xi_1\xi_2^3 Y_3^{m_2}(\hat{\xi}_2)Y_1^{m_1}(\hat{\xi}_1)$$
$$P^{40103}_{m_10m_3}(\boldsymbol{\xi}_j) = \xi_1\xi_3^3 Y_3^{m_3}(\hat{\xi}_3)Y_1^{m_1}(\hat{\xi}_1)$$
$$P^{40013}_{0m_2m_3}(\boldsymbol{\xi}_j) = \xi_2\xi_3^3 Y_3^{m_3}(\hat{\xi}_3)Y_1^{m_2}(\hat{\xi}_2)$$
$$P^{40031}_{0m_2m_3}(\boldsymbol{\xi}_j) = \xi_2^3\xi_3 Y_3^{m_2}(\hat{\xi}_2)Y_1^{m_3}(\hat{\xi}_3)$$
$$P^{40220}_{m_1m_20}(\boldsymbol{\xi}_j) = \xi_1^2\xi_2^2 Y_2^{m_1}(\hat{\xi}_1)Y_2^{m_2}(\hat{\xi}_2)$$
$$P^{40202}_{m_10m_3}(\boldsymbol{\xi}_j) = \xi_1^2\xi_3^2 Y_2^{m_1}(\hat{\xi}_1)Y_2^{m_3}(\hat{\xi}_3)$$
$$P^{40022}_{0m_2m_3}(\boldsymbol{\xi}_j) = \xi_2^2\xi_3^2 Y_2^{m_2}(\hat{\xi}_2)Y_2^{m_3}(\hat{\xi}_3)$$
$$P^{40211}_{m_1m_2m_3}(\boldsymbol{\xi}_j) = \xi_1^2\xi_2\xi_3 Y_2^{m_1}(\hat{\xi}_1)Y_1^{m_2}(\hat{\xi}_2)Y_1^{m_3}(\hat{\xi}_3)$$
$$P^{40121}_{m_1m_2m_3}(\boldsymbol{\xi}_j) = \xi_1\xi_2^2\xi_3 Y_1^{m_1}(\hat{\xi}_1)Y_2^{m_2}(\hat{\xi}_2)Y_1^{m_3}(\hat{\xi}_3)$$
$$P^{40112}_{m_1m_2m_3}(\boldsymbol{\xi}_j) = \xi_1\xi_2\xi_3^2 Y_1^{m_1}(\hat{\xi}_1)Y_1^{m_2}(\hat{\xi}_2)Y_2^{m_3}(\hat{\xi}_3)$$

$$k > 4$$

$$P^{70100}_{m_100}(\boldsymbol{\xi}_j) = (\xi_1\xi_2^6 - \tfrac{1}{3}\xi_1^7 + 3\xi_1^5\xi_2^2 - \tfrac{21}{5}\xi_1^3\xi_2^4)Y_1^{m_1}(\hat{\xi}_1)$$
$$P^{72100}_{m_100}(\boldsymbol{\xi}_j) = (\xi_1\xi_2^4\xi_3^2 + \tfrac{3}{7}\xi_1^5\xi_3^2 + \tfrac{6}{7}\xi_1^5\xi_2^2 - 2\xi_1^3\xi_3^2\xi_2^2 - \tfrac{3}{5}\xi_1^3\xi_2^4 - \tfrac{1}{7}\xi_1^7)Y_1^{m_1}(\hat{\xi}_1)$$
$$P^{80000}_{000}(\boldsymbol{\xi}_j) = \xi_1^8 + \xi_2^8 - 12\xi_1^6\xi_2^2 - 12\xi_1^2\xi_2^6 + \tfrac{126}{5}\xi_1^4\xi_2^4$$
$$P^{82000}_{000}(\boldsymbol{\xi}_j) = \tfrac{1}{3}\xi_1^8 - 3\xi_1^6\xi_2^2 - \xi_1^6\xi_3^2 - \xi_1^2\xi_2^6 + \xi_2^6\xi_3^2 + \tfrac{21}{5}\xi_1^4\xi_2^4 + 7\xi_1^4\xi_2^2\xi_3^2 - 7\xi_1^2\xi_2^4\xi_3^2$$
$$P^{84000}_{000}(\boldsymbol{\xi}_j) = \tfrac{5}{21}\xi_1^8 - \tfrac{10}{7}\xi_1^6\xi_2^2 - \tfrac{10}{7}\xi_1^6\xi_3^2 + \xi_2^4\xi_3^4 + \xi_1^4\xi_3^4 + \xi_1^4\xi_2^4 + \tfrac{20}{3}\xi_1^4\xi_2^2\xi_3^2 - \tfrac{10}{3}(\xi_1^2\xi_2^4\xi_3^2 + \xi_1^2\xi_2^2\xi_3^4)$$

APPENDIX B: S_4-SYMMETRIC STATES

We now give the results of the calculation of the number of S_4 symmetric states which occur having $L = 0$. We use character techniques throughout. The well-known formula is

$$a = \frac{1}{24}\sum_{\beta} N_\beta \mathrm{X}(\mathcal{C}_\beta). \tag{B1}$$

In this equation $\mathcal{C}_\beta$ is an operator of class β of the S_4 group, $\mathrm{X}(\mathcal{C}_\beta)$ is its character taken in the matrix representation of Sec. 5, and N_β is the number of operators belonging to class β.

One soon finds that certain traces are automatically zero for k and κ having particular values, e.g., $l_1 \neq l_2 \neq l_3$ implies that X[(123)] is zero.

We define

$$\textstyle\sum_1^3 l_j \equiv P, \tag{B2}$$

and use equations (70)–(74) in (B1) to obtain the following.

A. k-odd

$$N_{(k)}^{\binom{\mathrm{e}}{\mathrm{o}}} = a_1 + a_2 + a_3 + \textstyle\sum_1^4 b_j^{\binom{\mathrm{e}}{\mathrm{o}}}. \tag{B3}$$

In (B3) $N_{(k)}^{\binom{\mathrm{e}}{\mathrm{o}}}$ is the total number of symmetric states with

$$\frac{(k-1)}{2} = \begin{pmatrix}\text{even}\\ \text{odd}\end{pmatrix}. \tag{B4}$$

a_1 is the number of symmetric states having $l_1 \neq l_2 \neq l_3$ and $(P-1)/2$ even,

$$a_1 = \sum_P^{e} \left(\frac{\kappa}{2}+1\right) \Big/ 4\left\{\frac{P-3}{2}\begin{bmatrix}(\alpha-1)/2\\ \alpha/2\end{bmatrix} - 3\left(\begin{bmatrix}(\alpha-1)/2\\ \alpha/2\end{bmatrix}\right)^2 + \frac{1+(-)^{\alpha-1}-2\alpha}{4} + (-1)^{(P-9)/4}\left(\frac{1+(-)^{\left[\substack{(\alpha-3)/2\\(\alpha-2)/2}\right]}}{2}\right)\right\}. \quad \text{(B5)}$$

In (B5)

$$\alpha = \frac{P-1}{2} - \frac{1}{3}\left[P + \begin{pmatrix}2\\1\\0\end{pmatrix}\right] \quad \text{(B6)}$$

must be an integer, and by the matrix symbol [ ] is meant the integer term. The sum is over all P-values having $(P-1)/2$ even.

Likewise a_2 is the number having $l_1 \neq l_2 \neq l_3$ and $(P-1)/2$ odd:

$$a_2 = \sum_P^{o} \left(\frac{\kappa}{2}+1\right) \Big/ 4\left\{\frac{P-7}{2}\begin{bmatrix}(\alpha-1)/2+1\\ (\alpha-2)/2+1\end{bmatrix} - 3\begin{bmatrix}(\alpha-1)/2+1\\ (\alpha-2)/2+1\end{bmatrix}\begin{bmatrix}(\alpha-1)/2\\ (\alpha-2)/2\end{bmatrix} - \frac{1+(-)^{\alpha-1}+2\alpha}{4} + \frac{(-)^{(P-15)/4}}{2}\left[1+(-)^{\left[\substack{(\alpha-1)/2\\(\alpha-2)/2}\right]}\right]\right\}. \quad \text{(B7)}$$

The sum here is over $(P-1)/2$ odd, and α is given by (B6).

a_3 in (B3) is the number of symmetric states with $l_1 = l_2 = l_3$. Noting, however, that in this case X[(123)] is not zero and that

$$\mathrm{X}[(123)] = \sum_j A_{jj}, \quad \text{(B8)}$$

we do not calculate a_3 explicitly.

$b_1^{(\mathrm{e})}_{(\mathrm{o})}$ is the number having two l's equal, with $(P-1)/2$ even, $(k-1)/2$ $\binom{\text{even}}{\text{odd}}$, and $P/3$ not an integer. One has

$$b_1^{\mathrm{e}} = \frac{k-11}{2} + \sum_{r=0}^{\infty}\left\{\theta\left(\frac{k-21}{4}-r\right) \times \begin{bmatrix}(r+4)/2\\ (r+3)/2\end{bmatrix}\left(\frac{k-17}{4}-r\right) - \theta\left(\frac{k-13}{4}-3r\right) \times \begin{bmatrix}(3r+3)/2\\ (3r+2)/2\end{bmatrix}\left(\frac{k-9}{4}-3r\right)\right\}, \quad \text{(B9)}$$

$$b_1^{\mathrm{o}} = \sum_{r=1}^{\infty}\left\{\theta\left(\frac{k-7}{4}-r\right)\begin{bmatrix}r/2\\ (r+1)/2\end{bmatrix}\left(\frac{k-7}{4}-r+1\right) - \theta\left(\frac{k-7}{4}-3r-1\right) \times \begin{bmatrix}(3r+1)/2\\ (3r+2)/2\end{bmatrix}\left(\frac{k-7}{4}-3r\right)\right\} - \frac{k-7}{4}. \quad \text{(B10)}$$

$b_2^{(\mathrm{e})}_{(\mathrm{o})}$ is the number of symmetric, $L=0$, states of $l_i = l_j$ with $(P-1)/2$ odd, $(k-1)/2$ $\binom{\text{even}}{\text{odd}}$, $P/3$ not an integer:

$$b_2^{\mathrm{o}} = \frac{k-7}{4} + 2\theta\left(\frac{k-15}{4}\right)\left(\frac{k-11}{4}\right) + \sum_{r=1}^{\infty}\left\{\theta\left(\frac{k-15}{4}-r\right)\begin{bmatrix}(r+4)/2\\ (r+3)/2\end{bmatrix}\left(\frac{k-11}{4}-r\right) - \theta\left(\frac{k-7}{4}-3r\right)\begin{bmatrix}(3r+1)/2\\ (3r+2)/2\end{bmatrix}\left(\frac{k-3}{4}-3r\right)\right\}, \quad \text{(B11)}$$

$$b_2^{\mathrm{e}} = \frac{k-3}{2} + 2\theta\left(\frac{k-7}{2}\right)\left(\frac{k-5}{2}\right) + \sum_{r=1}^{\infty}\left\{\theta\left(\frac{k-7}{2}-r\right)\begin{bmatrix}(r+4)/2\\ (r+3)/2\end{bmatrix}\left(\frac{k-5}{2}-2\right) - \theta\left(\frac{k-3}{2}-3r\right)\begin{bmatrix}(3r+1)/2\\ (3r+2)/2\end{bmatrix}\left(\frac{k-1}{2}-3r\right)\right\}. \quad \text{(B12)}$$

$b_3^{(\mathrm{e})}_{(\mathrm{o})}$ is the number having $l_i = l_j$ with $P/3$ integer, $(P-1)/2$ even, but not including cases of $l_1 = l_2 = l_3$:

$$b_3^{\mathrm{e}} = \frac{k-21}{4} + \sum_{r=1}^{\infty}\theta\left(\frac{k-25}{4}-3r\right) \times \begin{bmatrix}(3r+3)/2\\ (3r+2)/2\end{bmatrix}\left(\frac{k-21}{4}-3r\right), \quad \text{(B13)}$$

$$b_3^{\mathrm{o}} = \frac{k-19}{4} + \sum_{r=1}^{\infty}\theta\left(\frac{k-23}{4}-3r\right) \times \begin{bmatrix}(3r+3)/2\\ (3r+2)/2\end{bmatrix}\left(\frac{k-19}{4}-3r\right). \quad \text{(B14)}$$

$b_4^{(\mathrm{e})}_{(\mathrm{o})}$ has $(P-1)/2$ odd, but all other conditions duplicate $b_3^{(\mathrm{e})}_{(\mathrm{o})}$:

$$b_4^{\mathrm{e}} = \frac{k-13}{4} + \sum_{r=1}^{\infty}\theta\left(\frac{k-17}{4}-3r\right) \times \begin{bmatrix}(3r+3)/2\\ (3r+2)/2\end{bmatrix}\left(\frac{k-13}{4}-3r\right), \quad \text{(B15)}$$

$$b_4^{\mathrm{o}} = \frac{k-15}{4} + \sum_{r=1}^{\infty}\theta\left(\frac{k-19}{4}-3r\right) \times \begin{bmatrix}(3r+3)/2\\ (3r+2)/2\end{bmatrix}\left(\frac{k-15}{4}-3r\right). \quad \text{(B16)}$$

F. k even, $k/2\,\binom{\text{even}}{\text{odd}}$

$$N_{(k)}^{\binom{e}{o}} = c_1 + c_2 + c_3 + \textstyle\sum_1^4 d_j^{\binom{e}{o}}. \tag{B17}$$

As for k odd we have:

C_1 is the number of symmetric states of $P/2$ odd, $l_1 \neq l_2 \neq l_3$,

$$\begin{aligned} C_1 = \sum_P{}^{\mathrm{o}} \left(\frac{\kappa}{2}+1\right)\Big/4\Bigg\{&\frac{P-4}{2}\begin{bmatrix}(\alpha-1)/2\\ \alpha/2\end{bmatrix} \\ &- 3\left(\begin{bmatrix}(\alpha-1)/2\\ \alpha/2\end{bmatrix}\right)^2 + \frac{1+(-)^{\alpha-1}-2\alpha}{4} \\ &+ (-)^{(P-2)/4}\cdot\frac{1+(-)^{\left[\substack{(\alpha-3)/2\\ (\alpha-2)/2}\right]}}{2}\Bigg\}, \end{aligned} \tag{B18}$$

where α is given by

$$\alpha = \frac{P}{2} - \frac{1}{3}\left[P + \begin{pmatrix}2\\1\\0\end{pmatrix}\right] \tag{B19}$$

and the sum is over all $P/2$ odd. C_2 is the number of symmetric states of $P/2$ even, $l_1 \neq l_2 \neq l_3$,

$$\begin{aligned} C_2 = \sum_P{}^{\mathrm{e}} \left(\frac{\kappa}{2}+1\right)\Big/4\Bigg\{&\frac{P-4}{2}\begin{bmatrix}(\alpha-1)/2+1\\ (\alpha-2)/2+1\end{bmatrix} \\ &- 3\begin{bmatrix}(\alpha-1)/2+1\\ (\alpha-2)/2+1\end{bmatrix}\begin{bmatrix}(\alpha-1)/2\\ (\alpha-2)/2\end{bmatrix} \\ &- \frac{1+(-)^{\alpha-1}+2\alpha}{4} \\ &+ \frac{(-)^{P/4}}{4}\left(-1-2(-)^{\left[\substack{(\alpha-1)/2\\ (\alpha-2)/2}\right]} + (-)^{\left[\substack{(\alpha-1)/2+1\\ (\alpha-2)/2}\right]}\right) \\ &+ (-)^{(P+4)/4}\cdot\frac{1+(-)^{\left[\substack{(\alpha-3)/2\\ (\alpha-2)/2}\right]}}{4}\Bigg\}, \end{aligned} \tag{B20}$$

where now we sum on $P/2$ even. As before, C_3 is the number for $l_1 = l_2 = l_3$, and must be calculated from (B8). $d_1^{\binom{e}{o}}$ is the number of symmetric states with $P/3$ not an integer, $P/2$ even, $k/2\,\binom{\text{even}}{\text{odd}}$, and $l_i = l_j$:

$$\begin{aligned} d_1^{\mathrm{e}} = \sum_{r=0}^{\infty}\Bigg\{&\theta\left(\frac{k}{4}-r-1\right)\begin{bmatrix}(r+3)/2\\ (r+2)/2\end{bmatrix}\left(\frac{k}{4}-r\right) \\ &- \theta\left(\frac{k}{4}-3r-3\right)\begin{bmatrix}(3r+4)/2\\ (3r+5)/2\end{bmatrix}\left(\frac{k}{4}-3r-2\right)\Bigg\}, \end{aligned} \tag{B21}$$

$$\begin{aligned} d_1^{\mathrm{o}} = \sum_{r=0}^{\infty}\Bigg\{&\theta\left(\frac{k-6}{4}-r\right)\begin{bmatrix}(r+3)/2\\ (r+2)/2\end{bmatrix}\left(\frac{k-2}{4}-r\right) \\ &- \theta\left(\frac{k-14}{4}-3r-3\right) \\ &\times \begin{bmatrix}(3r+4)/2\\ (3r+5)/2\end{bmatrix}\left(\frac{k-10}{4}-3r\right)\Bigg\}. \end{aligned} \tag{B22}$$

Similarly, $d_2^{\binom{e}{o}}$ is the number having $P/3$ not integer, $P/2$ odd, $k/2\,\binom{\text{even}}{\text{odd}}$, with $l_i = l_j$:

$$\begin{aligned} d_2^{\mathrm{e}} = \sum_{r=0}^{\infty}\Bigg\{&\theta\left(\frac{k-8}{4}-r\right)\begin{bmatrix}(r+2)/2\\ (r+1)/2\end{bmatrix}\left(\frac{k-4}{4}-r\right) \\ &- \theta\left(\frac{k-8}{4}-3r\right)\begin{bmatrix}(3r+1)/2\\ (3r+2)/2\end{bmatrix}\left(\frac{k-4}{4}-3r\right)\Bigg\}, \end{aligned} \tag{B23}$$

$$\begin{aligned} d_2^{\mathrm{o}} = \sum_{r=0}^{\infty}\Bigg\{&\theta\left(\frac{k-6}{4}-r\right)\begin{bmatrix}(r+2)/2\\ (r+1)/2\end{bmatrix}\left(\frac{k-2}{4}-r\right) \\ &- \theta\left(\frac{k-6}{4}-3r\right)\begin{bmatrix}(3r+1)/2\\ (3r+2)/2\end{bmatrix}\left(\frac{k-2}{4}-3r\right)\Bigg\}. \end{aligned} \tag{B24}$$

$d_3^{\binom{e}{o}}$ and $d_4^{\binom{e}{o}}$ have $l_i = l_j$ and $P/3$ an integer, with $P/2$ odd and $P/2$ even, respectively:

$$d_3^{\mathrm{e}} = \sum_{r=0}^{\infty}\theta\left(\frac{k-20}{4}-3r\right)\begin{bmatrix}(3r+2)/2\\ (3r+3)/2\end{bmatrix}\left(\frac{k-16}{4}-3r\right), \tag{B25}$$

$$d_3^{\mathrm{o}} = \sum_{r=0}^{\infty}\theta\left(\frac{k-26}{4}-3r\right)\begin{bmatrix}(3r+2)/2\\ (3r+3)/2\end{bmatrix}\left(\frac{k-22}{4}-3r\right), \tag{B26}$$

$$d_4^{\mathrm{e}} = \sum_{r=0}^{\infty}\theta\left(\frac{k-12}{4}-3r\right)\begin{bmatrix}(3r+2)/2\\ (3r+3)/2\end{bmatrix}\left(\frac{k-8}{4}-3r\right), \tag{B27}$$

$$d_4^{\mathrm{o}} = \sum_{r=0}^{\infty}\theta\left(\frac{k-14}{4}-3r\right)\begin{bmatrix}(3r+2)/2\\ (3r+3)/2\end{bmatrix}\left(\frac{k-10}{4}-3r\right). \tag{B28}$$

* This work supported in part by the National Science Foundation.
† National Science Foundation Predoctoral Fellow.

[1] A. J. Dragt, J. Math. Phys. 6, 533 (1965).

[2] J. M. Lévy-Leblond and F. Lurçat, J. Math. Phys. 6, 1565 (1965).

[3] This holds also for n-particle case. See F. T. Smith, J. Math. Phys. 3, 735 (1962).

[4] J. M. Lévy-Leblond, J. Math. Phys. 7, 2217 (1966).

[5] With notation (ij) of the interchange of particle i with j, $\exists$ a subgroup of S_4 of elements $\{i, (24)(13), (14)(23), (34)(12)\}$ denoted "V_4" such that if $v \in V_4$ and $l_{\alpha i}$ is an infinitesimal generator of O_3^α, then $v l_{\alpha i} v^{-1} = l_{\alpha i}$, which means that $O_3^1 \times O_3^2 \times O_3^3$ is a "V_4-democratic" subgroup of O_9. This was first shown by Lévy-Leblond, Ref. 4.

[6] For the case of unequal masses see Ref. 1.

[7] J. M. Lévy-Leblond and M. Lévy-Nahas, J. Math. Phys. 6, 1372 (1965).

[8] A. Chakrabarti, Ann. Inst. Henri Poincaré 1, 301 (1964).

[9] Such calculations have been done using a different set of functions by F. De La Ripelle, published in the "Proceedings of the International School of Theoretical Physics," Predeal-Rumania, 10–23 September, 1969.

[10] J. D. Louck, J. Math. Phys. 6, 1786 (1965).

Low-Frequency Scattering by Perfectly Conducting Obstacles*

J. S. ASVESTAS AND R. E. KLEINMAN
Department of Mathematics, University of Delaware, Newark, Delaware

(Received 14 May 1970)

Coupled Fredholm integral equations of the second kind are derived for the electric and magnetic fields scattered when a smooth, bounded, perfectly conducting three-dimensional obstacle is illuminated by a time harmonic, monochromatic, otherwise arbitrary incident field. The kernels of the equations are dyadics constructed from potential functions associated with the scattering surface, i.e., solutions of Laplace's equation satisfying particular boundary conditions. If the frequency of the incident field is sufficiently low, the integral equations may be solved in a standard Neumann series. This is demonstrated in an example, scattering of a plane wave by a sphere.

I. INTRODUCTION

The purpose of this paper is to present a new method for obtaining solutions to the exterior boundary value problem arising when a time harmonic electromagnetic wave is scattered by a perfectly conducting, three-dimensional, smooth, closed, bounded obstacle in the particular case when the wavelength of the incident radiation is large compared with the characteristic dimension of the scatterer. The surrounding medium is linear, isotropic, homogeneous, and of zero conductivity.

This problem was first investigated by Lord Rayleigh.[1] In his now classic paper he examined the scattering of both acoustical and electromagnetic waves by two- as well as three-dimensional obstacles. For three-dimensional electromagnetic problems, he showed that, in the limit as the wavenumber k tends to zero, the electric and magnetic scattered vectors in the near field region can be expressed in terms of solutions of standard potential problems. Furthermore, he was able to continue these solutions to the far field region and arrive at his famous fourth power of frequency law for the scattering cross section of objects whose characteristic dimension is small compared with the wavelength of the incident radiation.

Since that time considerable work has been done in obtaining higher-order terms in the low-frequency expansions of the scattered fields and in generalizing Lord Rayleigh's ideas. Kleinman[2] gives an extensive bibliography up to 1965. The major contribution to the subject came from Stevenson,[3] who showed that if the scattered electric and magnetic fields (denoted by $\mathbf{E}^s$ and $\mathbf{H}^s$, respectively) are written in power series of the form

$$\mathbf{E}^s = \sum_{m=0}^{\infty} k^m \mathbf{E}_m^s, \quad \mathbf{H}^s = \sum_{m=0}^{\infty} k^m \mathbf{H}_m^s, \tag{1}$$

then, by employing the Stratton–Chu formulation, the coefficients $\mathbf{E}_m^s$ and $\mathbf{H}_m^s$ in (1) can be written in the form

$$\mathbf{E}_m^s = \mathbf{F}_m + \nabla\phi_m, \quad \mathbf{H}_m^s = \mathbf{G}_m + \nabla\psi_m, \tag{2}$$

where $\mathbf{F}_m$ and $\mathbf{G}_m$ are known in terms of the previous coefficients in the expansions and ϕ_m and ψ_m are solutions of well-defined potential problems. The expansions of the fields thus obtained are valid in the near field region only, but they can be continued into the far field. This procedure as presented by Stevenson had certain weaknesses which were rectified by Kleinman.[4]

Inherent in all three-dimensional low-frequency techniques is the assumption that low-frequency expansions of the type (1) exist. That this is so was proved by Werner[5] who showed that, in the limit as $k \to 0$, the electric scattered field tends analytically to a corresponding electrostatic field. The same is true of the magnetic scattered field.[6]

In the present paper a new method is developed by means of which one may obtain as many terms as desired in the low-frequency expansions of the scattered fields by successive operations on two dyadic potential functions. These fundamental dyadics are derivable from solutions of the Laplace equation which satisfy certain boundary conditions on the scatterers. The class of surfaces for which the method applies is thus limited to those surfaces for which the requisite potential problems can be solved. This is the same limitation to which Stevenson's method is subject. The advantage of the present method over Stevenson's is that once the fundamental dyadics for a particular surface are determined, we can find successive terms in the low-frequency expansion by a straightforward iteration of a pair of coupled Fredholm integral equations of the second kind, doing away with the need to solve $3m$ boundary value problems[7] in the determination of $\mathbf{E}_m^s$ and $\mathbf{H}_m^s$ in (1). The present method, moreover, yields the fields everywhere in space, thus obviating the problem of continuation of the near field results to the far field. A disadvantage is that it applies (at least for the present) only to the case of perfectly conducting scatterers, while Stevenson's applies to dielectric and imperfectly conducting scatterers as well.

The main result of the paper is the following: If $\mathbf{E}^s(\mathbf{R}')$ and $\mathbf{H}^s(\mathbf{R}')$ are the electric and magnetic fields at $\mathbf{R}'$ scattered by a smooth surface S when illuminated by a time harmonic, monochromatic, but otherwise arbitrary incident electromagnetic field and if $\mathbf{e}$ and $\mathbf{h}$ denote $e^{-ikR'}\mathbf{E}^s$ and $e^{-ikR'}\mathbf{H}^s$, respectively, then[8]

$$\mathbf{h}(\mathbf{R}') = ik\int_V \{[Y\mathbf{e}(\mathbf{R}) + \hat{\mathbf{R}} \times \mathbf{h}(\mathbf{R})] \cdot \mathsf{E}_m^{(1)}(\mathbf{R} \mid \mathbf{R}') - \hat{\mathbf{R}} \cdot \mathbf{h}(\mathbf{R})\nabla' N^{(e)}(\mathbf{R} \mid \mathbf{R}')\}\, dv - \nabla' \int_S \hat{\mathbf{n}} \cdot \mathbf{h}(\mathbf{R}) N^{(e)}(\mathbf{R} \mid \mathbf{R}')\, ds, \tag{3}$$

$$\mathbf{e}(\mathbf{R}') = -ik\int_V \Big[[Z\mathbf{h}(\mathbf{R}) - \hat{\mathbf{R}} \times \mathbf{e}(\mathbf{R})] \cdot \mathsf{H}_e^{(1)}(\mathbf{R} \mid \mathbf{R}') + \hat{\mathbf{R}} \cdot \mathbf{e}(\mathbf{R})\nabla' \Big(G^{(e)}(\mathbf{R} \mid \mathbf{R}') - \frac{[U(\mathbf{R}) - 1]U(\mathbf{R}')}{4\pi C} \Big) \Big] + \int_S [\hat{\mathbf{n}} \times \mathbf{e}(\mathbf{R})] \cdot \mathsf{H}_e^{(1)}(\mathbf{R} \mid \mathbf{R}')\, ds. \tag{4}$$

$N^{(e)}$, $G^{(e)}$, and U are all potential functions and $\mathsf{E}_m^{(1)}$, $\mathsf{H}_e^{(1)}$, and C are defined in terms of them. If k is small, the right-hand sides are dominated by terms which are known through the boundary conditions on $\mathbf{E}^s$ and $\mathbf{H}^s$, thus providing a basis for the iteration process which is shown to yield the exact result in a particular example.

The plan for the development of the method is as follows: The dyadic form of Green's theorem is employed in Sec. II to derive two vector integral equations whose kernels are dyadic functions of position. In Sec. III it is shown that the requirements imposed on the dyadics are satisfied by certain of the coefficients in the low-frequency expansions of harmonically oscillating infinitesimal electric and magnetic dipoles. Moreover, these coefficients (dyadics) are shown to be derivable from standard potential functions. In Sec. IV we employ the expansion theorem of Wilcox[9] to show that the electric and magnetic fields of the scattering problem belong to the same class of vector functions as the unknowns of the integral equations, thus arriving (with the help of Maxwell's equations) at two coupled integral equations for the scattered fields. These equations may be iterated to produce a Neumann series for each of the fields. In Sec. V the results are applied to the problem of scattering by a sphere as a check and demonstration of the method.

II. NOTATION AND DERIVATION OF TWO INTEGRAL EQUATIONS

Let S denote a closed, bounded, regular surface in E^3, and denote the exterior by V and the interior by V_i. Let $\hat{\mathbf{n}}$ denote a unit normal vector directed from S into V_i. (Boldface denotes a vector, a caret denotes a unit vector, and sans serif denotes a dyadic.) Erect a Cartesian coordinate system with origin in V_i and let $\mathbf{R}$ be a position vector with spherical polar coordinates (R, θ, ϕ). The smoothness of the surface S is stipulated by requiring that S be described by an equation

$$R = g(\theta, \phi), \quad 0 \le \theta \le \pi, \quad 0 \le \phi \le 2\pi, \tag{5}$$

where g is continuously differentiable in θ and ϕ,

$$g(\theta, 0) = g(\theta, 2\pi), \tag{6}$$

$$\frac{\partial}{\partial \phi} g(0, \phi) = \frac{\partial}{\partial \phi} g(\pi, \phi) = 0, \tag{7}$$

and by requiring that $\hat{\mathbf{n}} \cdot \hat{\mathbf{R}}$ be uniformly Hölder continuous on S.

A vector-valued function $\mathbf{F}(\mathbf{R})$ will be called *regular* in the exterior domain V if[10]

$$\mathbf{F}(\mathbf{R}) \in C^2(V), \quad \mathbf{F}(\mathbf{R}) \in C^1(V \cup S), \tag{8}$$

$$\lim_{R\to\infty} |\hat{\mathbf{R}} \times \mathbf{F}(\mathbf{R})| < \infty, \tag{9}$$

$$\lim_{R\to\infty} |R\nabla \times \mathbf{F}(\mathbf{R})| < \infty. \tag{10}$$

The object of this section is to derive two integral equations for vector-valued functions regular in this sense. The kernels of the integral equations involve the following fundamental dyadic functions of two points:

$$\mathsf{E}_m^{(1)}(\mathbf{R} \mid \mathbf{R}') = \nabla \times \Big(-\frac{\mathsf{I}}{4\pi |\mathbf{R} - \mathbf{R}'|} \Big) + \mathsf{E}_{m_r}^{(1)}(\mathbf{R} \mid \mathbf{R}'), \tag{11}$$

where

$$\nabla \times \nabla \times \mathsf{E}_{m_r}^{(1)} = 0, \quad \mathbf{R}, \mathbf{R}' \in V, \tag{12}$$

$$\hat{\mathbf{n}} \times \mathsf{E}_m^{(1)} = 0, \quad \mathbf{R} \in S, \tag{13}$$

$$\lim_{R\to\infty} |R^2 \hat{R} \times \mathsf{E}_m^{(1)}| < \infty, \tag{14}$$

$$\lim_{R\to\infty} |R^3 \nabla \times \mathsf{E}_m^{(1)}| < \infty, \tag{15}$$

and

$$\mathsf{H}_e^{(1)}(\mathbf{R} \mid \mathbf{R}') = \nabla \times \Big(-\frac{\mathsf{I}}{4\pi |\mathbf{R} - \mathbf{R}'|} \Big) + \mathsf{H}_{e_r}^{(1)}(\mathbf{R} \mid \mathbf{R}'), \tag{16}$$

where

$$\nabla \times \nabla \times \mathsf{H}_{e_r}^{(1)} = 0, \quad \mathbf{R}, \mathbf{R}' \in V, \tag{17}$$

$$\hat{\mathbf{n}} \times \nabla \times \mathsf{H}_e^{(1)} = 0, \qquad \mathbf{R} \in S, \tag{18}$$

$$\lim_{R \to \infty} |R^2 \hat{\mathbf{R}} \times \mathsf{H}_e^{(1)}| < \infty, \tag{19}$$

$$\lim_{R \to \infty} |R^3 \nabla \times \mathsf{H}_e^{(1)}| < \infty. \tag{20}$$

∇ operates on the coordinates of the unprimed vector $\mathbf{R}$ (∇' and ∇_s will operate on $\mathbf{R}'$ and $\mathbf{R}_s$, respectively), where $\mathbf{R}_s$ denotes a radius vector to a point on the surface S. I is the identity dyadic which is given in terms of rectangular unit vectors by

$$\mathsf{I} = \hat{\mathbf{a}}_1\hat{\mathbf{a}}_1 + \hat{\mathbf{a}}_2\hat{\ }_2 + \hat{\mathbf{a}}_3\hat{\mathbf{a}}_3. \tag{21}$$

An explicit definition of these dyadics in terms of potential Green's functions for the surface S, as well as a physical interpretation as coefficients in an expansion of the fields due to infinitesimal electric and magnetic dipoles, will be deferred to the next section. The notation is motivated by this interpretation. Actually, the above properties of $\mathsf{E}_m^{(1)}$ and $\mathsf{H}_e^{(1)}$ do not uniquely define these dyadics; additional restrictions will be imposed in the next section. They are sufficient, however, to establish the following.

Theorem 1: If $\mathbf{F}(\mathbf{R})$ is a vector-valued function, regular in V, then

(a)

$$\nabla' \times \mathbf{F}(\mathbf{R}') = -\int_V [\nabla \times \nabla \times \mathbf{F}(\mathbf{R})] \cdot \mathsf{E}_m^{(1)}(\mathbf{R} \mid \mathbf{R}')\, dv + \int_S [\hat{\mathbf{n}} \times \mathbf{F}(\mathbf{R}_s)] \cdot \nabla_s \times \mathsf{E}_m^{(1)}(\mathbf{R}_s \mid \mathbf{R}')\, ds \tag{22}$$

and

(b)

$$\nabla' \times \mathbf{F}(\mathbf{R}') = -\int_V [\nabla \times \nabla \times \mathbf{F}(\mathbf{R})] \cdot \mathsf{H}_e^{(1)}(\mathbf{R} \mid \mathbf{R}')\, dv + \int_S \{\hat{\mathbf{n}} \times [\nabla_s \times \mathbf{F}(\mathbf{R}_s)]\} \cdot \mathsf{H}_e^{(1)}(\mathbf{R}_s \mid \mathbf{R}')\, ds, \tag{23}$$

where $\hat{\mathbf{n}}$ is the unit normal from S into its interior, directed away from V.

The proof of this theorem is based on the dyadic form of the divergence theorem for infinite domains; namely, if $\mathsf{A}(\mathbf{R})$ is continuously differentiable in V, then

$$\int_V \nabla \cdot \mathsf{A}(\mathbf{R})\, dv = \int_{S+S_\infty} \hat{\mathbf{n}} \cdot \mathsf{A}(\mathbf{R})\, ds, \tag{24}$$

where V is the volume exterior to S and interior to S_∞ and S_∞ is a large sphere whose radius will tend to infinity. The unit normal $\hat{\mathbf{n}}$ is always directed out of V.

This form of the divergence theorem follows immediately from the corresponding theorem for vector functions. Attention is drawn to the fact that the dot product in the surface integral is not necessarily commutative. By writing $\mathsf{A}(\mathbf{R}')$ as

$$\mathsf{A}(\mathbf{R}) = \mathbf{F}(\mathbf{R}) \times (\nabla \times \mathsf{P}) + [\nabla \times \mathbf{F}(\mathbf{R})] \times \mathsf{P}, \tag{25}$$

where $\mathbf{F}(\mathbf{R})$ is regular in V and P is a dyadic function of position, eventually to be identified as one of the fundamental dyadics $\mathsf{E}_m^{(1)}$ or $\mathsf{H}_e^{(1)}$, and employing (A3),[11] we obtain the following Green's identity:

$$\int_V [\nabla \times \nabla \times \mathbf{F}) \cdot \mathsf{P} - \mathbf{F} \cdot (\nabla \times \nabla \times \mathsf{P})]\, dv = \int_{S+S_\infty} \hat{\mathbf{n}} \cdot [\mathbf{F} \times (\nabla \times \mathsf{P}) + (\nabla \times \mathbf{F}) \times \mathsf{P}]\, ds. \tag{26}$$

Since P is to be identified with one of the fundamental dyadics, which have the same singularity at $\mathbf{R}' = \mathbf{R}$, it is necessary to delete from V a small sphere S' with origin at R' and radius r (see Fig. 1) and then let $r \to 0$, yielding, since $\nabla \times \nabla \times \mathsf{P} = 0$ in the remaining volume,

$$\int_V (\nabla \times \nabla \times \mathbf{F}) \cdot \mathsf{P}\, dv = \int_{S+S_\infty+S'} \hat{\mathbf{n}} \cdot [\mathbf{F} \times (\nabla \times \mathsf{P}) + (\nabla \times \mathbf{F}) \times \mathsf{P}]\, ds. \tag{27}$$

The behavior of P as $R \to \infty$ is given in (14) and (15)

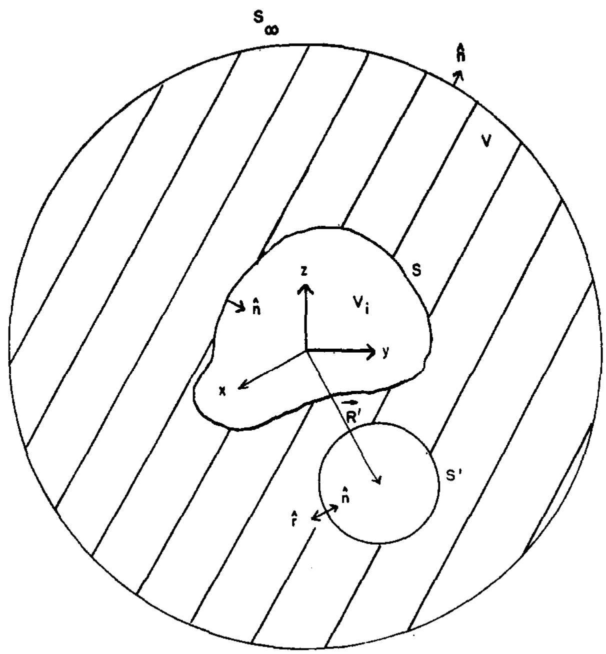


FIG. 1. Geometry for the application of Green's identity.

or (19) and (20). This, together with the behavior of F, (9), (10), and (A1), guarantees that

$$\lim_{R\to\infty}\int_{S_\infty}\hat{\mathbf{n}}\cdot[\mathbf{F}\times(\nabla\times\mathsf{P})+(\nabla\times\mathbf{F})\times\mathsf{P}]\,ds$$
$$=\lim_{R\to\infty}\int_0^{2\pi}d\phi\int_0^{\pi}d\theta R^2\sin\theta[(\hat{\mathbf{R}}\times\mathbf{F})\cdot\nabla\times\mathsf{P}-(\nabla\times\mathbf{F})\cdot(\hat{\mathbf{R}}\times\mathsf{P})]$$
$$=0. \tag{28}$$

Also, since the singularity of the fundamental dyadics is the same, regardless of whether P is identified with $\mathsf{E}_m^{(1)}$ or $\mathsf{H}_e^{(1)}$, the surface integral over S' is

$$I_{S'}=\lim_{r\to 0}\int_{S'}\hat{\mathbf{n}}\cdot[\mathbf{F}\times(\nabla\times\mathsf{P})+(\nabla\times\mathbf{F})\times\mathsf{P}]\,ds$$
$$=\lim_{r\to 0}\int_{S'}\hat{\mathbf{n}}\cdot\left\{\mathbf{F}\times\left[\nabla\times\nabla\times\left(\frac{-\mathsf{I}}{4\pi|\mathbf{R}-\mathbf{R}'|}\right)\right]+(\nabla\times\mathbf{F})\times\left[\nabla\times\left(\frac{-\mathsf{I}}{4\pi|\mathbf{R}-\mathbf{R}'|}\right)\right]\right\}ds, \tag{29}$$

where $r=|\mathbf{R}-\mathbf{R}'|$. By (A7) and the fact that, on S', $\hat{\mathbf{n}}=-\hat{\mathbf{r}}=(\mathbf{R}'-\mathbf{R})/|\mathbf{R}'-\mathbf{R}|$, this becomes

$$I_{S'}=-\lim_{r\to 0}\int_{S'}\hat{\mathbf{r}}\cdot\left\{\mathbf{F}\times\nabla\nabla\left(-\frac{1}{4\pi r}\right)+(\nabla\times\mathbf{F})\times\left[\nabla\times\left(\frac{-\mathsf{I}}{4\pi r}\right)\right]\right\}ds. \tag{30}$$

By means of (A5) this may be rewritten as

$$I_{S'}=-\lim_{r\to 0}\int_{S'}\hat{\mathbf{r}}\cdot\left\{(\nabla\times\mathbf{F})\nabla\left(\frac{-1}{4\pi r}\right)-\nabla\times\left[\mathbf{F}\nabla\left(-\frac{1}{4\pi r}\right)\right]+(\nabla\times\mathbf{F})\times\left[\nabla\times\left(-\frac{\mathsf{I}}{4\pi r}\right)\right]\right\}ds, \tag{31}$$

and thus the integral of the term involving $\hat{\mathbf{r}}\cdot\nabla\times$ vanishes by Stokes' theorem. With the help of Eqs. (A1), (A2), and (A6), the remaining terms in the integrand may be rewritten as follows:

$$\hat{\mathbf{r}}\cdot\{(\nabla\times\mathbf{F})\nabla(-1/4\pi r)+(\nabla\times\mathbf{F})\times[\nabla\times(-\mathsf{I}/4\pi r)]\}$$
$$=(1/4\pi r^2)\hat{\mathbf{r}}\cdot[(\nabla\times\mathbf{F})\hat{\mathbf{r}}+(\nabla\times\mathbf{F})\times(\hat{\mathbf{r}}\times\mathsf{I})]$$
$$=(1/4\pi r^2)\{\hat{\mathbf{r}}\cdot(\nabla\times\mathbf{F})\hat{\mathbf{r}}-(\nabla\times\mathbf{F})\cdot[\hat{\mathbf{r}}\times(\hat{\mathbf{r}}\times\mathsf{I})]\}$$
$$=(1/4\pi r^2)[\hat{\mathbf{r}}\cdot(\nabla\times\mathbf{F})\hat{\mathbf{r}}-(\nabla\times\mathbf{F})\cdot(\hat{\mathbf{r}}\hat{\mathbf{r}}-\mathsf{I})]$$
$$=(1/4\pi r^2)\nabla\times\mathbf{F}. \tag{32}$$

Substituting this result in (31) gives

$$I_{S'}=-\lim_{r\to 0}\int_{S'}ds\frac{\nabla\times\mathbf{F}}{4\pi r^2}=-\nabla'\times\mathbf{F}(\mathbf{R}'). \tag{33}$$

Using this result, as well as (28), in (27) yields

$$\nabla'\times\mathbf{F}(\mathbf{R}')=-\int_V(\nabla\times\nabla\times\mathbf{F})\cdot\mathsf{P}\,dv+\int_S\hat{\mathbf{n}}\cdot[\mathbf{F}\times(\nabla\times\mathsf{P})+(\nabla\times\mathbf{F})\times\mathsf{P}]\,ds \tag{34}$$

or, with (A1),

$$\nabla'\times\mathbf{F}(\mathbf{R}')=-\int_V(\nabla\times\nabla\times\mathbf{F})\cdot\mathsf{P}\,dv-\int_S\{\mathbf{F}\cdot[\hat{\mathbf{n}}\times(\nabla\times\mathsf{P})]+(\nabla\times\mathbf{F})\cdot(\hat{\mathbf{n}}\times\mathsf{P})\}\,ds. \tag{35}$$

If $\mathsf{P}=\mathsf{E}_m^{(1)}$, then the boundary condition (13) implies

$$\int_S(\nabla\times\mathbf{F})\cdot(\hat{\mathbf{n}}\times\mathsf{E}_m^{(1)})\,ds=0 \tag{36}$$

and yields

$$\nabla'\times\mathbf{F}(\mathbf{R}')=-\int_V(\nabla\times\nabla\times\mathbf{F})\cdot\mathsf{E}_m^{(1)}\,dv-\int_S\mathbf{F}\cdot(\hat{\mathbf{n}}\times\nabla\times\mathsf{E}_m^{(1)})\,ds \tag{37}$$

which, with (A1), establishes Theorem 1(a).

If $\mathsf{P}=\mathsf{H}_e^{(1)}$, then the boundary condition (18) implies

$$\int_S\mathbf{F}\cdot\hat{\mathbf{n}}\times\nabla\times\mathsf{H}_e^{(1)}\,ds=0 \tag{38}$$

and

$$\nabla'\times\mathbf{F}(R')=-\int_V(\nabla\times\nabla\times\mathbf{F})\cdot\mathsf{H}_e^{(1)}\,dv-\int_S(\nabla\times\mathbf{F})\cdot(\hat{\mathbf{n}}\times\mathsf{H}_e^{(1)})\,ds \tag{39}$$

which, with (A1), establishes Theorem 1(b).

The integral equations given in Theorem 1 constitute the basis for the solution of the low-frequency electromagnetic scattering problem. They will be employed in Sec. IV to derive integral equations for the scattered fields. In order to be soluble, it is necessary to have explicit expressions for the dyadic kernels in terms of potential functions, and this is accomplished in the next section.

III. THE FUNDAMENTAL DYADICS AND THE FIELDS OF INFINITESIMAL DIPOLES

The fundamental dyadics appearing in Theorem 1 are not uniquely specified by (11)–(20). The present section is devoted to defining physically meaningful dyadics which not only fulfill the requirements for use in Theorem 1 but also are expressed explicitly in terms of standard potential functions.

Using the definitions and notation of Sec. II, let S be a perfectly conducting, closed, bounded, regular surface immersed in a linear, isotropic, homogeneous, nonconducting medium of infinite extent. Let J be the volume dyadic current density, nonzero in a finite region of V, the exterior of S. A harmonic time variation ($e^{-i\omega t}$) associated with J is suppressed. The total time independent electromagnetic fields in dyadic form satisfy Maxwell's equations

$$\nabla \times \mathsf{E}(\mathbf{R}) = ikZ\mathsf{H}(\mathbf{R}), \qquad \mathbf{R} \in V, \tag{40}$$

$$\nabla \times \mathsf{H}(\mathbf{R}) = \mathsf{J}(\mathbf{R}) - ikY\mathsf{E}(\mathbf{R}), \quad \mathbf{R} \in V, \tag{41}$$

where $1/Y = Z = (\mu/\epsilon)^{\frac{1}{2}}$ is the characteristic impedance of the medium, the boundary conditions

$$\hat{\mathbf{n}} \times \mathsf{E}(\mathbf{R}_s) = 0, \quad \hat{\mathbf{n}} \cdot \mathsf{H}(\mathbf{R}_s) = 0, \tag{42}$$

and the Silver–Müller radiation condition. Specify two types of current density, namely

$$\mathsf{J}_e = -ik\mathsf{I}\delta(\mathbf{R} \mid \mathbf{R}') \tag{43}$$

and

$$\mathsf{J}_m = -Y\nabla \times [\mathsf{I}\delta(\mathbf{R} \mid \mathbf{R}')]. \tag{44}$$

The current distribution in (43) is that of three orthogonal harmonically oscillating infinitesimal electric dipoles situated at $\mathbf{R}'$, of dipole moment

$$\mathbf{P}_{ej} = \hat{\mathbf{a}}_j/c, \quad j = 1, 2, 3, \tag{45}$$

where c is the velocity of propagation in the exterior medium. Similarly, the current distribution (44) is that of three orthogonal harmonically oscillating infinitesimal magnetic dipoles at $\mathbf{R}'$, of dipole moment

$$\mathbf{P}_{mj} = -Y\hat{\mathbf{a}}_j, \quad j = 1, 2, 3. \tag{46}$$

Let E_e and H_e denote the fields due to the current distribution J_e [Eq. (43)], and expand them in powers of (ik):

$$\mathsf{E}_e = \sum_{n=0}^{\infty}(ik)^n\mathsf{E}_e^{(n)}, \quad \mathsf{H}_e = \sum_{n=0}^{\infty}(ik)^n\mathsf{H}_e^{(n)}. \tag{47}$$

Substituting these expressions together with (43) in (40)–(42) and equating like powers of ik yields

$$\nabla \times \mathsf{E}_e^{(0)} = 0, \tag{48}$$

$$\nabla \times \mathsf{E}_e^{(n)} = Z\mathsf{H}_e^{(n-1)}, \quad n > 0, \tag{49}$$

$$\nabla \times \mathsf{H}_e^{(0)} = 0, \tag{50}$$

$$\nabla \times \mathsf{H}_e^{(1)} = -\mathsf{I}\delta(\mathbf{R} \mid \mathbf{R}') - Y\mathsf{E}_e^{(0)}, \tag{51}$$

$$\nabla \times \mathsf{H}_e^{(n)} = -Y\mathsf{E}_e^{(n-1)}, \quad n > 1, \tag{52}$$

$$\hat{\mathbf{n}} \times \mathsf{E}_e^{(n)} = 0, \quad \hat{\mathbf{n}} \cdot \mathsf{H}_e^{(n)} = 0 \quad \text{on } S, \quad n \geq 0. \tag{53}$$

Similarly if E_m and H_m denote the fields due to the current distribution J_m [Eq. (44)] with expansions

$$\mathsf{E}_m = \sum_{n=0}^{\infty}(ik)^n\mathsf{E}_m^{(n)}, \quad \mathsf{H}_m = \sum_{n=0}^{\infty}(ik)^n\mathsf{H}_m^{(n)}, \tag{54}$$

then the same procedure leads to

$$\nabla \times \mathsf{E}_m^{(0)} = 0, \tag{55}$$

$$\nabla \times \mathsf{E}_m^{(n)} = Z\mathsf{H}_m^{(m-1)}, \quad n > 0, \tag{56}$$

$$\nabla \times \mathsf{H}_m^{(0)} = -Y\nabla \times [\mathsf{I}\delta(\mathbf{R} \mid \mathbf{R}')], \tag{57}$$

$$\nabla \times \mathsf{H}_m^{(n)} = -Y\mathsf{E}_m^{(n-1)}, \quad n > 0, \tag{58}$$

$$\hat{\mathbf{n}} \times \mathsf{E}_m^{(n)} = 0, \quad \hat{\mathbf{n}} \cdot \mathsf{H}_m^{(n)} = 0 \quad \text{on } S, \quad n \geq 0. \tag{59}$$

The terms $\mathsf{E}_m^{(1)}$ and $\mathsf{H}_e^{(1)}$ occurring in these equations, that is, the second terms in the expansions of the electric field due to the magnetic dipoles and the magnetic field due to the electric dipoles, are the fundamental dyadics to be used in Theorem 1. Explicit expressions for these terms, derived in Appendix B, are

$$\begin{aligned}\mathsf{E}_m^{(1)}(\mathbf{R} \mid \mathbf{R}') &= \nabla \times \left(-\frac{\mathsf{I}}{4\pi\,|\mathbf{R} - \mathbf{R}'|}\right) \\ &\quad + \sum_{j=1}^{3}\left(-\frac{1}{4\pi}\nabla \times \int_S \frac{\hat{\mathbf{a}}_j \cdot \nabla' N^{(e)}(\mathbf{R}_s \mid \mathbf{R}')}{|\mathbf{R}_s - \mathbf{R}|}\hat{\mathbf{n}}\,ds \right. \\ &\quad \left. + \nabla G_{mj}^{(e)}(\mathbf{R} \mid \mathbf{R}')\right)\hat{\mathbf{a}}_j,\end{aligned} \tag{60}$$

$$\begin{aligned}\mathsf{H}_e^{(1)}(\mathbf{R} \mid \mathbf{R}') &= \nabla \times \left(-\frac{\mathsf{I}}{4\pi\,|\mathbf{R} - \mathbf{R}'|}\right) \\ &\quad + \sum_{j=1}^{3}\left(-\frac{1}{4\pi}\nabla \times \int_S \frac{\hat{\mathbf{a}}_j \cdot \nabla' N_e^{(i)}(\mathbf{R}_s \mid \mathbf{R}')}{|\mathbf{R}_s - \mathbf{R}|}\hat{\mathbf{n}}\,ds \right. \\ &\quad \left. + \nabla N_{ej}^{(e)}(\mathbf{R} \mid \mathbf{R}')\right)\hat{\mathbf{a}}_j,\end{aligned} \tag{61}$$

where $N^{(e)}(\mathbf{R} \mid \mathbf{R}')$ is the exterior Green's function for the Neumann boundary conditions on S, i.e.,

$$N^{(e)}(\mathbf{R} \mid \mathbf{R}') = -\frac{1}{4\pi\,|\mathbf{R} - \mathbf{R}'|} + N_r^{(e)}(\mathbf{R} \mid \mathbf{R}'), \tag{62}$$

$$\nabla^2 N_r^{(e)} = 0, \quad \mathbf{R}, \mathbf{R}' \in V, \tag{63}$$

$$\frac{\partial N^{(e)}}{\partial n_s}(\mathbf{R}_s \mid \mathbf{R}') = 0, \tag{64}$$

$N^{(e)}$ regular at infinity in the sense of Kellogg[12]; $G_{mj}^{(e)}(\mathbf{R} \mid \mathbf{R}')$ are exterior Dirichlet potential functions for S, viz.,

$$\nabla^2 G_{mj}^{(e)}(\mathbf{R} \mid \mathbf{R}') = 0, \quad \mathbf{R}, \mathbf{R}' \in V, \quad j = 1, 2, 3, \tag{65}$$

$$\begin{aligned}\hat{\mathbf{n}} \times \nabla G_{mj}^{(e)}(\mathbf{R} \mid \mathbf{R}') &= \frac{1}{4\pi}\hat{\mathbf{n}} \times \nabla \times \left(\frac{\hat{\mathbf{a}}_j}{|\mathbf{R} - \mathbf{R}'|}\right. \\ &\quad \left. + \int_S \frac{\hat{\mathbf{a}}_j \cdot \nabla' N^{(e)}(\mathbf{R}_s \mid \mathbf{R}')}{|\mathbf{R}_s - \mathbf{R}|}\hat{\mathbf{n}}\,ds\right), \\ &\qquad \mathbf{R} \in S, \quad \mathbf{R}' \in V,\end{aligned} \tag{66}$$

$$\int_S \hat{\mathbf{n}} \cdot \nabla_s G_{mj}^{(e)}(\mathbf{R}_s \mid \mathbf{R}')\,ds = 0, \tag{67}$$

$G_{mj}^{(e)}(\mathbf{R} \mid \mathbf{R}')$ regular at infinity in the sense of Kellogg; $N_e^{(i)}(\mathbf{R} \mid \mathbf{R}')$ is an interior Neumann potential, viz.,

$$\nabla^2 N_e^{(i)}(\mathbf{R} \mid \mathbf{R}') = 0, \quad \mathbf{R} \in V_i, \quad \mathbf{R}' \in V, \tag{68}$$

$$\hat{\mathbf{n}}_s \cdot \nabla_s N_e^{(i)}(\mathbf{R}_s \mid \mathbf{R}') = -\hat{\mathbf{n}}_s \cdot \nabla_s \left(G^{(e)}(\mathbf{R}_s \mid \mathbf{R}') - \frac{U(\mathbf{R}_s) U(\mathbf{R}')}{4\pi C} \right); \tag{69}$$

$G^{(e)}(\mathbf{R} \mid \mathbf{R}')$ is the exterior Green's function for the Dirichlet boundary condition on S, i.e.,

$$G^{(e)}(\mathbf{R} \mid \mathbf{R}') = -(4\pi \lvert \mathbf{R} - \mathbf{R}' \rvert)^{-1} + G_r^{(e)}(\mathbf{R} \mid \mathbf{R}'), \tag{70}$$

$$\nabla^2 G_r^{(e)}(\mathbf{R} \mid \mathbf{R}') = 0, \quad \mathbf{R}, \mathbf{R}' \in V, \tag{71}$$

$$G^{(e)}(\mathbf{R}_s \mid \mathbf{R}') = 0, \quad \mathbf{R}_s \in S, \quad \mathbf{R}' \in V, \tag{72}$$

$G^{(e)}$ regular at infinity in the sense of Kellogg; $U(\mathbf{R})$ is the conductor potential for S, i.e.,

$$U(\mathbf{R}) = \int_S \frac{\partial}{\partial n_s} G^{(e)}(\mathbf{R}_s \mid \mathbf{R})\, ds, \quad \mathbf{R} \in V; \tag{73}$$

C is the capacity of S, i.e.,

$$C = \frac{1}{4\pi} \int_S \frac{\partial}{\partial n_s} U(\mathbf{R}_s)\, ds; \tag{74}$$

and the $N_{ej}^{(e)}(\mathbf{R} \mid \mathbf{R}')$ are exterior Neumann potential functions for S, viz.,

$$\nabla^2 N_{ej}^{(e)}(\mathbf{R} \mid \mathbf{R}') = 0, \quad \mathbf{R}, \mathbf{R}' \in V, \tag{75}$$

$$\hat{\mathbf{n}} \cdot \nabla N_{ej}^{(e)}(\mathbf{R} \mid \mathbf{R}') = \hat{\mathbf{n}} \cdot \nabla \times \left(\frac{\hat{\mathbf{a}}_j}{4\pi \lvert \mathbf{R} - \mathbf{R}' \rvert} + \frac{1}{4\pi} \int_S \frac{\hat{\mathbf{a}}_j \cdot \nabla' N_e^{(i)}(\mathbf{R}_s \mid \mathbf{R}')}{\lvert \mathbf{R}_s - \mathbf{R} \rvert} \hat{\mathbf{n}}\, ds \right), \quad \mathbf{R} \in S, \quad \mathbf{R}' \in V, \tag{76}$$

$N_{ej}^{(e)}(\mathbf{R} \mid \mathbf{R}')$ regular at infinity in the sense of Kellogg.

With $\mathsf{E}_m^{(1)}$ and $\mathsf{H}_e^{(1)}$ thus explicitly defined, it is straightforward to verify that (11)–(20) are satisfied; (11) and (16) are valid by inspection, (12) and (17) are valid by a direct calculation using the fact that the curl of a gradient vanishes, the boundary condition (13) is satisfied by virtue of (66) while the boundary condition (18) is seen to be fulfilled using (51), (B32), and (B35), and the conditions at infinity (14), (15), (19), and (20) are all fulfilled because both $\mathsf{E}_m^{(1)}$ and $\mathsf{H}_e^{(1)}$ behave at infinity as $[\mathbf{A}(\theta, \phi)/R^2] + O(1/R^3)$.

IV. INTEGRAL REPRESENTATIONS OF THE ELECTROMAGNETIC SCATTERED FIELDS

Using the same definitions and notation introduced previously, we now direct our attention to the scattering of a time harmonic monochromatic incident electromagnetic field by a perfectly conducting surface S. If $\mathbf{E}^i(\mathbf{R})$ and $\mathbf{H}^i(\mathbf{R})$ denote the incident electric and magnetic fields, respectively, the problem is one of determining the scattered fields $\mathbf{E}^s(\mathbf{R})$ and $\mathbf{H}^s(\mathbf{R})$ such that

$$\nabla \times \mathbf{E}^s(\mathbf{R}) = ikZ\mathbf{H}^s(\mathbf{R}), \quad \nabla \times \mathbf{H}^s(\mathbf{R}) = -ikY\mathbf{E}^s(\mathbf{R}), \quad \mathbf{R} \in V, \tag{77}$$

$$\hat{\mathbf{n}} \times \mathbf{E}^s(\mathbf{R}) = -\hat{\mathbf{n}} \times \mathbf{E}^i(\mathbf{R}), \quad \hat{\mathbf{n}} \cdot \mathbf{H}^s(\mathbf{R}) = -\hat{\mathbf{n}} \cdot \mathbf{H}^i(\mathbf{R}), \quad \mathbf{R} \in S, \tag{78}$$

and

$$\lim_{R \to \infty} R \left[\hat{\mathbf{R}} \times \nabla \times \begin{pmatrix} \mathbf{E}^s \\ \mathbf{H}^s \end{pmatrix} + ik \begin{pmatrix} \mathbf{E}^s \\ \mathbf{H}^s \end{pmatrix} \right] = 0, \quad \text{uniformly in } \hat{\mathbf{R}}. \tag{79}$$

Recall that $\hat{\mathbf{n}}$ is directed from the surface S into its interior V_i, away from the exterior V.

In addition to the Silver–Müller radiation condition (79), the scattered fields $\mathbf{E}^s$ and $\mathbf{H}^s$ satisfy the conditions (9) and (10), namely

$$\lvert \hat{\mathbf{R}} \times \mathbf{F} \rvert < \infty \quad \text{and} \quad \lvert R\nabla \times \mathbf{F} \rvert < \infty, \quad \text{as} \quad R \to \infty.$$

This follows from an expansion theorem due to Wilcox[9] which asserts: If $\mathbf{F}(\mathbf{R})$ is a vector radiation function [satisfies Maxwell's equations (77) and the radiation condition (79)] in an exterior region $R > c$, then $\mathbf{F}(\mathbf{R})$ has an expansion

$$\mathbf{F}(\mathbf{R}) = \frac{e^{ikR}}{R} \sum_{n=0}^{\infty} \frac{\mathbf{F}_n(\theta, \phi)}{R^n}, \tag{80}$$

valid for $R > c$, which converges absolutely and uniformly in the parameters R, θ, and ϕ provided that $R \geq c + \epsilon > c$. Furthermore, the series can be differentiated term by term any number of times with respect to R, θ, and ϕ, and the resulting series all converge absolutely and uniformly.

The scattered fields are thus regular vector valued functions suitable for use in Theorem 1. Letting $\mathbf{F}$ be $\mathbf{E}^s$ in Theorem 1(a), (22), and $\mathbf{H}^s$ in Theorem 1(b), (23), and making use of Maxwell's equations (77) and the boundary conditions (78), we obtain the following integral equations:

$$ikZ\mathbf{H}^s(\mathbf{R}') = -k^2 \int_V \mathbf{E}^s(\mathbf{R}) \cdot \mathsf{E}_m^{(1)}(\mathbf{R} \mid \mathbf{R}')\, dv - \int_S [\hat{\mathbf{n}} \times \mathbf{E}^i(\mathbf{R}_s)] \cdot [\nabla_s \times \mathsf{E}_m^{(1)}(\mathbf{R}_s \mid \mathbf{R}')]\, ds, \tag{81}$$

$$\mathbf{E}^s(\mathbf{R}') = -ikZ \int_V \mathbf{H}^s(\mathbf{R}) \cdot \mathsf{H}_e^{(1)}(\mathbf{R} \mid \mathbf{R}')\, dv - \int_S [\hat{\mathbf{n}} \times \mathbf{E}^i(\mathbf{R}_s)] \cdot \mathsf{H}_e^{(1)}(\mathbf{R}_s \mid \mathbf{R}')\, ds. \tag{82}$$

The first of these equations may be written more conveniently by taking into consideration the explicit form of $\mathsf{E}_m^{(1)}$. It is easily shown, by using (B1), (B11), and the definition of the exterior Neumann Green's function (62), that

$$\nabla \times \mathsf{E}_m^{(1)}(\mathbf{R} \mid \mathbf{R}') = -\nabla\nabla' N^{(e)}(\mathbf{R} \mid \mathbf{R}') - \mathsf{I}\delta(\mathbf{R} \mid \mathbf{R}'). \tag{83}$$

For $\mathbf{R} \in S$ and $\mathbf{R}' \in V$ the δ function does not contribute; hence this result, together with the identities (A1) and (A5), allows the integrand of the surface integral in (81) to be written

$$\begin{aligned}[\hat{\mathbf{n}} \times \mathbf{E}^i(\mathbf{R}_s)] \cdot \nabla_s \times \mathsf{E}_m^{(1)}(\mathbf{R}_s \mid \mathbf{R}') &= \hat{\mathbf{n}} \cdot \{[\nabla_s \times \mathbf{E}^i(\mathbf{R}_s)]\nabla' N^{(e)}(\mathbf{R}_s \mid \mathbf{R}') \\ &\quad - \nabla_s \times [\mathbf{E}^i(\mathbf{R}_s)\nabla' N^{(e)}(\mathbf{R}_s \mid \mathbf{R}')]\}. \end{aligned} \tag{84}$$

The $\hat{\mathbf{n}} \cdot \nabla_s \times$ term in this expression vanishes by Stokes' theorem when integrated over the closed surface S. Moreover, the incident field $\mathbf{E}^i$ in the remaining term must satisfy Maxwell's equation $\nabla \times \mathbf{E}^i = ikZ\mathbf{H}^i$. The surface integral in (81) can then be written

$$\int_S (\hat{\mathbf{n}} \times \mathbf{E}^i) \cdot (\nabla \times \mathsf{E}_m^{(1)})\, ds = -ikZ \int_S \hat{\mathbf{n}} \cdot \mathbf{H}^i \nabla' N^{(e)}\, ds. \tag{85}$$

Utilizing this result in (81) leads to the simplification incorporated in the following theorem.

Theorem 2: If $\mathbf{E}^s$ and $\mathbf{H}^s$ are electromagnetic fields scattered by S when illuminated by $\mathbf{E}^i$ and $\mathbf{H}^i$, i.e., if $\mathbf{E}^s$ and $\mathbf{H}^s$ satisfy (77)–(79), then

$$\begin{aligned}\mathbf{H}^s(\mathbf{R}') &= ikY \int_V \mathbf{E}^s(\mathbf{R}) \cdot \mathsf{E}_m^{(1)}(\mathbf{R} \mid \mathbf{R}')\, dv \\ &\quad + \nabla' \int_S [\hat{\mathbf{n}} \cdot \mathbf{H}^i(\mathbf{R}_s)] N^{(e)}(\mathbf{R}_s \mid \mathbf{R}')\, ds, \end{aligned} \tag{86}$$

$$\begin{aligned}\mathbf{E}^s(\mathbf{R}') &= -ikZ \int_V \mathbf{H}^s(\mathbf{R}) \cdot \mathsf{H}_e^{(1)}(\mathbf{R} \mid \mathbf{R}')\, dv \\ &\quad - \int_S [\hat{\mathbf{n}} \times \mathbf{E}^i(\mathbf{R}_s)] \cdot \mathsf{H}_e^{(1)}(\mathbf{R}_s \mid \mathbf{R}')\, ds, \end{aligned} \tag{87}$$

where $\mathsf{E}_m^{(1)}$ and $\mathsf{H}_e^{(1)}$ are the fundamental dyadics (60) and (61) and $N^{(e)}$ is the exterior Neumann Green's function for Laplace's equation [(62)–(64)].

At this point one might be tempted to solve this coupled system of integral equations for small k by iteration, using the surface integral terms, which do not have k as a factor, as the zeroth-order iterates. Such a procedure will, unfortunately, be unsuccessful. The reason for this lies in the fact that neither the surface integral term nor any of the iterates will contain $e^{ikR'}$ as a factor. However, Wilcox's theorem (80) makes it clear that the scattered fields should contain this factor. This leads to the conclusion that partial sums of the Taylor series expansion of $e^{ikR'}$ appear in the iterates. As the iteration proceeds, positive powers of R will appear in the volume integrals and these integrals will diverge. This, in fact, is the cause of the breakdown of Stevenson's special method after three terms in the expansion are found.[3b]

To avoid this difficulty, the exponential e^{ikR} is removed by introducing the following vector functions:

$$\mathbf{e}(\mathbf{R}) = e^{-ikR}\mathbf{E}^s(\mathbf{R}), \quad \mathbf{h}(\mathbf{R}) = e^{-ikR}\mathbf{H}^s(\mathbf{R}). \tag{88}$$

The motivation for doing so lies in Wilcox's expansion theorem. From (80) it is seen that, at least in the region where the expansion is valid, the new fields $\mathbf{e}$ and $\mathbf{h}$ do not contain the troublesome factor e^{ikR}. Furthermore, these fields are regular in the exterior region V and hence may be represented by using Theorem 1. This same device proved successful in an analogous treatment of scalar scattering problems at low frequencies.[14,15] Identifying $\mathbf{F}$ with $\mathbf{e}$ in Theorem 1(a) and with $\mathbf{h}$ in Theorem 1(b) leads to

$$\begin{aligned}\nabla' \times \mathbf{e}(\mathbf{R}') &= -\int_V [\nabla \times \nabla \times \mathbf{e}(\mathbf{R})] \cdot \mathsf{E}_m^{(1)}(\mathbf{R} \mid \mathbf{R}')\, dv \\ &\quad + \int_S [\hat{\mathbf{n}} \times \mathbf{e}(\mathbf{R})] \cdot \nabla \times \mathsf{E}_m^{(1)}(\mathbf{R} \mid \mathbf{R}')\, ds \end{aligned} \tag{89}$$

and

$$\begin{aligned}\nabla' \times \mathbf{h}(\mathbf{R}') &= -\int_V [\nabla \times \nabla \times \mathbf{h}(\mathbf{R})] \cdot \mathsf{H}_e^{(1)}(\mathbf{R} \mid \mathbf{R}')\, dv \\ &\quad + \int_S \{\hat{\mathbf{n}} \times [\nabla \times \mathbf{h}(\mathbf{R})]\} \cdot \mathsf{H}_e^{(1)}(\mathbf{R} \mid \mathbf{R}')\, ds. \end{aligned} \tag{90}$$

With the definition (88) of $\mathbf{e}$ and $\mathbf{h}$ and Maxwell's equations, it follows that

$$\nabla \times \mathbf{e} = ik(Z\mathbf{h} - \hat{\mathbf{R}} \times \mathbf{e}), \tag{91}$$

$$\nabla \times \mathbf{h} = -ik(Y\mathbf{e} + \hat{\mathbf{R}} \times \mathbf{h}), \tag{92}$$

and

$$\nabla \times \nabla \times \mathbf{e} = k^2(\mathbf{e} + Z\hat{\mathbf{R}} \times \mathbf{h}) - ik\nabla \times (\hat{\mathbf{R}} \times \mathbf{e}), \tag{93}$$

$$\nabla \times \nabla \times \mathbf{h} = k^2(\mathbf{h} - Y\hat{\mathbf{R}} \times \mathbf{e}) - ik\nabla \times (\hat{\mathbf{R}} \times \mathbf{h}). \tag{94}$$

With these results for $\mathbf{e}$ and $\mathbf{h}$, the following theorem may be established.

Theorem 3: If $\mathbf{E}^s$ and $\mathbf{H}^s$ are the electromagnetic fields scattered by the perfectly conducting surface S when illuminated by $\mathbf{E}^i$ and $\mathbf{H}^i$, i.e., $\mathbf{E}^s$ and $\mathbf{H}^s$ satisfy

(77)–(79), and $\mathbf{e} = e^{-ikR}\mathbf{E}^s$, $\mathbf{h} = e^{-ikR}\mathbf{H}^s$, then

$$\mathbf{h}(\mathbf{R}') = ik\int_V \{[Y\mathbf{e}(\mathbf{R}) + \hat{\mathbf{R}} \times \mathbf{h}(\mathbf{R})] \cdot \mathsf{E}_m^{(1)}(\mathbf{R} \mid \mathbf{R}') - \hat{\mathbf{R}} \cdot \mathbf{h}(\mathbf{R})\nabla' N^{(e)}(\mathbf{R} \mid \mathbf{R}')\}\, dv - \nabla' \int_S \hat{\mathbf{n}} \cdot \mathbf{h}(\mathbf{R}_s) N^{(e)}(\mathbf{R}_s \mid \mathbf{R}')\, ds \tag{95}$$

and

$$\mathbf{e}(\mathbf{R}') = -ik\int_V \Big[[Z\mathbf{h}(\mathbf{R}) - \hat{\mathbf{R}} \times \mathbf{e}(\mathbf{R})] \cdot \mathsf{H}_e^{(1)}(\mathbf{R} \mid \mathbf{R}') + [\hat{\mathbf{R}} \cdot \mathbf{e}(\mathbf{R})]\nabla' \Big(G^{(e)}(\mathbf{R} \mid \mathbf{R}') + \frac{U(\mathbf{R}') - U(\mathbf{R})U(\mathbf{R}')}{4\pi C} \Big) \Big]\, dv + \int_S [\hat{\mathbf{n}} \times \mathbf{e}(\mathbf{R}_s)] \cdot \mathsf{H}_e^{(1)}(\mathbf{R}_s \mid \mathbf{R}')\, ds, \tag{96}$$

where $\mathsf{E}_m^{(1)}$ and $\mathsf{H}_e^{(1)}$ are the fundamental dyadics (60) and (61), $G^{(e)}$ and $N^{(e)}$ are the exterior potential Green's functions for Dirichlet and Neumann conditions, respectively, (70)–(72) and (62)–(64), U is the conductor potential (73), and C is the electrostatic capacity of S [(74)].

The proof of this theorem rests on straightforward manipulation of Eqs. (89) and (90), the highlights of which are indicated in Appendix C.

These integral equations are in a form which permits solution for the unknown fields $\mathbf{e}$ and $\mathbf{h}$, provided that k is sufficiently small. There are two slightly different procedures, both of which depend on the fact that the surface integral terms do not involve the unknown fields since the boundary conditions (78) imply

$$\hat{\mathbf{n}} \cdot \mathbf{h}(\mathbf{R}_s) = -e^{-ikR_s}\hat{\mathbf{n}} \cdot \mathbf{H}^i(\mathbf{R}_s), \tag{97}$$

$$\hat{\mathbf{n}} \times \mathbf{e}(\mathbf{R}_s) = -e^{-ikR_s}\hat{\mathbf{n}} \times \mathbf{E}^i(\mathbf{R}_s). \tag{98}$$

One method of solution involves iteration of these coupled Fredholm equations using the known surface integral terms as the zeroth-order iterates.

The second method leads more directly to the low-frequency expansion. It involves expanding the known terms as well as the unknown fields in powers of k, then substituting these expressions in the integral equations, equating like powers of k, and obtaining recursion relations for the coefficients in the expansions. Explicitly, write

$$\mathbf{e} = \sum_{n=0}^{\infty} (ik)^n \mathbf{e}_n, \quad \mathbf{h} = \sum_{n=0}^{\infty} (ik)^n \mathbf{h}_n, \tag{99}$$

$$-\nabla' \int_S \hat{\mathbf{n}} \cdot \mathbf{h} N^e\, ds = \nabla' \int_S e^{-ikR_s}\hat{\mathbf{n}} \cdot \mathbf{H}^i(\mathbf{R}_s) N^e(\mathbf{R}_s \mid \mathbf{R}')\, ds = \sum_{n=0}^{\infty} (ik)^n \mathbf{f}_n(\mathbf{R}'), \tag{100}$$

$$\int_S (\hat{\mathbf{n}} \times \mathbf{e}) \cdot \mathsf{H}_e^{(n)}\, ds = -\int_S e^{-ikR_s}[\hat{\mathbf{n}} \times \mathbf{E}^i(\mathbf{R}_s)] \cdot \mathsf{H}_e^{(1)}(\mathbf{R}_s \mid \mathbf{R}')\, ds = \sum_{n=0}^{\infty} (ik)^n \mathbf{g}_n(\mathbf{R}'). \tag{101}$$

That $\mathbf{e}$ and $\mathbf{h}$ may be written in this form follows from the definition (88) and the work of Werner,[5,6] who showed that the electric and magnetic scattered fields tend analytically to corresponding electrostatic and magnetostatic fields as $k \to 0$. The power series representations of the surface integral terms follow from the analyticity in k of the incident fields whether they are dipoles or plane waves. Substituting (99)–(101) in the integral equations of Theorem 3 and equating like powers of ik yields the following recursion formulas:

$$\mathbf{h}_0(\mathbf{R}') = f_0(\mathbf{R}'), \tag{102}$$

$$\mathbf{h}_{n+1}(\mathbf{R}') = \int_V \{[Y\mathbf{e}_n(\mathbf{R}) + \hat{\mathbf{R}} \times \mathbf{h}_n(\mathbf{R})] \cdot \mathsf{E}_m^{(1)}(\mathbf{R} \mid \mathbf{R})' - \hat{\mathbf{R}} \cdot \mathbf{h}_n(\mathbf{R})\nabla' N^{(e)}(\mathbf{R} \mid \mathbf{R}')\}\, dv + \mathbf{f}_{n+1}(\mathbf{R}'), \tag{103}$$

$$\mathbf{e}_0(\mathbf{R}') = \mathbf{g}_0(\mathbf{R}'), \tag{104}$$

$$\mathbf{e}_{n+1}(\mathbf{R}') = -\int_V \Big[[Z\mathbf{h}_n(\mathbf{R}) - \hat{\mathbf{R}} \times \mathbf{e}_n(\mathbf{R})] \cdot \mathsf{H}_e^{(1)}(\mathbf{R} \mid \mathbf{R}') + [\hat{\mathbf{R}} \cdot \mathbf{e}_n(\mathbf{R})]\nabla' \Big(G^{(e)}(\mathbf{R} \mid \mathbf{R}') + \frac{U(\mathbf{R}') - U(\mathbf{R})U(\mathbf{R}')}{4\pi C} \Big) \Big]\, dv + \mathbf{g}_{n+1}(\mathbf{R}'). \tag{105}$$

V. SCATTERING OF A PLANE WAVE BY A PERFECTLY CONDUCTING SPHERE

The results of the previous section are applied here to the problem of scattering of a plane wave by a perfectly conducting sphere. The sphere is of radius a, and its center coincides with the origin of a rectangular coordinate system (x, y, z). According to the notation of the previous sections, V_i denotes the volume of the sphere, V the rest of space, and S the surface of the sphere. The unit normal $\hat{\mathbf{n}}$ is directed away from V and into V_i. The plane wave propagates in the direction of the negative z axis with its electric vector polarized along the positive x axis. A spherical coordinate system (R, θ, ϕ) will be used along with the rectangular coordinate system (x, y, z).

The starting point in the problem is to find the explicit forms of the dyadics $\mathsf{H}_e^{(1)}$ and $\mathsf{E}_m^{(1)}$. The derivation of these dyadics involves frequent use of the Dirichlet and Neumann static Green's functions, and these are given below.

The expansion of the free-space static Green's function in spherical harmonics is

$$-\frac{1}{4\pi|\mathbf{R}-\mathbf{R}'|} = -\frac{1}{4\pi}\sum_{n=0}^{\infty}\sum_{m=0}^{n}\epsilon_m\frac{(n-m)!}{(n+m)!}P_n^m(\cos\theta')P_n^m(\cos\theta) \times \cos m(\phi-\phi')\frac{R_<^n}{R_>^{n+1}}, \quad (106)$$

where $R_< = \min(R, R')$, $R_> = \max(R, R')$, and ϵ_m is the Neumann factor: $\epsilon_0 = 1$, $\epsilon_m = 2$ for $m = 1, 2, \cdots$. The functions P_n^m are the associated Legendre functions defined by

$$P_n^m(x) = \frac{(-1)^m}{2^m}\frac{(n+m)!}{m!(n-m)!}(1-x^2)^{m/2} \times {}_2F_1\left(1+m+n, m-n; 1+m; \frac{1-x}{2}\right), \quad -1 \le x \le 1. \quad (107)$$

This definition is according to Magnus *et al.*[16] and all the contiguous relations for these functions that will be used subsequently can be found there (Ref. 16, p. 171). The regular part of the exterior static Dirichlet Green's function [Eqs. (70)–(72)] for the sphere is given by

$$G_r^{(e)}(\mathbf{R}\mid\mathbf{R}') = \frac{1}{4\pi}\sum_{n=0}^{\infty}\sum_{m=0}^{n}\epsilon_m\frac{(n-m)!}{(n+m)!}P_n^m(\cos\theta')P_n^m(\cos\theta) \times \cos m(\phi-\phi')\frac{a^{2n+1}}{(RR')^{n+1}}, \quad (108)$$

while the regular part of the corresponding Neumann Green's function [Eqs. (62)–(64)] is given by

$$N_r^{(e)}(\mathbf{R}\mid\mathbf{R}') = -\frac{1}{4\pi}\sum_{n=0}^{\infty}\sum_{m=0}^{n}\epsilon_m\frac{n}{n+1}\frac{(n-m)!}{(n+m)!}p_n^m(\cos\theta') \times p_n^m(\cos\theta)\cos m(\phi-\phi')\frac{a^{2n+1}}{(RR')^{n+1}}. \quad (109)$$

The conductor potential for the sphere (73) is

$$U(\mathbf{R}) = \int_S \frac{\partial}{\partial n_s}G^{(e)}(\mathbf{R}_s\mid\mathbf{R})\,ds = -\int_S\frac{\partial}{\partial R_s}G^{(e)}(\mathbf{R}_s\mid\mathbf{R})\,ds = \frac{a}{R}, \quad (110)$$

and the electrostatic capacity (74) is

$$C = \frac{1}{4\pi}\int_S\frac{\partial}{\partial n_s}U(\mathbf{R}_s)\,ds = \frac{a}{4\pi}\int_0^{2\pi}d\phi\int_0^{\pi}d\theta = a. \quad (111)$$

To complete the explicit calculation of the fundamental dyadics, the following intermediate results are useful:

The integral of the exterior Neumann function in (60) is

$$-\frac{1}{4\pi}\nabla\times\int_S\frac{\hat{\mathbf{a}}_j\cdot\nabla'N^{(e)}(\mathbf{R}_s\mid\mathbf{R}')}{|\mathbf{R}_s-\mathbf{R}|}\hat{\mathbf{n}}_s\,ds = -\frac{1}{4\pi}\hat{\mathbf{a}}_j\cdot\nabla'\nabla\times\hat{\mathbf{R}}\sum_{n=1}^{\infty}\sum_{m=0}^{n}\frac{\epsilon_m(n-m)!\,a^{2n+1}}{(n+1)(n+m)!(RR')^{n+1}} \times p_n^m(\cos\theta)p_n^m(\cos\theta')\cos m(\phi-\phi')$$
$$= -\hat{\mathbf{a}}_j\cdot\nabla'\nabla\times\left(G_r^{(e)}(\mathbf{R}\mid\mathbf{R}') - \frac{U(\mathbf{R})U(\mathbf{R}')}{4\pi C} + N_r^{(e)}(\mathbf{R}\mid\mathbf{R}')\right)\hat{\mathbf{R}}, \quad (112)$$

the interior Neumann function [(68) and (69)] is

$$N_e^{(i)}(\mathbf{R}\mid\mathbf{R}') = \frac{1}{4\pi}\sum_{n=1}^{\infty}\sum_{m=0}^{n}\epsilon_m\frac{(2n+1)}{n}\frac{(n-m)!}{(n+m)!} \times P_n^m(\cos\theta)P_n^m(\cos\theta') \times \cos m(\phi-\phi')\frac{R^n}{R'^{n+1}} + \nu_e(\mathbf{R}'), \quad (113)$$

the integral of the interior Neumann function appearing in (61) is

$$-\frac{1}{4\pi}\nabla\times\int_S\frac{\hat{\mathbf{a}}_j\cdot\nabla'N_e^{(i)}(\mathbf{R}_s\mid\mathbf{R}')}{|\mathbf{R}_s-\mathbf{R}|}\hat{\mathbf{n}}_s\,ds = \frac{1}{4\pi}\hat{\mathbf{a}}_j\cdot\nabla'\nabla\times\hat{\mathbf{R}}\sum_{n=1}^{\infty}\sum_{m=0}^{n} \times\frac{\epsilon_m(n-m)!\,a^{2n+1}P_n^m(\cos\theta)P_n^m(\cos\theta')\cos m(\phi-\phi')}{n(n+m)!(RR')^{n+1}}, \quad (114)$$

the functions $G_{mj}^{(e)}$ [(65)–(67)] are

$$G_{m1}^{(e)}(\mathbf{R}\mid\mathbf{R}') = -\frac{1}{4\pi}\sum_{n=1}^{\infty}\frac{a^{2n+1}}{n(RR')^{n+1}} \times\left(\sum_{m=0}^{n-1}\frac{(n-m)!}{(n+m)!}P_n^m(\cos\theta') \times P_n^{m+1}(\cos\theta)\sin[(m+1)\phi-m\phi'] + \sum_{m=1}^{n}\frac{(n-m+1)!}{(n+m-1)!}P_n^m(\cos\theta') \times P_n^{m-1}(\cos\theta)\sin[(m-1)\phi-m\phi']\right), \quad (115)$$

$$G_{m2}^{(e)}(\mathbf{R}\,|\,\mathbf{R}') = \frac{1}{4\pi}\sum_{n=1}^{\infty}\frac{a^{2n+1}}{n(RR')^{n+1}} \times \left(\sum_{m=0}^{n-1}\frac{(n-m)!}{(n+m)!}P_n^m(\cos\theta') \times p_n^{m+1}(\cos\theta)\cos[(m+1)\phi - m\phi'] - \sum_{m=1}^{n}\frac{(n-m+1)!}{(n+m-1)!}P_n^m(\cos\theta') \times P_n^{m-1}(\cos\theta)\cos[(m-1)\phi - m\phi']\right), \quad (116)$$

$$G_{m3}^{(e)}(\mathbf{R}\,|\,\mathbf{R}') = -\frac{1}{2\pi}\sum_{n=1}^{\infty}\sum_{m=1}^{n}\frac{m(n-m)!}{n(n+m)!}P_n^m(\cos\theta') \times P_n^m(\cos\theta)\sin m(\phi-\phi')\frac{a^{2n+1}}{(RR')^{n+1}}, \quad (117)$$

and the functions $N_{ej}^{(e)}$ [(75) and (76)] are

$$N_{e1}^{(e)}(\mathbf{R}\,|\,\mathbf{R}') = \frac{1}{4\pi}\sum_{n=1}^{\infty}\frac{a^{2n+1}}{(n+1)(RR')^{n+1}} \times \left(\sum_{m=0}^{n-1}\frac{(n-m)!}{(n+m)!}P_n^m(\cos\theta') \times P_n^{m+1}(\cos\theta)\sin[(m+1)\phi - m\phi'] + \sum_{m=1}^{n}\frac{(n-m+1)!}{(n+m-1)!}P_n^m(\cos\theta') \times P_n^{m-1}(\cos\theta)\sin[(m-1)\phi - m\phi']\right), \quad (118)$$

$$N_{e2}^{(e)}(\mathbf{R}\,|\,\mathbf{R}') = -\frac{1}{4\pi}\sum_{n=1}^{\infty}\frac{a^{2n+1}}{(n+1)(RR')^{n+1}} \times \left(\sum_{m=0}^{n-1}\frac{(n-m)!}{(n+m)!}P_n^m(\cos\theta') \times p_n^{m+1}(\cos\theta)\cos[(m+1)\phi - m\phi'] - \sum_{m=1}^{n}\frac{(n-m+1)!}{(n+m-1)!}p_n^m(\cos\theta') \times p_n^{m-1}(\cos\theta)\cos[(m-1)\phi - m\phi']\right), \quad (119)$$

$$N_{e3}^{(e)}(\mathbf{R}\,|\,\mathbf{R}') = \frac{1}{2\pi}\sum_{n=1}^{\infty}\sum_{m=1}^{n}\frac{m}{n+1}\frac{(n-m)!}{(n+m)!}P_n^m(\cos\theta') \times P_n^m(\cos\theta)\sin m(\phi-\phi')\frac{a^{2n+1}}{(RR')^{n+1}}. \quad (120)$$

The derivation of these results is straightforward though tedious, with repeated use of the orthogonality of the spherical harmonics and the contiguous function relations mentioned previously. These results complete the definitions of $\mathsf{E}_m^{(1)}$ and $\mathsf{H}_e^{(1)}$ [(60) and (61)].

As mentioned at the beginning of this section, the incident field is a plane wave propagating along the negative z axis with its electric field polarized along the positive x axis. Explicitly,

$$\mathbf{E}^i = \hat{\mathbf{a}}_1 e^{-ikz}, \quad \mathbf{H}^i = -\hat{\mathbf{a}}_2 Y e^{-ikz}. \quad (121)$$

The scattered fields $\mathbf{E}^s$ and $\mathbf{H}^s$ satisfy (77)–(79) and are related to $\mathbf{e}$ and $\mathbf{h}$ by (88). To determine the coefficients in the expansions of $\mathbf{e}$ and $\mathbf{h}$ [Eq. (99)], it is necessary first to determine the functions $\mathbf{f}_n$ and $\mathbf{g}_n$ [(100) and (101)]. With the incident field (121),

$$\mathbf{f}_n(\mathbf{R}) = \frac{Y(-a)^n}{n!}\nabla\int_S \hat{\mathbf{R}}_s\cdot\hat{\mathbf{a}}_2(1+\cos\theta_s)^n N^{(e)}(\mathbf{R}_s\,|\,\mathbf{R})\,ds, \quad (122)$$

$$\mathbf{g}_n(\mathbf{R}) = \frac{(-a)^n}{n!}\int_S(1+\cos\theta_s)^n\hat{\mathbf{R}}_s\times\hat{\mathbf{a}}_1\cdot\mathsf{H}_e^{(1)}(\mathbf{R}_s\,|\,\mathbf{R})\,ds. \quad (123)$$

Expressions for $N^{(e)}$ and $\mathbf{H}_e^{(1)}$ have been given previously. Combining all these results in (102)–(105), the first few terms in the expansions of $\mathbf{e}$ and $\mathbf{h}$, after laborious but straightforward calculation, are found to be

$$\mathbf{e}_0(\mathbf{R}) = a^3\nabla[P_1^1(\cos\theta)\cos\phi/R^2] = (a^3/R^3)(\hat{\mathbf{R}}2\sin\theta\cos\phi - \hat{\boldsymbol{\theta}}\cos\theta\cos\phi + \hat{\boldsymbol{\phi}}\sin\phi), \quad (124)$$

$$\mathbf{h}_0(\mathbf{R}) = \tfrac{1}{2}Ya^3\nabla[P_1^1(\cos\theta)\sin\phi/R^2] = (Ya^3/R^3)(\hat{\mathbf{R}}\sin\theta\sin\phi - \hat{\boldsymbol{\theta}}\tfrac{1}{2}\cos\theta\sin\phi - \hat{\boldsymbol{\phi}}\tfrac{1}{2}\cos\phi), \quad (125)$$

$$\mathbf{e}_1(\mathbf{R}) = (a^3/R^2)[-\hat{\mathbf{R}}2\sin\theta\cos\phi + \hat{\boldsymbol{\theta}}(\tfrac{1}{2}+\cos\theta)\cos\phi - \hat{\boldsymbol{\phi}}(1+\tfrac{1}{2}\cos\theta)\sin\phi] - \tfrac{1}{6}a^5\nabla[P_2^1(\cos\theta)\cos\phi(1/R^3)], \quad (126)$$

$$\mathbf{h}_1(\mathbf{R}) = (Ya^3/R^2)[-\hat{\mathbf{R}}\sin\theta\sin\phi + \hat{\boldsymbol{\theta}}(1+\tfrac{1}{2}\cos\theta)\sin\phi + \hat{\boldsymbol{\phi}}(\tfrac{1}{2}+\cos\theta)\cos\phi] - \tfrac{1}{9}Ya^5\nabla[P_2^1(\cos\theta)\sin\phi(1/R^3)], \quad (127)$$

$$\mathbf{e}_2(\mathbf{R}) = (a^3/R)[-\hat{\boldsymbol{\theta}}(\tfrac{1}{2}+\cos\theta)\cos\phi + \hat{\boldsymbol{\phi}}(1+\tfrac{1}{2}\cos\theta)\sin\phi] + (a^5/R^3)[\hat{\mathbf{R}}(-\tfrac{3}{5}\sin\theta + \tfrac{3}{4}\sin 2\theta)\cos\phi + \hat{\boldsymbol{\theta}}(\tfrac{2}{15}\cos\theta - \tfrac{1}{2}\cos 2\theta)\cos\phi + \hat{\boldsymbol{\phi}}(-\tfrac{3}{10} + \tfrac{1}{2}\cos\theta + \tfrac{1}{6}\cos 2\theta)\sin\phi] + \tfrac{1}{45}a^7\nabla[P_3^1(\cos\theta)\cos\phi(1/R^4)], \quad (128)$$

$$\mathbf{h}_2(\mathbf{R}) = -(Ya^3/R)[\hat{\boldsymbol{\theta}}(1+\tfrac{1}{2}\cos\theta)\sin\phi + \hat{\boldsymbol{\phi}}(\tfrac{1}{2}+\cos\theta)\cos\phi] + (Ya^5/R^3)[\hat{\mathbf{R}}(\tfrac{3}{5}\sin\theta + \tfrac{1}{2}\sin 2\theta)\sin\phi + \hat{\boldsymbol{\theta}}(-\tfrac{11}{20}\cos\theta - \tfrac{1}{3}\cos 2\theta)\sin\phi + \hat{\boldsymbol{\phi}}(-\tfrac{3}{10} - \tfrac{1}{3}\cos\theta - \tfrac{1}{4}\cos 2\theta)\cos\phi] + \tfrac{1}{60}Ya^7\nabla[P_3^1(\cos\theta)\sin\phi(1/R^4)], \quad (129)$$

$$\begin{aligned}\mathbf{e}_3(\mathbf{R}) = {} & (a^5/R^2)[\hat{\mathbf{R}}(\tfrac{3}{5}\sin\theta - \tfrac{1}{4}\sin 2\theta)\cos\phi \\ & + \hat{\boldsymbol{\theta}}(\tfrac{3}{10} - \tfrac{2}{15}\cos\theta + \tfrac{1}{4}\cos 2\theta)\cos\phi \\ & + \hat{\boldsymbol{\phi}}(\tfrac{3}{10} - \tfrac{11}{20}\cos\theta - \tfrac{1}{6}\cos 2\theta)\sin\phi] \\ & - \tfrac{2}{3}a^6\nabla[P_1^1(\cos\theta)\cos\phi(1/R^2)] \\ & + (a^7/R^4)[\hat{\mathbf{R}}(-\tfrac{1}{30}\sin\theta \\ & - \tfrac{5}{56}\sin 2\theta - \tfrac{1}{6}\sin 3\theta)\cos\phi \\ & + \hat{\boldsymbol{\theta}}(\tfrac{1}{80} + \tfrac{1}{120}\cos\theta \\ & + \tfrac{9}{112}\cos 2\theta + \tfrac{1}{8}\cos 3\theta)\cos\phi \\ & + \hat{\boldsymbol{\phi}}(-\tfrac{1}{20} - \tfrac{69}{1120}\cos\theta \\ & - \tfrac{1}{12}\cos 2\theta - \tfrac{1}{32}\cos 3\theta)\sin\phi] \\ & - \tfrac{1}{420}a^9\nabla[P_4^1(\cos\theta)\cos\phi(1/R^5)]. \end{aligned} \tag{130}$$

The scattered fields themselves are

$$\mathbf{E}^s(\mathbf{R}) = e^{ikR}\left(\sum_{n=0}^{3}(ik)^n\mathbf{e}_n(\mathbf{R}) + O(k^4)\right), \tag{131}$$

$$\mathbf{H}^s(\mathbf{R}) = e^{ikR}\left(\sum_{n=0}^{2}(ik)^n\mathbf{h}_n(\mathbf{R}) + O(k^3)\right), \tag{132}$$

which in the far field ($R \to \infty$) become

$$\begin{aligned}\mathbf{E}^s(\mathbf{R}) = {} & (e^{ikR}/kR)\{(ka)^3[\hat{\boldsymbol{\theta}}(\tfrac{1}{2} + \cos\theta)\cos\phi \\ & - \hat{\boldsymbol{\phi}}(1 + \tfrac{1}{2}\cos\theta)\sin\phi] + O(k^5)\} \\ & + O(1/R^2), \end{aligned} \tag{133}$$

$$\begin{aligned}\mathbf{H}^s(\mathbf{R}) = {} & (Ye^{ikR}/kR)\{(ka)^3[\hat{\boldsymbol{\theta}}(1 + \tfrac{1}{2}\cos\theta)\sin\phi \\ & + \hat{\boldsymbol{\phi}}(\tfrac{1}{2} + \cos\theta)\cos\phi] + O(k^5)\} \\ & + O(1/R^2). \end{aligned} \tag{134}$$

These two results are in agreement with the ones obtained by Lord Rayleigh.[1]

CONCLUSIONS

To summarize, the main result of the paper was the derivation of coupled Fredholm integral equations of the second kind for the electric and magnetic fields scattered by a perfectly conducting surface when immersed in an arbitrary incident field. These integral equations are of such a form as to admit of solution in a standard Neumann series when k, the wavenumber, is sufficiently small. The technique is the electromagnetic analog of a recently developed method of solving acoustic scattering problems.[14,15] Here two dyadic potential functions play the role that the potential Green's functions had in the scalar case. The derivation and definition of these fundamental dyadics constitutes a large part of the present work.

No proof of convergence of the iterative solution of the integral equations has been given and this remains as an important subject for future work. Support for the conjecture that iteration does yield a sequence which converges to the correct result is provided by the application of the method to the specific problem of scattering by a sphere. Not only are the correct first few terms obtained in the low-frequency expansion of the scattered field, but also the calculation of the fourth term was carried out without the appearance of divergent integrals. This is significant since Stevenson has pointed out that his special method, as well as that of Tai,[17] breaks down at the fourth term.

Whenever the requisite potential problems can be solved, the present method offers a direct means by which the electromagnetic scattering problem can be solved at low frequencies. The method is more systematic than the corresponding method of Stevenson and produces expressions for the field directly which are valid in both near and far zone. While some of the calculations required in Stevenson's approach are eliminated, those remaining are by no means trivial. The question of whether the present method can be further simplified has importance from a practical as well as an aesthetic point of view since tractability of calculation, rather than availability of potential solutions, has proven to be the real limitations of Stevenson's method. Central to this question is a study of the fundamental dyadics in an attempt to express them in simpler form. The example considered in the present paper, e.g., Eq. (112), offers a hint of the simplification possible. No general results are yet available. Another open question is whether the present method can be extended to include scattering from dielectric bodies.

ACKNOWLEDGMENTS

This work was carried out while both authors were at the Radiation Laboratory, The University of Michigan. The advice, support, encouragement, and occasional nagging of their colleagues, especially C. T. Tai and C. F. White, was invaluable and is gratefully acknowledged. The work was supported by the U.S. Air Force and the National Science Foundation.

APPENDIX A: DYADIC RELATIONSHIPS

The following dyadic relationships have been used in this work.[18]

Multiplicative Relationships

$\mathbf{a}$ and $\mathbf{b}$ are vectors and A a dyadic:

$$(\mathbf{a}\times\mathbf{b})\cdot\mathsf{A} = \mathbf{a}\cdot(\mathbf{b}\times\mathsf{A}) = -\mathbf{b}\cdot(\mathbf{a}\times\mathsf{A}), \tag{A1}$$

$$\mathbf{a}\times(\mathbf{b}\times\mathsf{A}) = \mathbf{b}(\mathbf{a}\cdot\mathsf{A}) - \mathsf{A}(\mathbf{a}\cdot\mathbf{b}). \tag{A2}$$

Derived Relationships

I is the identity dyadic and ϕ a scalar function:

$$\nabla\cdot(\mathbf{a}\times\mathsf{A}) = (\nabla\times\mathbf{a})\cdot\mathsf{A} - \mathbf{a}\cdot(\nabla\times\mathsf{A}), \quad \text{(A3)}$$

$$\nabla\cdot(\mathbf{ab}) = (\nabla\cdot\mathbf{a})\mathbf{b} + \mathbf{a}\cdot\nabla\mathbf{b}, \quad \text{(A4)}$$

$$\nabla\times(\mathbf{ab}) = (\nabla\times\mathbf{a})\mathbf{b} - \mathbf{a}\times\nabla\mathbf{b}, \quad \text{(A5)}$$

$$\nabla\times(\phi\mathsf{I}) = \nabla\phi\times\mathsf{I}, \quad \text{(A6)}$$

$$\nabla\times\nabla\times(\phi\mathsf{I}) = \nabla\nabla\phi - \mathsf{I}\nabla^2\phi. \quad \text{(A7)}$$

APPENDIX B: EXPLICIT DETERMINATION OF THE FUNDAMENTAL DYADICS

To properly take into account the singularity due to the magnetic dipole source [(44)], in Eqs. (55)–(59) let

$$\mathsf{E}_m^{(1)} = \nabla\times(-\mathsf{I}/4\pi\,|\mathbf{R}-\mathbf{R}'|) + \mathsf{E}_{m_r}^{(1)}(\mathbf{R}\mid\mathbf{R}') \quad \text{(B1)}$$

and

$$\mathsf{H}_m^{(0)} = Y\nabla\times\nabla\times(-\mathsf{I}/4\pi\,|\mathbf{R}-\mathbf{R}'|) + \mathsf{H}_{m_r}^{(0)}(\mathbf{R}\mid\mathbf{R}'), \quad \text{(B2)}$$

where $\mathsf{E}_{m_r}^{(1)}$ and $\mathsf{H}_{m_r}^{(0)}$ are $C^2(V)$. Then (56) and (57) imply

$$\nabla\times\mathsf{H}_{m_r}^{(0)}(\mathbf{R}) = 0, \qquad \mathbf{R}\in V, \quad \text{(B3)}$$

$$\nabla\times\mathsf{E}_{m_r}^{(1)}(\mathbf{R}) = Z\mathsf{H}_{m_r}^{(0)}(\mathbf{R}), \quad \mathbf{R}\in V. \quad \text{(B4)}$$

The dyadic $\mathsf{H}_m^{(0)}$ is the magnetostatic field due to three orthogonal static magnetic dipoles, and its expression in terms of potential functions follows the corresponding treatment for a single dipole. Thus (B3) implies that

$$\mathsf{H}_{m_r}^{(0)} = -Y\sum_{j=1}^{3}\nabla\phi_{mj_r}^{(0)}\hat{\mathbf{a}}_j \quad \text{(B5)}$$

and, taking the divergence of (B4), we conclude that

$$\nabla^2\phi_{mj_r}^{(0)} = 0, \quad j = 1, 2, 3. \quad \text{(B6)}$$

Furthermore, the potential functions $\phi_{mj_r}^{(0)}$ are regular at infinity in the sense of Kellogg. This follows from examination of the low-frequency expansion of the Stratton–Chu integral representation of the scattered field [e.g., see Ref. 4a]. Substituting (B2) and (B5) in the boundary condition on $\mathsf{H}_m^{(0)}$ [(59)] yields

$$Y\hat{\mathbf{n}}_s\cdot\nabla_s\times\nabla_s\times\left(-\frac{\mathsf{I}}{4\pi\,|\mathbf{R}_s-\mathbf{R}'|}\right) - Y\sum_{j=1}^{3}\hat{\mathbf{n}}_s\cdot\nabla_s\phi_{mj_r}^{(0)}\hat{\mathbf{a}}_j = 0 \quad \text{(B7)}$$

or, with (A7),

$$\frac{\partial}{\partial n_s}\phi_{mj_r}^{(0)}(\mathbf{R}_s\mid\mathbf{R}') = \frac{\partial}{\partial n_s}\left[(\hat{\mathbf{a}}_j\cdot\nabla_s)\left(-\frac{1}{4\pi\,|\mathbf{R}_s-\mathbf{R}'|}\right)\right], \quad j = 1, 2, 3. \quad \text{(B8)}$$

In terms of $N^{(e)}(\mathbf{R}\mid\mathbf{R}')$, the exterior potential Green's function for Neumann boundary condition on S [see (62)–(64)], the potential functions $\phi_{mj_r}^{(0)}$ may be expressed as

$$\begin{aligned}\phi_{mj_r}^{(0)}(\mathbf{R}\mid\mathbf{R}') &= -\int_S N^{(e)}(\mathbf{R}_s\mid\mathbf{R})\frac{\partial}{\partial n_s}\phi_{mj_r}^{(0)}(\mathbf{R}_s\mid\mathbf{R}')\,ds \\ &= \hat{\mathbf{a}}_j\cdot\nabla'\int_S N^{(e)}(\mathbf{R}_s\mid\mathbf{R})\frac{\partial}{\partial n_s}\left(-\frac{1}{4\pi\,|\mathbf{R}_s-\mathbf{R}'|}\right)ds \\ &= \hat{\mathbf{a}}_j\cdot\nabla' N_r^{(e)}(\mathbf{R}\mid\mathbf{R}').\end{aligned} \quad \text{(B9)}$$

The fact that

$$\nabla_s(1/|\mathbf{R}_s-\mathbf{R}'|) = -\nabla'(1/|\mathbf{R}_s-\mathbf{R}'|) \quad \text{(B10)}$$

was employed in deriving (B9) and will be used frequently in succeeding calculations.

With (B5) and (B9), (B4) becomes

$$\nabla\times\mathsf{E}_{m_r}^{(1)}(\mathbf{R}) = -\sum_{j=1}^{3}[\nabla\phi_{mj_r}^{(0)}]\hat{\mathbf{a}}_j = -\nabla\nabla' N_r^{(e)}(\mathbf{R}\mid\mathbf{R}'). \quad \text{(B11)}$$

Stevenson[20] has shown that the necessary and sufficient conditions for the vector equations

$$\nabla\times\mathbf{E}_{mj_r}^{(1)}(\mathbf{R}) = -\nabla\phi_{mj_r}^{(0)} \quad \text{(B12)}$$

to have a solution are

$$\nabla^2\phi_{mj_r}^{(0)} = 0, \quad \mathbf{R}\in V, \quad \text{(B13)}$$

and

$$\int_S \hat{\mathbf{n}}_s\cdot\nabla_s\phi_{mj_r}^{(0)}\,ds = 0. \quad \text{(B14)}$$

The first condition is satisfied by virtue of (B6), and the second is satisfied since, with (B8),

$$\int_S \hat{\mathbf{n}}_s\cdot\nabla_s\phi_{mj_r}^{(0)}\,ds = -\hat{\mathbf{a}}_j\cdot\nabla'\int_S \hat{\mathbf{n}}_s\cdot\nabla_s\left(-\frac{1}{4\pi\,|\mathbf{R}_s-\mathbf{R}'|}\right)ds \quad \text{(B15)}$$

and, by the divergence theorem,

$$\int_S \hat{\mathbf{n}}_s\cdot\nabla_s\left(-\frac{1}{4\pi\,|\mathbf{R}_s-\mathbf{R}'|}\right)ds = 0, \quad \mathbf{R}'\in V. \quad \text{(B16)}$$

The complete solution of (B12) is

$$\mathbf{E}_{mj_r}^{(1)}(\mathbf{R}) = -\frac{1}{4\pi}\nabla\times\int_S\left(\frac{\phi_{mj_r}^{(0)}(\mathbf{R}_s\mid\mathbf{R}') - N_{mj}^{(i)}(\mathbf{R}_s\mid\mathbf{R}')}{|\mathbf{R}_s-\mathbf{R}|}\right)\hat{\mathbf{n}}_s\,ds + \nabla G_{mj}^{(e)}(\mathbf{R}\mid\mathbf{R}'). \quad \text{(B17)}$$

The first term on the right is a particular solution of (B12),[4a,20] where $N^{(i)}_{mj}$ is an interior Neumann potential function for S, i.e.,

$$\nabla^2 N^{(i)}_{mj}(\mathbf{R} \mid \mathbf{R}') = 0, \quad \mathbf{R} \in V_i, \quad \mathbf{R}' \in V, \quad \text{(B18)}$$

$$\hat{\mathbf{n}}_s \cdot \nabla_s N^{(i)}_{mj}(\mathbf{R}_s \mid \mathbf{R}') = \hat{\mathbf{n}}_s \cdot \nabla_s \phi^{(0)}_{mj_r}(\mathbf{R}_s \mid \mathbf{R}'), \quad \mathbf{R}_s \in S, \quad \mathbf{R}' \in V. \quad \text{(B19)}$$

This is a standard interior Neumann problem and has a solution provided that

$$\int_S \hat{\mathbf{n}}_s \cdot \nabla_s N^{(i)}_{mj}(\mathbf{R}_s \mid \mathbf{R}')\, ds = 0, \quad \text{(B20)}$$

a condition guaranteed by (B14) and (B19). In fact, with (B8) it is seen that the solution of this problem is simply[21]

$$N^{(i)}_{mj}(\mathbf{R} \mid \mathbf{R}') = \hat{\mathbf{a}}_j \cdot \nabla'(1/4\pi\, |\mathbf{R} - \mathbf{R}'|), \quad \mathbf{R} \in V_i, \quad \mathbf{R}' \in V. \quad \text{(B21)}$$

Thus, with (B9),

$$\mathbf{E}^{(1)}_{mj_r}(\mathbf{R}) = -\frac{1}{4\pi}(\hat{\mathbf{a}}_j \cdot \nabla')\nabla \times \int_S \frac{N^{(e)}(\mathbf{R}_s \mid \mathbf{R}')}{|\mathbf{R}_s - \mathbf{R}|}\hat{\mathbf{n}}_s\, ds + \nabla G^{(e)}_{mj}(\mathbf{R} \mid \mathbf{R}'), \quad \text{(B22)}$$

where $N^{(e)}$ is the exterior Neumann function [(62)–(64)]. The second term on the right in (B22) or (B17) is a solution of the homogeneous equation

$$\nabla \times \mathbf{E}^{(1)}_{mj_r}(\mathbf{R} \mid \mathbf{R}') = 0. \quad \text{(B23)}$$

That $G^{(e)}_{mj}$ are potential functions follows by taking the divergence of (B22) and noting that $\mathbf{E}^{(1)}_{mj_r}$ is the curl of a vector [(58)] and hence divergence free. That $G^{(e)}_{mj}$ is regular at infinity in the sense of Kellogg follows, as with $\phi^{(0)}_{mj_r}$, from an examination of the Stratton–Chu integral representation of $E^{(1)}_{mj_r}$. The behavior of $G^{(e)}_{mj}$ on S is determined from the boundary condition (59) on $\mathsf{E}^{(1)}_m$. Hence $G^{(e)}_{mj}$ are standard exterior Dirichlet potential functions, i.e.,

$$\nabla^2 G^{(e)}_{mj}(\mathbf{R} \mid \mathbf{R}') = 0, \quad \mathbf{R}, \mathbf{R}' \in V, \quad \text{(B24)}$$

$$\hat{\mathbf{n}} \times \nabla G^{(e)}_{mj}(\mathbf{R} \mid \mathbf{R}') = \frac{1}{4\pi}\hat{\mathbf{n}} \times \nabla \times \left(\frac{\hat{\mathbf{a}}_j}{|\mathbf{R} - \mathbf{R}'|} + (\hat{\mathbf{a}}_j \cdot \nabla')\int_S \frac{N^{(e)}(\mathbf{R}_s \mid \mathbf{R}')}{|\mathbf{R}_s - \mathbf{R}|}\hat{\mathbf{n}}_s\, ds\right), \quad \mathbf{R} \in S, \quad \mathbf{R}' \in V, \quad \text{(B25)}$$

$G^{(e)}_{mj}(\mathbf{R} \mid \mathbf{R}')$ regular at infinity.

To completely determine $G^{(e)}_{mj}$, the additional condition

$$\int_S \hat{\mathbf{n}}_s \cdot \nabla_s G^{(e)}_{mj}(\mathbf{R}_s \mid \mathbf{R}')\, ds = 0, \quad j = 1, 2, 3, \quad \text{(B26)}$$

is employed. This arises from the representation (B22), the fact that $\mathbf{E}^{(1)}_{mj_r}$ is expressible as a curl [(58)], and Stokes' theorem.

In summary

$$\mathsf{E}^{(1)}_m(\mathbf{R} \mid \mathbf{R}') = \nabla \times \left(-\frac{\mathsf{I}}{4\pi\, |\mathbf{R} - \mathbf{R}'|}\right) + \sum_{j=1}^{3}\left(-\frac{1}{4\pi}(\hat{\mathbf{a}}_j \cdot \nabla')\nabla \times \int_S \frac{N^{(e)}(\mathbf{R}_s \mid \mathbf{R}')}{|\mathbf{R}_s - \mathbf{R}|}\hat{\mathbf{n}}_s\, ds + \nabla G^{(e)}_{mj}(\mathbf{R} \mid \mathbf{R}')\right)\hat{\mathbf{a}}_j, \quad \text{(B27)}$$

where $N^{(e)}$ is an exterior Neumann potential function for S [(62)–(64)] and $G^{(e)}_{mj}$ are exterior Dirichlet potential functions for S [(B24)–(B26)].

The Dyadic $\mathsf{H}^{(1)}_e$ in Terms of Potential Functions

The procedure for finding $\mathsf{H}^{(1)}_e$ is similar except that the singularity enters in a different way hence different potential functions arise. Thus, let

$$\mathsf{H}^{(1)}_e(\mathbf{R} \mid \mathbf{R}') = \nabla \times (-\mathsf{I}/4\pi\, |\mathbf{R} - \mathbf{R}'|) + \mathsf{H}^{(1)}_{e_r}(\mathbf{R} \mid \mathbf{R}'), \quad \text{(B28)}$$

with (A7) and (51) and the fact that

$$\nabla^2(-1/4\pi\, |\mathbf{R} - \mathbf{R}'|) = \delta(\mathbf{R} \mid \mathbf{R}'), \quad \text{(B29)}$$

$$\nabla \times \mathsf{H}^{(1)}_{e_r}(\mathbf{R} \mid \mathbf{R}') = -\nabla\nabla(-1/4\pi\, |\mathbf{R} - \mathbf{R}'|) - Y\mathsf{E}^{(0)}_e(\mathbf{R} \mid \mathbf{R}'). \quad \text{(B30)}$$

Taking the curl of (B30) and using (48) yields

$$\nabla \times \nabla \times \mathsf{H}^{(1)}_{e_r}(\mathbf{R} \mid \mathbf{R}') = 0, \quad \text{(B31)}$$

which is the desired result (17). Furthermore, (48) implies that

$$\mathsf{E}^{(0)}_e(\mathbf{R} \mid \mathbf{R}') = -Z\sum_{j=1}^{3}[\nabla\phi^{(0)}_{ej}(\mathbf{R} \mid \mathbf{R}')]\hat{\mathbf{a}}_j, \quad \text{(B32)}$$

which in combination with the divergence of (B30) yields

$$\sum_{j=1}^{3}[\nabla^2\phi^{(0)}_{ej}(\mathbf{R} \mid \mathbf{R}')]\hat{\mathbf{a}}_j = \nabla\delta(\mathbf{R} \mid \mathbf{R}') \quad \text{(B33)}$$

or

$$\nabla^2\phi^{(0)}_{ej}(\mathbf{R} \mid \mathbf{R}') = \hat{\mathbf{a}}_j \cdot \nabla\delta(\mathbf{R} \mid \mathbf{R}'), \quad j = 1, 2, 3. \quad \text{(B34)}$$

From (53) and (B32), the boundary condition satisfied by these scalar functions is

$$\hat{\mathbf{n}} \times \nabla\phi^{(0)}_{ej}(\mathbf{R} \mid \mathbf{R}') = 0, \quad \mathbf{R} \in S, \quad j = 1, 2, 3, \quad \text{(B35)}$$

which, by Stokes' theorem for scalar fields, implies that $\phi_{ej}^{(0)}$ is a constant on S with respect to $\mathbf{R}$ though it may still depend on $\mathbf{R}'$. By (B34), $\phi_{ej}^{(0)}$ may be written in the form

$$\phi_{ej}^{(0)}(\mathbf{R} \mid \mathbf{R}') = \hat{\mathbf{a}}_j \cdot \nabla(-1/4\pi\,|\mathbf{R} - \mathbf{R}'|) + \phi_{ej_r}^{(0)}(\mathbf{R} \mid \mathbf{R}'), \tag{B36}$$

where

$$\nabla^2 \phi_{ej_r}^{(0)}(\mathbf{R} \mid \mathbf{R}') = 0, \tag{B37}$$

in which case the boundary condition reads

$$\phi_{ej_r}^{(0)}(\mathbf{R}_s \mid \mathbf{R}') = -\hat{\mathbf{a}}_j \cdot \nabla_s(-1/4\pi\,|\mathbf{R}_s - |\mathbf{R}) + C_{ej}, \quad \mathbf{R}_s \in S, \quad \mathbf{R}' \in V, \tag{B38}$$

where C_{ej} is the constant value of $\phi_{ej}^{(0)}$ on S. The functions $\phi_{ej_r}^{(0)}(\mathbf{R} \mid \mathbf{R}')$ are regular at infinity in the sense of Kellogg by the same argument as before (examination of the Stratton–Chu integral representation of $\mathbf{E}_{ej_r}^{(0)}$) and hence may be expressed everywhere in V in terms of values on S by the formula

$$\phi_{ej_r}^{(0)}(\mathbf{R} \mid \mathbf{R}') = \int_S \phi_{ej_r}^{(0)}(\mathbf{R}_s \mid \mathbf{R}') \frac{\partial}{\partial n_s} G^e(\mathbf{R}_s \mid \mathbf{R})\, ds, \tag{B39}$$

where $G^{(e)}$ is the exterior potential Green's function for Dirichlet boundary condition on S [(70)–(72)]. Substitution of (B38) in (B39) gives

$$\begin{aligned}\phi_{ej_r}^{(0)}&(\mathbf{R} \mid \mathbf{R}') \\ &= \hat{\mathbf{a}}_j \cdot \nabla' \int_S \left(-\frac{1}{4\pi\,|\mathbf{R}_s - \mathbf{R}'|}\right) \frac{\partial}{\partial n_s} G^{(e)}(\mathbf{R}_s \mid \mathbf{R})\, ds \\ &\quad + C_{ej} \int_S \frac{\partial}{\partial n_s} G^{(e)}(\mathbf{R}_s \mid \mathbf{R}')\, ds \\ &= -(\hat{\mathbf{a}}_j \cdot \nabla') G_r^{(e)}(\mathbf{R} \mid \mathbf{R}') \\ &\quad + C_{ej} \int_S \frac{\partial}{\partial n_s} G^{(e)}(\mathbf{R}_s \mid \mathbf{R}')\, ds.\end{aligned} \tag{B40}$$

The conductor potential $U(\mathbf{R})$ is defined as an exterior potential function, regular at infinity and taking on the value 1 on S (Ref. 13, p. 330). Expressing the conductor potential in terms of the Dirichlet Green's function yields

$$U(\mathbf{R}) = \int_S \frac{\partial}{\partial n_s} G^{(e)}(\mathbf{R}_s \mid \mathbf{R})\, ds. \tag{B41}$$

Thus (B40) may be rewritten

$$\phi_{ej_r}^{(0)}(\mathbf{R} \mid \mathbf{R}') = -(\hat{\mathbf{a}}_j \cdot \nabla') G_r^{(e)}(\mathbf{R} \mid \mathbf{R}') + C_{ej} U(\mathbf{R}). \tag{B42}$$

This equation, with (B36) and (70), leads to an expression for the full potential $\phi_{ej}^{(0)}$:

$$\phi_{ej}^{(0)} = -(\hat{\mathbf{a}}_j \cdot \nabla') G^{(e)}(\mathbf{R} \mid \mathbf{R}') + C_{ej} U(\mathbf{R}). \tag{B43}$$

The constants C_{ej} are determined from the condition

$$\int_S \hat{\mathbf{n}}_s \cdot \nabla_s \phi_{ej}^{(0)}(\mathbf{R}_s \mid \mathbf{R}')\, ds = 0, \tag{B44}$$

which is a consequence of (51), (B32), and Stokes' theorem. It is a mathematical statement of the physical fact that the total induced static charge in the perfectly conducting surface must be zero. Substitution of (B43) in this expression gives

$$C_{ej} = \hat{\mathbf{a}}_j \cdot \nabla' U(\mathbf{R}') \left(\int_S \frac{\partial}{\partial n_s} U(\mathbf{R}_s)\, ds \right)^{-1} \tag{B45}$$

The electrostatic capacity C of the surface S is defined (Ref. 13, p. 330) as

$$C = \frac{1}{4\pi} \int_S \frac{\partial}{\partial n_s} U(\mathbf{R}_s)\, ds, \tag{B46}$$

hence

$$C_{ej} = \hat{\mathbf{a}}_j \cdot \nabla' U(\mathbf{R}')/4\pi C \tag{B47}$$

and

$$\phi_{ej}^{(0)} = -(\hat{\mathbf{a}}_j \cdot \nabla')[G^{(e)}(\mathbf{R} \mid \mathbf{R}') - U(\mathbf{R}) U(\mathbf{R}')/4\pi C], \tag{B48}$$

$$\phi_{ej_r}^{(0)} = -(\hat{\mathbf{a}}_j \cdot \nabla')[G_r^{(e)}(\mathbf{R} \mid \mathbf{R}') - U(\mathbf{R}) U(\mathbf{R}')/4\pi C]. \tag{B49}$$

The electric field dyadic (B32) can then be written

$$\begin{aligned}\mathsf{E}_e^{(0)}&(\mathbf{R} \mid \mathbf{R}') \\ &= Z \sum_{j=1}^{3} \nabla \left((\hat{\mathbf{a}}_j \cdot \nabla') G^e(\mathbf{R} \mid \mathbf{R}') - \hat{\mathbf{a}}_j \cdot \nabla' \frac{U(\mathbf{R}') U(\mathbf{R})}{4\pi C} \right) \hat{\mathbf{a}}_j \\ &= Z \nabla\nabla' \left(G^{(e)}(\mathbf{R} \mid \mathbf{R}') - \frac{U(\mathbf{R}) U(\mathbf{R}')}{4\pi C} \right).\end{aligned} \tag{B50}$$

This is the electric field due to three orthogonally crossed static electric dipoles with moments defined by (45).

The dyadic $\mathsf{H}_e^{(1)}$ can now be found. Substituting (B50) in (B30) and making use of the definition in (70)–(72) of $G^{(e)}$ yields

$$\nabla \times \mathsf{H}_{e_r}^{(1)} = -\nabla\nabla'[G_r^{(e)}(\mathbf{R} \mid \mathbf{R}') - U(\mathbf{R}) U(\mathbf{R}')/4\pi C]. \tag{B51}$$

This dyadic equation can be broken into the three vector equations

$$\nabla \times \mathbf{H}_{ej_r}^{(1)} = -(\hat{\mathbf{a}}_j \cdot \nabla') \times \nabla[G_r^{(e)}(\mathbf{R} \mid \mathbf{R}') - U(\mathbf{R}) U(\mathbf{R}')/4\pi C], \quad j = 1, 2, 3, \tag{B52}$$

or, with (B49),

$$\nabla \times \mathbf{H}_{ej_r}^{(1)} = \nabla \phi_{ej_r}^{(0)}, \quad j = 1, 2, 3. \tag{B53}$$

The necessary and sufficient conditions for these equations to have solutions, (B13) and (B14), are satisfied by virtue of (B37), (B44), and (B16). The complete solution of (B53) is

$$\mathbf{H}^{(1)}_{ej_r}(\mathbf{R} \mid \mathbf{R}') = \frac{1}{4\pi} \nabla \times \int_S \left(\frac{\phi^{(0)}_{ej_r}(\mathbf{R}_s \mid \mathbf{R}') - N^{(i)}_{ej_r}(\mathbf{R}_s \mid \mathbf{R}')}{|\mathbf{R}_s - \mathbf{R}|} \right) \hat{\mathbf{n}}_s \, ds + \nabla N^{(e)}_{ej}(\mathbf{R} \mid \mathbf{R}'). \quad \text{(B54)}$$

Just as in the corresponding solution of the equation for $E^{(1)}_{mj_r}$ [(B17)], the first term on the right is a particular solution of (B53), and the second is a solution of the homogeneous equation

$$\nabla \times \mathbf{H}^{(1)}_{ej_r}(\mathbf{R} \mid \mathbf{R}') = 0, \quad \mathbf{R}, \mathbf{R}' \in V. \quad \text{(B55)}$$

The functions $N^{(i)}_{ej_r}(\mathbf{R}_s \mid \mathbf{R}')$ that appear in (B54) are interior potential functions satisfying the boundary condition

$$\begin{aligned} \hat{\mathbf{n}}_s \cdot \nabla_s N^{(i)}_{ej_r}(\mathbf{R}_s \mid \mathbf{R}') &= \hat{\mathbf{n}}_s \cdot \nabla_s \phi^{(0)}_{ej_r}(\mathbf{R}_s \mid \mathbf{R}') \\ &= -\hat{\mathbf{n}}_s \cdot \nabla_s (\hat{\mathbf{a}}_j \cdot \nabla') \\ &\quad \times [G^{(e)}_r(\mathbf{R}_s \mid \mathbf{R}') - U(\mathbf{R}_s) U(\mathbf{R}')/4\pi C], \\ &\quad \mathbf{R}_s \in S, \quad \mathbf{R}' \in V. \end{aligned} \quad \text{(B56)}$$

It is convenient to introduce an associated interior Neumann function $N^{(i)}_e(\mathbf{R} \mid \mathbf{R}')$ such that

$$\nabla^2 N^{(i)}_e(\mathbf{R} \mid \mathbf{R}') = 0, \quad \mathbf{R} \in V_i, \quad \mathbf{R}' \in V, \quad \text{(B57)}$$

$$\hat{\mathbf{n}}_s \cdot \nabla_s N^{(i)}_e(\mathbf{R}_s \mid \mathbf{R}') = -\hat{\mathbf{n}}_s \cdot \nabla_s [G^{(e)}(\mathbf{R}_s \mid \mathbf{R}') - U(\mathbf{R}_s) U(\mathbf{R}')/4\pi C]. \quad \text{(B58)}$$

The definitions of $U(\mathbf{R})$ [(B41)] and C [(B46)] suffice to show that

$$\int_S \hat{\mathbf{n}}_s \cdot \nabla_s N^{(i)}_e(\mathbf{R}_s \mid \mathbf{R}') \, ds = 0, \quad \text{(B59)}$$

which is a necessary condition for the existence of $N^{(i)}_e$. Since $1/4\pi |\mathbf{R} - \mathbf{R}'|$ is a solution of (B57), the functions $N^{(i)}_{ej_r}$ may be written as

$$N^{(i)}_{ej_r}(\mathbf{R} \mid \mathbf{R}') = \hat{\mathbf{a}}_j \cdot \nabla' [N^{(i)}_e(\mathbf{R} \mid \mathbf{R}') - 1/4\pi |\mathbf{R} - \mathbf{R}'|]; \quad \text{(B60)}$$

hence, the first integral on the right in (B54) can be written as

$$\begin{aligned} &\nabla \times \int_S \left(\frac{\phi^{(0)}_{ej_r}(\mathbf{R}_s \mid \mathbf{R}') - N^{(i)}_{ej_r}(\mathbf{R}_s \mid \mathbf{R}')}{|\mathbf{R}_s - \mathbf{R}|} \right) \hat{\mathbf{n}}_s \, ds \\ &= \nabla \times \int_S (\hat{\mathbf{a}}_j \cdot \nabla') \\ &\quad \times \left(-G^{(e)}_r(\mathbf{R}_s \mid \mathbf{R}') + \frac{U(\mathbf{R}_s) U(\mathbf{R}')}{4\pi C} - N^{(i)}_e(\mathbf{R}_s \mid \mathbf{R}') + \frac{1}{4\pi |\mathbf{R}_s - \mathbf{R}'|} \right) \frac{\hat{\mathbf{n}}_s}{|\mathbf{R}_s - \mathbf{R}|} \, ds \\ &= -\nabla \times \int_S \hat{\mathbf{a}}_j \cdot \nabla' \frac{N^{(i)}_e(\mathbf{R}_s \mid \mathbf{R}')}{|\mathbf{R}_s - \mathbf{R}|} \hat{\mathbf{n}}_s \, ds. \end{aligned} \quad \text{(B61)}$$

This last expression was obtained through the boundary conditions on $G^{(e)}$ and U and using the fact that Gauss' theorem for scalar fields implies

$$\nabla \times \int_S \frac{\hat{\mathbf{n}}_s}{|\mathbf{R}_s - \mathbf{R}|} \, ds = 0. \quad \text{(B62)}$$

This fact also indicates that, while (B57)–(B59) specify $N^{(i)}_e(\mathbf{R} \mid \mathbf{R}')$ only to within an arbitrary function of $\mathbf{R}'$, this arbitrary function does not contribute to (B61). Thus, (B54) may be written

$$\mathbf{H}^{(1)}_{ej_r} = -\frac{1}{4\pi} (\hat{\mathbf{a}}_j \cdot \nabla') \nabla \times \int_S \frac{N^{(i)}_e(\mathbf{R}_s \mid \mathbf{R}')}{|\mathbf{R}_s - \mathbf{R}|} \hat{\mathbf{n}}_s \, ds + \nabla N^{(e)}_{ej}(\mathbf{R} \mid \mathbf{R}'). \quad \text{(B63)}$$

The functions $N^{(e)}_{ej}(\mathbf{R} \mid \mathbf{R}')$ are exterior potential functions determined so that the boundary condition (53) on $\mathbf{H}^{(1)}_e$ will be satisfied. Thus,

$$\nabla^2 N^{(e)}_{ej}(\mathbf{R} \mid \mathbf{R}') = 0, \quad \mathbf{R}, \mathbf{R}' \in V, \quad \text{(B64)}$$

$N^{(e)}_{ej}(\mathbf{R} \mid \mathbf{R}')$ regular at infinity in the sense of Kellogg, and

$$\begin{aligned} \hat{\mathbf{n}} \cdot \nabla N^{(e)}_{ej}(\mathbf{R} \mid \mathbf{R}') &= \hat{\mathbf{n}} \cdot \nabla \times \left(\frac{\hat{\mathbf{a}}_j}{4\pi |\mathbf{R} - \mathbf{R}'|} + \frac{1}{4\pi} (\hat{\mathbf{a}}_j \cdot \nabla') \times \int_S \frac{N^{(i)}_e(\mathbf{R}_s \mid \mathbf{R}')}{|\mathbf{R}_s - \mathbf{R}|} \hat{\mathbf{n}}_s \, ds \right), \\ &\quad \mathbf{R} \in S, \quad \mathbf{R}' \in V. \end{aligned} \quad \text{(B65)}$$

In terms of $N^{(e)}$, the Green's function for the Neumann problem, we have

$$N^{(e)}_{ej}(\mathbf{R} \mid \mathbf{R}') = -\int_S N^{(e)}(\mathbf{R}_s \mid \mathbf{R}) \hat{\mathbf{n}}_s \cdot \nabla_s N^{(e)}_{ej}(\mathbf{R}_s \mid \mathbf{R}') \, ds. \quad \text{(B66)}$$

In summary

$$\mathsf{H}_e^{(1)}(\mathbf{R}\mid\mathbf{R}') = \nabla\times\left(-\frac{\mathsf{I}}{4\pi|\mathbf{R}-\mathbf{R}'|}\right) + \sum_{j=1}^{3}\left(-\frac{1}{4\pi}(\hat{\mathbf{a}}_j\cdot\nabla')\nabla\times\int\frac{N_e^{(i)}(\mathbf{R}_s\mid\mathbf{R}')}{|\mathbf{R}_s-\mathbf{R}|}\hat{\mathbf{n}}_s\,ds + \nabla N_{ej}^{(e)}(\mathbf{R}\mid\mathbf{R}')\right)\hat{\mathbf{a}}_j, \quad \text{(B67)}$$

where $N_e^{(i)}$ is an interior Neumann potential function [(B57)–(B59)] and $N_{ej}^{(e)}$ are exterior Neumann potential functions [(B64) and (B65)].

APPENDIX C: PROOF OF THEOREM 3

Substitution of (91)–(94) in (89) and (90) leads to the following equations[22]:

$$ik[Z\mathbf{h}(\mathbf{R}') - \hat{\mathbf{R}}'\times\mathbf{e}(\mathbf{R}')] = -k^2\int_V(\mathbf{e}+Z\hat{\mathbf{R}}\times\mathbf{h})\cdot\mathsf{E}_m^{(1)}\,dV + ik\int_V[\nabla\times(\hat{\mathbf{R}}\times\mathbf{e})]\cdot\mathsf{E}_m^{(1)}\,dV + \int_S(\hat{\mathbf{n}}\times\mathbf{e})\cdot(\nabla\times\mathsf{E}_m^{(1)})\,dS \quad \text{(C1)}$$

and

$$-ik[Y\mathbf{e}(\mathbf{R}') + \hat{\mathbf{R}}'\times\mathbf{h}(\mathbf{R}')] = -k^2\int_V(\mathbf{h}-Y\hat{\mathbf{R}}\times\mathbf{e})\cdot\mathsf{H}_e^{(1)}\,dV + ik\int_V[\nabla\times(\hat{\mathbf{R}}\times\mathbf{h})]\cdot\mathsf{H}_e^{(1)}\,dV - ik\int_S[\hat{\mathbf{n}}\times(Y\mathbf{e}+\hat{\mathbf{R}}\times\mathbf{h})]\cdot\mathsf{H}_e^{(1)}\,dS. \quad \text{(C2)}$$

By the identity (A3) the second volume integral of (C1) can be written in the following form:

$$\int_V[\nabla\times(\hat{\mathbf{R}}\times\mathbf{e})]\cdot\mathsf{E}_m^{(1)}\,dV = \int_V\nabla\cdot[(\hat{\mathbf{R}}\times\mathbf{e})\times\mathsf{E}_m^{(1)}]\,dV + \int_V(\hat{\mathbf{R}}\times\mathbf{e})\cdot(\nabla\times\mathsf{E}_m^{(1)})\,dV. \quad \text{(C3)}$$

By (B1) and (B11) and application of the divergence theorem (24) to the first integral on the right, (C3) becomes

$$\int_V[\nabla\times(\hat{\mathbf{R}}\times\mathbf{e})]\cdot\mathsf{E}_m^{(1)}\,dV = \int_{S+S'}\hat{\mathbf{n}}\cdot[(\hat{\mathbf{R}}\times\mathbf{e})\times\mathsf{E}_m^{(1)}]\,dS - \int_V(\hat{\mathbf{R}}\times\mathbf{e})\cdot\nabla\nabla' N^{(e)}\,dV. \quad \text{(C4)}$$

By the boundary condition $\hat{\mathbf{n}}\times\mathsf{E}_m^{(1)}=0$ and (A1), the integral over S in (C4) vanishes. Application of (A4) to the last integral results in

$$\int_V[\nabla\times(\hat{\mathbf{R}}\times\mathbf{e})]\cdot\mathsf{E}_m^{(1)}\,dv = \int_{S'}\hat{\mathbf{n}}\cdot[(\hat{\mathbf{R}}\times\mathbf{e})\times\mathsf{E}_m^{(1)}]\,ds - \int_V\nabla\cdot[(\hat{\mathbf{R}}\times\mathbf{e})\nabla' N^{(e)}]\,dv + \int_V\nabla\cdot(\hat{\mathbf{R}}\times\mathbf{e})\nabla' N^{(e)}\,dv, \quad \text{(C5)}$$

which can be further written as

$$\int_V[\nabla\times(\hat{\mathbf{R}}\times\mathbf{e})]\cdot\mathsf{E}_m^{(1)}\,dv = \int_{S'}\hat{\mathbf{n}}\cdot[(\hat{\mathbf{R}}\times\mathbf{e})\times\mathsf{E}_m^{(1)} - (\hat{\mathbf{R}}\times\mathbf{e})\nabla' N^{(e)}]\,ds + \int_S\hat{\mathbf{R}}\cdot[(\hat{\mathbf{n}}\times\mathbf{e})\nabla' N^{(e)}]\,ds - ik\int_V\hat{\mathbf{R}}\cdot(Z\mathbf{h}-\hat{\mathbf{R}}\times\mathbf{e})\nabla' N^{(e)}\,dv, \quad \text{(C6)}$$

by means of the divergence theorem, standard vector identities, and (91). The integral over S' can be evaluated in the standard manner (see Sec. II), the result being $-\hat{\mathbf{R}}'\times\mathbf{e}(\mathbf{R}')$; thus Eq. (C6) becomes

$$\int_V[\nabla\times(\hat{\mathbf{R}}\times\mathbf{e})]\cdot\mathsf{E}_m^{(1)}\,dv = -\hat{\mathbf{R}}'\times\mathbf{e}(\mathbf{R}') + \int_S\hat{\mathbf{R}}\cdot[(\hat{\mathbf{n}}\times\mathbf{e})\nabla' N^{(e)}]\,ds - ikZ\int_V(\hat{\mathbf{R}}\cdot\mathbf{h})\nabla' N^{(e)}\,dv. \quad \text{(C7)}$$

Substitution of this expression in (C1) yields the result

$$ikZ\mathbf{h}(\mathbf{R}') = -k^2\int_V(\mathbf{e}+Z\hat{\mathbf{R}}\times\mathbf{h})\cdot\mathsf{E}_m^{(1)}\,dv + k^2Z\nabla'\int_V(\hat{\mathbf{R}}\cdot\mathbf{h})N^{(e)}\,dv + ik\int_S\hat{\mathbf{R}}\cdot[(\hat{\mathbf{n}}\times\mathbf{e})\nabla' N^{(e)}]\,ds + \int_S(\hat{\mathbf{n}}\times\mathbf{e})\cdot\nabla\times\mathsf{E}_m^{(1)}\,ds. \quad \text{(C8)}$$

Following the same treatment used in deriving (85) and using (92) rather than Maxwell's equations, we can write the last integral of the above expression as

$$\int_S(\hat{\mathbf{n}}\times\mathbf{e})\cdot\nabla\times\mathsf{E}_m^{(1)}\,ds = -ikZ\int_S(\hat{\mathbf{n}}\cdot\mathbf{h})\nabla' N^{(e)}\,ds - ik\int_S\hat{\mathbf{R}}\cdot[(\hat{\mathbf{n}}\times\mathbf{e})\nabla' N^{(e)}]\,ds. \quad \text{(C9)}$$

With this substitution (C8) becomes

$$\mathbf{h}(\mathbf{R}') = ik\int_V (Y\mathbf{e} + \hat{\mathbf{R}} \times \mathbf{h}) \cdot \mathsf{E}_m^{(1)}\, dv - ik\nabla' \int_V (\hat{\mathbf{R}} \cdot \mathbf{h}) N^{(e)}\, dv - \nabla' \int_S (\hat{\mathbf{n}} \cdot \mathbf{h}) N^{(e)}\, ds, \qquad \text{(C10)}$$

which is the desired result.

The second volume integral in (C2) can be modified in an analogous manner as the corresponding integral in (C1) to give

$$\begin{aligned}\int_V \{\nabla \times [\hat{\mathbf{R}} \times \mathbf{h}(\mathbf{R})]\} \cdot \mathsf{H}_e^{(1)}(\mathbf{R} \mid \mathbf{R}')\, dv &= -\hat{\mathbf{R}}' \times \mathbf{h}(\mathbf{R}') + \int_S \hat{\mathbf{n}}_s \cdot \{[\hat{\mathbf{R}}_s \times \mathbf{h}(\mathbf{R}_s)] \times \mathsf{H}_e^{(1)}(\mathbf{R}_s \mid \mathbf{R}')\}\, ds \\ &\quad - \int_S \hat{\mathbf{n}} \cdot [\hat{\mathbf{R}}_s \times \mathbf{h}(\mathbf{R}_s)] \times \nabla' \left(G^{(e)}(\mathbf{R}_s \mid \mathbf{R}') - \frac{U(\mathbf{R}_s)U(\mathbf{R}')}{4\pi C}\right) ds \\ &\quad + ikY\nabla' \int_V \hat{\mathbf{R}} \cdot \mathbf{e}(\mathbf{R}) \times \left(G^{(e)}(\mathbf{R} \mid \mathbf{R}') - \frac{U(\mathbf{R})U(\mathbf{R}')}{4\pi C}\right) dv. \qquad \text{(C11)}\end{aligned}$$

Since $G^{(e)}(\mathbf{R}_s \mid \mathbf{R}') = 0$ and $U(\mathbf{R}_s) = 1$, this becomes

$$\begin{aligned}\int_V \{\nabla \times [\hat{\mathbf{R}} \times \mathbf{h}]\} \cdot \mathsf{H}_e^{(1)}\, dv &= -\hat{\mathbf{R}}' \times \mathbf{h}(\mathbf{R}') + \int_S \hat{\mathbf{n}}_s \cdot [(\hat{\mathbf{R}}_s \times \mathbf{h}) \times \mathsf{H}_e^{(1)}]\, ds \\ &\quad + \frac{\nabla' U(\mathbf{R}')}{4\pi C} \int_S \hat{\mathbf{n}}_s \cdot (\hat{\mathbf{R}}_s \times \mathbf{h})\, ds \\ &\quad + ikY\nabla' \int_V \hat{\mathbf{R}} \cdot \mathbf{e}\left(G^{(e)} - \frac{U(\mathbf{R})U(\mathbf{R}')}{4\pi C}\right) dv. \qquad \text{(C12)}\end{aligned}$$

With the divergence theorem and (92), this can be written as

$$\begin{aligned}\int_V [\nabla \times (\hat{\mathbf{R}} \times \mathbf{h})] \cdot \mathsf{H}_e^{(1)}\, dv &= -\hat{\mathbf{R}}' \times \mathbf{h}(\mathbf{R}') + \int_S \hat{\mathbf{n}}_s \cdot [(\hat{\mathbf{R}}_s \times \mathbf{h}) \times \mathsf{H}_e^{(1)}]\, ds \\ &\quad + ik \frac{\nabla' U(\mathbf{R}')}{4\pi C} \int_V Y\hat{\mathbf{R}} \cdot \mathbf{e}(\mathbf{R})\, dv \\ &\quad + ikY\nabla' \int_V \hat{\mathbf{R}} \cdot \mathbf{e}\left(G^{(e)} - \frac{U(\mathbf{R})U(\mathbf{R}')}{4\pi C}\right) dv. \qquad \text{(C13)}\end{aligned}$$

Substitution of this expression in (C2) yields the desired result

$$\begin{aligned}\mathbf{e}(\mathbf{R}') &= -ik\int_V (Z\mathbf{h} - \hat{\mathbf{R}} \times \mathbf{e}) \cdot \mathsf{H}_e^{(1)}\, dv \\ &\quad - ik\nabla' \int_V \hat{\mathbf{R}} \cdot \mathbf{e}\left(G^{(e)} + \frac{U(\mathbf{R}') - U(\mathbf{R})U(\mathbf{R}')}{4\pi C}\right) dv \\ &\quad + \int_S (\hat{\mathbf{n}}_s \times \mathbf{e}) \cdot \mathsf{H}_e^{(1)}\, ds. \qquad \text{(C14)}\end{aligned}$$

* This work was supported by the U.S. Air Force and a National Science Foundation Grant.

[1] Lord Rayleigh, Phil. Mag. **44**, 28 (1897).

[2] R. E. Kleinman, Proc. IEEE **53**, 848 (1965).

[3] (a) A. F. Stevenson, J. Appl. Phys. **24**, 1134 (1953); (b) *ibid.*, 1143 (1953).

[4] (a) R. E. Kleinman, "Low Frequency Solutions of Electromagnetic Scattering Problems," in *Electromagnetic Wave Theory* (Delft Symposium) (Pergamon, New York, 1967); (b) J. Appl. Sci. Res. **18**, 108 (1967).

[5] P. Werner, J. Math. Anal. Appl. **7**, 348 (1963).

[6] P. Werner, J. Math. Anal. Appl. **15**, 447 (1966).

[7] In addition to the $2m$ exterior potential problems indicated in (2), there is a subsidiary interior problem to be solved at each stage.

[8] Y and Z are the characteristic admittance and impedance of the medium.

[9] C. H. Wilcox, Commun. Pure Appl. Math. **9**, 115 (1956).

[10] $C^n(D)$ denotes the space of functions with continuous nth-order derivatives at all points of D.

[11] All dyadic relations employed in this work are collected in Appendix A.

[12] A function is said to be regular at infinity in the sense of Kellogg[13] if $|R\,f(\mathbf{R})| < \infty$ and $|R^2\, \partial f(\mathbf{R})/\partial R| < \infty$ as $R \to \infty$.

[13] O. D. Kellogg, *Foundations of Potential Theory* (Springer-Verlag, Berlin, 1929), p. 217.

[14] R. E. Kleinman, Arch. Ratl. Mech. Anal. **18**, 205 (1965).

[15] E. Ar and R. E. Kleinman, Arch. Ratl. Mech. Anal. **23**, 218 (1966).

[16] W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer-Verlag, Berlin, 1966).

[17] C. T. Tai, IRE Professional Group Antennas Propagation **1**, 13 (1952).

[18] The source of all formulas in this appendix is Van Bladel.[19]

[19] J. Van Bladel, *Electromagnetic Fields* (McGraw-Hill, New York, 1964).

[20] A. F. Stevenson, Quart. J. Appl. Math. **12**, 194 (1954).

[21] Actually, an arbitrary function of R' may be added to $N_{mj}^{(i)}(\mathbf{R} \mid \mathbf{R}')$; however, this will not contribute to (B22); see (B62).

[22] In the discussion that follows, the notation of Sec. II will be used: S' denotes a spherical surface over the singularity at $\mathbf{R}$, and the volume enclosed by the surface is excluded from V. Integrals over the surface at infinity (S_∞) will be omitted if it is clear that they vanish. A knowledge of the properties of the dyadics $\mathsf{E}_m^{(1)}$ and $\mathsf{H}_e^{(1)}$ and the notation of Sec. III will also be assumed.

Functional Calculus Theory for Incompressible Fluid Turbulence*

GERALD ROSEN
Drexel University, Philadelphia, Pa. 19104

(Received 14 October 1970)

A functional integral representation for the space-time Hopf characteristic functional is derived from the probability theory for a statistical ensemble of velocity fields that satisfy the Navier-Stokes equation for boundary-free incompressible fluid flow. The functional integral representation involves a pair of real vector field integration variables denoted by **u** and **v**, and the evaluation of the integral is performed in two steps. First, the integration over the field variable **u** is effected exactly in the general case by applying methods of explicit functional integration. Second, the resulting functional integral over the field variable **v** is reduced to a form amenable to specialized analysis by applying a suitable transformation of the integration field variable $\mathbf{v} \rightarrow \mathbf{z}$. Specializing to mathematically defined "*C*-dominant turbulence," the final functional integration over the field variable **z** is effected exactly and yields a characteristic functional of Gaussian form. The two-point velocity correlation tensor for *C*-dominant turbulence is then obtained from the characteristic functional.

I. INTRODUCTION

All modern approaches to the theory of incompressible fluid turbulence take the statistical ensemble hypothesis of Taylor[1,2] as a fundamental postulate and have the common objective of computing physically significant velocity correlation tensors. However, there have emerged two essentially different lines of research which aim at the formulation and solution of a useful statistical theory for incompressible fluid turbulence.

The first line of research has its genesis in the work of Chandrasekhar,[3] the original author to apply an ad hoc closure approximation scheme to the infinite hierarchy of correlation tensor equations and to obtain a nonlinear integro-differential equation for the two-point velocity correlation tensor. Authors have followed Chandrasekhar with more sophisticated ad hoc closure approximation schemes which lead to sets of coupled integro-differential equations for velocity correlation tensors. To what extent such ad hoc closure approximation schemes are meaningful in the context of the exact complete theory has been investigated by Wyld.[4] The detailed analysis of Wyld makes it evident that what is left out is not necessarily small compared to what is included by the various ad hoc closure approximation schemes.

The second line of research has its genesis in the work of Hopf,[5] the original author to derive a functional differential equation for the dynamical evolution of the probability distribution over the statistical ensemble of velocity fields and to obtain a rigorous (closed and complete) mathematical formulation, free of any ad hoc statistical approximation. Authors have followed Hopf with partially successful attempts[6,7] to solve the functional differential equation and to develop the more tractable space-time version of the probability theory.[8] There is no physical deficiency in the Hopf formulation, but mathematical difficulties have been associated with this second line of research owing to the underdeveloped state of functional differential equation theory.

The present paper reports recent mathematical results that advance the second line of research. We have obtained a functional integral representation for the general solution to the Hopf functional differential equation in the space-time version of the theory, and have shown that this integral representation can be evaluated by applying methods of functional calculus.[9] Details of this work are given here.

The organization of the paper is as follows. In Sec. II we fix notation and recast the Navier-Stokes equation for boundary-free incompressible fluid flow into the form of the integral equation (15) which incorporates a generically prescribed initial velocity field. The latter Navier-Stokes integral equation plays a central role in Sec. III, where we develop a general theory for space-time probability distributions and characteristic functionals associated with a statistical ensemble of velocity fields. The functional integral representation (49) for the space-time characteristic functional is derived from the probability theory. In Sec. IV we perform a general partial evaluation of the functional integral (49), the integration over the field variable **u** in (49) being effected exactly in the general case by applying methods of explicit functional integration. The resulting functional integral over the field variable **v** is reduced to a form amenable to specialized analysis by applying a suitable transformation of the integration field variable $\mathbf{v} \rightarrow \mathbf{z}$, and we arrive at expression (69) for the characteristic functional. In Sec. V we complete the evaluation of the characteristic functional for the special case of "*C*-dominant turbulence," defined by the mathematical condition (71). The final functional integration over the field variable $\mathbf{z}$ is effected exactly and yields the characteristic functional (79), which manifests a Gaussian form. We then obtain the two-point velocity correlation tensor (89) for C-dominant turbulence with the forms (81) and (82) for the disposable physical quantities. In the Appendix we derive the general expression (A5) for the two-point velocity correlation tensor associated with weak turbulence.

II. NAVIER-STOKES EQUATION FOR BOUNDARY-FREE FLOW

We consider an unbounded spatial domain with Cartesian coordinates $\mathbf{x} = (x_1, x_2, x_3)$ and a semi-infinite temporal domain with the time coordinate $t \geq 0$. The pressure term can be eliminated from the Navier–Stokes equation for the free flow of an incompressible fluid, and we have the governing dynamical law

$$\mathfrak{L}_x \mathbf{u} + (\mathbf{u} \cdot \nabla \mathbf{u})^{\mathrm{tr}} = 0, \tag{1}$$

with

$$\mathfrak{L}_x \equiv \frac{\partial}{\partial t} - \nu \nabla^2 \tag{2}$$

and the velocity field

$$\mathbf{u} = (u_1(\mathbf{x}, t), u_2(\mathbf{x}, t), u_3(\mathbf{x}, t))$$

satisfying the subsidiary condition that expresses incompressibility of the fluid,

$$\nabla \cdot \mathbf{u} = 0. \tag{3}$$

In Eq. (1), "tr" denotes the transverse (solenoidal) part of the inertial term, the transverse part of a generic vector field $\mathbf{w} = \mathbf{w}(\mathbf{x}, t)$ being defined as

$$\mathbf{w}^{\mathrm{tr}} = \mathbf{w}(\mathbf{x}, t)^{\mathrm{tr}} = \mathbf{w} - \nabla[\nabla^{-2}(\nabla \cdot \mathbf{w})] \equiv \mathbf{w}(\mathbf{x}, t) + \nabla \int \frac{\nabla \cdot \mathbf{w}(\mathbf{y}, t)}{4\pi\, |\mathbf{x} - \mathbf{y}|} d^3\mathbf{y}. \tag{4}$$

Because we have

$$\nabla \cdot [(\mathbf{u} \cdot \nabla)\mathbf{u}]^{\mathrm{tr}} \equiv 0,$$

the subsidiary condition (3) is compatible with the integro-differential dynamical equation (1) for all $t \geq 0$. Assuming that the velocity field is prescribed at $t = 0$,

$$\mathbf{u}(\mathbf{x}, 0) \equiv \hat{\mathbf{u}} = \hat{\mathbf{u}}(\mathbf{x}) \tag{5}$$

with $\hat{\mathbf{u}}$ solenoidal,

$$\nabla \cdot \hat{\mathbf{u}} = 0, \tag{6}$$

it is convenient to recast (1) in the form of an integral equation which incorporates the initial data (5),

$$u_\mu(\mathbf{x}, t) - \int_0^t \int G_{\mu\alpha\beta}(\mathbf{x} - \mathbf{y}, t - s) u_\alpha(\mathbf{y}, s) u_\beta(\mathbf{y}, s) d^3y\, ds = \tilde{u}_\mu(\mathbf{x}, t), \tag{7}$$

where Greek subscript indices (referring to the Cartesian spatial axes) run 1, 2, 3 and the summation convention applies to repeated indices. In (7) we have introduced the quantities

$$\tilde{u}_\mu(\mathbf{x}, t) \equiv (4\pi\nu t)^{-\frac{3}{2}} \int \hat{u}_\mu(\mathbf{y}) \exp(-|\mathbf{x} - \mathbf{y}|^2/4\nu t) d^3y = \pi^{-\frac{3}{2}} \int \hat{u}_\mu(\mathbf{x} + 2(\nu t)^{\frac{1}{2}}\boldsymbol{\lambda}) e^{-|\boldsymbol{\lambda}|^2} d^3\lambda \tag{8}$$

and

$$G_{\mu\alpha\beta}(\mathbf{x}, t) \equiv -\frac{1}{2}\left(\frac{\partial G_{\mu\alpha}(\mathbf{x}, t)}{\partial x_\beta} + \frac{\partial G_{\mu\beta}(\mathbf{x}, t)}{\partial x_\alpha}\right) \tag{9}$$

in which

$$G_{\alpha\beta}(\mathbf{x}, t) \equiv \delta_{\alpha\beta} G(r, t) + \frac{\partial^2 H(r, t)}{\partial x_\alpha \partial x_\beta}, \tag{10}$$

$$G(r, t) \equiv (4\pi\nu t)^{-\frac{3}{2}}[\exp(-r^2/4\nu t)], \quad \text{for} \quad t > 0, \qquad \equiv 0, \quad \text{for} \quad t \leq 0, \tag{11}$$

$$H(r, t) \equiv -\nabla^{-2} G(r, t) = \frac{2\nu t}{r} \int_0^r G(s, t)\, ds = \frac{1}{4\pi^{\frac{3}{2}}(\nu t)^{\frac{1}{2}}} \sum_{n=0}^{\infty} \frac{(-r^2/4\nu t)^n}{n!\,(2n + 1)}, \quad \text{for} \quad t > 0, \qquad = 0, \quad \text{for} \quad t \leq 0, \tag{12}$$

where $r \equiv |\mathbf{x}|$. To prove that (7) is equivalent to Eq. (1) subject to (5), we note that the Green's function (11) satisfies the inhomogeneous diffusion equation

$$\mathfrak{L}_x G(r, t) = \delta(\mathbf{x})\delta(t), \tag{13}$$

while the vector field (8), solenoidal as a consequence of (6), is the solution to the homogeneous diffusion equation

$$\mathfrak{L}_x \tilde{\mathbf{u}} = 0 \tag{14}$$

subject to the initial value $\tilde{\mathbf{u}}(\mathbf{x}, 0) = \hat{\mathbf{u}}(\mathbf{x})$; hence, we obtain Eq. (1) by applying the differential operator (2) to Eq. (7).

There are two important notational simplifications that facilitate analysis based on Eq. (7). First, we abbreviate the space–time coordinates by $x = (\mathbf{x}, t)$ and the space–time infinitesimal volume element by $dx = dx_1\, dx_2\, dx_3\, dt$, so that Eq. (7) takes the form

$$u_\mu(x) - \int G_{\mu\alpha\beta}(x - y) u_\alpha(y) u_\beta(y)\, dy = \tilde{u}_\mu(x), \tag{15}$$

with the space–time integral understood to be over all $\mathbf{y}$ and over the semi-infinite interval for the time component of y, the three-index Green's function (9) vanishing for negative values of the time argument.

Second, we may express Eq. (7) or (15) symbolically as

$$\mathbf{u} - \mathbf{G}:\mathbf{uu} = \tilde{\mathbf{u}}, \tag{16}$$

where the colon denotes a double contraction on tensor indices together with the associated integration over space–time in (15). The symbolic notation in (16) makes it possible to display many of the equations of the theory in a transparent form. For example, the iteration solution series to Eq. (7) or (15) is exhibited neatly as

$$\begin{aligned}\mathbf{u} = \tilde{\mathbf{u}} &+ \mathbf{G}:\tilde{\mathbf{u}}\tilde{\mathbf{u}} + 2\mathbf{G}:\tilde{\mathbf{u}}(\mathbf{G}:\tilde{\mathbf{u}}\tilde{\mathbf{u}}) \\ &+ 4\mathbf{G}:\tilde{\mathbf{u}}(\mathbf{G}:\tilde{\mathbf{u}}(\mathbf{G}:\tilde{\mathbf{u}}\tilde{\mathbf{u}})) + \mathbf{G}:(\mathbf{G}:\tilde{\mathbf{u}}\tilde{\mathbf{u}})(\mathbf{G}:\tilde{\mathbf{u}}\tilde{\mathbf{u}}) \\ &+ (\text{terms of higher order in } \tilde{\mathbf{u}}),\end{aligned} \tag{17}$$

a formal solution which is valid if the $\mathbf{u}$-to-$\tilde{\mathbf{u}}$ correspondence provided by (15) is one-to-one and useful if the quantity $\mathbf{G}:\tilde{\mathbf{u}}\tilde{\mathbf{u}}$ is small compared to $\tilde{\mathbf{u}}$ for all $\mathbf{x}$ and all $t \geq 0$, as for the velocity fields in a "weak turbulence" statistical ensemble (see Appendix A).

III. PROBABILITY DISTRIBUTIONS AND CHARACTERISTIC FUNCTIONALS

Underlying the mathematical description of incompressible fluid turbulence is the postulate that a statistical ensemble of velocity fields can be evoked for the theoretical prediction of observable averages (expectation values) of velocity field components. The initial velocity fields (5) are prescribed statistically in terms of a probability distribution $P_0[\hat{\mathbf{u}}]$, a nonnegative real functional of $\hat{\mathbf{u}} = \hat{\mathbf{u}}(\mathbf{x})$ concentrated on solenoidal fields in accordance with (6), so that

$$(\nabla \cdot \hat{\mathbf{u}})P_0[\hat{\mathbf{u}}] = 0$$

for any arbitrary real vector field $\hat{\mathbf{u}} = \hat{\mathbf{u}}(\mathbf{x})$. With $D(\hat{\mathbf{u}})$ denoting a displacement-invariant infinitesimal volume element in the $\hat{\mathbf{u}}$ function space, the probability of finding the initial velocity field with the specific form $\hat{\mathbf{u}} = \hat{\mathbf{u}}(\mathbf{x})$ is given by $P_0[\hat{\mathbf{u}}]D(\hat{\mathbf{u}})$. A heuristic way of expressing the infinitesimal volume element is

$$D(\hat{\mathbf{u}}) \equiv (\text{const}) \prod_{\text{all } \mathbf{x}} [d\hat{u}_1(\mathbf{x})\, d\hat{u}_2(\mathbf{x})\, d\hat{u}_3(\mathbf{x})],$$

but the property of displacement invariance

$$D(\hat{\mathbf{u}} + \mathbf{w}) = D(\hat{\mathbf{u}}),$$

for any $\mathbf{w} = \mathbf{w}(\mathbf{x})$ independent of $\hat{\mathbf{u}}$, is adequate to fix $D(\hat{\mathbf{u}})$ in practical computations to within a constant numerical prefactor which is subsequently determined by the probability normalization condition

$$\int P_0[\hat{\mathbf{u}}]D(\hat{\mathbf{u}}) = 1. \tag{18}$$

Rigorous mathematical meaning obtains for the left side of (18) as a functional integral over all $\hat{\mathbf{u}}$ with displacement-invariant measure.[10] Likewise, rigorous mathematical meaning obtains for the generic expectation value formula

$$\langle F[\hat{\mathbf{u}}]\rangle \equiv \int F[\hat{\mathbf{u}}]P_0[\hat{\mathbf{u}}]D(\hat{\mathbf{u}}) \tag{19}$$

for a generic functional $F[\hat{\mathbf{u}}]$.

The probability distribution $P_0[\hat{\mathbf{u}}]$ for the statistical ensemble of initial velocity fields at $t = 0$ induces a probability distribution for the statistical ensemble of vector fields (8) in space–time, because $\tilde{\mathbf{u}} = \tilde{\mathbf{u}}(\mathbf{x}, t) \equiv \tilde{\mathbf{u}}(x)$ and $\hat{\mathbf{u}} = \hat{\mathbf{u}}(\mathbf{x})$ are related generically in a one-to-one fashion by (8). To compute the expectation value of an arbitrary functional of $\tilde{\mathbf{u}}$, we simply evoke (8) and formula (19). Thus, for homogeneous turbulence the probability distribution $P_0[\hat{\mathbf{u}}]$ is invariant under translations of space $\mathbf{x} \to \mathbf{x} + \mathbf{a}$, and we find the two-point correlation tensor

$$\begin{aligned} S_{\mu\nu}(x', x'') &\equiv \langle \tilde{u}_\mu(x')\tilde{u}_\nu(x'')\rangle \\ &= \pi^{-3}\left(\frac{\partial^2}{\partial x'_\mu \partial x'_\nu} - \delta_{\mu\nu}\nabla^2_{\mathbf{x}'}\right) \\ &\quad \times \int h(\mathbf{x}' - \mathbf{x}'' + 2(\nu t')^{\frac{1}{2}}\boldsymbol{\lambda}' - 2(\nu t'')^{\frac{1}{2}}\boldsymbol{\lambda}'') \\ &\quad \times e^{-|\boldsymbol{\lambda}'|^2 - |\boldsymbol{\lambda}''|^2} d^3\lambda' d^3\lambda'' \\ &= [4\pi\nu(t' + t'')]^{-\frac{3}{2}}\left(\frac{\partial^2}{\partial x'_\mu \partial x'_\nu} - \delta_{\mu\nu}\nabla^2_{\mathbf{x}'}\right) \\ &\quad \times \int h(\mathbf{x}' - \mathbf{x}'' + \mathbf{y}) \exp\,[-|\mathbf{y}|^2/4\nu(t' + t'')]d^3y, \end{aligned} \tag{20}$$

if the two-point correlation tensor at $t' = t'' = 0$ is prescribed as

$$\langle \hat{u}_\mu(\mathbf{x}')\hat{u}_\nu(\mathbf{x}'')\rangle \equiv \left(\frac{\partial^2}{\partial x'_\mu \partial x'_\nu} - \nabla^2_{\mathbf{x}'}\delta_{\mu\nu}\right) h(\mathbf{x}' - \mathbf{x}''). \tag{21}$$

It should be noted that the real scalar function $h(\mathbf{x})$ in (20) and (21) is required to be nonnegative convex, since we have

$$\begin{aligned}\left\langle \left| \int \hat{\mathbf{u}}(\mathbf{x})f(\mathbf{x})d^3x \right|^2 \right\rangle &= \int \langle \hat{u}_\mu(\mathbf{x}')\hat{u}_\mu(\mathbf{x}'')\rangle f(\mathbf{x}')f(\mathbf{x}'')d^3x'd^3x'' \\ &= 2\int h(\mathbf{x}' - \mathbf{x}'')\frac{\partial f(\mathbf{x}')}{\partial x'_\mu}\frac{\partial f(\mathbf{x}'')}{\partial x''_\mu}\,d^3x'd^3x'' \geq 0\end{aligned}$$

for all arbitrary real scalar functions $f(\mathbf{x})$ that vanish at spatial infinity. Indirectly then, via (8) and (19), the space–time vector fields $\tilde{\mathbf{u}}$ are endowed with a probability distribution $\hat{\mathbb{P}}[\tilde{\mathbf{u}}]$, a nonnegative real functional of $\tilde{\mathbf{u}} = \tilde{\mathbf{u}}(x)$ concentrated on solenoidal fields satisfying the homogeneous diffusion equation (14),

$$(\nabla \cdot \tilde{\mathbf{u}}(x))\hat{\mathbb{P}}[\tilde{\mathbf{u}}] = 0 \quad \text{and} \quad [\mathfrak{L}_x \tilde{\mathbf{u}}(x)]\hat{\mathbb{P}}[\tilde{\mathbf{u}}] = 0$$

for any arbitrary real vector field $\tilde{\mathbf{u}} = \tilde{\mathbf{u}}(x)$. The probability of finding the vector field (8) with the specific form $\tilde{\mathbf{u}} = \tilde{\mathbf{u}}(\mathbf{x}, t)$ is given by $\hat{\mathbb{P}}[\tilde{\mathbf{u}}]\mathfrak{D}(\tilde{\mathbf{u}})$, where $\mathfrak{D}(\tilde{\mathbf{u}})$ denotes a displacement-invariant infinitesimal volume element in the function space of the $\tilde{\mathbf{u}}$'s. In heuristic symbolic notation, we have

$$\mathfrak{D}(\tilde{\mathbf{u}}) = (\text{const}) \prod_{\text{all } \mathbf{x}} [d\tilde{u}_1(\mathbf{x})\, d\tilde{u}_2(\mathbf{x})\, d\tilde{u}_3(\mathbf{x})]$$

with $\mathfrak{D}(\tilde{\mathbf{u}} + \mathbf{w}) = \mathfrak{D}(\tilde{\mathbf{u}})$ for any $\mathbf{w} = \mathbf{w}(x)$ independent of $\tilde{\mathbf{u}}$, and the probability normalization condition is

$$\int \hat{\mathbb{P}}[\tilde{\mathbf{u}}]\mathfrak{D}(\tilde{\mathbf{u}}) = 1, \tag{22}$$

a functional integral over all $\tilde{\mathbf{u}}$. From the one-to-one correspondence provided by (8), it follows that

$$\hat{\mathbb{P}}[\tilde{\mathbf{u}}]\mathfrak{D}(\tilde{\mathbf{u}}) = P[\hat{\mathbf{u}}]D(\hat{\mathbf{u}}) \tag{23}$$

for a pair of fields related by (8). Hence, the expectation value of an arbitrary functional of $\tilde{\mathbf{u}}$ can be computed by evoking (8) and (19), or, alternatively, by using the generic formula

$$\langle F[\tilde{\mathbf{u}}]\rangle \equiv \int F[\tilde{\mathbf{u}}]\hat{\mathbb{P}}[\tilde{\mathbf{u}}]\mathfrak{D}(\tilde{\mathbf{u}}). \tag{24}$$

In particular, we have the characteristic functional associated with $\hat{\mathbb{P}}[\tilde{\mathbf{u}}]$ given by

$$\hat{\Phi}[\mathbf{v}] \equiv \left\langle \exp i \int \mathbf{v}(x) \cdot \tilde{\mathbf{u}}(x)\, dx \right\rangle = \int \left(\exp i \int \mathbf{v}(x) \cdot \tilde{\mathbf{u}}(x)\, dx\right) \hat{\mathbb{P}}[\tilde{\mathbf{u}}]\mathfrak{D}(\tilde{\mathbf{u}}), \tag{25}$$

where the ordinary space–time integration in the exponential is understood to be over all $\mathbf{x}$ and all $t \geq 0$. The characteristic functional (25) is therefore a functional Fourier transform of the probability distribution $\hat{\mathbb{P}}[\tilde{\mathbf{u}}]$. Expectation values (24) are extracted from (25) by functional differentiation,

$$\langle F[\tilde{\mathbf{u}}]\rangle = (F[-i\delta/\delta\mathbf{v}]\hat{\Phi}[\mathbf{v}])_{\mathbf{v}=0}. \tag{26}$$

Thus, for example, the two-point correlation tensor (20) is obtained from $\hat{\Phi}[\mathbf{v}]$ as

$$S_{\mu\nu}(x', x'') = -\left(\frac{\delta^2\hat{\Phi}[\mathbf{v}]}{\delta v_\mu(x')\delta v_\nu(x'')}\right)_{\mathbf{v}=0}. \tag{27}$$

Because the vector fields $\tilde{\mathbf{u}} = \tilde{\mathbf{u}}(x)$ that make a finite contribution to the functional integral (25) satisfy (14) and are solenoidal, we have the characteristic functional (25) satisfying the equations

$$\mathfrak{L}_x \frac{\delta\hat{\Phi}[\mathbf{v}]}{\delta v_\mu(x)} = 0, \tag{28}$$

$$\frac{\partial}{\partial x_\mu} \frac{\delta\hat{\Phi}[\mathbf{v}]}{\delta v_\mu(x)} = 0. \tag{29}$$

In addition to being a solution to Eqs. (28) and (29), the characteristic functional (25) must satisfy certain holonomic and nonholonomic conditions, such as

$$\hat{\Phi}[\mathbf{0}] = 1, \quad |\hat{\Phi}[\mathbf{v}]| \leq 1, \quad \hat{\Phi}[\mathbf{v}]^* = \hat{\Phi}[-\mathbf{v}], \quad \text{etc.},$$

which stem from the fact that $\hat{\mathbb{P}}[\tilde{\mathbf{u}}]$ is real, nonnegative, and normalized with respect to $\mathfrak{D}(\tilde{\mathbf{u}})$ according to (22). Corresponding to a Gaussian probability distribution, we have a characteristic functional of the form

$$\hat{\Phi}[\mathbf{v}] = \exp -\tfrac{1}{2}\iint v_\alpha(x')S_{\alpha\beta}(x', x'')v_\beta(x'')\, dx'\, dx'', \tag{30}$$

in which $S_{\mu\nu}(x', x'') \equiv S_{\nu\mu}(x'', x')$ is a real symmetric tensor. Since the functional derivative of (30) is

$$\frac{\delta\hat{\Phi}[\mathbf{v}]}{\delta v_\mu(x)} = -\int S_{\mu\beta}(x, x')v_\beta(x')\, dx' \hat{\Phi}[\mathbf{v}], \tag{31}$$

it follows that (28) and (29) are satisfied if

$$\frac{\partial S_{\mu\nu}(x', x'')}{\partial x'_\mu} = 0 = \frac{\partial S_{\mu\nu}(x', x'')}{\partial x''_\nu} \tag{32}$$

and

$$\mathfrak{L}_{x'} S_{\mu\nu}(x', x'') = 0 = \mathfrak{L}_{x''} S_{\mu\nu}(x', x''). \tag{33}$$

The solenoidal property of $S_{\mu\nu}(x', x'')$ displayed by (32) shows that the characteristic functional (30) depends only on the transverse part of $\mathbf{v}$, $\hat{\Phi}[\mathbf{v}] \equiv \hat{\Phi}[\mathbf{v}^{tr}]$ for all $\mathbf{v}$, while (33) shows that $\hat{\Phi}[\mathbf{v} + \mathfrak{L}_x^\dagger \mathbf{w}] = \hat{\Phi}[\mathbf{v}]$ for all $\mathbf{w} = \mathbf{w}(x)$ with $\mathfrak{L}_x^\dagger$ denoting the formal adjoint of the operator (2). For homogeneous turbulence the $S_{\mu\nu}(x', x'')$ in (30) equals the two-point correlation tensor (20) by virtue of (27), and it is readily verified that the final member in (20) satisfies Eqs. (32) and (33). Moreover, $S_{\mu\nu}(x', x'')$ is then a positive-definite matrix kernel with respect to solenoidal $\mathbf{v}$ ($\equiv \mathbf{v}^{tr}$) in (30), because $h(\mathbf{x})$ is a nonnegative convex real function in (20).

The probability distribution $P_0[\hat{\mathbf{u}}]$ for the statistical ensemble of initial velocity fields at $t = 0$ also induces a probability distribution for the statistical ensemble of physical velocity fields $\mathbf{u} = \mathbf{u}(x)$ in space–time, $\mathbb{P}[\mathbf{u}]$, a probability distribution concentrated on solenoidal fields satisfying the Navier–Stokes equation (1),

$$(\nabla \cdot \mathbf{u})\mathbb{P}[\mathbf{u}] = 0 \quad \text{and} \quad \{\mathfrak{L}_x \mathbf{u} + [(\mathbf{u} \cdot \nabla)\mathbf{u}]^{tr}\}\mathbb{P}[\mathbf{u}] = 0$$

for any arbitrary real vector field $\mathbf{u} = \mathbf{u}(x)$. By definition, the probability of finding the velocity field with the specific form $\mathbf{u} = \mathbf{u}(x)$ is $\mathbb{P}[\mathbf{u}]\mathfrak{D}(\mathbf{u})$. If for a prescribed $\hat{\mathbf{u}} = \hat{\mathbf{u}}(\mathbf{x})$ a unique $\mathbf{u} = \mathbf{u}(x)$ exists as a solution to Eq. (1) for all $t \geq 0$, or as a solution to Eq. (15) for all $t \geq 0$ with $\tilde{\mathbf{u}} = \tilde{\mathbf{u}}(x)$ prescribed, then the correspondence between $\mathbf{u}$ and $\tilde{\mathbf{u}}$ is one-to-one, and it follows that the probability $\mathbb{P}[\mathbf{u}]\mathfrak{D}(\mathbf{u})$ equals the probability $\hat{\mathbb{P}}[\tilde{\mathbf{u}}]\mathfrak{D}(\tilde{\mathbf{u}})$ for a pair of fields related by (15). The mathematical theory for the Navier–Stokes initial value problem[11] does indeed suggest existence of a unique solution for all $t \geq 0$ if the initial velocity field at $t = 0$ is suitably smooth. However, the precise form of a smoothness condition on the initial velocity field $\hat{\mathbf{u}} = \hat{\mathbf{u}}(\mathbf{x})$ for existence of a unique solution $\mathbf{u} = \mathbf{u}(x)$ for all $t \geq 0$ has not been established, and therefore it is unknown whether a unique solution is always associated with an initial velocity field realizable in nature. Taking into account the possibility of nonuniqueness, which would feature a local breakdown of regularity at a finite value of t and bifurcation of certain solutions to the Navier–Stokes initial value problem, we write

$$e^{A[\mathbf{u}]}\mathbb{P}[\mathbf{u}]\mathfrak{D}(\mathbf{u}) = \hat{\mathbb{P}}[\tilde{\mathbf{u}}]\mathfrak{D}(\tilde{\mathbf{u}}), \qquad (34)$$

where $A[\mathbf{u}]$ is a real functional of $\mathbf{u}$ that vanishes identically if and only if the $\mathbf{u}$-to-$\tilde{\mathbf{u}}$ correspondence provided by (15) is one-to-one. More generally, the functional $A[\mathbf{u}]$ must be indefinite in sign as $\mathbf{u}$ ranges over the statistical ensemble of velocity fields in order to admit the normalization condition

$$\int \mathbb{P}[\mathbf{u}]\mathfrak{D}(\mathbf{u}) = 1 \qquad (35)$$

with $\hat{\mathbb{P}}[\tilde{\mathbf{u}}]$ normalized according to (22) and requiring

$$\int e^{A[\mathbf{u}]}\mathbb{P}[\mathbf{u}]\mathfrak{D}(\mathbf{u}) \equiv \langle e^{A[\mathbf{u}]}\rangle = 1. \qquad (36)$$

The displacement-invariant infinitesimal volume elements in (34) are related by a function-space determinant

$$\mathfrak{D}(\tilde{\mathbf{u}}) = [\det(\delta\tilde{u}_\mu(x')/\delta u_\nu(x''))]\mathfrak{D}(\mathbf{u}), \qquad (37)$$

in which $[\det(\delta\tilde{u}_\mu(x')/\delta u_\nu(x''))]$ is defined as the product of all eigenvalues of the matrix kernel

$$\delta\tilde{u}_\mu(x')/\delta u_\nu(x'') = \delta_{\mu\nu}\delta(x' - x'') - 2G_{\mu\nu\alpha}(x' - x'')u_\alpha(x''). \qquad (38)$$

Since the three-index Green's function (9) vanishes if the time-argument is not positive, we have

$$G_{\mu\nu\alpha}(x' - x'') = 0 \quad \text{for} \quad t' \leq t'', \qquad (39)$$

and the function-space determinant of the matrix kernel (38) equals unity, as shown in convenient symbolic notation by the elementary calculation

$$\begin{aligned}
&\det(\delta\tilde{u}_\mu(x')/\delta u_\nu(x'')) \\
&= \det(\mathbb{1} - 2\mathbf{G}\cdot\mathbf{u}) \\
&= \det\{\exp[\ln(\mathbb{1} - 2\mathbf{G}\cdot\mathbf{u})]\} \\
&= \exp\{\mathrm{sp}[\ln(\mathbb{1} - 2\mathbf{G}\cdot\mathbf{u})]\} \\
&= \exp\{\mathrm{sp}[-2\mathbf{G}\cdot\mathbf{u} - 2(\mathbf{G}\cdot\mathbf{u})^2 + O((\mathbf{G}\cdot\mathbf{u})^3)]\} \\
&= 1 - \mathrm{sp}(\mathbf{G}\cdot\mathbf{u}) + 2[\mathrm{sp}(\mathbf{G}\cdot\mathbf{u})]^2 - 2\,\mathrm{sp}(\mathbf{G}\cdot\mathbf{u})^2 \\
&\quad + [\text{higher order terms in sp}(\mathbf{G}\cdot\mathbf{u}), \\
&\qquad \mathrm{sp}(\mathbf{G}\cdot\mathbf{u})^2, \mathrm{sp}(\mathbf{G}\cdot\mathbf{u})^3, \cdots] = 1. \qquad (40)
\end{aligned}$$

In (40), "sp" denotes the function–space spur (trace) of a matrix kernel, obtained by contracting the indices, setting the space–time coordinate arguments equal, and integrating over the space–time region (all $\mathbf{x}$ and all $t \geq 0$); thus, for example, we have

$$\mathrm{sp}(\mathbf{G}\cdot\mathbf{u}) \equiv \int G_{\mu\mu\alpha}(0)u_\alpha(x)\,dx = 0, \qquad (41)$$

and

$$\mathrm{sp}(\mathbf{G}\cdot\mathbf{u})^2 \equiv \int G_{\mu\nu\alpha}(x' - x'')u_\alpha(x'') \times G_{\nu\mu\beta}(x'' - x')u_\beta(x')\,dx'\,dx'' = 0. \qquad (42)$$

Hence, (37) reduces to an equality of the infinitesimal volume elements, $\mathfrak{D}(\mathbf{u}) = \mathfrak{D}(\tilde{\mathbf{u}})$, and (34) becomes

$$\begin{aligned}
e^{A[\mathbf{u}]}\mathbb{P}[\mathbf{u}] &= \hat{\mathbb{P}}[\tilde{\mathbf{u}}] = \hat{\mathbb{P}}[\mathbf{u} - G:\mathbf{uu}] \\
&= \int \delta[\tilde{\mathbf{u}} - \mathbf{u} + \mathbf{G}:\mathbf{uu}]\hat{\mathbb{P}}[\tilde{\mathbf{u}}]\mathfrak{D}(\tilde{\mathbf{u}}), \qquad (43)
\end{aligned}$$

where $\delta[\mathbf{w}]$ is the δ-functional with respect to the infinitesimal volume element $\mathfrak{D}(\mathbf{w})$,

$$\delta[\mathbf{w}] = 0 \quad \text{for} \quad \mathbf{w} = \mathbf{w}(x) \not\equiv 0, \qquad (44)$$

$$\int \delta[\mathbf{w}]\mathfrak{D}(\mathbf{w}) = 1. \qquad (45)$$

The functional integration is over all real vector fields in (43), (45), and in the functional integral representation of the δ-functional

$$\delta[\mathbf{w}] = (\text{const})\int\left(\exp i\int \mathbf{v}(x)\cdot\mathbf{w}(x)\,dx\right)\mathfrak{D}(\mathbf{v}). \qquad (46)$$

In (46) and subsequent equations, it is understood that ordinary space–time integrations with the infinitesimal volume element $dx \equiv dx_1\,dx_2\,dx_3\,dt$ are over all $\mathbf{x}$ and all $t \geq 0$. By putting (46) into the final member of (43) and recalling definition (25), we obtain

$$\mathbb{P}[\mathbf{u}] = (\text{const}) \times e^{-A[\mathbf{u}]} \int \left(\exp i \int \mathbf{v}(x) \cdot (-\mathbf{u} + \mathbf{G}:\mathbf{uu})(x)\, dx\right) \times \hat{\Phi}[\mathbf{v}]\mathfrak{D}(\mathbf{v}). \quad (47)$$

It follows from (47) that the characteristic functional associated with $\mathbb{P}[\mathbf{u}]$,

$$\Phi[\mathbf{y}] \equiv \left\langle \exp i \int \mathbf{y}(x) \cdot \mathbf{u}(x)\, dx \right\rangle = \int \left(\exp i \int \mathbf{y}(x) \cdot \mathbf{u}(x)\, dx\right) \mathbb{P}[\mathbf{u}] D(\mathbf{u}), \quad (48)$$

can be expressed as

$$\Phi[\mathbf{y}] = \iint \left[\exp\left(i \int \{[\mathbf{y}(x) - \mathbf{v}(x)] \cdot \mathbf{u}(x) + \mathbf{v}(x) \cdot (\mathbf{G}:\mathbf{uu})(x)\}\, dx - A[\mathbf{u}]\right)\right] \times \hat{\Phi}[\mathbf{v}]\mathfrak{D}(\mathbf{u})\mathfrak{D}(\mathbf{v}), \quad (49)$$

where a numerical prefactor constant has been absorbed into the product of the displacement-invariant infinitesimal volume elements with the over-all normalization of (49) fixed by (35) as

$$\Phi[0] = 1. \quad (50)$$

Because the probability distribution $\mathbb{P}[\mathbf{u}]$ is concentrated on solutions to Eqs. (1) and (3), the characteristic functional $\Phi[\mathbf{y}]$ must satisfy the Hopf equation

$$\mathfrak{L}_x \frac{\delta\Phi[\mathbf{y}]}{\delta y_\mu(x)} - i\left(\frac{\partial}{\partial x_\alpha} \frac{\delta^2\Phi[\mathbf{y}]}{\delta y_\alpha(x)\delta y_\mu(x)}\right)^{\text{tr}} = 0 \quad (51)$$

subject to the subsidiary condition

$$\frac{\partial}{\partial x_\alpha} \frac{\delta\Phi[\mathbf{y}]}{\delta y_\alpha(x)} = 0. \quad (52)$$

That the functional integral representation of $\Phi[\mathbf{y}]$ given by (49) satisfies Eq. (51) is verified by computation of the functional derivatives and an application of the functional integration by parts lemma[12]; that (49) satisfies (52) (or equivalently, that the characteristic functional depends only on the transverse part of $\mathbf{y}$, so that $\Phi[\mathbf{y}] \equiv \Phi[\mathbf{y}^{\text{tr}}]$ for all $\mathbf{y}$) is verified by noting the solenoidal quality of (9),

$$\frac{\partial G_{\mu\alpha\beta}(\mathbf{x}, t)}{\partial x_\mu} = 0, \quad (53)$$

the property implied by (29) $\hat{\Phi}[\mathbf{v}] \equiv \hat{\Phi}[\mathbf{v}^{\text{tr}}]$ for all $\mathbf{v}$, and the displacement-invariance of the infinitesimal volume element $\mathfrak{D}(\mathbf{v})$. Since the two-point velocity correlation tensor

$$R_{\mu\nu}(x', x'') \equiv \langle u_\mu(x')u_\nu(x'')\rangle \equiv \int u_\mu(x')u_\nu(x'')\mathbb{P}[\mathbf{u}]\mathfrak{D}(\mathbf{u}) = -\left(\frac{\delta^2\Phi[\mathbf{y}]}{\delta y_\mu(x')\delta y_\nu(x'')}\right)_{\mathbf{y}=0} \quad (54)$$

and all other expectation values

$$\langle F[\mathbf{u}]\rangle \equiv \int F[\mathbf{u}]\mathbb{P}[\mathbf{u}]\mathfrak{D}(\mathbf{u}) = (F[-i\delta/\delta\mathbf{y}]\Phi[\mathbf{y}])_{\mathbf{y}=0} \quad (55)$$

are obtainable immediately from a closed-form expression for the characteristic functional (48), the closed-form evaluation of the functional integral representation (49) enables one to predict all observable averages that are associated with a turbulent fluid velocity field.

IV. EVALUATION OF THE CHARACTERISTIC FUNCTIONAL INTEGRAL REPRESENTATION

By introducing the symmetric tensor field

$$M_{\alpha\beta}(x) \equiv \int v_\mu(x')G_{\mu\alpha\beta}(x' - x)\, dx' = \frac{1}{2}\int \left(\frac{\partial v_\alpha^{\text{tr}}(x')}{\partial x'_\beta} + \frac{\partial v_\beta^{\text{tr}}(x')}{\partial x'_\alpha}\right) G(x' - x)\, dx', \quad (56)$$

the functional integral representation (49) takes the form

$$\Phi[\mathbf{y}] = \iint \left[\exp\left(i \int \{[y_\alpha(x) - v_\alpha(x)]u_\alpha(x) + M_{\alpha\beta}(x)u_\alpha(x)u_\beta(x)\}\, dx - A[\mathbf{u}]\right)\right] \times \hat{\Phi}[\mathbf{v}]\mathfrak{D}(\mathbf{u})\mathfrak{D}(\mathbf{v}). \quad (57)$$

A study of the Navier–Stokes initial value problem[11] suggests that the generic form

$$A[\mathbf{u}] = \int C_{\alpha\beta}(x)u_\alpha(x)u_\beta(x)\, dx - \ln k \quad (58)$$

is a plausible approximation for $A[\mathbf{u}]$ in (57), where $C_{\alpha\beta}(x)$ is a real symmetric matrix field that either vanishes identically (in case the $\mathbf{u}$-to-$\tilde{\mathbf{u}}$ correspondence provided by (15) is one-to-one) or is positive-definite for all x, and k (≥ 1) is a real constant prescribed by (36) as

$$k \equiv \left\langle \exp \int C_{\alpha\beta}(x)u_\alpha(x)u_\beta(x)\, dx \right\rangle. \quad (59)$$

With $A[\mathbf{u}]$ of the form (58), the functional integration over $\mathbf{u}$ in (57) can be performed exactly to yield

$$\Phi[\mathbf{y}] = k\int\Big(\exp -\tfrac{1}{4}\int[y_\alpha(x) - v_\alpha(x)] \times Q^{-1}_{\alpha\beta}(x)[y_\beta(x) - v_\beta(x)]\,dx\Big)\Delta[\mathbf{v}]\hat{\Phi}[\mathbf{v}]\mathfrak{D}(\mathbf{v}), \tag{60}$$

in which $Q^{-1}_{\alpha\beta}(x)$ is the inverse of the complex symmetric matrix

$$Q_{\alpha\beta}(x) \equiv C_{\alpha\beta}(x) - iM_{\alpha\beta}(x), \tag{61}$$

and the quantity

$$\Delta[\mathbf{v}] \equiv \Big(\prod_{\text{all } x}|Q_{\alpha\beta}(x)|\Big)^{-\frac{1}{2}} \tag{62}$$

is defined in terms of a formal product of the determinant of $Q_{\alpha\beta}(x)$ at all x, with a multiplying constant absorbed into $\mathfrak{D}(\mathbf{v})$. Let us now transform the integration variable in (60) from the real-valued field $\mathbf{v} = \mathbf{v}(x)$ to the complex-valued field $\mathbf{z} = \mathbf{z}(x)$, where

$$z_\alpha(x) \equiv Q^{-\frac{1}{2}}_{\alpha\beta}(x)[v_\beta(x) - y_\beta(x)], \tag{63}$$

in which $Q^{-\frac{1}{2}}_{\alpha\beta}(x)$ is a complex symmetric matrix that squares to the inverse of the complex symmetric matrix (61),

$$Q^{-\frac{1}{2}}_{\alpha\beta}(x) \equiv Q^{-\frac{1}{2}}_{\beta\alpha}(x), \quad Q^{-\frac{1}{2}}_{\alpha\gamma}(x)Q^{-\frac{1}{2}}_{\beta\gamma}(x) \equiv Q^{-1}_{\alpha\beta}(x). \tag{64}$$

We evaluate the function–space determinant

$$\det[\delta z_\mu(x')/\delta v_\nu(x'')] = \det\Big(Q^{-\frac{1}{2}}_{\mu\nu}(x')\delta(x' - x'') + \frac{\delta Q^{-\frac{1}{2}}_{\mu\beta}(x')}{\delta v_\nu(x'')}[v_\beta(x') - y_\beta(x')]\Big) \tag{65}$$

by first noting that (61) and (56) gives

$$\frac{\delta Q_{\alpha\beta}(x')}{\delta v_\nu(x'')} = -i\frac{\delta M_{\alpha\beta}(x')}{\delta v_\nu(x'')} = -iG_{\nu\alpha\beta}(x'' - x') = 0 \quad \text{for} \quad t'' \le t'. \tag{66}$$

Thus, the algebraic relationship between $Q^{-\frac{1}{2}}_{\alpha\beta}(x)$ and $Q_{\alpha\beta}(x)$ implies that

$$\frac{\delta Q^{-\frac{1}{2}}_{\mu\beta}(x')}{\delta v_\nu(x'')} = 0 \quad \text{for} \quad t'' \le t',$$

and therefore, by performing an elementary calculation similar to (40), we find that (65) produces

$$\det[\delta z_\mu(x')/\delta v_\nu(x'')] = \prod_{\text{all } x}|Q^{-\frac{1}{2}}_{\mu\nu}(x)| = \Delta[\mathbf{v}]. \tag{67}$$

Hence we have

$$\mathfrak{D}(\mathbf{z}) = [\det(\delta z_\mu(x')/\delta v_\nu(x''))]\mathfrak{D}(\mathbf{v}) = \Delta[\mathbf{v}]\mathfrak{D}(\mathbf{v}), \tag{68}$$

from which it follows that (60), expressed in terms of the integration variable (63), is given by

$$\Phi[\mathbf{y}] = k\int_{\mathbb{K}}\Big(\exp -\tfrac{1}{4}\int z_\alpha(x)z_\alpha(x)\,dx\Big)\hat{\Phi}[\mathbf{y} + \mathbf{Q}^{\frac{1}{2}}\mathbf{z}]\mathfrak{D}(\mathbf{z}). \tag{69}$$

In (69), $\mathbf{Q}^{\frac{1}{2}} = [Q^{\frac{1}{2}}_{\alpha\beta}(x)]$ is a complex symmetric matrix that squares to the complex symmetric matrix (61), and $\mathbb{K}$ is the class of all complex-valued $\mathbf{z} = \mathbf{z}(x)$ such that $\mathbf{Q}^{\frac{1}{2}}\mathbf{z} = (\mathbf{Q}^{\frac{1}{2}}\mathbf{z})^*$ is real. By combining Eqs. (61), (56), and (63), we obtain a quadratic integral equation for $\mathbf{Q}^{\frac{1}{2}}$ in terms of $\mathbf{y}$ and $\mathbf{z}$,

$$Q^{\frac{1}{2}}_{\alpha\gamma}(x)Q^{\frac{1}{2}}_{\beta\gamma}(x) \equiv Q_{\alpha\beta}(x) = C_{\alpha\beta}(x) - i\int[Q^{\frac{1}{2}}_{\mu\nu}(x')z_\nu(x') + y_\mu(x')] \times G_{\mu\alpha\beta}(x' - x)\,dx'. \tag{70}$$

Solution of (70) for $\mathbf{Q}^{\frac{1}{2}}$ allows the functional integrand in (69) to be expressed explicitly in terms of $\mathbf{y}$ and $\mathbf{z}$, and then the functional integral can be evaluated by exact or approximate techniques.

A detailed study shows that the general solution to the integral equation (70) is unobtainable in closed form. Thus it is necessary to consider special approximate solutions for $\mathbf{Q}^{\frac{1}{2}}$, associated with distinct generic cases for $C_{\alpha\beta}(x)$, in order to proceed with the final evaluation of (69).

V. C-DOMINANT TURBULENCE

Here we consider the most immediate special approximate solutions admitted by (70), namely those valid in cases for which the $C_{\alpha\beta}(x)$ term is dominant over the integral term on the right side of (70) for the significant range of values of the components of $\mathbf{z}$ and $\mathbf{y}$ in (69). Such "C-dominant turbulence" is characterized by the approximate solution to (70)

$$Q_{\alpha\beta}(x) \cong C_{\alpha\beta}(x), \tag{71}$$

and the positive-definite nature of $C_{\alpha\beta}(x)$ then implies that $Q^{\frac{1}{2}}_{\alpha\beta}(x) \cong C^{\frac{1}{2}}_{\alpha\beta}(x)$ is real and can be taken to be positive definite. More explicitly, by expressing $C_{\alpha\beta}(x)$ in terms of its real positive eigenvalues $\lambda_i(x)$ and real eigenvectors $\epsilon_{i\alpha}(x)$,

$$C_{\alpha\beta}(x) = \sum_{i=1}^{3}\lambda_i(x)\epsilon_{i\alpha}(x)\epsilon_{i\beta}(x), \tag{72}$$

we have

$$Q^{\frac{1}{2}}_{\alpha\beta}(x) \cong \sum_{i=1}^{3}(\lambda_i(x))^{\frac{1}{2}}\epsilon_{i\alpha}(x)\epsilon_{i\beta}(x). \tag{73}$$

It follows that the class $\mathbb{K}$ of $\mathbf{z}$ for the functional integration in (69) is the function space of all real fields $\mathbf{z} = \mathbf{z}^*$. Putting the form (30) into (69), we obtain

$$\Phi[\mathbf{y}] \cong k\left(\exp -\tfrac{1}{2}\iint y_\alpha(x')S_{\alpha\beta}(x', x'') \times y_\beta(x'')\, dx'\, dx''\right)I[\mathbf{y}], \quad (74)$$

in which the functional integral that remains is

$$\begin{aligned} I[\mathbf{y}] &\equiv \int\Big[\exp\Big(-\tfrac{1}{2}\iint z_\alpha(x')K_{\alpha\beta}(x', x'')z_\beta(x'')\, dx'\, dx'' \\ &\quad - \iint y_\alpha(x')S_{\alpha\beta}(x', x'') \times C^{\frac{1}{2}}_{\beta\gamma}(x'')z_\gamma(x'')\, dx'\, dx''\Big)\Big]\mathfrak{D}(\mathbf{z}) \\ &= (\text{const})\{\det[K_{\alpha\beta}(x', x'')]\}^{-\frac{1}{2}} \\ &\quad \times \left(\exp \tfrac{1}{2}\iint y_\alpha(x')T_{\alpha\beta}(x', x'')y_\beta(x'')\, dx'\, dx''\right). \end{aligned} \quad (75)$$

We have evaluated the functional integral in (75) by evoking the displacement-invariance of the infinitesimal volume element to reduce the integral to a standard form,[13] with the real symmetric kernels that appear in the final member defined by

$$K_{\alpha\beta}(x', x'') \equiv \tfrac{1}{2}\delta_{\alpha\beta}\delta(x' - x'') + C^{\frac{1}{2}}_{\alpha\gamma}(x')S_{\gamma\delta}(x', x'')C^{\frac{1}{2}}_{\delta\beta}(x'') \quad (76)$$

and

$$T_{\alpha\beta}(x', x'') \equiv \iint S_{\alpha\gamma}(x', x''')C^{\frac{1}{2}}_{\gamma\delta}(x''')K^{-1}_{\delta\epsilon}(x''', x'''') \times C^{\frac{1}{2}}_{\epsilon\eta}(x'''')S_{\eta\beta}(x'''', x'')\, dx'''\, dx''''. \quad (77)$$

In the latter definition the real symmetric kernel $K^{-1}_{\alpha\beta}(x', x'')$ is the inverse of (76), defined implicitly by the equation

$$\int K^{-1}_{\alpha\gamma}(x', x)K_{\gamma\beta}(x, x'')\, dx = \delta_{\alpha\beta}\delta(x' - x''). \quad (78)$$

The product of all eigenvalues of the symmetric kernel (76) appears in (75) as $\{\det[K_{\alpha\beta}(x', x'')]\}$, a quantity independent of $\mathbf{y}$ by virtue of the fact that (76) is independent of $\mathbf{y}$. Hence, the substitution of (75) into (74) and the normalization condition (50) produces the result

$$\Phi[\mathbf{y}] \cong \exp\left(-\tfrac{1}{2}\iint y_\alpha(x')[S_{\alpha\beta}(x', x'') - T_{\alpha\beta}(x', x'')] \times y_\beta(x'')\, dx'\, dx''\right), \quad (79)$$

which shows that the probability distribution is approximately Gaussian with the two-point velocity correlation tensor (54) given by

$$R_{\mu\nu}(x', x'') = S_{\mu\nu}(x', x'') - T_{\mu\nu}(x', x''). \quad (80)$$

We now specialize to the simplest possible forms for $C_{\alpha\beta}(x)$ and for $h(\mathbf{x})$ in (20) and (21), namely,

$$C_{\alpha\beta}(x) = \lambda\delta_{\alpha\beta}, \quad (81)$$

$$h(\mathbf{x}) = \xi\delta(\mathbf{x}), \quad (82)$$

in which λ and ξ are positive constant physical parameters. With the forms (81) and (82), we find that (76), (77), and (20) become

$$K_{\alpha\beta}(x', x'') = \tfrac{1}{2}\delta_{\alpha\beta}\delta(x' - x'') + \lambda S_{\alpha\beta}(x', x''), \quad (83)$$

$$T_{\alpha\beta}(x', x'') = \lambda\int S_{\alpha\gamma}(x', x''')K^{-1}_{\gamma\delta}(x''', x'''') \times S_{\delta\beta}(x'''', x'')\, dx'''\, dx'''', \quad (84)$$

$$\begin{aligned} S_{\mu\nu}(x', x'') &= \xi\left(\frac{\partial^2}{\partial x'_\mu \partial x'_\nu} - \delta_{\mu\nu}\nabla^2_{\mathbf{x}'}\right)G(|\mathbf{x}' - \mathbf{x}''|, t' + t'') \\ &= (4\pi)^{-\frac{3}{2}}\xi\{[\nu(t' + t'')]^{-\frac{5}{2}}\delta_{\mu\nu} + \tfrac{1}{4}[\nu(t' + t'')]^{-\frac{7}{2}} \\ &\quad \times [(x'_\mu - x''_\mu)(x'_\nu - x''_\nu) - (x'_\sigma - x''_\sigma)(x'_\sigma - x''_\sigma)\delta_{\mu\nu}]\} \\ &\quad \times \{\exp[-(x'_\sigma - x''_\sigma)(x'_\sigma - x''_\sigma)/4\nu(t' + t'')]\}, \end{aligned} \quad (85)$$

where definition (11) is recalled for the second member in the two-point correlation tensor (85). It follows from (85) that

$$\begin{aligned} &\int S_{\alpha\gamma}(x', x)S_{\gamma\beta}(x, x'')\, dx \\ &= \xi^2\left(\frac{\partial^2}{\partial x'_\mu \partial x'_\nu} - \delta_{\mu\nu}\nabla^2_{\mathbf{x}'}\right)(-\nabla^2_{\mathbf{x}'}) \times \int G(|\mathbf{x}' - \mathbf{x}|, t' + t)G(|\mathbf{x} - \mathbf{x}''|, t + t'')\, dx \\ &= \xi^2\left(\frac{\partial^2}{\partial x'_\mu \partial x'_\nu} - \delta_{\mu\nu}\nabla^2_{\mathbf{x}'}\right)(-\nabla^2_{\mathbf{x}'}) \times \int_0^\infty G(|\mathbf{x}' - \mathbf{x}''|, t' + t'' + 2t)\, dt \\ &= \xi^2\left(\frac{\partial^2}{\partial x'_\mu \partial x'_\nu} - \delta_{\mu\nu}\nabla^2_{\mathbf{x}'}\right)(-\nu)^{-1} \times \int_0^\infty [\partial G(|\mathbf{x}' - \mathbf{x}''|, t' + t'' + 2t)/\partial t']\, dt \\ &= \xi^2\left(\frac{\partial^2}{\partial x'_\mu \partial x'_\nu} - \delta_{\mu\nu}\nabla^2_{\mathbf{x}'}\right)(2\nu)^{-1}G(|\mathbf{x}' - \mathbf{x}''|, t' + t'') \\ &= \xi(2\nu)^{-1}S_{\alpha\beta}(x', x''), \end{aligned} \quad (86)$$

with use being made of the convolution equation

$$\int G(|\mathbf{x}' - \mathbf{x}|, t' + t)G(|\mathbf{x} - \mathbf{x}''|, t + t'')d^3x = G(|\mathbf{x}' - \mathbf{x}''|, t' + t'' + 2t)$$

and the equation (13) satisfied by (11). Because (86) shows that the iterated two-point correlation tensor is simply proportional to (85), the inverse of (83) prescribed by (78) is

$$K^{-1}_{\alpha\beta}(x', x'') = 2\delta_{\alpha\beta}\delta(x' - x'') - 4\lambda\nu(\lambda\xi + \nu)^{-1}S_{\alpha\beta}(x', x''). \quad (87)$$

Hence, (84) reduces to

$$T_{\alpha\beta}(x', x'') = \lambda\xi(\lambda\xi + \nu)^{-1}S_{\alpha\beta}(x', x'') \quad (88)$$

with the substitution of (87) and (86). We finally obtain the two-point velocity correlation tensor (80) as

$$R_{\mu\nu}(x', x'') \cong \nu(\lambda\xi + \nu)^{-1}S_{\mu\nu}(x', x''). \quad (89)$$

Rigorous for C-dominant turbulence with the forms (81) and (82), the remarkably simple result (89) should be compared with the general expression (A5) in the Appendix for the two-point velocity correlation tensor associated with weak turbulence. The C-dominant turbulence features a decay law for the specific kinetic energy

$$\begin{aligned}\sigma = \sigma(t) &\equiv \tfrac{1}{2}\langle u_\mu(x)u_\mu(x)\rangle = \tfrac{1}{2}R_{\mu\mu}(x, x)\\ &= \tfrac{1}{2}\nu(\lambda\xi + \nu)^{-1}S_{\mu\mu}(x, x)\\ &= 3\cdot 2^{-\frac{13}{2}}(\pi\nu)^{-\frac{3}{2}}\xi(\lambda\xi + \nu)^{-1}t^{-\frac{5}{2}}, \end{aligned} \quad (90)$$

by virtue of the contracted expression obtained from (89) and (85) with $x' = x'' = x$. It is interesting to note that the decay law (90), $\sigma \propto t^{-\frac{5}{2}}$, is associated with the final period for wind-tunnel turbulence generated in the usual fashion by a square-mesh grid. Whether a two-point velocity correlation tensor with the form (89) and the decay law (90) are related to other varieties of physical turbulence for the entire duration of decay is a question that will be answered by future experimental measurements. In any event, the exactly solvable theory of C-dominant turbulence is a mathematical prototype for the evaluation of (69) subject to (70).

APPENDIX: TWO-POINT VELOCITY CORRELATION TENSOR FOR WEAK TURBULENCE

The formulation given in Sec. III leads to an immediate expression for the two-point velocity correlation tensor

$$R_{\mu\nu}(x', x'') \equiv \langle u_\mu(x')u_\nu(x'')\rangle \quad (A1)$$

in the case of "weak turbulence," represented by a statistical ensemble of velocity fields with $\mathbf{G}:\tilde{\mathbf{u}}\tilde{\mathbf{u}}$ small compared to $\tilde{\mathbf{u}}$ for all $\mathbf{x}$ and all $t \geq 0$, so that only the leading low-order terms in the iteration solution series (17) are significant. Assuming that the probability distribution $\hat{\mathbb{P}}[\tilde{\mathbf{u}}]$ is Gaussian and invariant under translations of space, we have a two-point correlation tensor of the form (20), a characteristic functional (30), and the expectation values

$$\langle \tilde{u}_\mu(x')\tilde{u}_\nu(x'')\rangle = S_{\mu\nu}(x', x''), \quad (A2)$$

$$\langle \tilde{u}_\mu(x')\tilde{u}_\nu(x'')\tilde{u}_\rho(x''')\rangle = 0, \quad (A3)$$

$$\begin{aligned}\langle \tilde{u}_\mu(x')\tilde{u}_\nu(x'')\tilde{u}_\rho(x''')\tilde{u}_\sigma(x'''')\rangle \\ = S_{\mu\nu}(x', x'')S_{\rho\sigma}(x''', x'''') + S_{\mu\rho}(x', x''')S_{\nu\sigma}(x'', x'''') \\ + S_{\mu\sigma}(x', x'''')S_{\nu\rho}(x'', x'''). \end{aligned} \quad (A4)$$

Then by putting (17) into (A1) we obtain

$$\begin{aligned}R_{\mu\nu}(x', x'') = S_{\mu\nu}(x', x'') \\ + 2\int G_{\mu\alpha\beta}(x' - y')G_{\nu\gamma\delta}(x'' - y'') \\ \times S_{\alpha\gamma}(y', y'')S_{\beta\delta}(y', y'')\, dy'\, dy'' \\ + 4\int G_{\mu\alpha\beta}(x' - y')G_{\beta\gamma\delta}(y' - y'') \\ \times S_{\nu\gamma}(x'', y'')S_{\alpha\delta}(y', y'')\, dy'\, dy'' \\ + 4\int G_{\nu\alpha\beta}(x'' - y'')G_{\beta\gamma\delta}(y'' - y') \\ \times S_{\mu\gamma}(x', y')S_{\alpha\delta}(y'', y')\, dy'\, dy'' + O(S^3), \end{aligned} \quad (A5)$$

where the space–time coordinate arguments of the three-index Green's function (9) are abbreviated in $G_{\mu\alpha\beta}(x) \equiv G_{\mu\alpha\beta}(\mathbf{x}, t)$, and the property of the autocorrelation tensor $S_{\mu\nu}(x, x) =$ (function of t alone) has been employed in order to eliminate three terms which vanish. The integral terms appearing in (A5) can ordinarily be evaluated by analytical or numerical procedures for specialized two-point correlation tensors (20) with $h(\mathbf{x})$ prescribed.

* Work supported by a National Science Foundation grant.

[1] G. I. Taylor, Proc. Roy. Soc. **151A**, 421 (1935).
[2] G. K. Batchelor, *Homogeneous Turbulence* (Cambridge University, New York, 1960).
[3] S. Chandrasekhar, Proc. Roy. Soc. **229A**, 1 (1955); Phys. Rev. **102**, 941 (1956).
[4] H. W. Wyld, Jr., Ann. Phys. (N.Y.) **14**, 143 (1961).
[5] E. Hopf, J. Ratl. Mech. Anal. **1**, 87 (1952).
[6] G. Rosen, Phys. Fluids **3**, 519, 525 (1960).
[7] I. Hosokawa, J. Math. Phys. **8**, 221 (1967) and works cited therein.
[8] G. Rosen, Phys. Fluids **10**, 2614 (1967) and works cited therein; "Dynamics of Probability Distributions over Classical Fields," preprint, Drexel University Physics Dept., 1970.
[9] G. Rosen, Phys. Letters **31A**, 142 (1970).
[10] For a recent introduction to field-theoretic functional integration methods see: G. Rosen, *Formulations of Classical and Quantum Dynamical Theory* (Academic, New York, 1969), Chap. 4.
[11] G. Rosen, Phys. Fluids **13**, 2891 (1970).
[12] Reference 10, Appendix D.
[13] See, for example: K. O. Friedrichs, H. N. Shapiro, J. Schwartz, T. Seidman, and B. Wendroff, *Seminar on Integration of Functionals* (Courant Institute of Mathematical Sciences, New York University, 1957), p. I-19.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Multiplicities in the Classical Groups. II*

B. GRUBER, B. H. HAN,† AND J. A. SALDANHA†
Department of Physics, Saint Louis University, St. Louis, Missouri 63103

(Received 16 January 1970)

In a previous paper, a method has been developed which allows a calculation in the same manner of all multiplicities occurring in the theory of linear representations of the classical compact Lie groups. This method is applied to the group $SO(9)$ and some of its subgroups which are of interest in physics. It is demonstrated with the group $SO(9)$ as example that this method of calculating the multiplicities allows us to accumulate in a condensed way a large amount of information and, moreover, that this method can be easily applied to any other (classical, compact) rank-4 group and its semisimple subgroups.

I. INTRODUCTION

In a previous paper[1] a simple method has been developed by means of which the multiplicities arising in the theory of linear (irreducible) representations of the classical (compact) Lie groups can be obtained in the *same* manner from a given pattern (containing a certain number of representations). For each of the multiplicities—multiplicity of weights ("inner multiplicity"), multiplicity arising in the decomposition of the inner direct product of two representations ("outer multiplicity"), and branching multiplicities—a different diagram is applied to the pattern according to essentially the same rules. Through the application of these diagrams to the pattern, the desired multiplicity is obtained. Subsequently, this method for the calculation of the multiplicities was applied to the classical groups of rank 2, resulting in a complete compilation of all multiplicities for a limited number of representations of these groups.

The present paper deals with the application of the results of Ref. 1 to classical groups of rank 3 and 4. The motivation for this program is twofold: Several of the rank-3 and rank-4 groups are used in physics. For example, Wigner's $SU(4)$,[2] the chains of groups

$$SO(7) \supset G_2 \supset SO(3),$$
$$SO(2l+1) \supset SO(3),$$
$$SU(2l+1) \supset SO(2l+1), \quad \text{etc.},$$

in atomic spectroscopy,[3,4] and others. Moreover, groups and subgroups not used at the present time might some day become of interest to physics. Thus, the knowledge of the multiplicities associated with these groups is certainly of interest. It is one of the aims of this paper to demonstrate that, whenever the multiplicities of a (set of) representation(s) of some rank-3 or rank-4 group become of interest, it presents no difficulty to set up patterns and diagrams by means of which these multiplicities can be easily obtained. (Strictly speaking, this is true only for representations with not too high dimensionality; otherwise, the pattern becomes unmanageable. In physics, however, usually only the first few representations of a group are used. Since representations with dimensionalities well in the ten thousands can still be handled with relative ease, it is assumed that all representations of the rank-3 and rank-4 groups of interest to physics fall into the domain of the applicability of this method.)

The second purpose of this paper is to demonstrate that a pattern (together with the diagrams) amounts almost to a "tabulation" of *all* multiplicities for a set of representations of a group. Even more, a pattern amounts as well to a "tabulation" of *all* multiplicities of all those representations of the subgroups of a group which are contained in the pattern. And this holds for all semisimple (Lie) subgroups of the group. Thus, a pattern together with the diagram amounts to an implicit compilation ("tabulation") of inner multiplicities, outer multiplicities, and branching multiplicities for all those representations of the group and all those representations of its (semisimple) subgroups which are contained in the pattern. Thus, any single pattern may contain data which are otherwise widely scattered throughout the literature—if at all known.

It would be unnecessary to give patterns and diagrams for all the rank-4 groups and their semisimple subgroups. Not all of these groups are of immediate interest. Instead the group $SO(9)$ will be taken as an example. The choice of $SO(9)$ as example is motivated by the fact that, on the one hand, the group $SO(9)$ and some of its subgroups are of interest to physics, while on the other hand the group $SO(9)$ is as good as any other rank-4 group to demonstrate the two points made above. Namely, the patterns and diagrams which are given for $SO(9)$ will demonstrate (a) the relative ease with which the multiplicities of any (classical) rank-3 or rank-4 group (and its subgroups) can be obtained and (b) the amount of information contained in just a few patterns. In all, two patterns and four diagrams will be given. Pattern P_1 of Fig. 5 and pattern P_2 of Fig. 6, corresponding to the defining representations (for all definitions and rules refer to Ref. 1) $D(2,2,0,0)$ and $D[\frac{1}{2}(5,5,1,1)]$ of $SO(9)$ respectively, contain the following information (the patterns are deliberately kept small in order to serve as illustrative examples).

(1) The weight diagrams of four representations of G_2, of 12 representations of $SO(5)$, of 16 representations of $SO(7)$, and parts of 18 representations of $SO(9)$. Moreover, there are a number of representations of $SO(3)$ contained in the pattern. Which representations of $SO(3)$ [or any other semisimple subgroup G' of $SO(9)$] occur in the pattern depends on which subgroup $SO(3)$ (which semisimple subgroup G') is considered.

The weight diagrams are not always given explicitly, but in the language of d.w. numbers.[1] Nevertheless, the explicit weight diagrams can be easily obtained from the pattern as follows from its construction (Ref. 1). For the particular representations included in the pattern, see Sec. II.

(2) The multiplicities of weights for all these representations of the groups G_2, $SO(5)$, $SO(7)$, and $SO(9)$.

(3) The Clebsch–Gordan series for all these representations of the groups G_2, $SO(5)$, and $SO(7)$ with any other representation of these groups.

(4) The branching multiplicities

$$SO(7) \to G_2, \quad SO(9) \to SO(3), \quad SO(7) \to SO(3),$$
$$SO(5) \to SO(3), \quad \text{and} \quad G_2 \to SO(3)$$

for the representations included in the patterns. [Here $SO(3)$ is the so-called principal $SO(3)$ subgroup[5] used in atomic spectroscopy (see Sec. III).] The branching multiplicity of any of these groups with respect to any simple or semisimple subgroup can be obtained from these patterns too, once the mapping onto the subgroup and the diagram corresponding to the subgroup have been determined. [It might, however, for the branchings of the $SO(9)$ representations, be necessary to enlarge or complete the $SO(9)$ pattern. This will happen if a part of the pattern which is not given should be mapped onto weights of the subgroup which participate in the determination of the branching multiplicity.]

II. THE GROUP $SO(9)$

In this section all information pertaining to the patterns and diagrams, Figs. 1–6, will be given. However, no information given in Ref. 1 will be unnecessarily repeated. Thus, Ref. 1 is essential for an understanding of this paper. (All rules for obtaining the different multiplicities are given in Ref. 1.) The reader is therefore explicitly referred to Ref. 1.

The group $SO(9)$: Simple negative roots are

$$\beta_1 = (0, 0, 0, -1), \quad \beta_2 = (0, 0, -1, 1),$$
$$\beta_3 = (0, -1, 1, 0), \quad \beta_4 = (-1, 1, 0, 0),$$
$$R = \tfrac{1}{2}(7, 5, 3, 1).$$

The weights are

$$m = (m_1, m_2, m_3, m_4),$$
$$m_i \text{ all integers or all half-integers.}$$

Sets of equivalent weights: Weights which go over into another by a permutation of their components with or without a change of sign of some of their components belong to the same set of equivalent weights. Thus, the order of the Weyl group is $2^4 4!$

Dimensionality of representations:

$$\begin{aligned}\dim D(M) = {} & (M_1 - M_2 + 1)\tfrac{1}{2}(M_1 - M_3 + 2) \\ & \times \tfrac{1}{3}(M_1 - M_4 + 3)(M_2 - M_3 + 1) \\ & \times \tfrac{1}{2}(M_2 - M_4 + 2)(M_3 - M_4 + 1) \\ & \times \tfrac{1}{6}(M_1 + M_2 + 6)\tfrac{1}{5}(M_1 + M_3 + 5) \\ & \times \tfrac{1}{4}(M_1 + M_4 + 4)\tfrac{1}{4}(M_2 + M_3 + 4) \\ & \times \tfrac{1}{3}(M_2 + M_4 + 3)\tfrac{1}{2}(M_3 + M_4 + 2) \\ & \times \tfrac{1}{7}(2M_1 + 7)\tfrac{1}{5}(2M_2 + 5) \\ & \times \tfrac{1}{3}(2M_3 + 3)(2M_4 + 1),\end{aligned}$$

with

$$M_1 \geq M_2 \geq M_3 \geq M_4 \geq 0 \quad \text{(d.w. condition).}$$

Diagrams: In Fig. 1, the diagram for the multiplicity of weights for the groups $SO(9)$, $SO(7)$, $SO(5)$, and $SO(3)$ is given. It should be noted that the diagram is not complete but adapted to the pattern (elements which cannot contribute have been deleted).

In Fig. 2, the diagram for the branching multiplicity $SO(7) \to G_2$ is given. The diagram for the multiplicity of weights for G_2 is obtained from this diagram by deleting the element (0, 0; 1) and by changing the sign $\delta_8 \to -\delta_8$ for all other elements. [The resulting diagram is different from the diagram given for G_2 in Ref. 1. The reason is that the simple roots which arise through the mapping L of $SO(7)$ onto G_2 are $(-1, 1, 0)$ and $(1, -2, 1)$. Thus, the two simple roots are not only different from the ones used in Ref. 1 but, moreover, one of them is a positive root. This latter property has the effect that now also *negative* k_i values can appear. This has, however, no consequences if the counting process is adapted in the obvious way (a negative value k_i means counting "backward"].

In Fig. 3, the diagram for the Clebsch–Gordan series for the groups $SO(7)$ and $SO(5)$ is given.

In Fig. 4, the diagram for the Clebsch–Gordan series for the group G_2 is given. What was said about negative values of k_i for Fig. 2 holds also for this case.

The diagram for restrictions to the $SO(3)$ subgroups is given by the two elements $(j; 1)$ and $(j + 1; -1)$,

0	1	2	3	4		
	1				000	
1			-1		001	
	-1			1	002	
			1	-1	003	SO(5)
1	-1				010	
-1					012	
	1				013	
-1			1		021	
1					022	SO(7)
1	-1				100	
-1			1		101	
	1			-1	102	
			-1	1	103	
-1	1				120	
1					123	
	-1				124	
-1	1				210	
1					212	
	-1				213	
1	-1				220	
-1					223	
	1				224	SO(9)

FIG. 1

0	1	2	3	4	5	6	7	8	9	10	
-1	1										-1
1				-1							0
	-1					1					1
				1					-1		3
						-1				1	4
									1	-1	5

FIG. 2

0	...	q+1	...	p+4	...	p+2q+6	...	2p+q+9	...	2(p+q)+10	
-1		1									1/3(-p+q)-1
1				-1							0
		-1				1					q+1
				1				-1			1/3(2p+q)+3
						-1				1	2/3(p+2q)+4
								1		-1	p+q+5

FIG. 4

0	$2M_3+1$	$2M_2+3$	$2(M_2+M_3+2)$		
1	-1			0	0
-1		1		0	M_2-M_3+1
	1		-1	0	M_2+M_3+2
		-1	1	0	$2M_2+3$

0	$2M_3+1$	$2M_2+3$	$2(M_2+M_3+2)$	$2M_1+5$	$2(M_1+M_3+3)$		
-1	1					M_1-M_2+1	0
1				-1		M_1-M_2+1	M_1-M_3+2
	-1				1	M_1-M_2+1	M_1+M_3+3
				1	-1	M_1-M_2+1	$2M_1+5$

0	$2M_3+1$	$2M_2+3$	$2(M_2+M_3+2)$	$2M_1+5$	$2(M_1+M_3+3)$	$2(M_1+M_2+4)$		
1		-1					M_1-M_3+2	M_2-M_3+1
-1				1			M_1-M_3+2	M_1-M_3+2
		1				-1	M_1-M_3+2	$M_1+2M_2-M_3+5$
				-1		1	M_1-M_3+2	$2M_1+M_2-M_3+6$

0	$2M_3+1$	$2M_2+3$	$2(M_2+M_3+2)$	$2M_1+5$	$2(M_1+M_3+3)$	$2(M_1+M_2+4)$	$2(M_1+M_2+M_3)+9$		
	-1		1					M_1+M_3+3	M_2+M_3+2
	1				-1			M_1+M_3+3	M_1+M_3+3
			-1				1	M_1+M_3+3	$M_1+2M_2+M_3+6$
					1		-1	M_1+M_3+3	$2M_1+M_2+M_3+7$

0	$2M_3+1$	$2M_2+3$	$2(M_2+M_3+2)$	$2M_1+5$	$2(M_1+M_3+3)$	$2(M_1+M_2+4)$	$2(M_1+M_2+M_3)+9$		
		1	-1					M_1+M_2+4	$2M_2+3$
		-1				1		M_1+M_2+4	$M_1+2M_2-M_3+5$
			1				-1	M_1+M_2+4	$M_1+2M_2+M_3+6$
						-1	1	M_1+M_2+4	$2(M_1+M_2+4)$

0	$2M_3+1$	$2M_2+3$	$2(M_2+M_3+2)$	$2M_1+5$	$2(M_1+M_3+3)$	$2(M_1+M_2+4)$	$2(M_1+M_2+M_3)+9$		
				-1	1			$2M_1+5$	$2M_1+5$
				1		-1		$2M_1+5$	$2M_1+M_2-M_3+6$
					-1		1	$2M_1+5$	$2M_1+M_2+M_3+7$
						1	-1	$2M_1+5$	$2(M_1+M_2+4)$

FIG. 3

FIG. 1. Diagram for the multiplicity of weights for $SO(9)$, $SO(7)$, $SO(5)$, and $SO(3)$. This diagram is to be applied to dominant weights only.

FIG. 2. Diagram for multiplicity of weights of G_2. See text. Diagram for branching multiplicity $SO(7) \to G_2$.

FIG. 3. Diagrams for Clebsch–Gordan series for the groups $SO(7)$ and $SO(5)$.

FIG. 4. Diagram for Clebsch–Gordan series for the group G_2. Note the negative value for k_2. See text.

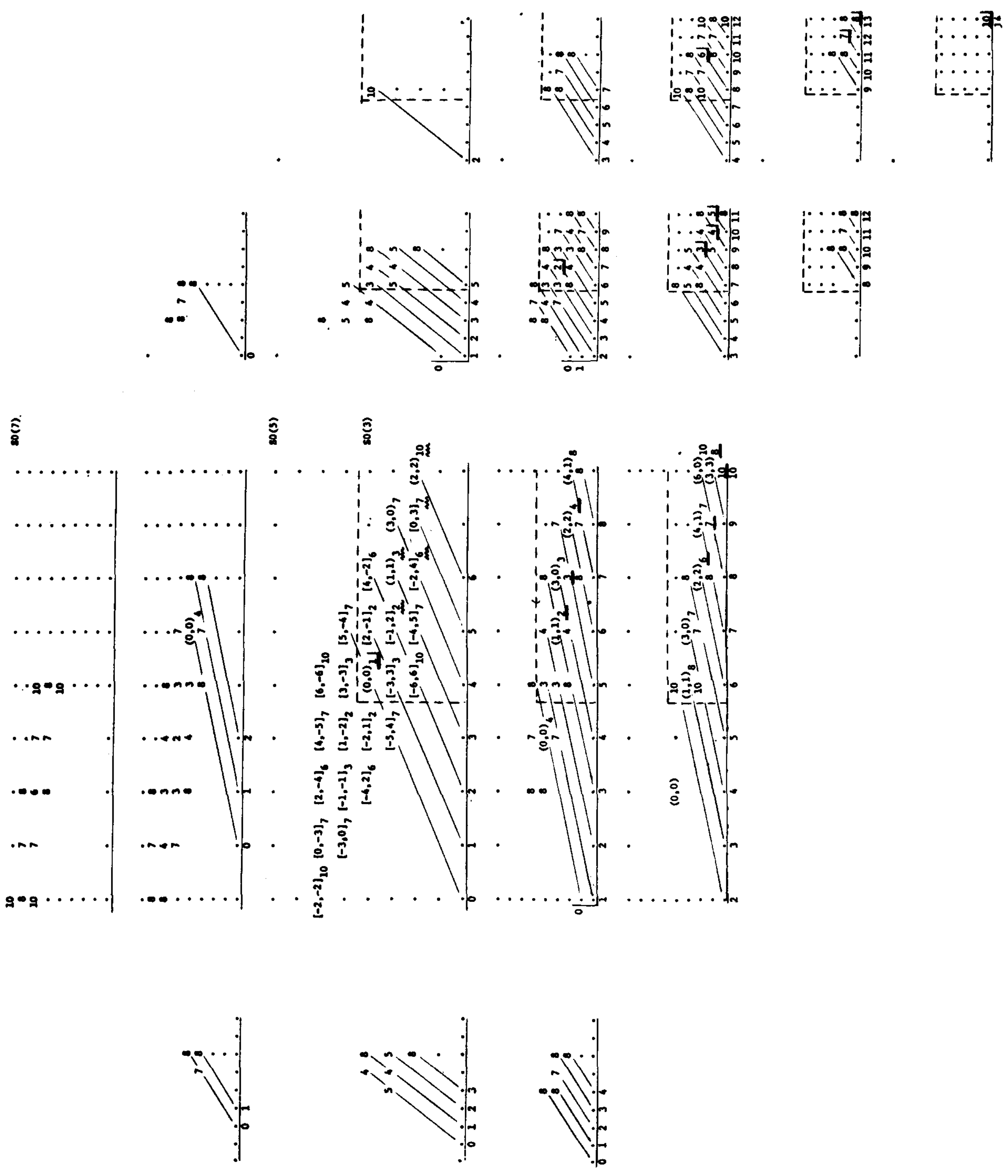


FIG. 5. Pattern P_1, defining representation $D(2, 2, 0, 0)$. Multiplicity of weights for $SO(9)$, $SO(7)$, $SO(5)$, and G_2; Clebsch-Gordan series for $SO(7)$, $SO(5)$, and G_2; branching multiplicities $SO(9)$, $SO(7)$, $SO(5)$, $G_2 \to SO(3)_2$, and $SO(7) \to G_2$. The subgroup $SO(3)_2$ is the principal $SO(3)$.

where j stands for a highest weight (angular momentum) of $SO(3)$. [Here the same notation is used as in Ref. 1, pp. 3085 and 3090. This notation is not to be confused with the similar notation $(k_1, \cdots, k_4; \delta_s)$ for the elements of the diagram. Here j *is* the angular momentum under consideration. The notation implies that $\tilde{\nu}(j)$ is obtained by summing up the multiplicities of all weights mapped onto j and $j + 1$, respectively, and by subtracting the number obtained for $j + 1$ from the number obtained for j.]

Patterns: For pattern P_1, Fig. 5, and pattern P_2, Fig. 6, the following conventions are made.

(a) The d.w. numbers refer to weights of $SO(9)$. The d.w. numbers for weights of the subgroups $SO(7)$, $SO(5)$, and $SO(3)$ are assigned through the

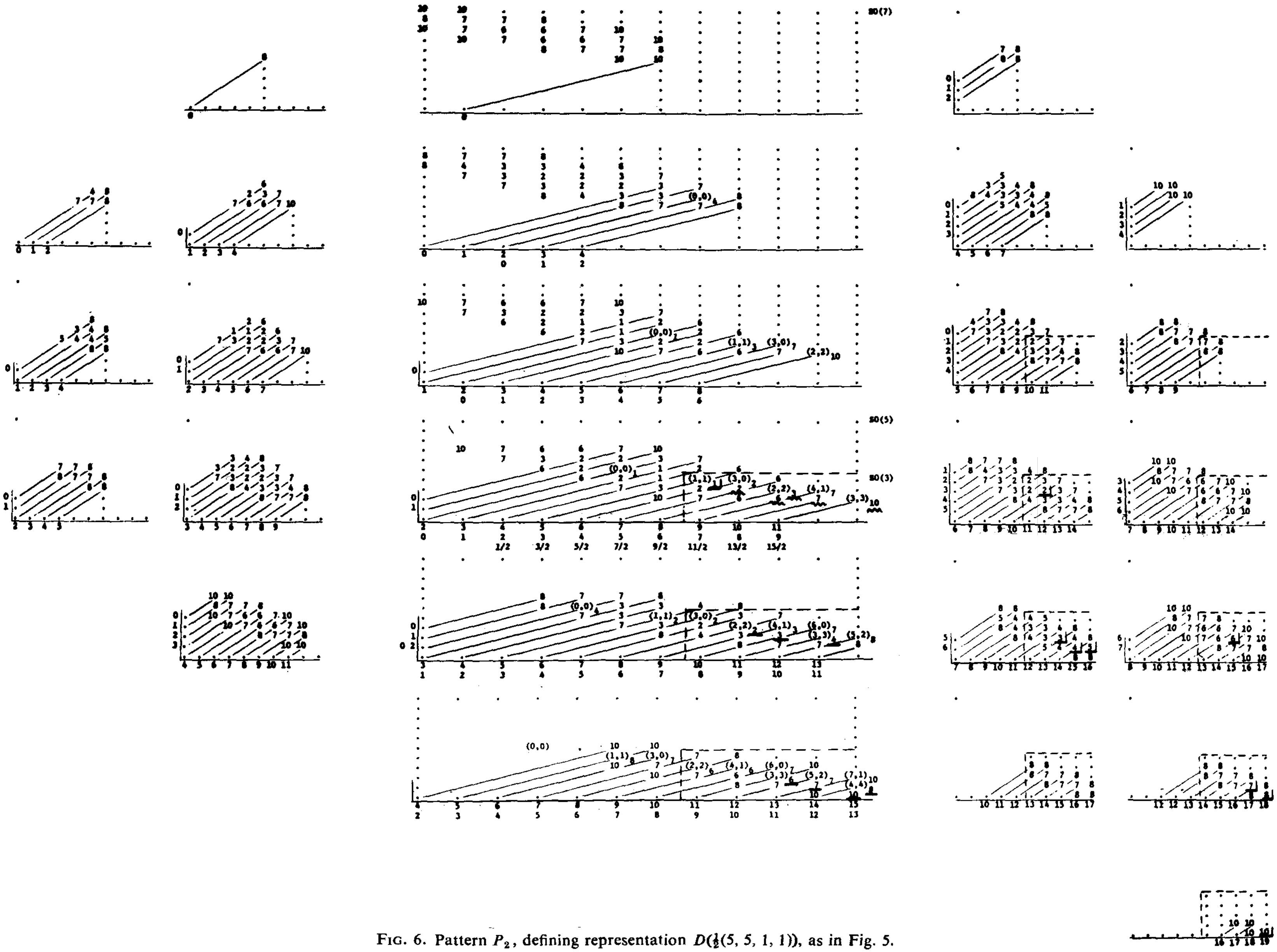


FIG. 6. Pattern P_2, defining representation $D(\frac{1}{2}(5, 5, 1, 1))$, as in Fig. 5.

mapping L. This means that the set of weights obtained through subsequent mappings of an $SO(9)$ weight onto some $SO(7)$, $SO(5)$, and $SO(3)$ weights are assigned the same d.w. number as the weight of $SO(9)$. (See the list of weights below.) Thus, the weights of the rank-2 groups will have assigned to them in general different d.w. numbers as the ones assigned to them in Ref. 1.

(b) Weights which have been written out explicitly are weights of G_2. Their d.w. number, however, refers to weights of $SO(9)$. [See (a) above.] These weights of G_2 have been obtained from the $SO(7)$ weight at that position in the pattern through the mapping L of $SO(7)$ onto G_2. If a weight is written with normal brackets, it is a dominant G_2 weight; if it is written with square brackets, then it is a nondominant G_2 weight. Moreover, the G_2 weight (0, 0) has been added to other blocks in order to serve as point of reference for application of the G_2 diagrams.

(c) The highest weight of a set of equivalent weights (the members of which have assigned to them the same d.w. number) is indicated for $SO(9)$ by an underlined numeral with a vertical line to its right, for $SO(7)$ by an underlined numeral, and for $SO(5)$ by a numeral underlined by a curved line. The highest weight of the set of weights characterized by the d.w. number 1 is the highest weight of this set for all three groups $SO(9)$, $SO(7)$, and $SO(5)$.[6] [For $SO(3)$, the highest weight of a set of equivalent weights is not indicated. The highest weights are given, in natural sequence, by the weights of the line containing **1**, when moving from **1** to the right.]

(d) All weights when written out explicitly (if not solely represented by their d.w. number) have brackets, apart from the $SO(3)$ weights which are plotted "outside" the pattern (i.e., below and on the side of a block, separated from it by a line). Thus, if in the pattern a number appears without brackets, then it is a d.w. number. On the other hand, if a weight appears without a d.w. number, then it does not belong to the $SO(9)$ pattern (including subgroup pattern).

(e) The mappings L of $SO(9)$, $SO(7)$, $SO(5)$, and G_2 onto $SO(3)_2$ are indicated by straight lines. All weights along a straight line are mapped onto the same weight of the subgroup $SO(3)_2$.

TABLE I.

$SO(9)$	$SO(7)$	$SO(5)$	$SO(3)$	G_2
${}_{24}(2, 2, 0, 0)_{10}$	${}_{12}(2, 2, 0)_{10}$	${}_{4}(2, 2)_{10}$		${}_{6}(2, 2)$
${}_{96}(2, 1, 1, 0)_{8}$	${}_{24}(2, 1, 1)_{8}$			${}_{6}(3, 0)$
${}_{48}(2, 1, 0, 0)_{7}$	${}_{12}(2, 1, 0)_{7}$	${}_{8}(2, 1)_{7}$		${}_{6}(1, 1)$
${}_{8}(2, 0, 0, 0)_{6}$	${}_{6}(2, 0, 0)_{6}$	${}_{4}(2, 0)_{6}$	$(2)_6$	${}_{1}(0, 0)$
${}_{16}(1, 1, 1, 1)_{5}$				
${}_{32}(1, 1, 1, 0)_{4}$	${}_{8}(1, 1, 1)_{4}$			
${}_{24}(1, 1, 0, 0)_{3}$	${}_{12}(1, 1, 0)_{3}$	${}_{4}(1, 1)_{3}$		
${}_{8}(0, 0, 0, 0)_{2}$	${}_{6}(1, 0, 0)_{2}$	${}_{4}(1, 0)_{2}$	$(1)_2$	
${}_{1}(1, 0, 0, 0)_{1}$	${}_{1}(0, 0, 0)_{1}$	${}_{1}(0, 0)_{1}$	$(0)_1$	

TABLE II.

$SO(9)$	$SO(7)$	$SO(5)$	$SO(3)$
${}_{96}(5, 5, 1, 1)_{10}$	${}_{24}(5, 5, 1)_{10}$	${}_{4}(5, 5)_{10}$	
${}_{192}(5, 3, 3, 1)_{8}$	${}_{24}(5, 3, 3)_{8}$		
${}_{192}(5, 3, 1, 1)_{7}$	${}_{48}(5, 3, 1)_{7}$	${}_{8}(5, 3)_{7}$	
${}_{64}(5, 1, 1, 1)_{6}$	${}_{24}(5, 1, 1)_{6}$	${}_{8}(5, 1)_{6}$	$(5)_6$
${}_{16}(3, 3, 3, 3)_{5}$			
${}_{64}(3, 3, 3, 1)_{4}$	${}_{8}(3, 3, 3)_{4}$		
${}_{96}(3, 3, 1, 1)_{3}$	${}_{24}(3, 3, 1)_{3}$	${}_{4}(3, 3)_{3}$	
${}_{64}(3, 1, 1, 1)_{2}$	${}_{24}(3, 1, 1)_{2}$	${}_{8}(3, 1)_{2}$	$(3)_2$
${}_{16}(1, 1, 1, 1)_{1}$	${}_{8}(1, 1, 1)_{1}$	${}_{4}(1, 1)_{1}$	$(1)_1$

With these conventions, the d.w. numbers of the entire pattern describe (pieces of) representations of $SO(9)$, the d.w. numbers of the block column containing **1** describe representations of $SO(7)$, the d.w. numbers of the block containing **1** describe representations of $SO(5)$, the d.w. numbers of the line containing **1** describe representations of $SO(3)$, and, finally, the *weight vectors* of the block containing **1** describe representations of G_2. The only other elements of the pattern are vectors (0, 0) (without d.w. numbers) which serve as point of reference for the G_2 diagrams.

It should be pointed out that the blocks to the left of the "block column" containing **1** are needed for the calculation of the branching multiplicity with respect to the restriction to $SO(3)_2$ subgroup only. Their relative position to the rest of the pattern is therefore irrelevant.

The Groups and Their Representations

The representations of the groups $SO(9)$, $SO(7)$, $SO(5)$, $SO(3)$, and G_2 contained in pattern P_1 are listed in Table I.

The representations contained in pattern P_2 (the weights given have to be divided by 2) are listed in Table II.

The representations $D(j)$ of the principal $SO(3)$ subgroup $[SO(3)_2]$ contained in patterns P_1 and P_2 are given for the values $j = 0, 1, \cdots, 14$, and $j = 0, 1, \cdots, 19$, respectively.

III. SUBGROUPS

The mappings L and simple root systems are given below.

Branchings for patterns P:

$$SO(7) \to G_2 : Lm = (m_2 + m_4, m_3 - m_4, -m_2 - m_3)$$
$$= (m_2 - m_3 + 2m_4, m_2 + 2m_3 - m_4) \quad \text{in } (p, q) \text{ notation,}$$
$$\beta = (-1, 1, 0), \quad \alpha = (1, -2, 1).$$

The (p, q) notation has been used in the pattern:

$$G_2 \to SO(3)_2 : L(p, q) = \tfrac{1}{3}(4p + 5q) = 3m_2 + 2m_3 + m_4, \quad \beta = -1,$$

$$SO(9) \to SO(3)_2 : L(m_1, m_2, m_3, m_4) = 4m_1 + 3m_2 + 2m_3 + m_4,$$

$$SO(7) \to SO(3)_2 : L(m_2, m_3, m_4) = 3m_2 + 2m_3 + m_4,$$

$$SO(5) \to SO(3)_2 : L(m_3, m_4) = 2m_3 + m_4,$$

$\beta = -1$ in all the three cases.

IV. RULES

Some minor adaptions of the rules given for obtaining the various multiplicities, in Ref. 1, have to be made.

(A) *Multiplicity of Weights:* To each group considered there corresponds a particular part of the pattern P, as noted before in Sec. II. For instance, the representations of the group $SO(5)$ are given by the block containing the d.w. number **1**. It can then be noted that, due to the way in which the d.w. numbers have been assigned to the weights of the subgroups, in general not all numbers $p, p - 1, \cdots, 1$ are d.w. numbers of the representation labeled by **p**. Some of the numbers of such a sequence may not appear at all. This is, however, of no concern. The modified rule is that all weights characterized by d.w. numbers smaller than p (and greater than zero) belong to the representation **p**, whenever such a number appears. If a number does not appear, then there exists no weight of the group which is characterized by this number.

For reasons of convenience, that part of the pattern which contributes to the multiplicity of weights has been indicated by dashed lines and lies below and to the right of these lines. These areas correspond to the blocks referred to in the "counting process" (see Ref. 1, Sec. IIIA) and take over their role in the determination of the multiplicity γ.

(B) *Multiplicity in the direct product $D(M) \otimes D(M')$:* What was said in Sec. IVA for the d.w. numbers holds obviously also in this case. Once a group has been chosen [$SO(7)$, $SO(5)$, G_2] and the part of the pattern P corresponding to it has been identified, the d.w. numbers of a representation $D(M)$ characterized by the d.w. number **p** are given by those d.w. numbers p', $1 \le p' \le p$, which *occur* in that part of the pattern.

For groups whose pattern consists of d.w. numbers, as it is the case for the groups $SO(7)$ and $SO(5)$ in pattern P, it becomes essential to determine the argument $M' + m$ of $\bar{\gamma}(M' + m)$. The d.w. numbers, namely, do not distinguish among the weights of a set of equivalent weights.

As for the case in which the elements of the pattern are the weights themselves, in this case the multiplicity $\bar{\gamma}(M' + m)$, $m \in D(M)$, of the direct product $D(M) \otimes D(M')$ is also determined successively. If **p** is the d.w. number which corresponds to the weight M, the multiplicities are calculated successively beginning with **p** [for which the result is the multiplicity $\bar{\gamma}(M' + M) = 1$]. The succession in which the multiplicities

$$\bar{\gamma}(M' + m) = \bar{\gamma}(M' + M + k_1\beta_1 + \cdots + k_4\beta_4), \quad m \in D(M),$$

are calculated is in the order

$$k_1 = 0, 1, 2, \cdots, \quad k_2 = k_3 = k_4 = 0,$$

followed by

$$k_1 = 0, 1, 2, \cdots, \quad k_2 = 1, k_3 = k_4 = 0,$$

and so on, exhausting all weights of the representation $D(M)$ [and only weights of $D(M)$]. This implies that when the diagram is applied to some element of the pattern belonging to $D(M)$ from which **p** is reached by the counting process (k_1, k_2, k_3, k_4), then the argument $M' + m$ for the multiplicity $\bar{\gamma}$ is for that element given by $M' + M + k_1\beta_1 + \cdots + k_4\beta_4$.

(C) *Branching Multiplicity:* For a branching of a representation resulting from the restriction $G \to G'$, that part of the pattern is relevant which corresponds to representations of G. Inside that part of the pattern the mapping L of weights of G onto weights of G' is considered. The dominant weights arising through that mapping are given by weight *vectors* in regular brackets (inside the pattern) and by the numbers j (highest weights) for the $SO(3)$ subgroups (without brackets, plotted "outside" the pattern).

The first layer of $SO(3)$ weights corresponds to the restriction $SO(9) \to SO(3)$, the second to $SO(7) \to SO(3)$, and the third layer to $SO(5) \to SO(3)$.

* Supported in part by the National Science Foundation Grant GP 9623.

† Saint Louis University Graduate School.

[1] B. Gruber, J. Math. Phys. **11**, 3077 (1970).
[2] E. P. Wigner, Phys. Rev. **51**, 106 (1937).
[3] B. R. Judd, *Operator Techniques in Atomic Spectroscopy* (McGraw-Hill, New York, 1963).
[4] M. Hamermesh, *Group Theory and Its Applications to Physical Problems* (Addison-Wesley, Reading, Mass., 1964).
[5] E. B. Dynkin, Am. Math. Soc. Transl., Ser. 2, **6**, 111 (1957).

[6] In the text a boldface numeral refers to the highest weight of a set of equivalent weights (all of which are characterized by the same dominant weight number) in order to distinguish the highest weight from the set of weights equivalent to it. In the figures, however, different symbols (depending on the group) are used to identify the highest weight of a set of equivalent weights.

Geometric Properties of Neutrino Fields in Curved Space–Time

J. WAINWRIGHT
Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada

(Received 14 September 1970)

Classical neutrino fields in curved space–time are studied subject to the condition that the neutrino energy tensor T_{ab} satisfies $T_{ab}u^a u^b \neq 0$ for all timelike vectors u^a. It is shown that the principal null congruence of these neutrino fields is geodesic and that its shear and twist are restricted. In addition, there exists a canonical null tetrad with respect to which T_{ab} assumes a simple form. These conditions, in fact, characterize this class of neutrino fields. In addition, it is shown that if T_{ab} satisfies the stronger condition that $T_{ab}u^b$ be a timelike or null vector for all timelike vectors u^a, then the principal null congruence is also shear-free. Comparisons are made with the well-known properties of the electromagnetic energy tensor.

1. INTRODUCTION

In the 2-spinor formalism, a classical neutrino field in curved space–time is described by a 2-spinor field $\phi_A(x^a)$ which satisfies[1]

$$\sigma^{aAA'}\phi_{A;a} = 0. \tag{1.1}$$

The quantities $\sigma^{aAA'}$ determine the metric tensor according to[2]

$$\sigma^{aAA'}\sigma^{bBB'}\epsilon_{AB}\epsilon_{A'B'} = g^{ab}, \tag{1.2}$$

and the covariant derivative $\phi_{A;a}$ is defined in terms of the usual spinor connection coefficients.[3]

At each event a neutrino field $\phi_A(x^a)$ defines a direction tangent to the null cone, given (up to a scale factor) by

$$j^a = \sigma^{aAA'}\phi_A\phi_{A'}, \tag{1.3}$$

called *the principal null direction (p.n.d.) of the field* at that event, in the terminology of Penrose.[4] The congruence of null curves to which the vector field j^a is tangent is called *the principal null congruence (p.n.c.) of the field*. We are going to show that there is a close relationship between the geometric properties of this congruence and certain physically relevant restrictions on the energy tensor of the neutrino field, which is defined (with suitable units) by[5]

$$T_{ab} = i[\sigma_{aAA'}(\phi^A\phi^{A'}{}_{;b} - \phi^A{}_{;b}\phi^{A'}) + \sigma_{bAA'}(\phi^A\phi^{A'}{}_{;a} - \phi^A{}_{;a}\phi^{A'})]. \tag{1.4}$$

As motivation for the particular restrictions which we are going to consider in connection with the neutrino energy tensor, we briefly mention some aspects of the physical interpretation of energy tensors in general. First, *the energy density of a field*[6] at any event P with respect to an observer whose world line contains P is defined by[7]

$$E(u) = T_{ab}u^a u^b, \tag{1.5}$$

where u^a is the (unit) *future-pointing* velocity of the observer (at P). In addition, the vector

$$Q_a(u) = T_{ab}u^b \tag{1.6}$$

describes *the flow of energy in the field* with respect to this observer, in the following sense. Let dS_2 be a small 2-surface carried along by the observer, with unit normal n^a in the observer's instantaneous rest space (so that $u^a n_a = 0$). Then the energy flux across dS_2 per unit time per unit area is given by[7,8]

$$Q_a(u)\, n^a.$$

In terms of these concepts we define the following.

Definition: A field is said to satisfy *the strong energy condition* if its energy tensor T_{ab} satisfies the conditions[9]

$$E(u) > 0 \tag{1.7}$$

and

$$Q^a(u) \text{ is a future-pointing timelike or null vector,} \tag{1.8}$$

for all observers at each event P for which $T_{ab} \neq 0$.

For example, a distribution of perfect fluid, with

$$T_{ab} = (A + p)v_a v_b - pg_{ab},$$

where A, p, and v_a are the rest energy density, the pressure, and velocity of the fluid, respectively, satisfies the strong energy condition, provided that $A^2 - p^2 \geq 0$, as is easily shown. Secondly, the electromagnetic energy tensor is given by

$$T_{ab} = F_{ac}F_b{}^c - \tfrac{1}{4}g_{ab}F_{cd}F^{cd}, \tag{1.9}$$

in terms of the electromagnetic field tensor F_{ab}, and satisfies the conditions[10]

$$E(u) \geq k, \tag{1.10}$$

$$T_a{}^c T_c{}^b = k^2\delta_a^b, \tag{1.11}$$

where

$$k = \tfrac{1}{4}[(F_{ab}F^{ab})^2 + (F_{ab}F^{*ab})^2]^{\frac{1}{2}} \geq 0. \quad (1.12)$$

Equality in (1.10) can be achieved for some observer if and only if $k \neq 0$. Since (1.8) is a consequence of (1.11), the strong energy condition is again satisfied.

For the classical neutrino field, however, it is well known that the condition (1.7) can be violated by even the simplest of solutions of (1.1), as we briefly illustrate in *flat* space–time. For the neutrino field

$$\phi^A(x^a) = \psi^A \exp(i\epsilon p_a x^a), \quad (1.13)$$

in terms of preferred coordinates with ψ^A a constant 2-spinor and

$$p_a = \sigma_{aAA'}\psi^A\psi^{A'}, \quad \epsilon = \pm 1,$$

the energy tensor (1.4) assumes the form

$$T_{ab} = 4\epsilon p_a p_b \Rightarrow E(u) = 4\epsilon(p_a u^a)^2.$$

Thus the strong energy condition is not valid for the field (1.13) with $\epsilon = -1$.

We thus weaken the requirements (1.7) and (1.8) in the following definitions.

Definition: A field is said to satisfy *the weak energy condition* E_1 if its energy tensor satisfies

$$E(u) \neq 0 \quad (1.14)$$

for all observers at each event P for which $T_{ab} \neq 0$.

Definition: A field is said to satisfy *the weak energy condition* E_2 if its energy tensor satisfies[11]

$$Q^a(u) \text{ is a timelike or null vector} \quad (1.15)$$

for all observers at each event P for which $T_{ab} \neq 0$.

Note that (1.15) implies (1.14). For the sake of brevity these fields will be referred to as *fields of class* E_1 and E_2, respectively.

We are going to show that the generic neutrino field does not satisfy E_1 (and hence E_2). In fact, if the p.n.c. of the field is nongeodesic, the condition E_1 is nonvalid. Furthermore, if the p.n.c. of a neutrino field of class E_1 has nonzero shear, then E_2 is nonvalid.

In Sec. 2 the Newman–Penrose formalism[12] is used to write the neutrino energy tensor in a form which explicitly displays its relationship to the p.n.c. of the field. The main results concerning the energy conditions are stated in Sec. 3, and comparisons are made with the electromagnetic field. The theorems of Sec. 3 are proved in Sec. 4, and in Sec. 5 the properties of the minimal polynomial of the neutrino energy tensor are briefly described for fields of class E_1.

2. NEUTRINO ENERGY TENSOR IN NEWMAN–PENROSE FORM

Let o^A and ι^A be a pair of spinor fields on space–time normalized so that

$$o_A\iota^A = 1. \quad (2.1)$$

The set $\{o^A, \iota^A\}$ is called a *dyad*[12] or *spin frame*. We adapt the dyad to the neutrino field by requiring that o^A be parallel to the neutrino field spinor ϕ^A, i.e.,

$$\phi^A(x^a) = \phi(x^a)o^A, \quad (2.2)$$

where ϕ is a complex function. The remaining freedom in the choice of dyad is described by[13]

$$o^{A*} = R^{\frac{1}{2}}e^{\frac{1}{2}iS}o^A,$$
$$\iota^{A*} = e^{-\frac{1}{2}iS}R^{-\frac{1}{2}}(\iota^A - RTo^A), \quad (2.3)$$

where $R > 0$ and S are arbitrary real functions and T is complex.

The dyad gives rise, in the usual manner,[12] to a null tetrad $\{k^a, n^a, m^a, \bar{m}^a\}$, with

$$k^a = \sigma^a{}_{AA'}o^A o^{A'}, \quad n^a = \sigma^a{}_{AA'}\iota^A\iota^{A'},$$
$$m^a = \sigma^a{}_{AA'}o^A\iota^{A'}, \quad \bar{m}^a = \sigma^a{}_{AA'}\iota^A o^{A'}. \quad (2.4)$$

The null vector field k^a is thus tangent to the p.n.c. of the neutrino field. The normalization (2.1) implies

$$k^a n_a = -m^a\bar{m}_a = 1,$$

all other contractions being zero. The freedom in the choice of tetrad corresponding to (2.3) is[13]

$$k^{a*} = Rk^a,$$
$$n^{a*} = R^{-1}n^a - Tm^a - \bar{T}\bar{m}^a + RT\bar{T}k^a, \quad (2.5)$$
$$m^{a*} = e^{iS}(m^a - R\bar{T}k^a).$$

The Newman–Penrose spin coefficients[12] are defined according to

$$\begin{aligned}
\kappa &= o^A o^{A'} o^B \nabla_{AA'} o_B, & \nu &= \iota^A\iota^{A'}\iota^B\nabla_{AA'}\iota_B,\\
\sigma &= o^A\iota^{A'}o^B\nabla_{AA'}o_B, & \lambda &= \iota^A o^{A'}\iota^B\nabla_{AA'}\iota_B,\\
\rho &= \iota^A o^{A'}o^B\nabla_{AA'}o_B, & \mu &= o^A\iota^{A'}\iota^B\nabla_{AA'}\iota_B,\\
\tau &= \iota^A\iota^{A'}o^B\nabla_{AA'}o_B, & \pi &= o^A o^{A'}\iota^B\nabla_{AA'}\iota_B,\\
\epsilon &= o^A o^{A'}\iota^B\nabla_{AA'}o_B, & \gamma &= \iota^A\iota^{A'}o^B\nabla_{AA'}\iota_B,\\
\beta &= o^A\iota^{A'}\iota^B\nabla_{AA'}o_B, & \alpha &= \iota^A o^{A'}o^B\nabla_{AA'}\iota_B,
\end{aligned} \quad (2.6)$$

where

$$\nabla_{AA'}o_B = \sigma^a{}_{AA'}o_{B;a}, \quad \text{etc.}$$

The spin coefficients κ, σ, and ρ describe certain geometric properties of the congruence k^a (which by construction is the p.n.c. of the neutrino field). In fact, this congruence is geodesic[12] if and only if $\kappa = 0$.

Furthermore, if $\kappa = 0$, then σ is the complex shear while $\rho - \bar{\rho}$ and $\rho + \bar{\rho}$ describe respectively the twist and expansion of this congruence.[14]

The neutrino field equations (1.1) are equivalent to a set of two equations involving the function ϕ [defined by (2.2)], the spin coefficients, and the Newman–Penrose differential operators[12]

$$D\phi = k^a \frac{\partial \phi}{\partial x^a}, \quad \Delta\phi = n^a \frac{\partial \phi}{\partial x^a},$$
$$\delta\phi = m^a \frac{\partial \phi}{\partial x^a}, \quad \bar{\delta}\phi = \bar{m}^a \frac{\partial \phi}{\partial x^a}. \tag{2.7}$$

A straightforward calculation yields

$$D\phi = (\rho - \epsilon)\phi, \quad \delta\phi = (\tau - \beta)\phi. \tag{2.8}$$

These are the neutrino field equations (1.1) in Newman–Penrose form when the spinor dyad is adapted to the neutrino field according to (2.2).

By means of a similar calculation, one can expand the energy tensor (1.4) in terms of the null tetrad vectors. After using the field equations (2.8) to eliminate $D\phi$ and $\delta\phi$, one obtains[15]

$$\begin{aligned} T_{ab} = {} & Ak_a k_b + B(4k_{(a}n_{b)} - g_{ab}) + Cm_a m_b + \bar{C}\bar{m}_a \bar{m}_b \\ & + 2Dn_{(a}m_{b)} + 2\bar{D}n_{(a}\bar{m}_{b)} + 2Ek_{(a}m_{b)} \\ & + 2\bar{E}k_{(a}\bar{m}_{b)}, \end{aligned} \tag{2.9}$$

where

$$A = 2i[\phi\Delta\bar{\phi} - \bar{\phi}\Delta\phi - \phi\bar{\phi}(\gamma - \bar{\gamma})], \tag{2.10}$$
$$B = -i\phi\bar{\phi}(\rho - \bar{\rho}), \quad C = 2i\phi\bar{\phi}\bar{\sigma}, \tag{2.11}$$
$$D = -i\phi\bar{\phi}\bar{\kappa}, \qquad E = i\bar{\phi}(\bar{\delta}\phi + \alpha\phi - 2\bar{\tau}\phi). \tag{2.12}$$

From the remarks following Eqs. (2.6), it is clear that the form of T_{ab} depends strongly on the properties of the p.n.c. In fact, *the p.n.c. of the neutrino field is geodesic if and only if the energy tensor satisfies* $T_{ab}k^a m^b = 0$. In addition, *if the p.n.c. is geodesic, it is shear-free if and only if* $T_{ab}m^a m^b = 0$ *and twist-free if and only if* $T_{ab}k^a n^b = 0$. The property

$$T_{ab}k^a k^b = 0 \tag{2.13}$$

ensures that these conditions are invariant under the tetrad freedom (2.5), as is easily verified.

3. THE ENERGY CONDITIONS

In this section we characterize those neutrino fields which satisfy the various energy conditions of Sec. 1. The proofs of the theorems are based on the expression (2.9) for the energy tensor, but involve only elementary algebra. Since they are somewhat tedious, they are postponed until Sec. 4 in order not to interrupt the discussion.

The first theorem shows that *the existence of curvature in the p.n.c. of a neutrino field prohibits the validity of the condition* E_1.

Theorem 3.1: The p.n.c. of any neutrino field of class E_1 is geodesic and its shear σ and twist $\omega \equiv \frac{1}{2}i(\rho - \bar{\rho})$ are restricted by

$$\sigma\bar{\sigma} - 4\omega^2 \leq 0.$$

We next assert the existence of a canonical null tetrad for this class of neutrino fields.

Theorem 3.2: For any neutrino field of class E_1 there exists a null tetrad $\{k^a, n^a, m^a, \bar{m}^a\}$ with k^a tangent to the p.n.c., such that

$$T_{ab}n^a m^b = 0, \quad (T_{ab}n^a n^b)(T_{cd}n^c k^d) \geq 0.$$

The canonical tetrad is determined uniquely [up to scale and phase factors R and S in Eq. (2.5)] provided that

$$\sigma\bar{\sigma} - 4\omega^2 \neq 0.$$

The conditions of the preceding theorems, in fact, characterize the neutrino fields of class E_1.

Theorem 3.3: A neutrino field is of class E_1 if and only if there exists a null tetrad with respect to which its energy tensor assumes the form

$$\begin{aligned} T_{ab} = {} & Ak_a k_b - 2\phi\bar{\phi}\omega(-g_{ab} + 4k_{(a}n_{b)}) \\ & + 2i\phi\bar{\phi}(\bar{\sigma}m_a m_b - \sigma\bar{m}_a\bar{m}_b), \end{aligned} \tag{3.1}$$

with

$$\sigma\bar{\sigma} - 4\omega^2 \leq 0, \quad A\omega \leq 0, \tag{3.2}$$

where A is given by (2.10).

Further insight into the role of the twist ω and the quantity A is provided by the following.

Theorem 3.4: For neutrino fields of class E_1 the energy density $E(u)$ satisfies

$$\operatorname{sgn} E(u) = -\operatorname{sgn} \omega, \tag{3.3}$$

provided that $\omega \neq 0$, and

$$|E(u)| \geq 2\phi\bar{\phi}\,|\omega|. \tag{3.4}$$

The lower bound is attained by an observer if and only if his velocity is a (timelike) eigenvector of the energy tensor.[16] Such observers exist if and only if $A = 0$ with respect to a canonical tetrad.

Corollary: If the p.n.c. of a neutrino field of class E_1 is twist-free ($\Rightarrow \sigma = 0, A \neq 0$), then

$$\operatorname{sgn} E(u) = \operatorname{sgn} A, \quad (3.5)$$

and there exist observers for whom $|E(u)|$ is arbitrarily small.

Comparison of Eqs. (3.4) and (1.10) suggests that the quantity $2\phi\bar{\phi}\,|\omega|$ is the analog for neutrino fields of class E_1 of the quantity k, defined by (1.12), for electromagnetic fields. We thus conclude that as far as energy considerations are concerned *only those neutrino fields of class E_1 whose p.n.c. is twist-free (and hence shear-free) are the analog of the null electromagnetic fields*[17] (which are characterized by $k = 0$). In this case the energy tensor (3.1) reduces to[18]

$$T_{ab} = Ak_a k_b.$$

The energy tensor of a null electromagnetic field is of precisely this form,[19] with k^a being tangent to the *repeated* p.n.c. of the field and A being a *positive* function. From a geometrical point of view, however, the analogy is not complete, since for a null electromagnetic field the repeated p.n.c. will not necessarily have simple geometric properties. However, for a *source-free* field, the repeated p.n.c. is geodesic and shear-free,[20] and for the simplest such solutions in flat space–time, namely the plane-fronted[21] and spherical[22] electromagnetic waves, it is also twist-free.

We now characterize the neutrino fields of class E_2.

Theorem 3.5: A neutrino field is of class E_2 if and only if there exists a null tetrad with respect to which its energy tensor assumes the form

$$T_{ab} = Ak_a k_b - 2\phi\bar{\phi}\omega(-g_{ab} + 4k_{(a}n_{b)}), \quad (3.6)$$

with

$$A\omega \leq 0. \quad (3.7)$$

As an immediate consequence of this theorem and Eqs. (3.3) and (3.5), we have the following.

Corollary: A neutrino field satisfies the strong energy condition if and only if there exists a null tetrad with respect to which its energy tensor assumes the form (3.6) with

$$A \geq 0, \quad \omega \leq 0,$$

but not both zero.

One additional consequence of Theorem 3.4 is of interest. For the generic electromagnetic field ($k \neq 0$), it is well known that the energy tensor admits a timelike eigenvector.[16] Equivalently, this means that there exist observers for whom the energy flow vector $Q^a(u)$ [as defined by Eq. (1.6)] is tangent to their world lines, so that they measure zero flow of energy in the field. In contrast to this, Theorem 3.4 asserts that *for the generic neutrino field of class E_1* (for which $A \neq 0$) *all observers measure a nonzero flow of energy in the field.* The same holds for the generic neutrino field of class E_2.

4. PROOFS OF THEOREMS

Proof of Theorem 3.1

We assume, with E_1 satisfied, that the p.n.c. of the neutrino field is nongeodesic and satisfies $\sigma\bar{\sigma} - 4\omega^2 > 0$ (in some region of space–time), and prove that the equation

$$T_{ab}u^a u^b = 0, \quad (4.1)$$

subject to $u^a u_a = 1$ and $T_{ab} \neq 0$, has a nontrivial solution for u^a (in that region). The theorem then follows by contradiction. The detailed proof falls naturally into two parts.

Part 1: Assume, with E_1 satisfied, that the p.n.c. of the neutrino field is nongeodesic in some region, which implies $D \neq 0$ by Eq. (2.12). The tetrad freedom (2.5) can now be used to transform

$$C = 0, \quad D = 1. \quad (4.2)$$

An arbitrary timelike unit vector field u^a can be expressed in terms of a null tetrad as

$$u^a = pk^a + qn^a + s(e^{i\theta}m^a + e^{-i\theta}\bar{m}^a), \quad (4.3)$$

where $p, q \neq 0$, $s > 0$, and θ are real functions which satisfy

$$2pq - 2s^2 = 1. \quad (4.4)$$

By means of Eqs. (2.9), and (4.2)–(4.4), Eq. (4.1) can be written as

$$Aq^3 - 4q^2 s\,|E| \cos(\psi - \theta) + qB(1 + 4s^2) - 2s(1 + 2s^2)\cos\theta = 0, \quad (4.5)$$

where $E = |E|\,e^{i\psi}$.

If $A \neq 0$, we choose $s \neq 0$ and $\cos\theta = 1$. Then (4.5) is a cubic in q with nonzero constant term and hence has at least one real nonzero solution for q.

If $A = 0$ but $E \neq 0$, choose θ such that

$$\cos(\psi - \theta) \neq 0.$$

Then (4.5) is a quadratic in q with discriminant

$$B^2(1 + 4s^2)^2 - 32s^2(1 + 2s^2)\,|E| \cos(\psi - \theta)\cos\theta,$$

which can be made nonnegative by choosing s sufficiently small. (Choose $s = 0$, q arbitrary if $B = 0$.)

If $A = E = 0$ but $B \neq 0$, choose $s \neq 0$ and $\cos\theta \neq 0$. Then (4.5) yields a unique nonzero solution for q.

Thus in all cases we have constructed a nontrivial solution u^a of (4.1) as required.

Part 2: Assume that the p.n.c. is geodesic but satisfies $\sigma\bar{\sigma} - 4\omega^2 > 0$, which by Eqs. (2.11) and (2.12) implies

$$D = 0, \quad C\bar{C} - 4B^2 > 0 \Rightarrow C \neq 0. \qquad (4.6)$$

By choosing

$$R = 1, \quad S = 0, \quad T = (2BE - C\bar{E})/(C\bar{C} - 4B^2)$$

in the tetrad transformation (2.5), we achieve

$$E = 0. \qquad (4.7)$$

By means of Eqs. (4.3), (4.4), (4.6), and (4.7), Eq. (4.1) yields

$$Aq^2 + B(1 + 4s^2) + 2\,|C|\,s^2 \cos(2\theta - \psi) = 0, \quad (4.8)$$

where we have written $C = |C|\,e^{i\psi}$.

If $AB < 0$, then $s = 0$, and $q = (-B/A)^{\frac{1}{2}}$ is a solution of (4.8). If $AB \geq 0$ and $A \neq 0$, we choose $s \neq 0$ and $\cos(2\theta - \psi) = -\operatorname{sgn} B$. Then

$$q^2 = [2s^2(|C| - 2\,|B|) - |B|]/|A|.$$

By virtue of (4.6), we can ensure that there is a real solution for q by choosing s sufficiently large.

If $A = 0$ and $B \neq 0$, we choose $\cos(2\theta - \psi) = -\operatorname{sgn} B$. Then, q arbitrary, $s = [|B|/2(|C| - 2\,|B|)]^{\frac{1}{2}}$ is a solution. (Choose $s = 0$, q arbitrary if $B = 0$.)

Thus in each case we have constructed a nontrivial solution u^a of (4.1) as required, and the proof is complete.

Proof of Theorem 3.2

Part 1: Proof by contradiction. Assume, with E_1 satisfied, that $T_{ab}n^am^b \neq 0$, i.e.,

$$E \neq 0, \qquad (4.9)$$

for all null tetrads with k^a tangent to the p.n.c. This means that $E \neq 0$ is preserved under the tetrad transformation (2.5) with $R = 1$, $S = 0$, and T arbitrary. Under such a transformation,

$$E^* = E + \bar{T}C + 2TB.$$

Thus $E^* = 0 \Rightarrow \bar{E}C - 2BE = T(4B^2 - C\bar{C})$. Clearly, we must have $4B^2 - C\bar{C} = 0$; otherwise, there exists a unique solution for T. On the other hand, if $4B^2 - C\bar{C} = 0$ *and* $\bar{E}C - 2BE = 0$, it is easily seen that there exists a (nonunique) solution for T, unless $B = C = 0$.

The assumption (4.9) thus implies[23]

(a) $4B^2 - C\bar{C} = 0, \quad \bar{E}C - 2BE \neq 0,$

or

(b) $B = C = 0. \qquad (4.10)$

Case (a): Under a transformation (2.5) with $R = 1$ and $S = 0$, the quantity $2AB - E\bar{E}$ transforms, since $4B^2 - C\bar{C} = 0 = D$, according to

$$(2AB - E\bar{E})^* = 2AB - E\bar{E} + T(2B\bar{E} - E\bar{C}) + \bar{T}(2BE - \bar{E}C).$$

Thus one may use (2.5) to achieve $2AB - E\bar{E} = 0$. The remaining freedom in (2.5) enables us to transform $E = 1$. We thus have the conditions

$$D = 0, \quad E = 1, \quad 4B^2 - C\bar{C} = 0,$$
$$2AB = 1, \quad C - 2B \neq 0. \qquad (4.11)$$

With u^a given by (4.3), Eq. (4.1) assumes the form

$$Aq^2 - 4qs\cos\theta + \epsilon\,|B|\,\{1 + 4s^2[1 + \epsilon\cos(\psi - 2\theta)]\} = 0,$$

where we have written $C = |C|\,e^{i\psi}$, $B = \epsilon\,|B|$, and $\epsilon = \pm 1$. Since $A \neq 0$, this is a quadratic in q, with discriminant

$$\Delta = 2(-1 + 4s^2\chi), \qquad (4.12)$$

where

$$\chi = \cos 2\theta - \epsilon\cos(\psi - 2\theta).$$

If $\epsilon = 1$, we choose $2\theta = \pi + \psi \Rightarrow \chi = 1 - \cos\psi \geq 0$. But $\chi = 0 \Leftrightarrow \cos\psi = 1 \Rightarrow C = 2B$, which contradicts (4.11). Thus $\chi > 0$ for all permissible ψ. Similarly, if $\epsilon = -1$, we choose $2\theta = \psi \Rightarrow \chi = 1 + \cos\psi$, which is positive for all permissible ψ. Thus since $\chi > 0$ in all cases, we can ensure $\Delta \geq 0$ in (5.12) by choosing s sufficiently large, and a solution for q exists.

Case (b): Here Eq. (4.1) reduces to

$$Aq^2 - 4qs\cos\theta = 0,$$

after we transform $E = 1$. With appropriate choices for s and θ (depending on whether $A = 0$ or $A \neq 0$), this equation admits a nonzero solution for q.

Thus the assumption (4.9) contradicts E_1.

Part 2: Proof by contradiction. Assume, with E_1 satisfied, that there exists a null tetrad with $T_{ab}n^am^b = 0$ but $(T_{ab}n^an^b)(T_{cd}n^ck^d) < 0$. We thus have the conditions

$$D = E = 0, \quad AB < 0.$$

It is sufficient to consider vectors u^a of the form (4.3 with $s = 0$. Equation (4.1) then reduces to

$$Aq^2 + B = 0,$$

which clearly has a real nonzero solution for q, contradicting E_1.

The proof is thus complete.

Proof of Theorem 3.3

That E_1 implies the form (3.1), (3.2) for T_{ab} is a consequence of Theorems 3.1 and 3.2. We prove the converse.

For the energy tensor given by (3.1) and (3.2), Eqs. (2.11), (4.3), and (4.4) enable us to write

$$\epsilon T_{ab}u^a u^b = |A|\, q^2 + |B| + 2s^2[2\,|B| + \epsilon\,|C| \cos(2\theta - \psi)], \quad (4.13)$$

where $C = Ce^{i\psi}$, $A = \epsilon\,|A|$, $B = \epsilon\,|B|$, and $\epsilon = \pm 1$. The condition $\sigma\sigma - 4\omega^2 \leq 0$ (which is equivalent to $4B^2 - C\bar{C} \geq 0$) implies $|B| \geq \frac{1}{2}\,|C|$. We thus obtain

$$\epsilon T_{ab}u^a u^b \geq |A|\, q^2 + |B| \geq |B| \geq 0, \quad (4.14)$$

for all u^a. The right-hand side of (4.13) is a sum of three nonnegative terms, and q is nonzero. This means that $\epsilon T_{ab}u^a u^b$ can vanish only if $T_{ab} = 0$. Thus E_1 is satisfied, and the proof is complete.

Proof of Theorem 3.4

It is an immediate consequence of (4.13) and (2.11) that

$$\operatorname{sgn} E(u) = \epsilon = \operatorname{sgn} B = -\operatorname{sgn} \omega,$$

as required. In addition, (4.14) asserts that

$$|E(u)| \geq |B| \equiv 2\phi\bar{\phi}\,|\omega|.$$

Furthermore, for T_{ab} of the form (3.1), the equations $T_{ab}u^b = \chi u_a$, with u^a given by (4.3) and (4.4), are equivalent to

$$B - \chi = 0, \quad A = 0,$$
$$s[2B + |C| \cos(2\theta - \psi)] = 0, \quad s \sin(2\theta - \psi) = 0, \quad (4.15)$$

since $p, q \neq 0$. Thus T_{ab} admits a timelike eigenvector if and only if $A = 0$. Furthermore, if $\epsilon T_{ab}u^a u^b = |B|$, it follows from (4.13) that

$$|A|\, q^2 = 0 = 2s^2[2\,|B| + \epsilon\,|C| \cos(2\theta - \psi)],$$

since both terms are nonnegative. If $s = 0$, Eq. (4.15) implies that u^a is an eigenvector. If $s \neq 0$, we obtain

$$4B^2 - |C|^2 = -|C|^2 \sin^2(2\theta - \psi).$$

Since $4B^2 - |C|^2 \geq 0$, we conclude that $\sin(2\theta - \psi) = 0$, so that u^a is again an eigenvector of T_{ab}.

Conversely, if u^a is a timelike eigenvector of T_{ab}, it follows immediately from (4.13) and (4.15) that $\epsilon T_{ab}u^a u^b = |B|$, and the proof is complete.

Proof of Theorem 3.5

Assume that E_2, and hence E_1, are satisfied. Thus by Theorem 3.3, T_{ab} can be expressed in the form (3.1), with the conditions (3.2). A straightforward calculation using (4.3) and (4.4) yields

$$Q^a(u)Q_a(u) = 2ABq^2 + B^2 - 2\,|C|\, s^2(|C| + 2B\cos 2\theta),$$

where we have used (2.5) to transform $C = \bar{C}$. Assume that the p.n.c. has nonzero shear, i.e., $C \neq 0$. Choose $\cos 2\theta = 0$. Then, for any u^a with s sufficiently large, we will have $Q^a(u)Q_a(u) < 0$, so that E_2 is violated. Thus E_2 implies that the p.n.c. has zero shear, and T_{ab} assumes the form (3.6), with (3.7).

Conversely, with T_{ab} of the form (3.6), (3.7), we obtain

$$Q^a(u)Q_a(u) = 2ABq^2 + B^2 \geq 0,$$

so that E_2 is satisfied. This completes the proof.

5. MINIMAL POLYNOMIAL OF THE ENERGY TENSOR

In this section we compare the neutrino and electromagnetic energy tensors from a different and purely algebraic point of view. As was mentioned in the introduction, null electromagnetic fields can be characterized by the vanishing of the scalar k defined by Eq. (1.12). One can also differentiate between null and nonnull electromagnetic fields by considering the minimal polynomial[24] of the energy tensor $T_a{}^b$ regarded as a 4×4 matrix. From Eq. (1.11) one obtains for nonnull fields the polynomial

$$m(x) = (x - k)(x + k), \quad k \neq 0,$$

which has distinct roots. For null fields, on the other hand, this polynomial reduces to

$$m(x) = x^2,$$

which has repeated roots.

As regards neutrino fields of class E_1 [with energy tensor (3.1)], the situation is more complicated. In fact, for the most general such field, the minimal polynomial $m(x)$ attains its maximum degree, namely 4. The various possibilities for $m(x)$ are distinguished by the vanishing or nonvanishing of the quantities A, $\Delta \equiv 4\omega^2 - \sigma\bar{\sigma}$, and σ, all of which have significance in connection with the theorems of Sec. 3. The different polynomials, which are arrived at by a straightforward calculation based on Eq. (3.1), are listed in Table I. Note that by Theorem 3.4 the energy tensor (3.1) admits a timelike eigenvector if and only if the minimal polynomial $m(x)$ has no repeated factors.

TABLE I. The minimal polynomial $m(x)$ of the neutrino energy tensor for fields of class E_1. The quantities A, B, and C are defined by Eqs. (2.10) and (2.11), while $\Delta \equiv 4\omega^2 - \sigma\bar{\sigma}$ depends on the twist ω and shear σ of the p.n.c.

A	Δ	σ	$m(x)$
$\neq 0$	0	0	x^2
0	$\neq 0$	0	$(x - B)(x + B)$
$\neq 0$	$\neq 0$	0	$(x - B)^2(x + B)$
0	0	$\neq 0$	$(x - B)(x + 3B)$
$\neq 0$	0	$\neq 0$	$(x - B)^2(x + 3B)$
0	$\neq 0$	$\neq 0$	$(x - B)[(x + B)^2 - C\bar{C}]$
$\neq 0$	$\neq 0$	$\neq 0$	$(x - B)^2[(x + B)^2 - C\bar{C}]$

6. CONCLUSION

In flat space–time one is interested not so much in the (local) energy tensor of a field as in the total energy and momentum which is obtained by integrating certain components of the energy tensor over spacelike hypersurfaces.[25] In curved space–time, however, it is essential that the (local) energy tensor of a field be meaningful since it acts as a source of the gravitational field and directly affects the curvature of space–time through the Einstein field equations and the Bianchi identities. It is in this connection that the energy conditions studied in this paper are perhaps of most relevance.

We thus conclude with some remarks on gravitational fields which have a neutrino or electromagnetic field as source. First, if the neutrino or electromagnetic field admits a geodesic and shear-free p.n.c. (and no other sources are present), then it is a straightforward consequence of the generalized Goldberg–Sachs theorem[26] that this preferred null congruence is a repeated p.n.c. of the Weyl tensor, so that the gravitational field is algebraically special.[27] Despite this analogy, one would suspect, however, that, in general,[28] the gravitational fields which have a neutrino field as source differ considerably from those which have an electromagnetic field as source. The reasons for this are the prominent role played by the twist of the p.n.c. for neutrino fields of class E_1, in comparison with the case of the electromagnetic field (see Sec. 3), and the greater algebraic complexity of the neutrino energy tensor (see Sec. 5). In addition, this conjecture is supported by an example of Brill and Cohen[18] which shows that certain homogeneous purely gravitational universes can be modified to include a homogeneous electromagnetic field but *not* a homogeneous neutrino field. This matter requires further investigation, however.

Note added in proof: Results equivalent to Theorems 3.1 and 3.2 of Sec. 3 have been obtained independently by J. B. Griffiths and R. A. Newing. ["Geometrical Aspects of the Two-Component Neutrino Field in General Relativity," J. Physics (to appear).]

ACKNOWLEDGMENTS

The author is indebted to R. Penrose for valuable discussions and suggestions in connection with this work. He also wishes to thank D. Trim for assistance with calculations and for pointing out an error in an earlier draft of this paper. This work was supported in part by a grant from the National Research Council of Canada.

[1] One can equivalently use the 4-component spinor formalism. See D. Brill and J. A. Wheeler, Rev. Mod. Phys. **29**, 465 (1957). For our purposes it is essential to use the 2-spinor formalism. The 2-spinor formalism has also been used to study neutrino fields in curved space–time by O. Bergmann, J. Math. Phys. **1**, 172 (1960), and J. B. Griffith and R. A. Newing, J. Phys. A **3**, 269 (1970).

[2] Spinor indices are denoted $A, B, \cdots$ and $A', B', \cdots$, and assume the values 1, 2. Our conventions as regards the 2-spinor formalism are those of R. Penrose, Ann. Phys. (N.Y.) **10**, 171 (1960). The metric tensor has signature $(- - - +)$, so that timelike vectors satisfy $v^a v_a > 0$.

[3] F. A. E. Pirani, *Brandeis Summer Institute, 1964* (Prentice-Hall, Englewood Cliffs, N.J., 1965), pp. 324–25.

[4] R. Penrose, Proc. Roy. Soc. (London) **A284**, 159 (1965). See p. 164.

[5] This expression can be derived from a Lagrangian using the $\sigma_{aAA'}$ as gravitational field variables, as in J. Wainwright, Tensor N. S. **19**, 265 (1968). See p. 271, Eq. (3.20) with $\epsilon = 1$ and $\phi^A = 0$. The sign of T_{ab} is chosen so that when $\epsilon = 1$ in the simple plane wave solution (1.13), corresponding to the propagation vector ϵp_a being future-pointing, this energy density is positive. It is a consequence of (1.1) that the energy tensor satisfies $T^a_a = 0$, $T_a{}^b{}_{;b} = 0$.

[6] By the word "field" we mean any distribution of matter–energy which acts as a source for gravitational fields, e.g., a distribution of perfect fluid, an electromagnetic field, a neutrino field, a scalar meson field, etc.

[7] J. L. Synge, *Relativity: The Special Theory* (North-Holland, Amsterdam, 1956), pp. 286–87. We have written the relevant formulas in covariant form.

[8] F. A. E. Pirani, Phys. Rev. **105**, 1089 (1957), esp. p. 1093. In this paper the projection of $Q_a(u)$ into the instantaneous rest space of the observer is considered. This leads, however, to the same expression for energy flow.

[9] The conditions (1.7) and (1.8) are not independent. In fact, (1.8) implies (1.7) while (1.7) implies that $Q^a(u)$ is necessarily future-pointing. Here it is essential that the velocity u^a be future-pointing.

[10] See Ref. 7, footnote on p. 324 and pp. 326–27.

[11] The use of this condition is the present context was suggested to me by R. Penrose.

[12] E. T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).

[13] See Ref. 3, pp. 338–39.

[14] For the geometric interpretation of these quantities, see, for example, R. Penrose, *Battelle Rencontres*, edited by C. de Witt and J. Wheeler (Benjamin, New York, 1968), p. 166.

[15] This expression was partially derived by J. B. Griffiths and R. A. Newing, J. Phys. A **3**, 269 (1970), Eq. (2.3).

[16] Such an observer is said to be "following the field" in the terminology of Pirani. See Ref. 8, p. 1093.

[17] This is a subclass of the neutrino fields (in flat space–time), tentatively identified as null neutrino fields by R. Penrose, J. Math. Phys. **10**, 38 (1969).

[18] Neutrino fields whose energy tensor is of the form $T_{ab} = f(x)p_a p_b$, where p_a is (necessarily) null, have been considered by D. R. Brill and J. M. Cohen, J. Math. Phys. **7**, 238 (1966), esp. p. 242, and J. B. Griffiths and R. A. Newing, Ref. 15. It is a consequence of (2.13) that p_a is necessarily tangent to the p.n.c. of the field so that T_{ab} is of the form (3.6). This means that the p.n.c. is geodesic, shear-free, and twist-free (see end of Sec. 2) in agreement with a result of Ref. 15.

[19] See Ref. 7, p. 343.
[20] This is the so-called Mariot–Robinson theorem. See Ref. 3, p. 344.
[21] W. Kundt, Z. Physik **163**, 77 (1961), pp. 82–83.
[22] I. Robinson and A. Trautman, Proc. Roy. Soc. London **A265**, 463 (1962), pp. 463–64.
[23] Note that when $4B^2 - C\bar{C} = 0$, the quantity $\bar{E}C - 2BE$ is invariant under a tetrad transformation (2.5) with $R = 1$ and $S = 0$.
[24] G. Birkhoff and S. MacLane, *A Survey of Modern Algebra* (Macmillan, New York, 1953), rev. ed., p. 316.
[25] P. Roman, *The Theory of Elementary Particles* (North-Holland, Amsterdam, 1960), p. 226.
[26] See, for example, P. Szekeres, J. Math. Phys. **7**, 751 (1966), who proves this result for the case of gravitational fields with null electromagnetic fields as source (pp. 754–55).
[27] See, for example, Ref. 14, p. 162.
[28] The particular case of a null neutrino field and a null electromagnetic field appears as an exception, since there are gravitational fields which can be interpreted as having either a null neutrino field or a null electromagnetic field as source. See Refs. 15 and 18.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Arbitrary-Spin Wave Equations and Lorentz Invariance

M. Seetharaman, J. Jayaraman, and P. M. Mathews
Division of Theoretical Physics, University of Madras, Madras-25, India

(Received 15 June 1970)

In an earlier paper it was shown that invariance of the arbitrary-spin wave equation $i(\partial\psi/\partial t) = H\psi$ [with ψ transforming according to the representation $D(0, s) \oplus D(s, 0)$] under rotations, T, C, and P, and boosts in the direction of the momentum, permits four infinite classes of solutions for H. Here we show that when invariance under boosts transverse to the momentum also is imposed, just two possible forms for H survive (in the c-number formalism). When we go over to the q-number theory, one of these forms leads to a consistent theory for half-integer spins only and the other for integer spins.

1. INTRODUCTION

Considerable attention has been bestowed in recent years on formalisms[1–6] employing $2(2s + 1)$-component fields for the description of particles (and antiparticles) of spin s. Relativistic wave equations involving such fields have the advantage of not requiring any supplementary conditions since the field has only the minimum number of components dictated by the spin value. Such equations are obtained either directly from a knowledge of the unitary irreducible representation of the Poincaré group, by assuming the field to transform according to a direct sum of two such irreducible representations corresponding to spin s and mass $\pm m$, as in Foldy's work,[1] or by one or other of a variety of approaches starting with the requirement that the field should transform according to the representation $D(0, s) \oplus D(s, 0)$ of the homogeneous Lorentz group.[3–6] While it might appear that the equation resulting from this last requirement would be just a special one of an infinity of alternative possibilities which exist for given spin, actually it is quite general since the various types of wave equations for spin s can all be reduced to this form (in the absence of interactions) by making use of the relativistically invariant relations that can be written down[7,8] between different irreducible representations of the homogeneous Lorentz group. An explicit verification of this statement in the case $s = \frac{3}{2}$ may be found in the work of Shay, Song, and Good,[9] who have shown how to reduce the manifestly covariant Rarita–Schwinger,[10] Dirac,[11] Fierz and Pauli,[12] Bargmann and Wigner[13] equations to the form obtained in Refs. 3 and 5 (which is not manifestly covariant). The use of the representation $D(0, s) \oplus D(s, 0)$ then amounts simply to a convenient standard choice. This is especially appropriate for investigations aimed at unraveling the roles of the different invariance conditions (and other requirements like quantizability of the theory) in determining the admissible forms of wave equations and other aspects of the structure of theories of particles of given arbitrary spin. For, unlike the requirements in the more conventional approaches in which extra conditions (such as that the covariance be *manifest*) impose unintended restrictions in subtle ways,[14,15] in the derivation of wave equations involving the representation $D(0, s) \oplus D(s, 0)$ the various requirements can be imposed independently and at will; and the resulting class of equations encompasses the conventional equations too, in the sense already indicated.

In the course of an investigation of this type, initiated by one of us some time ago,[5] it was shown[6] that if the wave equation

$$i\frac{\partial\psi}{\partial t} = H\psi, \tag{1}$$

for a particle of spin s, is to be invariant under (i) rotations, (ii) pure Lorentz transformations parallel

[19] See Ref. 7, p. 343.
[20] This is the so-called Mariot-Robinson theorem. See Ref. 3, p. 344.
[21] W. Kundt, Z. Physik **163**, 77 (1961), pp. 82-83.
[22] I. Robinson and A. Trautman, Proc. Roy. Soc. London **A265**, 463 (1962), pp. 463-64.
[23] Note that when $4B^2 - C\bar{C} = 0$, the quantity $\bar{E}C - 2BE$ is invariant under a tetrad transformation (2.5) with $R = 1$ and $S = 0$.
[24] G. Birkhoff and S. MacLane, *A Survey of Modern Algebra* (Macmillan, New York, 1953), rev. ed., p. 316.
[25] P. Roman, *The Theory of Elementary Particles* (North-Holland, Amsterdam, 1960), p. 226.
[26] See, for example, P. Szekeres, J. Math. Phys. **7**, 751 (1966), who proves this result for the case of gravitational fields with null electromagnetic fields as source (pp. 754-55).
[27] See, for example, Ref. 14, p. 162.
[28] The particular case of a null neutrino field and a null electromagnetic field appears as an exception, since there are gravitational fields which can be interpreted as having either a null neutrino field or a null electromagnetic field as source. See Refs. 15 and 18.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Arbitrary-Spin Wave Equations and Lorentz Invariance

M. Seetharaman, J. Jayaraman, and P. M. Mathews
Division of Theoretical Physics, University of Madras, Madras-25, India

(Received 15 June 1970)

In an earlier paper it was shown that invariance of the arbitrary-spin wave equation $i(\partial\psi/\partial t) = H\psi$ [with ψ transforming according to the representation $D(0, s) \oplus D(s, 0)$] under rotations, T, C, and P, and boosts in the direction of the momentum, permits four infinite classes of solutions for H. Here we show that when invariance under boosts transverse to the momentum also is imposed, just two possible forms for H survive (in the c-number formalism). When we go over to the q-number theory, one of these forms leads to a consistent theory for half-integer spins only and the other for integer spins.

1. INTRODUCTION

Considerable attention has been bestowed in recent years on formalisms[1-6] employing $2(2s + 1)$-component fields for the description of particles (and antiparticles) of spin s. Relativistic wave equations involving such fields have the advantage of not requiring any supplementary conditions since the field has only the minimum number of components dictated by the spin value. Such equations are obtained either directly from a knowledge of the unitary irreducible representation of the Poincaré group, by assuming the field to transform according to a direct sum of two such irreducible representations corresponding to spin s and mass $\pm m$, as in Foldy's work,[1] or by one or other of a variety of approaches starting with the requirement that the field should transform according to the representation $D(0, s) \oplus D(s, 0)$ of the homogeneous Lorentz group.[3-6] While it might appear that the equation resulting from this last requirement would be just a special one of an infinity of alternative possibilities which exist for given spin, actually it is quite general since the various types of wave equations for spin s can all be reduced to this form (in the absence of interactions) by making use of the relativistically invariant relations that can be written down[7,8] between different irreducible representations of the homogeneous Lorentz group. An explicit verification of this statement in the case $s = \frac{3}{2}$ may be found in the work of Shay, Song, and Good,[9] who have shown how to reduce the manifestly covariant Rarita-Schwinger,[10] Dirac,[11] Fierz and Pauli,[12] Bargmann and Wigner[13] equations to the form obtained in Refs. 3 and 5 (which is not manifestly covariant). The use of the representation $D(0, s) \oplus D(s, 0)$ then amounts simply to a convenient standard choice. This is especially appropriate for investigations aimed at unraveling the roles of the different invariance conditions (and other requirements like quantizability of the theory) in determining the admissible forms of wave equations and other aspects of the structure of theories of particles of given arbitrary spin. For, unlike the requirements in the more conventional approaches in which extra conditions (such as that the covariance be *manifest*) impose unintended restrictions in subtle ways,[14,15] in the derivation of wave equations involving the representation $D(0, s) \oplus D(s, 0)$ the various requirements can be imposed independently and at will; and the resulting class of equations encompasses the conventional equations too, in the sense already indicated.

In the course of an investigation of this type, initiated by one of us some time ago,[5] it was shown[6] that if the wave equation

$$i\frac{\partial\psi}{\partial t} = H\psi, \tag{1}$$

for a particle of spin s, is to be invariant under (i) rotations, (ii) pure Lorentz transformations parallel to the momentum direction (longitudinal boosts), and (iii) T, C, and P, then the Hamiltonian H must belong to one of four infinite classes, each class being characterized by specific commutation relations among the representatives of the discrete operations. On account of the complexity of the equations expressing invariance under boosts transverse to the momentum directions, these were not considered in general (though it was verified that the special case of Ref. 5 possesses this invariance too). Instead, the question of quantizability of the theory was investigated,[16] and it was found that for any given spin a unique H is singled out (out of the four classes mentioned above) by the condition that the commutator or anticommutator of fields should vanish at spacelike separations. Further the usual spin–statistics connection follows as a consequence of this same microcausality condition—a remarkable result insofar as the wave equation was not even required to be invariant under the proper Lorentz group (invariance under transverse boosts was not imposed).

In the present paper we return to the c-number theory and investigate the question of invariance under transverse boosts. We show that when invariance of wave equation (1) is demanded with respect to all proper Lorentz transformations as well as T, C, and P, there are still *two* possible forms of H for any given spin. When the further condition of quantizability is imposed, it follows as a special case of the earlier treatment[16] that one of the two forms is applicable to integer spin particles alone and the other to half-integer spin particles.

2. DETERMINATION OF THE HAMILTONIAN

The conditions which determine H in the c-number theory[5] are

$$[H, P_\mu] = 0, \quad [H, \mathbf{J}] = 0, \quad [H, \mathbf{K}] = i\mathbf{P}, \tag{2a}$$

required for invariance of (1) under the Poincaré group, and

$$PH = HP, \quad TH = HT, \quad CH = -HC, \tag{2b}$$

for invariance under P, T, and C. The boost invariance condition [the last of Eq. (2a)] can be reexpressed as

$$[H, 2\boldsymbol{\lambda}] = [H, -i[\mathbf{x}, H]] \equiv [H, \nabla_p H], \tag{3}$$

where $i\boldsymbol{\lambda}$ is the "spin" part of the boost generator $\mathbf{K}$ in the representation $D(0, s) \oplus D(s, 0)$. It is shown in Refs. 5 and 6 that the set of Eqs. (2), excluding the boost invariance condition, has general solutions of the following two types:

$$H = \sum_{\nu\geq 0} c_\nu C_\nu + \sigma \sum_{\nu\geq 0} b'_\nu B_\nu \tag{4}$$

and

$$H = \sum_{\nu\geq 0} c_\nu C_\nu + \sigma \sum_{\nu\geq 0} c'_\nu C_\nu. \tag{5}$$

Here B_ν and C_ν are, respectively, even and odd functions (polynomials) of $\lambda_p = (\boldsymbol{\lambda}\cdot\mathbf{p}/p)$ and are defined by

$$B_\nu = \Lambda_\nu + \Lambda_{-\nu}, \quad C_\nu = \Lambda_\nu - \Lambda_{-\nu}, \tag{6}$$

where Λ_ν is the projection operator to the eigenvalue ν, $\nu = -s, -s+1, \cdots, s$, of λ_p. The matrix σ represents the parity operation $P\psi(\mathbf{x}, t) = \sigma\psi(-\mathbf{x}, t)$, and may quite generally be taken as the Pauli matrix ρ_1, which would effect the interchange of the $D(0, s)$ and $D(s, 0)$ parts of ψ. The coefficients c_ν, b'_ν, and c'_ν are functions of $p = |\mathbf{p}|$ on which the condition $H^2 = \mathbf{p}^2 + m^2$ (for unique mass m) imposes the constraints[5]

$$c_\nu^2 + b_\nu'^2 = E^2, \quad c_\nu^2 - c_\nu'^2 = E^2. \tag{7}$$

Incidentally, by virtue of the same condition which, on differentiation, gives $[H(\nabla_p H) + (\nabla_p H)H] = 2\mathbf{p}$, we may rewrite (3) as

$$H\nabla_p H = [H, \boldsymbol{\lambda}] + \mathbf{p}. \tag{8}$$

It is through this equation that we will impose boost invariance.

In determining the effect of (8) on H we will confine our attention to the form (4). [The treatment of (5) is quite similar, and we will give only the final results.] It is convenient, as in Ref. 5, to consider, along with the expression (4), also its explicit polynomial form

$$H = \sum_{l=1}^{2s} f_l(p)(\boldsymbol{\lambda}\cdot\mathbf{p})^l + \sigma \sum_{L=0}^{2s} g_L(p)(\boldsymbol{\lambda}\cdot\mathbf{p})^L. \tag{9}$$

In Eq. (9) the sums are to be taken over all odd (or even) integers in the specified range according as the summation index is lower case (or capital). We shall use Greek indices for integers which range over odd *and* even values. These conventions will be followed throughout this paper. The relations between the coefficients in (4) and (9) are

$$c_\nu = \sum_{l=1}^{2s} (\nu p)^l f_l, \quad b'_\nu = \sum_{L=0}^{2s} (\nu p)^L g_L. \tag{10}$$

These follow from the spectral representations of powers of $(\boldsymbol{\lambda}\cdot\mathbf{p})$:

$$(\boldsymbol{\lambda}\cdot\mathbf{p})^l = \sum_{\nu\geq 0} (\nu p)^l C_\nu, \quad (\boldsymbol{\lambda}\cdot\mathbf{p})^L = \sum_{\nu\geq 0} (\nu p)^L B_\nu. \tag{11}$$

It is known[6] that if only the longitudinal part of (8), obtained by taking the scalar product with $\mathbf{p}$, is employed, one is led to ordinary differential equations which yield expressions containing an undetermined constant of integration for c_ν and b'_ν. When the full power of Eq. (8) is brought to bear on the Hamiltonian, c_ν and b'_ν get determined completely; but, to see this, one has to go through a rather involved analysis which we now outline. Fuller details may be found in Appendix C.

The first step is to substitute H in the form (9) into (8) and evaluate both sides explicitly. For finding $[H, \boldsymbol{\lambda}]$ as well for pulling to one end the factor $\boldsymbol{\lambda}$ in

$$\nabla_p(\boldsymbol{\lambda}\cdot\mathbf{p})^\alpha = \sum_{\mu=0}^{\alpha-1}(\boldsymbol{\lambda}\cdot\mathbf{p})^\mu\boldsymbol{\lambda}(\boldsymbol{\lambda}\cdot\mathbf{p})^{\alpha-\mu-1} \tag{12}$$

(which appears in the left-hand side), we use the formula

$$[(\boldsymbol{\lambda}\cdot\mathbf{p})^\beta, \boldsymbol{\lambda}] = p^\beta \sum_{m=1}^{\beta}\binom{\beta}{m}\lambda_p^{\beta-m} i\rho_3 \boldsymbol{\tau} - p^\beta\sum_{M=2}^{\beta}\binom{\beta}{M}\lambda_p^{\beta-M}\boldsymbol{\lambda} + p^{\beta-1}\sum_{M=2}^{\beta}\binom{\beta}{M}\lambda_p^{\beta-M+1}\mathbf{p}, \tag{13}$$

where

$$\boldsymbol{\tau} = \boldsymbol{\lambda}\times\mathbf{p}/p \tag{14}$$

and ρ_3 is the third Pauli matrix which here has the effect of multiplying the $D(0, s)$ and $D(s, 0)$ parts of the wavefunction by 1 and -1, respectively. The proof of (13) is outlined in Appendix A.

With the use of Eqs. (12), (13), and (11) we analyze both sides of (8) into linearly independent terms whose coefficients on the two sides are then equated. The resulting equations contain the f_l and g_L through sums of precisely the form (10), which can thus be replaced by c's and b''s. Finally, we obtain the set of relations

$$c_\nu(c_{\nu+1} - c_{\nu-1}) + b'_\nu(b'_{\nu+1} - b'_{\nu-1}) = p(2c_\nu - c_{\nu+1} - c_{\nu-1}), \tag{15a}$$

$$2E^2 - c_\nu(c_{\nu+1} + c_{\nu-1}) - b'_\nu(b'_{\nu+1} + b'_{\nu-1}) = p(c_{\nu+1} - c_{\nu-1}), \tag{15b}$$

$$b'_\nu(c_{\nu+1} - c_{\nu-1}) - c_\nu(b'_{\nu+1} - b'_{\nu-1}) = p(2b'_\nu + b'_{\nu+1} + b'_{\nu-1}), \tag{15c}$$

$$b'_\nu(2c_\nu - c_{\nu+1} - c_{\nu-1}) - c_\nu(2b'_\nu - b'_{\nu+1} - b'_{\nu-1}) = p(b'_{\nu-1} - b'_{\nu+1}), \tag{15d}$$

$$2p\left(b'_\nu\frac{dc_\nu}{dp} - c_\nu\frac{db'_\nu}{dp}\right) - \nu b'_\nu(c_{\nu+1} - c_{\nu-1}) + \nu c_\nu(b'_{\nu+1} - b'_{\nu-1}) = \nu p(2b'_\nu - b'_{\nu+1} - b'_{\nu-1}). \tag{15e}$$

In the case when ν has its minimum value ν_0, the symbols c_{ν_0-1} and b'_{ν_0-1} are themselves undefined; but in (15) they simply stand for $\sum_l [(\nu_0 - 1)p]^l f_l$ and $\sum_L [(\nu_0 - 1)p]^L g_L$, respectively, which reduce to $-c_1$ and b'_1 in the integer spin case ($\nu_0 = 0$) and to $-c_{\frac{1}{2}}$ and $b'_{\frac{1}{2}}$ in the half-integer spin case ($\nu_0 = \frac{1}{2}$). When ν has its maximum value $\nu = s$, the linear independence conditions are different from those for $\nu \neq s$ (see Appendix B), and consequently the unwanted c_{s+1} and b'_{s+1} drop out altogether. The equations in this case turn out to be of exactly the same form as the following recurrence relations which one obtains from Eqs. (15a)-(15d):

$$E^2 - c_\nu c_{\nu-1} - b'_\nu b'_{\nu-1} = p(c_\nu - c_{\nu-1}), \tag{16a}$$

$$c_\nu b'_{\nu-1} - b'_\nu c_{\nu-1} = p(b'_\nu + b'_{\nu-1}). \tag{16b}$$

The b''s can be eliminated from (16) with the aid of the first of Eqs. (7). The result is that either $c_\nu = c_{\nu-1}$ or

$$\frac{c_\nu}{E} = \frac{(c_{\nu-1}/E) + \tanh 2\theta}{1 + (c_{\nu-1}/E)\tanh 2\theta}, \quad \nu = s, s-1, \cdots, \nu_0 + 1, \tag{17}$$

where θ is defined by

$$E = m\cosh\theta, \quad p = m\sinh\theta. \tag{18}$$

If we take $c_\nu = c_{\nu-1} = \cdots = c_{\nu_0}$, on substituting the value of c_{ν_0} (see below) into (16), we are led to a contradiction with the mass conditions (7). Thus, only the recurrence relation (17) is to be considered, and its solution is immediately obtained as

$$c_\nu = E\tanh 2\nu\theta, \quad \nu = s, s-1, \cdots, \nu_0. \tag{19}$$

There exists an alternative solution of (17), namely $c_\nu = E\coth 2\nu\theta$, but it is inconsistent with the special case $\nu = \nu_0$ of Eqs. (15). In the half-integer spin case, for example, one gets for $\nu = \frac{1}{2}$

$$c_{\frac{1}{2}}^2 - b'^2_{\frac{1}{2}} = 2pc_{\frac{1}{2}} - E^2, \tag{20}$$

which, together with (7), gives $c_{\frac{1}{2}} = p = E\tanh\theta$, confirming (19) and ruling out the other solution of (17). The same thing happens in the integer spin case too. The solution is thus given by (19) together with

$$b'_\nu = E\,\mathrm{sech}\,2\nu\theta, \tag{21}$$

which is a consequence of (19) and (7). Equation (15e) has not been made use of in the above analysis, but one may verify that our solution satisfies this equation too. We conclude then that the only Hamiltonian of the type (4) which is consistent with boost invariance is

$$H = \sum_\nu E\tanh 2\nu\theta\, C_\nu + \sigma\sum_\nu E\,\mathrm{sech}\,2\nu\theta\, B_\nu. \tag{22}$$

An analysis on the same lines as above, starting with the form (5) instead of (4), shows that there exists one admissible Hamiltonian of that type, namely

$$H = \sum_\nu E\coth 2\nu\theta\, C_\nu + \sigma\sum_\nu E\,\mathrm{csch}\,2\nu\theta\, C_\nu. \tag{23}$$

It may be noted that (22) is the special case corresponding to $\eta_v = 0$ of the class of Hamiltonians comprised under case (i) of Ref. 6, which is associated with the commutation rules $TP = PT$ and $CP = -PC$ between the discrete symmetry operators. Similarly, (23) belongs to case (iii), characterized by $TP = PT$ and $CP = PC$. The other two classes are now completely ruled out.

3. DISCUSSION

In determining the forms of H allowed by Poincaré and T, C, P invariance, we have steered clear of the trammels of nonessential criteria like *manifest* covariance and other considerations like locality which are not immediately pertinent to the question of Lorentz invariance. From earlier work[17,6] we know that with either of the choices (22) or (23) for H there are two possibilities for the metric operator M, one positive definite (M_1) and the other indefinite (M_2), such that $\int \psi^\dagger M\psi d^3x$ is Lorentz invariant. Thus, the freedom in the c-number theory is fourfold. If now one demands also locality (in the sense that observables of the field, like total energy and momentum, be expressible as space integrals of local functions of the field), then only two possibilities survive,[16] namely, Eq. (22) taken with M_1 and Eq. (23) with M_2. On the other hand, if quantizability of the theory is demanded (with the condition that either the commutator or the anticommutator of ψ and ψ^* should vanish at spacelike separations), then it follows under much more general conditions (without requiring locality as above or even invariance under transverse boosts) that Eq. (22) is the only one applicable to half-integer spin particles (which have to be fermions) and that Eq. (23) is applicable only to integer spin bosons.[16,18] Also, the requirement that field observables be generators of Poincaré transformations then leads to the association of M_1 with fermions and M_2 with bosons.[16] As a consequence, one gets locality as a byproduct in the quantized theory.

As the next step in the program, to delineate clearly the roles of the various conditions which go into the formulation of field theories of arbitrary-spin particles, we have investigated the consequence of relaxing the discrete invariance requirements. The results will be reported separately.

ACKNOWLEDGMENTS

Two of us (M. S. and J. J.) are grateful to the Department of Atomic Energy and Council of Scientific and Industrial Research, respectively, for financial assistance in the form of Research Fellowships.

APPENDIX A

To establish that the commutator $[(\boldsymbol{\lambda}\cdot\mathbf{p})^\beta, \boldsymbol{\lambda}]$ is given by the expression (13) for any positive integer β, we use the method of induction, starting with

$$[(\boldsymbol{\lambda}\cdot\mathbf{p}), \boldsymbol{\lambda}] = pi\rho_3\boldsymbol{\tau}, \tag{A1}$$

$$[(\boldsymbol{\lambda}\cdot\mathbf{p})^2, \boldsymbol{\lambda}] = 2p^2\lambda_p i\rho_3\boldsymbol{\tau} - p^2\boldsymbol{\lambda} + p\lambda_p\mathbf{p}. \tag{A2}$$

These relations are deduced directly from the fact that $\boldsymbol{\lambda} = \rho_3\mathbf{S}$ and $[S_i, S_j] = i\epsilon_{ijk}S_k$. Assume now that Eq. (13) is valid for some odd integer $\beta = l$, in which case it reads

$$[(\boldsymbol{\lambda}\cdot\mathbf{p})^l, \boldsymbol{\lambda}] = p^l\sum_{m=1}^{l}\binom{l}{m}\lambda_p^{l-m}i\rho_3\boldsymbol{\tau} - p^l\sum_{M=2}^{l-1}\binom{l}{M}\lambda_p^{l-M}\boldsymbol{\lambda} + p^{l-1}\sum_{M=2}^{l-1}\binom{l}{M}\lambda_p^{l-M+1}\mathbf{p}. \tag{A3}$$

Then, for $\beta = l + 2$, we get, with the aid of (A2),

$$\begin{aligned}
[(\boldsymbol{\lambda}\cdot\mathbf{p})^{l+2}, \boldsymbol{\lambda}]
&= (\boldsymbol{\lambda}\cdot\mathbf{p})^l[(\boldsymbol{\lambda}\cdot\mathbf{p})^2, \boldsymbol{\lambda}] + [(\boldsymbol{\lambda}\cdot\mathbf{p})^l, \boldsymbol{\lambda}](\boldsymbol{\lambda}\cdot\mathbf{p})^2 \\
&= p^{l+2}\Bigg[\sum_{m=1}^{l}\binom{l}{m}\lambda_p^{l-m+2} + \sum_{m=1}^{l}\binom{l}{m}\lambda_p^{l-m} + 2\sum_{M=2}^{l-1}\binom{l}{M}\lambda_p^{l-M+1} + 2\lambda_p^{l+1}\Bigg]i\rho_3\boldsymbol{\tau} \\
&\quad - p^{l+2}\Bigg[\sum_{M=2}^{l-1}\binom{l}{M}\lambda_p^{l-M+2} + \sum_{M=2}^{l-1}\binom{l}{M}\lambda_p^{l-M} + 2\sum_{m=1}^{l}\binom{l}{m}\lambda_p^{l-m+1} + \lambda_p^l\Bigg]\boldsymbol{\lambda} \\
&\quad + p^{l+2}\Bigg[\sum_{M=2}^{l-1}\binom{l}{M}\lambda_p^{l-M+2} + \sum_{M=2}^{l-1}\binom{l}{M}\lambda_p^{l-M} + 2\sum_{m=1}^{l}\binom{l}{m}\lambda_p^{l-m+1} + \lambda_p^l\Bigg]\lambda_p\mathbf{p}.
\end{aligned} \tag{A4}$$

When the summation indices are suitably redefined and the relation

$$\binom{\beta}{n} + \binom{\beta}{n-1} = \binom{\beta+1}{n} \tag{A5}$$

between binomial coefficients is used, (A4) reduces to the form (A3) with l replaced by $l + 2$. It follows then that if Eq. (A3) is true for an odd integer l, it is also true for the next higher one, $(l + 2)$. However, Eq. (A3) is independently verified to be true for $l = 1$ [Eq. (A1)]; hence, it is valid for all odd-integer values of β. A similar induction procedure starting with (A2) shows that the relation

$$[(\boldsymbol{\lambda}\cdot\mathbf{p})^\beta, \boldsymbol{\lambda}] = p^\beta\sum_{m=1}^{\beta-1}\binom{\beta}{m}\lambda_p^{\beta-m}i\rho_3\boldsymbol{\tau} - p^\beta\sum_{M=2}^{\beta}\binom{\beta}{M}\lambda_p^{\beta-M}\boldsymbol{\lambda} + p^{\beta-1}\sum_{M=2}^{\beta}\binom{\beta}{M}\lambda_p^{\beta-M+1}\mathbf{p}, \tag{A6}$$

which is the specific form of (13) for an even value of β, holds for all even β.

Equations (A3) and (A6) can be used directly for evaluating $\nabla_p(\boldsymbol{\lambda}\cdot\mathbf{p})^\beta$. For an odd value, $\beta = l$, for instance, we have

$$\nabla_p(\boldsymbol{\lambda}\cdot\mathbf{p})^l = \sum_{\alpha=0}^{l-1} (\boldsymbol{\lambda}\cdot\mathbf{p})^\alpha \boldsymbol{\lambda} (\boldsymbol{\lambda}\cdot\mathbf{p})^{l-\alpha-1}. \quad \text{(A7)}$$

By pulling $\boldsymbol{\lambda}$ to the right, using(A3), then re-expressing powers of $(\boldsymbol{\lambda}\cdot\mathbf{p})$ in terms of B_ν and C_ν, using (11), and by observing that the sum over α becomes just a binomial series, we finally obtain the following explicit form, which will be used in Appendix C:

$$\begin{aligned}\nabla_p(\boldsymbol{\lambda}\cdot\mathbf{p})^l &= \tfrac{1}{2}\sum_\nu \{[(\nu+1)p]^l - [(\nu-1)p]^l\}\frac{B_\nu\boldsymbol{\lambda}}{p} \\ &\quad - \tfrac{1}{2}\sum_\nu \{[(\nu+1)p]^l + [(\nu-1)p]^l - 2(\nu p)^l\}\frac{C_\nu i\rho_3\boldsymbol{\tau}}{p} \\ &\quad - \tfrac{1}{2}\sum_\nu \{\nu[(\nu+1)p]^l - \nu[(\nu-1)p]^l - 2l(\nu p)^l\}\frac{C_\nu\mathbf{p}}{p^2}. \end{aligned} \quad \text{(A8)}$$

For the gradient of an even power of $(\boldsymbol{\lambda}\cdot\mathbf{p})$, the use of (A6) gives an expression identical to (A8) except that B_ν and C_ν are interchanged and, of course, l is replaced by L:

$$\begin{aligned}\nabla_p(\boldsymbol{\lambda}\cdot\mathbf{p})^L &= \tfrac{1}{2}\sum_\nu \{[(\nu+1)p]^l - [(\nu-1)p]^L\}\frac{C_\nu\boldsymbol{\lambda}}{p} \\ &\quad - \tfrac{1}{2}\sum_\nu \{[(\nu+1)p]^l + [(\nu-1)p]^L - 2(\nu p)^L\}\frac{B_\nu i\rho_3\boldsymbol{\tau}}{p} \\ &\quad - \tfrac{1}{2}\sum_\nu \{\nu[(\nu+1)p]^L - \nu[(\nu-1)p]^L \\ &\qquad - 2L(\nu p)^L\}\frac{B_\nu\mathbf{p}}{p^2}. \end{aligned} \quad \text{(A9)}$$

APPENDIX B

The reduction of the boost invariance condition (8) to the set of Eqs. (15)—which is carried out in detail in Appendix C—depends on the fact that the quantities $iC_\nu\rho_3\boldsymbol{\tau}$, $B_\nu\boldsymbol{\lambda}$, and $C_\nu\mathbf{p}/p$ are linearly independent for $\nu \neq s$ while, for $\nu = s$,

$$C_s i\rho_3\boldsymbol{\tau} - B_s\boldsymbol{\lambda} + sC_s\mathbf{p}/p = 0. \quad \text{(B1)}$$

We prove these results here.

Let us suppose that

$$u_\nu C_\nu i\rho_3\boldsymbol{\tau} - v_\nu B_\nu\boldsymbol{\lambda} + w_\nu C_\nu\mathbf{p}/p = 0, \quad \text{(B2)}$$

where u_ν, v_ν, and w_ν are undetermined coefficients. On taking two successive cross products of (B2) with $\mathbf{p}$, we find

$$-u_\nu C_\nu i\rho_3\boldsymbol{\tau} - v_\nu B_\nu\boldsymbol{\lambda} + \nu v_\nu C_\nu\mathbf{p}/p = 0, \quad \text{(B3)}$$

which shows that w_ν and v_ν in (B2) are related by

$$w_\nu = -\nu v_\nu. \quad \text{(B4)}$$

Now take the cross product of (B2) with $\boldsymbol{\lambda}$ and use the fact that

$$\boldsymbol{\tau}\times\boldsymbol{\lambda} = (\boldsymbol{\lambda}\times\mathbf{p}/p)\times\boldsymbol{\lambda} = i\rho_3\boldsymbol{\tau} - \lambda_p\boldsymbol{\lambda} + s(s+1)\mathbf{p}/p. \quad \text{(B5)}$$

We then get, with the use of (B4),

$$(\nu v_\nu - u_\nu)C_\nu\boldsymbol{\tau} + (v_\nu - \nu u_\nu)i\rho_3 B_\nu\boldsymbol{\lambda} + u_\nu s(s+1)i\rho_3 C_\nu\mathbf{p}/p = 0. \quad \text{(B6)}$$

The dot product of this equation with $\mathbf{p}/p$ yields

$$\{[s(s+1) - \nu^2]u_\nu + \nu v_\nu\}i\rho_3 C_\nu = 0, \quad \text{(B7)}$$

so that

$$u_\nu = -\nu v_\nu/[s(s+1) - \nu^2]. \quad \text{(B8)}$$

On the other hand, on taking the cross product of (B2) with $\mathbf{p}$ once and then multiplying it by $i\rho_3 C_\nu$, we can get

$$v_\nu C_\nu i\rho_3\boldsymbol{\tau} + u_\nu B_\nu\boldsymbol{\lambda} - \nu u_\nu C_\nu\mathbf{p}/p = 0. \quad \text{(B9)}$$

Comparing (B9) with (B6), we can easily write down [similar to Eq. (B8)]

$$v_\nu = -\nu u_\nu/[s(s+1) - \nu^2]. \quad \text{(B10)}$$

Finally, we combine Eqs. (B8) and (B10) to obtain

$$u_\nu[\nu(\nu+1) - s(s+1)][\nu(\nu-1) - s(s+1)] = 0. \quad \text{(B11)}$$

This condition is automatically satisfied when $\nu = s$; then (B2) reduces to (B1) by virtue of (B4) and (B8). For all other $\nu \geq 0$, (B11) requires that u_ν should vanish and hence v_ν and w_ν too, so that no nontrivial linear relation exists among the three quantities we were concerned with. This completes our proof.

APPENDIX C

We shall outline here how the results proved in Appendices A and B can be used to solve Eq. (8), namely,

$$H\nabla_p H = [H, \boldsymbol{\lambda}] + \mathbf{p}. \quad \text{(C1)}$$

We observe first that the formula (A8), applied to the form (9) of H, yields

$$\begin{aligned}\nabla_p H = &\tfrac{1}{2}\sum_\nu (c_{\nu+1} - c_{\nu-1})\frac{B_\nu \boldsymbol{\lambda}}{p}\\ &+ \tfrac{1}{2}\sum_\nu (b'_{\nu+1} - b'_{\nu-1})\frac{\sigma C_\nu \boldsymbol{\lambda}}{p}\\ &- \tfrac{1}{2}\sum_\nu (c_{\nu+1} + c_{\nu-1} - 2c_\nu)\frac{C_\nu i\rho_3 \boldsymbol{\tau}}{p}\\ &- \tfrac{1}{2}\sum_\nu (b'_{\nu+1} + b'_{\nu-1} - 2b'_\nu)\frac{\sigma B_\nu i\rho_3 \boldsymbol{\tau}}{p}\\ &- \tfrac{1}{2}\sum_\nu \left(\nu(c_{\nu+1} - c_{\nu-1}) - 2p\frac{dc_\nu}{dp}\right)\frac{C_\nu \mathbf{p}}{p^2}\\ &- \tfrac{1}{2}\sum_\nu \left(\nu(b'_{\nu+1} - b'_{\nu-1}) - 2p\frac{db'_\nu}{dp}\right)\frac{\sigma B_\nu \mathbf{p}}{p^2}.\end{aligned} \quad \text{(C2)}$$

The coefficients c_ν and b'_ν, which occur in the form (4) of H, enter here through relations of the form (10). The sums over ν in (C2), which are to be carried from $\nu = \nu_0$ (0 or $\frac{1}{2}$) to $\nu = s$, bring in symbols c_{s+1}, b'_{s+1}, c_{ν_0-1}, and b'_{ν_0-1}, which are undefined in the sense that they do not occur in H. They are used here (as already explained in the text) as a formal notation for series like

$$\sum_l [(s+1)p]^l f_l = c_{s+1},$$
$$\sum [(\nu_0 - 1)p]^L g_L = b'_{\nu_0-1}, \text{ etc.} \quad \text{(C3)}$$

Now, multiplying out H [Eq. (4)] and $\nabla_p H$ [Eq. (C2)] and making use of the fact that

$$C_\mu C_\nu = B_\mu B_\nu = B_\mu \delta_{\mu\nu},\ B_\mu C_\nu = C_\mu B_\nu = C_\mu \delta_{\mu\nu}, \quad \text{(C4)}$$

we find the left-hand side of Eq. (C1) to be

$$\begin{aligned}H\nabla_p H = &\tfrac{1}{2}\sum_\nu [c_\nu(c_{\nu+1} - c_{\nu-1}) + b'_\nu(b'_{\nu+1} - b'_{\nu-1})]\frac{C_\nu \boldsymbol{\lambda}}{p}\\ &- \tfrac{1}{2}\sum_\nu [c_\nu(c_{\nu+1} + c_{\nu-1} - 2c_\nu)\\ &+ b'_\nu(b'_{\nu+1} + b'_{\nu-1} - 2b'_\nu)]\frac{B_\nu i\rho_3\boldsymbol{\tau}}{p}\\ &- \tfrac{1}{2}\sum_\nu \left[c_\nu\left(\nu(c_{\nu+1} - c_{\nu-1}) - 2p\frac{dc_\nu}{dp}\right)\right.\\ &\left.+ b'_\nu\left(\nu(b'_{\nu+1} - b'_{\nu-1}) - 2p\frac{db'_\nu}{dp}\right)\right]\frac{B_\nu \mathbf{p}}{p^2}\\ &+ \tfrac{1}{2}\sum_\nu [b'_\nu(c_{\nu+1} - c_{\nu-1})\\ &\qquad - c_\nu(b'_{\nu+1} - b'_{\nu-1})]\frac{\sigma B_\nu \boldsymbol{\lambda}}{p}\\ &+ \tfrac{1}{2}\sum_\nu [c_\nu(b'_{\nu+1} + b'_{\nu-1} - 2b'_\nu)\\ &- b'_\nu(c_{\nu+1} + c_{\nu-1} - 2c_\nu)]\frac{\sigma C_\nu i\rho_3\boldsymbol{\tau}}{p}\\ &+ \tfrac{1}{2}\sum_\nu \left[c_\nu\left(\nu(b'_{\nu+1} - b'_{\nu-1}) - 2p\frac{db'_\nu}{dp}\right)\right.\\ &\left.- b'_\nu\left(\nu(c_{\nu+1} - c_{\nu-1}) - 2p\frac{dc_\nu}{dp}\right)\right]\frac{\sigma C_\nu \mathbf{p}}{p^2}.\end{aligned} \quad \text{(C5)}$$

The right-hand side is evaluated, starting with the polynomial form of H, with the aid of Eqs. (A3) and (A6), and is reduced, by using Eq. (10), to

$$\begin{aligned}[H, \boldsymbol{\lambda}] + \mathbf{p} = &\tfrac{1}{2}\sum_\nu (2c_\nu - c_{\nu+1} - c_{\nu-1})C_\nu\boldsymbol{\lambda}\\ &+ \tfrac{1}{2}\sum_\nu (c_{\nu+1} - c_{\nu-1})B_\nu i\rho_3\boldsymbol{\tau}\\ &+ \tfrac{1}{2}\sum_\nu [\nu(c_{\nu+1} + c_{\nu-1} - 2c_\nu) + 2p]\frac{B_\nu\mathbf{p}}{p}\\ &+ \tfrac{1}{2}\sum_\nu (2b'_\nu + b'_{\nu+1} + b'_{\nu-1})\sigma B_\nu\boldsymbol{\lambda}\\ &- \tfrac{1}{2}\sum_\nu (b'_{\nu+1} - b'_{\nu-1})\sigma C_\nu i\rho_3\boldsymbol{\tau}\\ &+ \tfrac{1}{2}\sum_\nu [\nu(2b'_\nu - b'_{\nu+1} - b'_{\nu-1})]\frac{\sigma C_\nu\mathbf{p}}{p}.\end{aligned} \quad \text{(C6)}$$

Now, it follows, from what we have proved in Appendix B and Eq. (C4), that $C_\nu\boldsymbol{\lambda}$, $B_\nu i\rho_3\boldsymbol{\tau}$, and $B_\nu\mathbf{p}/p$ are linearly independent for all $\nu \neq s$, and it can then be seen trivially that $\sigma B_\nu\boldsymbol{\lambda}$, $\sigma C_\nu i\rho_3\boldsymbol{\tau}$, and $\sigma C_\nu\mathbf{p}/p$ are linearly independent of the first set as well as among themselves. Consequently, coefficients of these operators in (C5) and (C6), which form the two sides of Eq. (C1), can be separately equated. The result is the set of Eqs. (15) given in the text. Finally, for the case $\nu = s$, the relation (B1) shows that the coefficients of c_{s+1} (and b'_{s+1}) in (C5) and (C6) are identically zero, so that they vanish out, and then one simply gets Eqs. (16) with $\nu = s$.

[1] L. L. Foldy, Phys. Rev. **102**, 568 (1956).
[2] S. Weinberg, Phys. Rev. **133**, B1318 (1964).
[3] D. L. Weaver, C. L. Hammer, and R. H. Good, Jr., Phys. Rev. **135**, B241 (1964).
[4] R. A. Berg, J. Math. Phys. **6**, 34 (1965).
[5] P. M. Mathews, Phys. Rev. **143**, 978 (1966).
[6] P. M. Mathews and S. Ramakrishnan, Nuovo Cimento **50**, 339 (1967).
[7] D. L. Pursey, Ann. Phys. (N.Y.) **32**, 157 (1965).
[8] S. Weinberg, Phys. Rev. **181**, 1893 (1969).
[9] D. Shay, H. S. Song, and R. H. Good, Jr., Nuovo Cimento, Suppl., **3**, 455 (1965).
[10] W. Rarita and J. Schwinger, Phys. Rev. **60**, 61 (1941).
[11] P. A. M. Dirac, Proc. Roy. Soc. (London) **A155**, 447 (1936).
[12] M. Fierz and W. Pauli, Helv. Phys. Acta **12**, 211 (1939).
[13] V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. (U.S.) **34**, 211 (1948).
[14] When a special manifestly covariant form is assumed for a wave equation, it may happen, for instance, that invariance under the proper group is automatically accompanied by invariance under discrete transformations (reflections) too, so that one does not have the opportunity of imposing the conditions separately. See Ref. 15.
[15] P. M. Mathews, J. Math. Phys. **11**, 1360 (1970).
[16] P. M. Mathews, J. Math. Phys. Sci. (Madras) **1**, 197 (1967).
[17] P. M. Mathews, Phys. Rev. **143**, 985 (1966).
[18] P. M. Mathews, in *Lectures in Theoretical Physics* (Boulder Summer Institute, 1969) (Gordon and Breach, New York, 1970).

$SO(4, 2)$-Formulation of the Symmetry Breaking in Relativistic Kepler Problems with or without Magnetic Charges*

A. O. BARUT AND G. L. BORNZIN
Institute for Theoretical Physics and Department of Physics
University of Colorado, Boulder, Colorado

(Received 3 August 1970)

The relativistic Kepler problems in Dirac and Klein–Gordon forms are solved by dynamical group methods for particles having both electric and magnetic charges (dyons). The explicit forms of the $O(4, 2)$-algebra and two special $O(2, 1)$-algebras (which coincide in the symmetry limit) are given, and a new group-theoretical form of the symmetry breaking is pointed out. The Klein–Gordon $O(2, 1)$-algebra also solves the dynamics in the case of very strong coupling constants (attractive singular potential), if the principal series of representations is used instead of the discrete series.

1. INTRODUCTION

It is well known by now that the nonrelativistic Schrödinger theory for the Kepler problem can be treated completely algebraically in an irreducible unitary representation of the dynamical group $O(4, 2)$. In the Appendix, to which we shall refer frequently, we give a new version of this treatment. It is also clear that the relativistic Kepler problem (Klein–Gordon and Dirac equations) does not have the $O(4)$-symmetry of the nonrelativistic problem. If the two particles forming the atom have both electric and magnetic charges, then the $O(4)$-symmetry is broken even in the nonrelativistic Kepler problem. The main purpose of this paper is to show the remarkable group-theoretical way the $O(4)$-symmetry is broken in the above cases. All the above problems are actually exactly soluble, though some of these solutions have not yet been reported in the literature. We hope also to demonstrate the power of the method of dynamical groups in solving these problems, including the strong coupling case.

For the ordinary relativistic Dirac problem, the correspondence between the bound-state spectrum and an $O(4, 1)$-representation was given by Kiefer and Fradkin[1] and Pratt and Jordan.[2] The spectrum-correspondence is not the complete solution of the problem and the operators given in Ref. 1 are extremely complicated, because at that time the importance of the tilted states (see Appendix) was not recognized. The role of the $O(4)$-symmetry of the relativistic hydrogen atom (no spins) in covariant theories based on the Bethe–Salpeter equation was studied in Refs. 3 and 4. Although the use of the dynamical group $O(2, 1)$ for the radial wave equation of the Klein–Gordon and second-order Dirac equations is also known,[5] the complete dynamical group has not been given before.

Early studies of the group property and solution of the Kepler problems with both electric and magnetic charges are due to Fierz[6] and Banderet.[7] More recently Hurst[8] related the Dirac quantization condition[9] to the condition of integrability of the Lie-algebra to the Lie group. Zwanziger[10] has solved a related nonrelativistic, Kepler problem with magnetic charges plus an extra particular $1/r^2$ potential by using the $O(4)$-symmetry. This case is particularly simple, as we shall observe again. The relativistic Kepler problem with magnetic monopoles, as far as spectrum is concerned, was studied recently by Berrondo and McIntosh.[11] It was then recognized that the Kepler problem with magnetic charges realizes a different representation of the dynamical group $O(4, 2)$ than the ordinary Kepler problem, and a new quantum number μ arises.[12] With this a connection is established to the $O(4, 2)$-models of hadrons and to a theory of electromagnetic origin of strong interactions.[12]

Thus the motivation to complete the study of the Kepler problem with magnetic charges is threefold:

(1) to give the solutions of the Schrödinger, Klein–Gordon, and Dirac forms of the Kepler problem in the case of particles with both electric and magnetic charges and to treat the case of the very large coupling constant;

(2) to exhibit the dynamical group $O(4, 2)$ for these cases and the nature of symmetry breaking, because the type of symmetry breaking may be applicable to other symmetry-breaking processes;

(3) to have results applicable to the theory of strong interaction phenomena based on the concept of magnetic charges.

We should mention that the spin–orbit symmetry breaking of the relativistic atom has also been treated in the context of the covariant infinite-dimensional wave equations. In the spinless case, the relevant infinite component wave equation for the H atom contains also correctly the recoil effects.[13] In the case of spin, one can use the basic $O(4, 2)$-group (enlarged by Dirac matrices to account for the spins), but one adds suitable new terms in the wave equation to describe the spin–orbit interactions.[14]

2. GROUP THEORETICAL SOLUTIONS

A. Hamiltonians

It will be convenient to treat Schrödinger, Klein–Gordon, and Dirac forms in a parallel fashion, as we go along.

We consider a particle with electric charge e and magnetic charge g, or simply with charge $\mathbf{q} = (e, g)$. The electromagnetic field is described by the vector potentials $A_\mu = (A_0, -\mathbf{A}) = (\varphi_E, -\mathbf{A}_B)$ and $\tilde{A}_\mu = (\tilde{A}_0, -\tilde{\mathbf{A}}) = (\varphi_B, +\mathbf{A}_E)$. The relativistic Lagrangian of the spinless particle in the field is

$$L = mc\sqrt{u^2} + (e/c)A_\mu u^\mu + (g/c)\tilde{A}_\mu u^\mu. \tag{2.1}$$

Hence we have the canonical momentum

$$p_\mu = mcu_\mu + (e/c)A_\mu + (g/c)\tilde{A}_\mu \tag{2.2}$$

and, from the Euler–Lagrange equation, the Minkowski force

$$K_\mu = [(e/c)F_{\mu\nu} + (g/c)\tilde{F}_{\mu\nu}]u^\nu, \tag{2.3}$$

where

$$F_{\mu\nu} = A_{\nu;\mu} - A_{\mu;\nu},$$
$$\tilde{F}_{\mu\nu} = \tilde{A}_{\nu;\mu} - \tilde{A}_{\mu;\nu} = \tfrac{1}{2}\epsilon_{\mu\nu\lambda\rho}F^{\lambda\rho}.$$

From the spatial components of K_μ we find, as desired,

$$\mathbf{F} = e\mathbf{E} - (e/c)(\mathbf{B} \times \mathbf{v}) + g\mathbf{B} + (g/c)(\mathbf{E} \times \mathbf{v}). \tag{2.3'}$$

Because $u_\mu u^\mu = 1$, we obtain from (2.2)

$$[p_\mu - (e/c)A_\mu - (g/c)\tilde{A}_\mu]^2 = m^2c^2. \tag{2.4}$$

Consequently,

$$H^{(KG)} \equiv cp_0 = eA_0 + g\tilde{A}_0 + [m^2c^4 + (c\mathbf{p} - e\mathbf{A} - g\tilde{\mathbf{A}})^2]^{\frac{1}{2}}. \tag{2.5}$$

This is the desired Hamiltonian in the Klein–Gordon form. To obtain the Hamiltonian in the Schrödinger form, we expand formally the square root and subtract the rest energy (physically this expansion is meaningful only if $\tilde{\mathbf{A}}$ is small because g is very large!) with the following result:

$$H^{(S)} \equiv eA_0 + g\tilde{A}_0 + (1/2m)[\mathbf{p} - (e/c)\mathbf{A} - (g/c)\tilde{\mathbf{A}}]^2. \tag{2.6}$$

Finally, the Dirac form of the Hamiltonian is obtained by linearizing (2.5) with Dirac matrices:

$$H_I^{(D)} \equiv eA_0 + g\tilde{A}_0 + \boldsymbol{\alpha} \cdot (c\mathbf{p} - e\mathbf{A} - g\tilde{\mathbf{A}}) + \gamma^0 mc^2. \tag{2.7}$$

We also give the second-order Dirac Hamiltonian

$$H_{II}^{(D)} \equiv eA_0 + g\tilde{A}_0 + [(c\mathbf{p} - e\mathbf{A} - g\tilde{\mathbf{A}})^2 + m^2c^4 - e\hbar c(\boldsymbol{\sigma} \cdot \mathbf{B} - i\boldsymbol{\alpha} \cdot \mathbf{E}) - g\hbar c(-\boldsymbol{\sigma} \cdot \mathbf{E} - i\boldsymbol{\alpha} \cdot \mathbf{B})]^{\frac{1}{2}}. \tag{2.8}$$

B. Two-Body System

Let the particle of charge $\mathbf{q}_1 = (e_1, g_1)$ move now in the field of another particle of charge $\mathbf{q}_2 = (e_2, g_2)$ situated at the origin and thought to be heavy. In Eqs. (2.5)–(2.8), we replace (e, g) by (e_1, g_1) and (in Gaussian units) let

$$A_0 = \frac{e_2}{r}, \quad \tilde{A}_0 = \frac{g_2}{r}, \quad \mathbf{A} = +g_2\mathbf{D}(\mathbf{r}), \quad \tilde{\mathbf{A}} = -e_2\mathbf{D}(\mathbf{r}),$$

where

$$\mathbf{D}(\mathbf{r}) = \frac{\mathbf{r} \times \hat{\mathbf{n}}(\mathbf{r} \cdot \hat{\mathbf{n}})}{r[r^2 - (\mathbf{r} \cdot \hat{\mathbf{n}})^2]}, \tag{2.9}$$

where $\hat{n}$ is an arbitrary unit vector. $\mathbf{D}(\mathbf{r})$ has the desired property $\boldsymbol{\nabla} \times \mathbf{D}(\mathbf{r}) = \hat{\mathbf{r}}/r^2$. We then obtain, with the abbreviations ($\hbar = c = 1$)

$$\alpha = -(e_1e_2 + g_1g_2), \tag{2.10}$$

$$\mu = (e_1g_2 - g_1e_2), \tag{2.11}$$

$$\boldsymbol{\pi} = \mathbf{p} - \mu\mathbf{D}(\mathbf{r}), \tag{2.12}$$

the Hamiltonians

$$\begin{aligned} H^{(S)} &\equiv (1/2m)\pi^2 - \alpha/r, \\ H^{(KG)} &\equiv [\pi^2 + m^2]^{\frac{1}{2}} - \alpha/r, \\ H_I^{(D)} &\equiv \boldsymbol{\alpha} \cdot \boldsymbol{\pi} + \gamma^0 m - \alpha/r, \\ H_{II}^{(D)} &\equiv [\pi^2 + m^2 - (\mu\boldsymbol{\sigma} + i\alpha\boldsymbol{\alpha}) \cdot \hat{\mathbf{r}}/r^2]^{\frac{1}{2}} - \alpha/r. \end{aligned} \tag{2.13}$$

C. The Dynamical Group $O(4, 2)$ and the Two $O(2, 1)$-Algebras

Consider the following generalized operators which reduce to those of the usual hydrogen atom, (A1) and (A14), in the special case $\mu = 0$:

$$\begin{aligned} \mathbf{J} &= \mathbf{r} \times \boldsymbol{\pi} - \mu\hat{\mathbf{r}}, \\ \mathbf{A} &= \tfrac{1}{2}\mathbf{r}\pi^2 - \boldsymbol{\pi}(\mathbf{r} \cdot \boldsymbol{\pi}) + (\mu/r)\mathbf{J} + \frac{\mu^2}{2r^2}\mathbf{r} - \tfrac{1}{2}\mathbf{r}, \\ \mathbf{M} &= \tfrac{1}{2}\mathbf{r}\pi^2 - \boldsymbol{\pi}(\mathbf{r} \cdot \boldsymbol{\pi}) + (\mu/r)\mathbf{J} + \frac{\mu^2}{2r^2}\mathbf{r} + \tfrac{1}{2}\mathbf{r}, \\ \boldsymbol{\Gamma} &= r\boldsymbol{\pi}, \\ \Gamma_0 &= \tfrac{1}{2}(r\pi^2 + r + \mu^2/r), \\ \Gamma_4 &= \tfrac{1}{2}(r\pi^2 - r + \mu^2/r), \\ T &= \mathbf{r} \cdot \boldsymbol{\pi} - i. \end{aligned} \tag{2.14}$$

These operators also satisfy the commutation relations of the Lie algebra of $O(4, 2)$ as before, as can be verified by direct, though laborious, calculation. This fact is more remarkable than it appears at first glance, for the generalized momenta $\boldsymbol{\pi} = \mathbf{p} - \mu \mathbf{D}(\mathbf{r})$ [see Eqs. (2.12) and (2.9)] no longer commute among themselves as do the canonical momenta $\mathbf{p}$; rather we find $[\pi_i, \pi_j] = i\mu e_{ijk} x_k / r^3$. The Casimir operator of the $O(2, 1)$-subgroup generated by Γ_0, Γ_4, T is again as in (A3)

$$Q^2 = \Gamma_0^2 - \Gamma_4^2 - T^2 = J^2. \qquad (2.15)$$

Thus, if we had an associated Hamiltonian

$$H^a \equiv \pi^2/2m - \alpha/r + \mu^2/2mr^2, \qquad (2.16)$$

we would have

$$\Theta = r(H^a - E)$$
$$= (1/2m)(\Gamma_0 + \Gamma_4) - E(\Gamma_0 - \Gamma_4) - \alpha,$$

i.e., precisely the same equation as (A4); hence all the equations up to (A13) would equally apply to this case. Instead of Eq. (A15) the Casimir operators would be[15]

$$Q_2 = 3(\mu^2 - 1),$$
$$Q_3 = 0, \quad Q_4 = \mu^2(1 - \mu^2). \qquad (2.17)$$

In particular, we would have $O(4)$ symmetry, the Balmer formula (A8), etc. This is indeed the case studied by Zwanziger,[10] but not the case we want to solve. There is no physical reason to assume the extra scalar potential $\mu^2/2mr^2$ in (2.16). Instead we want to use the Hamiltonians (2.13) which include no extra scalar potential.

We notice that the three operators

$$\Gamma_0' = \tfrac{1}{2}(r\pi^2 + r),$$
$$\Gamma_4' = \tfrac{1}{2}(r\pi^2 - r), \qquad (2.18)$$
$$T' = T$$

also generate an $O(2, 1)$-algebra with the Casimir operator

$$Q'^2 = J^2 - \mu^2. \qquad (2.19)$$

For our Hamiltonian $H^{(S)} \equiv \pi^2/2m - \alpha/r$, we have

$$\Theta = r(H^{(S)} - E)$$
$$= (1/2m)(\Gamma_0' + \Gamma_4') - E(\Gamma_0' - \Gamma_4') - \alpha, \quad (2.20)$$

again an equation of exactly the same type as (A4). Thus, in terms of the spectrum of Γ_0' and Γ_4', we can immediately use the solutions (A8) and (A11). The only thing we do not know *a priori* is the range of $(J^2 - \mu^2)$, Eq. (2.19), that is contained in the spectrum of $H^{(S)}$. For the ordinary atom, a single representation of the full dynamical group $O(4, 2)$ [Eqs. (A14) and (A15)] determines the spectrum of the Casimir operator Q^2 of the $O(2, 1)$-subgroup and hence J^2. Now, however, the primed generators (2.18) cannot be completed to an $O(4, 2)$-algebra as the unprimed ones given in (2.14). Indeed, if they could be, we would still get an $O(4)$-symmetry which we know we do not have. Thus, we have *two* $O(2, 1)$-algebras, each commuting with $\mathbf{J}$, whose Casimir operators, (2.15) and (2.19), are related by

$$Q^2 - Q'^2 = \mu^2. \qquad (2.21)$$

We can indeed view μ^2 as the parameter of symmetry breaking; for $\mu^2 = 0$ we get back the results of the Appendix. It is important to note that the "unsymmetrical" case is also exactly soluble; this is because we know the range of J^2 from the $O(4, 2)$-representation (2.14), and we know the spectrum of Γ_0' and Γ_4' from the value of Q'^2. Thus, for $E < 0$, we solve $\Theta\tilde{\Phi} = 0$ (2.20) by analogy with (A7) and (A8), and immediately have

$$[(-2E/m)^{\frac{1}{2}}\Gamma_0' - \alpha]\Phi = 0$$

and

$$E_{n'} = -\tfrac{1}{2}(m\alpha^2/n'^2), \qquad (2.22)$$

where n' is the (discrete) spectrum of Γ_0'. From (2.19), letting

$$Q'^2 = j(j+1) - \mu^2 = \varphi'(\varphi' + 1),$$

we find

$$\varphi' = -\tfrac{1}{2} \pm [(j + \tfrac{1}{2})^2 - \mu^2]^{\frac{1}{2}}. \qquad (2.23)$$

Hence in the D_+-representation of $O(2, 1)$—which is bounded below—the spectrum of Γ_0' has the range

$$n' = -\varphi', -\varphi' + 1, -\varphi' + 2, \cdots$$
$$= \tfrac{1}{2} + [(j + \tfrac{1}{2})^2 - \mu^2]^{\frac{1}{2}}, \tfrac{3}{2} + [(j + \tfrac{1}{2})^2 - \mu^2]^{\frac{1}{2}}, \cdots.$$

(For comparison the range of the eigenvalues of Γ_0 is $n = j + 1, j + 2, \cdots$.) Consequently, Eq. (2.22) can be written as

$$E_s = -\tfrac{1}{2}m\alpha^2\{s + \tfrac{1}{2} + [(j + \tfrac{1}{2})^2 - \mu^2]^{\frac{1}{2}}\}^{-2},$$
$$s = 0, 1, 2, 3, \cdots. \quad (2.24)$$

For $\mu = 0$, we recover the Balmer formula. For fixed $\mu \neq 0$, we see from the $O(4, 2)$-representation (2.14)–(2.17) that again, for each $n(\Gamma_0)$, the range of j is

$$j: |\mu|, |\mu| + 1, |\mu| + 2, \cdots n - 1, \qquad (2.25)$$

which completes the specification of the spectrum.[15]

In the case of the Klein–Gordon Hamiltonian, the $O(4, 2)$-representation (2.14) remains the same. But instead of (2.18), we see that

$$\Gamma_0' = \tfrac{1}{2}(r\pi^2 + r - \alpha^2/r),$$
$$\Gamma_4' = \tfrac{1}{2}(r\pi^2 - r - \alpha^2/r), \qquad (2.26)$$
$$T' = T$$

also form the Lie algebra of an $O(2, 1)$-group with the Casimir operator

$$Q'^2 = \Gamma_0'^2 - \Gamma_4'^2 - T^2 = J^2 - \mu^2 - \alpha^2 = \varphi'(\varphi' + 1),$$

so that

$$\varphi' = -\tfrac{1}{2} \pm [(j + \tfrac{1}{2})^2 - \mu^2 - \alpha^2]^{\frac{1}{2}}. \quad (2.27)$$

The Lie algebra (2.26) solves the square of $H^{(KG)}$ given in (2.13) in the sense that

$$\begin{aligned}\Theta &= r[(H^{(KG)} + \alpha/r)^2 - (E + \alpha/r)^2] \\ &= r\pi^2 - (E^2 - m^2)r - 2\alpha E - \alpha^2/r \\ &= \Gamma_0' + \Gamma_4' - (E^2 - m^2)(\Gamma_0' - \Gamma_4') - 2\alpha E, \quad (2.28)\end{aligned}$$

which is again an equation of the type (A4) or (2.20). The equation $\Theta\tilde{\Phi} = 0$ can again easily be solved by putting $\tilde{\Phi} = e^{i\theta T}\Phi$ and by choosing $\tanh \theta = (E^2 - m^2 + 1)/(E^2 - m^2 - 1)$. Then

$$\{[-4(E^2 - m^2)]^{\frac{1}{2}}\Gamma_0' - 2\alpha E\}\Phi = 0. \quad (2.29)$$

From (2.27), the spectrum of Γ_0' is given by

$$n' = s + \tfrac{1}{2} + [(j + \tfrac{1}{2})^2 - \mu^2 - \alpha^2]^{\frac{1}{2}}, \qquad s = 0, 1, 2, \cdots.$$

Hence, the energy spectrum becomes

$$E_s = m(1 + \alpha^2\{s + \tfrac{1}{2} + [(j + \tfrac{1}{2})^2 - \mu^2 - \alpha^2]^{\frac{1}{2}}\}^{-2})^{-\frac{1}{2}} \qquad s = 0, 1, 2, \cdots. \quad (2.30)$$

Finally, in the case of the second-order Dirac equation, we introduce instead of (2.26) the $O(2, 1)$-algebra

$$\begin{aligned}\Gamma_0' &= \tfrac{1}{2}\{r\pi^2 + r + (1/r)[-\alpha^2 - (\mu\boldsymbol{\sigma} + i\alpha\boldsymbol{\alpha}) \cdot \hat{\mathbf{r}}]\}, \\ \Gamma_4' &= \tfrac{1}{2}\{r\pi^2 - r + (1/r)[-\alpha^2 - (\mu\boldsymbol{\sigma} + i\alpha\boldsymbol{\alpha}) \cdot \hat{\mathbf{r}}]\}, \\ T' &= T, \quad (2.31)\end{aligned}$$

with the Casimir operator

$$Q'^2 = J^2 - \mu^2 - \alpha^2 - (\mu\boldsymbol{\sigma} + i\alpha\boldsymbol{\alpha}) \cdot \hat{\mathbf{r}}. \quad (2.32)$$

The operator

$$\Gamma = \boldsymbol{\sigma} \cdot \mathbf{J} + (\mu\boldsymbol{\sigma} + i\alpha\boldsymbol{\alpha}) \cdot \hat{\mathbf{r}} + 1 \quad (2.33)$$

has the property that

$$\Gamma^2 = (\mathbf{J} + \tfrac{1}{2}\boldsymbol{\sigma})^2 - \mu^2 - \alpha^2 + \tfrac{1}{4}.$$

Let

$$\mathfrak{J} = \mathbf{J} + \tfrac{1}{2}\boldsymbol{\sigma};$$

then

$$\Gamma^2 = \mathfrak{J}^2 - \mu^2 - \alpha^2 + \tfrac{1}{4} = j(j + 1) - \mu^2 - \alpha^2 + \tfrac{1}{4}. \quad (2.34)$$

Note that j now denotes the total angular momentum of the spin-$\frac{1}{2}$ particle in the atom. The eigenvalues of Γ are then

$$\Gamma : \gamma = \pm[(j + \tfrac{1}{2})^2 - \mu^2 - \alpha^2]^{\frac{1}{2}}. \quad (2.35)$$

Now, from (2.32),

$$Q'^2 = \Gamma^2 - \Gamma = \gamma^2 - \gamma = \varphi'(\varphi' + 1)$$

or

$$\varphi' = -\gamma \quad \text{or} \quad \gamma - 1. \quad (2.36)$$

For the Dirac Hamiltonian (2.13), we get then from the second-order equation

$$\begin{aligned}\Theta &= r[(H^{(D)} + \alpha/r)^2 - (E + \alpha/r)^2] \\ &= r\pi^2 - (E^2 - m^2)r - 2\alpha E - (1/r) \\ &\quad \times [\alpha^2 + (\mu\boldsymbol{\sigma} + i\alpha\boldsymbol{\alpha}) \cdot \hat{\mathbf{r}}] \\ &= \Gamma_0' + \Gamma_4' \mp (E^2 - m^2)(\Gamma_0' - \Gamma_4') - 2\alpha E, \quad (2.37)\end{aligned}$$

i.e., the same equation as (2.28). Thus we can immediately write down the energy spectrum in analogy to the previous case:

$$E_s = m[1 + \alpha^2\{s + [(j + \tfrac{1}{2})^2 - \mu^2 - \alpha^2]^{\frac{1}{2}}\}^{-2}]^{-\frac{1}{2}}, \qquad s = 0, 1, 2, 3, \cdots. \quad (2.38)$$

This differs from the Klein–Gordon spectrum (2.30) in the additive term $\frac{1}{2}$ after s and in the eigenvalue j of total angular momentum which here includes spin.

D. The Case of Large Coupling Constants

Equations (2.30) and (2.38) hold only for a small coupling constant

$$\alpha^2 < (j + \tfrac{1}{2})^2 - \mu^2,$$

because then φ' [which is associated with the Casimir operator Q'^2 of $O(2, 1)$] is real [Eq. (2.27)], and we obtain the D^+-representations of the discrete series.

If α^2 is large, however, as is the case for magnetic charges, ($\alpha = 137/4$ instead of $1/137$ for ordinary atoms!)[12] we must use for φ' a value corresponding to the principal series of representations

$$\varphi' = -\tfrac{1}{2} + i\lambda, \quad \lambda \text{ real}. \quad (2.39)$$

Then

$$Q'^2 = \varphi'(\varphi' + 1) = -\lambda^2 - \tfrac{1}{4}.$$

In the case of the Klein–Gordon equation, for example, from Eq. (2.27)

$$Q'^2 = J^2 - \mu^2 - \alpha^2,$$

and we obtain

$$\lambda = \pm[\alpha^2 + \mu^2 - (j + \tfrac{1}{2})^2]^{\frac{1}{2}}. \quad (2.40)$$

Thus we have a particular representation in the principal series.

In the case of the principal series of representations of the Lie algebra of $O(2, 1)$, the spectrum of Γ_0 ranges from $-\infty$ to $+\infty$, i.e., it is not bounded below. Moreover, a new invariant quantum number E_0 occurs in addition to the invariant φ.[16] The spectrum of Γ_0' is then

$$\Gamma_0': E_0 + s, \quad s = 0, \pm 1, \pm 2, \cdots. \tag{2.41}$$

We have then from (2.29)

$$E_s = m[1 + \alpha^2(s + E_0)^{-2}]^{-\frac{1}{2}}. \tag{2.42}$$

The new quantum number is fixed within the $O(2, 1)$-subgroup; it can be determined only within the representation of the big group $O(4, 2)$.

The physical reason for the drastic change in the case of a large coupling constant for relativistic equations is that we now have a large *attractive* singular potential at $r = 0$. In the case of attractive singular potentials we cannot use the usual boundary conditions of the Schrödinger treatment; the solutions form an overcomplete set, and one needs indeed a new quantum number to characterize the problem completely.[17]

APPENDIX

The operators

$$\begin{aligned} \Gamma_0 &= \tfrac{1}{2}(rp^2 + r), \\ \Gamma_4 &= \tfrac{1}{2}(rp^2 - r), \\ T &= \mathbf{r}\cdot\mathbf{p} - i \end{aligned} \tag{A1}$$

satisfy the commutation relations of the Lie algebra of the group $O(2, 1)$:

$$[\Gamma_0, \Gamma_4] = iT, \quad [\Gamma_4, T] = -i\Gamma_0, \quad [T, \Gamma_0] = i\Gamma_4. \tag{A2}$$

The Casimir operator is given by

$$Q^2 = \Gamma_0^2 - \Gamma_4^2 - T^2 = (\mathbf{r} \times \mathbf{p})^2 = J^2. \tag{A3}$$

Consequently, from the Hamiltonian $H = p^2/2m - \alpha/r$, we obtain ($\hbar = c = 1$)

$$\Theta \equiv r(H - E) = 1/2m(\Gamma_0 + \Gamma_4) - E(\Gamma_0 - \Gamma_4) - \alpha. \tag{A4}$$

The equation

$$\Theta\tilde{\Phi} = 0 \tag{A5}$$

can be solved as follows. Let

$$\tilde{\Phi} = e^{i\theta T}\Phi, \tag{A6}$$

and choose $\tanh\theta = (E + 1/2m)/(E - 1/2m)$; then Eq. (A5) reduces to

$$[(-2E/m)^{\frac{1}{2}}\Gamma_0 - \alpha]\Phi = 0. \tag{A7}$$

Thus Φ's are the eigenstates of Γ_0 with discrete eigenvalues n if $E < 0$. Hence,

$$E_n = -\tfrac{1}{2}(m\alpha^2/n^2). \tag{A8}$$

For $E > 0$, Γ_0 cannot be diagonalized; we go back to Eqs. (A4) and (A5), and let

$$\tilde{\Phi} = e^{i\theta' T}\Phi' \tag{A9}$$

and choose $\tanh\theta' = (E - 1/2m)/(E + 1/2m)$. Then again, from (A5),

$$[(2E/m)^{\frac{1}{2}}\Gamma_4 - \alpha]\Phi' = 0. \tag{A10}$$

Now Γ_4 has a continuous real spectrum λ. Hence

$$E = \tfrac{1}{2}(m\alpha^2/\lambda^2). \tag{A11}$$

The states $\tilde{\Phi}$ must be normalized as follows,

$$\langle\tilde{\Phi}|\,(\Gamma_0 - \Gamma_4)\,|\tilde{\Phi}\rangle = 1, \tag{A12}$$

and are not identical with the Schrödinger wavefunctions ψ. The physical normalized solutions of (A5) are then

$$\tilde{\Phi} = (1/n)e^{i\theta T}\,|n\rangle, \tag{A13}$$

where $|n\rangle$ is a basis of the discrete unitary irreducible $O(2, 1)$ representation D_+^j with Casimir operator given in Eq. (A3): $Q^2 = j(j + 1) = \varphi(\varphi + 1)$, $\varphi < 0$. Hence $\varphi = -j - 1$. Therefore, for each j, $n = j + 1, j + 2, \cdots$. Similar equations hold for the continuous spectrum.

The treatment above does not tell us yet what values of j occur; it is yet incomplete. The complete solution is as follows. The operators (A1) together with

$$\begin{aligned} \mathbf{J} &= \mathbf{r} \times \mathbf{p}, \\ \mathbf{A} &= \tfrac{1}{2}\mathbf{r}p^2 - \mathbf{p}(\mathbf{r}\cdot\mathbf{p}) - \tfrac{1}{2}\mathbf{r}, \\ \mathbf{M} &= \tfrac{1}{2}\mathbf{r}p^2 - \mathbf{p}(\mathbf{r}\cdot\mathbf{p}) + \tfrac{1}{2}\mathbf{r}, \\ \mathbf{\Gamma} &= r\mathbf{p} \end{aligned} \tag{A14}$$

satisfy the commutation relations of the Lie algebra of $O(4, 2)$; $\mathbf{J}$ and $\mathbf{A}$ (Runge–Lenz vector) together generate a compact $O(4)$ subgroup that commutes with Γ_0. (Note then that $\mathbf{J}$ and "tilted $\mathbf{A}$" commute with Θ.) The Casimir operators of the Lie algebra of this $O(4, 2)$ are

$$Q_2 = \mathbf{J}^2 + \mathbf{A}^2 - \mathbf{M}^2 - \mathbf{\Gamma}^2 + \Gamma_0^2 - \Gamma_4^2 - T^2 = -3, \quad Q_3 = 0, \quad Q_4 = 0. \tag{A15}$$

In an irreducible representation of $SO(4, 2)$, for each $j = 0, 1, 2, 3, \cdots$ we have $n = j + 1, j + 2, \cdots$. Or, for each $n = 1, 2, \cdots$ we have $j = 0, 1, \cdots n - 1$. The energy levels depend only on n^2, which is the basis of the $O(4)$-symmetry.

* Supported in part by the Air Force Office of Scientific Research.

[1] H. M. Kiefer and D. M. Fradkin, Phys. Rev. **180**, 1282 (1969).
[2] R. H. Pratt and T. Jordan, Phys. Rev. **188**, 2534 (1969).

[3] A. O. Barut, P. Budini, and C. Fronsdal, Proc. Roy. Soc. (London) **A291**, 106 (1966).
[4] J. E. Mandula, Phys. Rev. **185**, 1774 (1969).
[5] H. Bacry and J. L. Richard, J. Math. Phys. **8**, 2231 (1967).
[6] M. Fierz, Helv. Phys. Acta **17**, 27 (1944).
[7] P. Banderet, Helv. Phys. Acta **19**, 503 (1946).
[8] C. A. Hurst, Ann. Phys. **50**, 51 (1968).
[9] P. A. M. Dirac, Phys. Rev. **74**, 817 (1948).
[10] D. Zwanziger, Phys. Rev. **176**, 1480 (1968).
[11] M. Berrondo and H. V. McIntosh, J. Math. Phys. **11**, 125 (1970).
[12] A. O. Barut, *Proceedings of the Coral Gables Conference on Fundamental Interactions* (Gordon and Breach, N.Y., 1970), Vol. II, p. 199–220 and Phys. Rev. D 3 (April 15, 1971) (to appear). See also A. O. Barut, "Atoms with Magnetic Charges as Models of Hadrons," in *Topics in Modern Physics—A Tribute to E. U. Condon* (Colorado Associated University Press, Boulder, Colorado, 1971).
[13] A. O. Barut and A. Baiquni, Phys. Rev. **184**, 1342 (1969).
[14] A. O. Barut and A. Baiquni, Phys. Letters **30A**, 352 (1969).
[15] These are the so-called oscillator representations of $O(4,2)$-algebra with $\mu = 0, \pm\frac{1}{2}, \pm 1, \cdots$. They are characterized by a representation relation, and have particularly simple reduction properties with respect to $SO(3,2)$- and $SO(4,1)$-subgroups see A. O. Barut and A. Böhm, J. Math. Phys. **11**, 2938 (1970) for multiplicity patterns and other details.
[16] A. O. Barut and C. Fronsdal, Proc. Roy. Soc. (London) **A287**, 532 (1965).
[17] K. M. Case, Phys. Rev. **80**, 797 (1950).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

$SL(2, C)$ Representations in Explicitly "Energy-Dependent" Basis. I

K. Szegö and K. Tóth
Central Research Institute for Physics, Budapest, Hungary

(Received 14 July 1970)

Unitary and nonunitary representations of the $SL(2, C)$ group are investigated in such a basis, in which the subgroup diagonalized is that one which in the four-dimensional representation leaves invariant the 4-vector $p_\mu = (\frac{1}{2}(1 + v), 0, 0, \frac{1}{2}(1 - v))$ for an arbitrary real value of $p_\mu^2 = v$. The split of the representation space into irreducible subspaces changes smoothly when varying the value of v. The formalism is of importance in physical theories which postulate analyticity requirements and Lorentz invariance simultaneously (e.g., Regge and Lorentz pole theory). In this paper we construct explicit basis functions of the representation spaces.

1. INTRODUCTION

The representation theory of the $SL(2, C)$ group is of great importance in physics, and a lot of work has been devoted to construct its representations explicitly. It is, however, surprising that attention has hardly been paid to constructing and investigating them in an explicitly "analytically continuable" form. We mean the following: The representations of the $SL(2, C)$ group are usually given in an $SU(2)$, $SU(1, 1)$, or $E(2)$ basis, i.e., the representation space is given as a direct sum (integral) of subspaces invariant with respect to the little groups of the 4-vectors $(1, 0, 0, 0)$, $(1, 0, 0, 1)$, and $(0, 0, 0, 1)$, respectively. Physical theories, which postulate analyticity requirements together with Lorentz invariance, necessitate the construction of $SL(2, C)$ representations over such spaces, which are split into subspaces invariant with respect to the little group of an appropriately chosen 4-vector, e.g., $p_\mu = (\frac{1}{2}(1 + v), 0, 0, \frac{1}{2}(1 - v))$. Its length $p_\mu^2 = v$ is kept a free parameter. Moreover, we want the representations to be analytic in this variable v in the sense that the split of the representation space into irreducible subspaces changes smoothly when we vary the value of this parameter.

In this paper we will explicitly construct the basis states for such representations. We shall apply a standard procedure.[1] This method consists first of choosing a subgroup which one wants to be diagonal in the basis to be constructed and second of determining the eigenfunctions of the Casimir operator of this subgroup.

For this purpose, we must obviously specify such a subgroup of $SL(2, C)$ which, depending on the value of a suitable parameter, becomes deformed from $SU(2)$ through $E(2)$ to $SU(1, 1)$. Then one must determine the representation matrix elements of this group, which is the second point of the previous program above. These problems have already been treated[2,3] but without embedding this group into $SL(2, C)$. [The term "interpolating group (IG)" was introduced for this group[3]; we are going to use it in this paper as well.]

After having constructed the basis with the above specified properties in the $SL(2, C)$ representation space, we naturally examine the matrix elements of finite $SL(2, C)$ transformations and the problem of the transformation coefficients between different basis sets.[4] Here we give them only in integral forms, as the explicit calculations can be found in a separate paper.[5]

In Sec. 2, we shall summarize the results of the

[3] A. O. Barut, P. Budini, and C. Fronsdal, Proc. Roy. Soc. (London) **A291**, 106 (1966).
[4] J. E. Mandula, Phys. Rev. **185**, 1774 (1969).
[5] H. Bacry and J. L. Richard, J. Math. Phys. **8**, 2231 (1967).
[6] M. Fierz, Helv. Phys. Acta **17**, 27 (1944).
[7] P. Banderet, Helv. Phys. Acta **19**, 503 (1946).
[8] C. A. Hurst, Ann. Phys. **50**, 51 (1968).
[9] P. A. M. Dirac, Phys. Rev. **74**, 817 (1948).
[10] D. Zwanziger, Phys. Rev. **176**, 1480 (1968).
[11] M. Berrondo and H. V. McIntosh, J. Math. Phys. **11**, 125 (1970).
[12] A. O. Barut, *Proceedings of the Coral Gables Conference on Fundamental Interactions* (Gordon and Breach, N.Y., 1970), Vol. II, p. 199–220 and Phys. Rev. D 3 (April 15, 1971) (to appear). See also A. O. Barut, "Atoms with Magnetic Charges as Models of Hadrons," in *Topics in Modern Physics—A Tribute to E. U. Condon* (Colorado Associated University Press, Boulder, Colorado, 1971).
[13] A. O. Barut and A. Baiquni, Phys. Rev. **184**, 1342 (1969).
[14] A. O. Barut and A. Baiquni, Phys. Letters **30A**, 352 (1969).
[15] These are the so-called oscillator representations of $O(4, 2)$-algebra with $\mu = 0, \pm\frac{1}{2}, \pm 1, \cdots$. They are characterized by a representation relation, and have particularly simple reduction properties with respect to $SO(3, 2)$- and $SO(4, 1)$-subgroups see A. O. Barut and A. Böhm, J. Math. Phys. **11**, 2938 (1970) for multiplicity patterns and other details.
[16] A. O. Barut and C. Fronsdal, Proc. Roy. Soc. (London) **A287**, 532 (1965).
[17] K. M. Case, Phys. Rev. **80**, 797 (1950).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

$SL(2, C)$ Representations in Explicitly "Energy-Dependent" Basis. I

K. Szegö and K. Tóth
Central Research Institute for Physics, Budapest, Hungary

(Received 14 July 1970)

Unitary and nonunitary representations of the $SL(2, C)$ group are investigated in such a basis, in which the subgroup diagonalized is that one which in the four-dimensional representation leaves invariant the 4-vector $p_\mu = (\frac{1}{2}(1 + v), 0, 0, \frac{1}{2}(1 - v))$ for an arbitrary real value of $p_\mu^2 = v$. The split of the representation space into irreducible subspaces changes smoothly when varying the value of v. The formalism is of importance in physical theories which postulate analyticity requirements and Lorentz invariance simultaneously (e.g., Regge and Lorentz pole theory). In this paper we construct explicit basis functions of the representation spaces.

1. INTRODUCTION

The representation theory of the $SL(2, C)$ group is of great importance in physics, and a lot of work has been devoted to construct its representations explicitly. It is, however, surprising that attention has hardly been paid to constructing and investigating them in an explicitly "analytically continuable" form. We mean the following: The representations of the $SL(2, C)$ group are usually given in an $SU(2)$, $SU(1, 1)$, or $E(2)$ basis, i.e., the representation space is given as a direct sum (integral) of subspaces invariant with respect to the little groups of the 4-vectors $(1, 0, 0, 0)$, $(1, 0, 0, 1)$, and $(0, 0, 0, 1)$, respectively. Physical theories, which postulate analyticity requirements together with Lorentz invariance, necessitate the construction of $SL(2, C)$ representations over such spaces, which are split into subspaces invariant with respect to the little group of an appropriately chosen 4-vector, e.g., $p_\mu = (\frac{1}{2}(1 + v), 0, 0, \frac{1}{2}(1 - v))$. Its length $p_\mu^2 = v$ is kept a free parameter. Moreover, we want the representations to be analytic in this variable v in the sense that the split of the representation space into irreducible subspaces changes smoothly when we vary the value of this parameter.

In this paper we will explicitly construct the basis states for such representations. We shall apply a standard procedure.[1] This method consists first of choosing a subgroup which one wants to be diagonal in the basis to be constructed and second of determining the eigenfunctions of the Casimir operator of this subgroup.

For this purpose, we must obviously specify such a subgroup of $SL(2, C)$ which, depending on the value of a suitable parameter, becomes deformed from $SU(2)$ through $E(2)$ to $SU(1, 1)$. Then one must determine the representation matrix elements of this group, which is the second point of the previous program above. These problems have already been treated[2,3] but without embedding this group into $SL(2, C)$. [The term "interpolating group (IG)" was introduced for this group[3]; we are going to use it in this paper as well.]

After having constructed the basis with the above specified properties in the $SL(2, C)$ representation space, we naturally examine the matrix elements of finite $SL(2, C)$ transformations and the problem of the transformation coefficients between different basis sets.[4] Here we give them only in integral forms, as the explicit calculations can be found in a separate paper.[5]

In Sec. 2, we shall summarize the results of the representation theory of the $SL(2, C)$ group and the most important results of the theory of bilinear functionals[6] constructed from $SL(2, C)$ representation functions. In Sec. 3, we explicitly construct the basis functions which can be analytically continued in the above specified manner. In Sec. 4, we discuss methods of calculating matrix elements and transformation coefficients between any type of basis functions, and in Sec. 5, we apply these methods to finding the norm of the basis functions.

2. SUMMARY OF SOME RESULTS OF $SL(2, C)$ REPRESENTATIONS

A. Basic Definitions

From the general theory of[6] $SL(2, C)$, we know that it can be represented in the space of infinitely differentiable functions $\phi(z, \bar{z})$, where $z = x + iy$ and the bar denotes complex conjugation; x and y are two real variables. [In the following we use the notation $\phi(z, \bar{z}) = \phi(z)$.] A general representation $A(g)$ of an element g of $SL(2, C)$ acts onto such functions as

$$A(g)\phi(z) = (\beta z + \delta)^{-j_0+\sigma-1}\overline{(\beta z + \delta)}^{j_0+\sigma-1}\phi\left(\frac{\alpha z + \gamma}{\beta z + \delta}\right), \tag{2.1}$$

where

$$g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad \alpha\delta - \beta\gamma = 1. \tag{2.2}$$

We allow any integer or half-integer value for j_0 and any complex one for σ. For a given $\chi = (j_0, \sigma)$ all the infinitely differentiable functions $\phi(z)$, satisfying Eq. (2.1), form an infinite-dimensional linear space D_χ which is generally irreducible with respect to the transformation (2.1). Hence, a $\chi = (j_0, \sigma)$ pair characterizes the representation as well. If $\sigma - j_0$ is a nonnegative integer, D_χ splits into two irreducible subspaces, a finite-dimensional and an infinite-dimensional one.

The $A(g)$ operators are unitary and the space D_χ is a Hilbert space when:

(i) j_0 is integer or half-integer and σ is imaginary—principal series;

(ii) $j_0 = 0$, σ is real, $0 < |\sigma| < 1$—supplementary series;

(iii) $j_0 = 0$, $\sigma = 1$—trivial representation.

In all the other cases the representations are nonunitary.

B. The Bilinear Functional

Our method for finding explicit basis functions (BF's in the following) is to solve a system of homogeneous partial differential equations, and it is up to us as to how we want to specify normalization factors. Since we are going to treat nonunitary representations as well, we shall consider the so-called invariant bilinear functional (IBF in the following) $B(\varphi, \psi)$ instead of the Hermitian, positive-definite invariant scalar product. The latter exists only for the unitary representations. (We use the terminology of the quoted literature.[6]) Nevertheless, we shall comment on the differences between the normalizations given by the different methods in the case of the unitary representations.

Let D_{χ_1} and D_{χ_2} be two representation spaces with functions $\{\varphi\} \in D_{\chi_1}$ and $\{\psi\} \in D_{\chi_2}$. $B(\varphi, \psi)$ is defined as a functional as follows:

(i) linear both in φ and ψ;

(ii) continuous in φ and ψ in the $D_{\chi_1} \oplus D_{\chi_2}$ direct sum space;

(iii) invariant with respect to $SL(2, C)$ transformations. The last property means that

$$B(\varphi, \psi) = B(T_{\chi_1}(g)\varphi, T_{\chi_2}(g)\psi),$$

where g is an element of $SL(2, C)$ and the operators $T_{\chi_1}(g)$ and $T_{\chi_2}(g)$ are its representations over D_{χ_1} and D_{χ_2}, respectively.

A suitable integral form of the IBF is

$$B(\varphi, \psi) = \tfrac{1}{2}i\int dz\, d\bar{z}\varphi_{\chi_1}(z)\psi_{\chi_2}(z), \tag{2.3}$$

where $\chi_1 = (j_0, \sigma)$ and $\chi_2 = -\chi_1$.

In the case of the unitary representations there exists a positive-definite invariant scalar product for $\varphi, \psi \in D_\chi$, with χ specified above. For the principal series it reads as

$$(\varphi, \psi) = \tfrac{1}{2}i\int dz\, d\bar{z}\bar{\varphi}(z)\psi(z), \tag{2.4}$$

for the supplementary one as

$$(\varphi, \psi) = (\tfrac{1}{2}i)^2\int dz_1\, d\bar{z}_1 dz_2\, d\bar{z}_2 |z_1 - z_2|^{-2\sigma-2}\varphi(z_1)\psi(z_2). \tag{2.5}$$

Some other question will be discussed in Sec. 4.

3. CONSTRUCTION OF BF's IN D_χ

A. The Construction of the Generators

As usual, we introduce the six generators M_i and N_i of the $SL(2, C)$, commuting as

$$[M_i, M_j] = -[N_i, N_j] = i\epsilon_{ijk}M_k, \quad [M_i, N_j] = i\epsilon_{ijk}N_k. \tag{3.1}$$

Their 2×2 matrix realization is, e.g.,

$$M_i = \tfrac{1}{2}\sigma_i, \quad N_i = \tfrac{1}{2}i\sigma_i, \tag{3.2}$$

where σ_i stands for the well-known Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.3}$$

A realization of the generators M_i and N_i as operators on the spaces D_χ can be found by defining the one-parameter elements of $SL(2, C)$ as $\exp(-i\alpha M_i)$ or $\exp(-i\alpha N_i)$ and by making use of Eqs. (2.1) and (3.2). Instead of writing down the results for M_i and N_i, first we introduce some linear combinations of them more suitable for our purpose:

$$\begin{aligned} S_1(v) &= \tfrac{1}{2}[(1+v)M_2 + (1-v)N_1], \\ S_2(v) &= \tfrac{1}{2}[-(1+v)M_1 + (1-v)N_2], \\ S_3(v) &= M_3, \end{aligned} \tag{3.4}$$

where v is a real parameter. The v dependence of the S_i operators will be suppressed.

Now, one can find out that

$$-2iS_1 = \cos\varphi S_r - \sin\varphi S_\varphi, \tag{3.5a}$$

$$-2iS_2 = \sin\varphi S_r + \cos\varphi S_\varphi, \tag{3.5b}$$

$$S_3 = -i\frac{\partial}{\partial\varphi} + j_0, \tag{3.5c}$$

$$-iN_3 = r\frac{\partial}{\partial r} - \sigma + 1, \tag{3.5d}$$

where

$$S_r = (1 + vr^2)\frac{\partial}{\partial r} + 2vr(1-\sigma),$$

$$S_\varphi = \frac{1 - vr^2}{r}\frac{\partial}{\partial\varphi} + 2ivrj_0, \tag{3.6}$$

and

$$r = (x^2 + y^2)^{\frac{1}{2}} = (z\bar{z})^{\frac{1}{2}},$$
$$\varphi = \arctan y/x = \ln (z/\bar{z})^{\frac{1}{2}}.$$

B. The $S_i(v)$ Algebra

We choose the basis functions in the D_χ space as the solution of the differential equations

$$(S_1^2 + S_2^2 + vS_3^2)\phi(z) = vj(j+1)\phi(z), \tag{3.7a}$$

$$S_3\phi(z) = m\phi(z). \tag{3.7b}$$

To make this choice plausible, we give a short review of the properties of the operators S_i.[2,3,7]

The S_i operators form a closed subalgebra of the $SL(2, C)$ algebra for any value of the parameter v:

$$[S_1, S_2] = ivS_3, \quad [S_2, S_3] = iS_1, \quad [S_3, S_1] = iS_2. \tag{3.8}$$

This algebra is exactly the $SU(2)$ algebra of the generators M_i for $v = 1$ and isomorphic to it when $v > 0$. At $v = 0$, (3.8) appears to be the $E(2)$ Lie algebra. Finally, for $v < 0$, (3.8) is isomorphic to the $SU(1, 1)$ algebra; at $v = -1$ the $S_i(-1)$ generators are exactly the N_1, N_2, and M_3 ones and form an $SU(1, 1)$ Lie algebra. Hence, the S_i operators are the generators of the IG's.

The operator $C = S_1^2 + S_2^2 + vS_3^2$ commutes with the elements of the S_i algebra:

$$[C, S_i] = 0, \quad i = 1, 2, 3; \tag{3.9}$$

it is the Casimir operator of the (3.8) algebra. It follows that, on a linear space which is irreducible under the IG transformations, the operator C acts like the unit operator, up to a fixed number, which is characteristic to the representation

$$C = vj(j+1)I. \tag{3.10}$$

(As for the precise statement about the irreducibility of the representation space when j is integer or half-integer, see, e.g., Ref. 6.) The representations of the (3.8) algebra can easily be constructed when $v \neq 0$. Some care is needed if we want to reach the point $v = 0$ continuously.[3] We only mention here that for the nontrivial Hermitian representations of the $E(2)$ algebra

$$\lim_{v\to 0} vj(j+1) = \tfrac{1}{4}\epsilon^2 \tag{3.11}$$

is a positive number.

C. Finite IG Elements

The operators given by the definition

$$\exp(-i\xi_1 S_3)\exp(-i\xi_2 S_2)\exp(-i\xi_3 S_3) \tag{3.12}$$

form a subgroup of the $SL(2, C)$ group, which changes its structure as v is varying. Concerning the different regions of v, similar remarks hold good as for the algebra of the S_i. We call this subgroup an IG. Their elements are enumerated if the parameters run over the values

$$0 \le \xi_1, \xi_3 \le 2\pi \quad \begin{cases} 0 \le \xi_2 \le \pi v^{-\frac{1}{2}} & \text{for } v > 0 \\ 0 \le \xi_2 < \infty & \text{for } v \le 0 \end{cases}. \tag{3.12'}$$

In the two-dimensional representation

$$\exp(-i\xi_1 S_3) = \begin{pmatrix} \exp(-\frac{1}{2}i\xi_1) & 0 \\ 0 & \exp(+\frac{1}{2}i\xi_1) \end{pmatrix}, \tag{3.13}$$

$$\exp(-i\xi_2 S_2) = \begin{pmatrix} \cos(\frac{1}{2}(v)^{\frac{1}{2}}\xi_2), & i(v)^{\frac{1}{2}}\sin(\frac{1}{2}(v)^{\frac{1}{2}}\xi_2) \\ i\sin(\frac{1}{2}(v)^{\frac{1}{2}}\xi_2)/(v)^{\frac{1}{2}} & \cos(\frac{1}{2}(v)^{\frac{1}{2}}\xi_2) \end{pmatrix}. \tag{3.14}$$

The last representation is a nonunitary one, even in the $v > 0$, $v \neq 1$ region. We come back to this in the Appendix.

It is easy to see that the elements of the IG leave invariant the 4-vector

$$p_\mu = (\tfrac{1}{2}(1 + v), 0, 0, \tfrac{1}{2}(1 - v)), \quad (3.15)$$

the length of which is equal to $p_\mu^2 = v$.

D. The Explicit Form of the BF's

Now we turn to Eqs. (3.7). From the above it is clear that the BF's $\phi(z)$ are representation functions of the IG's as well. Equation (3.7b) can be satisfied by

$$\phi(z) = \exp [i(m - j_0)\varphi]\tilde{\phi}_{mj_0}(r), \quad (3.16)$$

and (3.7a) is an equation only for $\tilde{\phi}_{mj_0}(r)$. The assumption

$$\tilde{\phi}_{mj_0}(r) = (1 + vr^2)^{\sigma-1}\phi_{mj_0}(r) \quad (3.17)$$

puts Eq. (3.7a) in the form

$$\left(-\frac{(1 + vr^2)^2}{4}\frac{d^2}{dr^2} - \frac{(1 + vr^2)^2}{4r}\frac{d}{dr} + \frac{1 + vr^2}{4r^2} \times [(m - j_0)^2 + vr^2(m + j_0)^2]\right)\phi_{mj_0}(r) = vj(j + 1)\phi_{mj_0}(r). \quad (3.18)$$

Although this equation is familiar from the literature,[1,3] we discuss its solutions with special attention to the fact that we are going to construct functions to represent $SL(2, C)$.

We start with the simplest case $v > 0$. It is easy to verify that the nonnormalized solution of (3.18), regular at $r = 0$, is

$$\phi_{mj_0}(r) = r^\alpha(1 + vr^2)^\beta \times F(-j - \beta, j + 1 - \beta; 1 + \alpha; vr^2/1 + vr^2). \quad (3.19)$$

This is valid for any complex j. We used the notations $\alpha = |m - j_0|$, $\beta = -\tfrac{1}{2}(|m + j_0| + |m - j_0|)$. In the following, it will always be assumed that $m \geq j_0 \geq 0$; then (3.19) looks like

$$\phi_{mj_0}(r) = r^{m-j_0}(1 + vr^2)^{-m} \times F(-j + m, j + 1 + m; 1 + m - j_0; vr^2/1 + vr^2). \quad (3.20)$$

The case $v < 0$ is more complicated. The complication appears because a solution of Eq. (3.18), similar to the $SU(2)$-type function, is not one-valued over the entire complex z plane. Instead, one can find a solution, regular at $r = 0$, one-valued inside the circle $1 + vr^2 = 0$,

$$\phi^+_{mj_0}(r) = r^\alpha(1 + vr^2)^{-j} \times F(-j - \beta, -j + \alpha + \beta; 1 + \alpha; -vr^2), \quad (3.21)$$

and another one, regular at $r = \infty$, one-valued outside the circle $1 + vr^2 = 0$,

$$\phi^-_{mj_0}(r) = !r^{-\alpha}(1 + 1/vr^2)^{-j} \times F(-j - \beta, -j + \alpha + \beta; 1 + \alpha; -1/vr^2). \quad (3.22)$$

The "!" in Eq. (3.22) is to call attention to the fact that, when calculating α and β, instead of m one should write $-m$ into them. In Sec. 3F, we shall give ϕ^+ and ϕ^- for $m \geq j_0 \geq 0$.

To get BF's for all values of r, we add the one-valued part of (3.21) and (3.22):

$$\phi_{mj_0}(r) = \phi^+_{mj_0}(r)\theta(1 + vr^2) + \phi^-_{mj_0}(r)\theta(-1 - vr^2). \quad (3.23)$$

In the special case $v = -1$, this proposal coincides with that of Ref. 1.

It is interesting to examine how the functions (3.19), (3.21), and (3.22) fit into our plan to construct BF's which change continuously their structure with v. Otherwise, can we get all the BF's, if we have a solution of Eq. (3.18) and if we continue it analytically in v at fixed r?

If we start, e.g., with the BF (3.19) we have no problem in the $v > 0$ region. Later we will show that the $v = 0$ point can be reached analytically. Applying the standard formula[8]

$$F(a, b; c; z) = (1 - z)^{-a}F(a, c - b; c; z/z - 1),$$

we see that we can arrive to (3.21), if $1 + vr^2 > 0$. Problems arise, as expected, only when $1 + vr^2 < 0$. In this case, the analytic continuation leads onto the cut of the function (3.19), where it is multivalued. Even restricting ourselves to one Riemann sheet, we have different functions on the upper and lower edge of the cut. Just as an observation, we mention that denoting, by $\phi_{mj_0}(r, v)$, (3.19), continued analytically in v, one can obtain (3.22) as

$$\phi^-_{mj_0}(r) = \phi_{mj_0}(r, |v|\, e^{i\pi}) \mp \phi_{mj_0}(r, |v|\, e^{-i\pi}). \quad (3.24)$$

The minus sign is to be taken when j_0 is integer; the plus sign when j_0 is half-integer.

It is interesting that the functions $\phi^\pm_{mj_0}$ have direct group-theoretical interpretation as well, which can be found out from the following observation: Any given point of the z plane can be reached from any other one by some IG transformation if $v > 0$:

$$z' = \alpha z + \gamma/\beta z + \delta, \quad g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \text{IG}.$$

When $v < 0$, the z plane is split into two disjoint regions with respect to the above mapping; the interior and exterior of the circle $1 + vr^2 = 0$ are mapped onto themselves. Consequently, the ϕ^+ and ϕ^- functions are separately suitable to represent the $SU(1, 1)$ or isomorphic groups. This phenomenon of the $SL(2, C)$ representation theory is well known, and it is termed that the $SU(1, 1)$ representations appear with multiplicity two in the representations of $SL(2, C)$.

Finishing the discussion of the solutions of Eq. (3.18), we turn to the point $v = 0$. First we mention that, since $F(a, b; c; z)$ is analytic inside the circle $|z| = 1$, the form of (3.21) assures

$$\lim_{v\to+0} \phi_{mj_0}(r, v) = \lim_{v\to-0} \phi_{mj_0}(r, v). \qquad (3.25)$$

Many aspects of the reaching of the point $v = 0$ have already been discussed,[3] and here we just mention that, if, together with $\lim v = 0$, $\lim vj(j + 1) = \frac{1}{4}\epsilon^2$, then Eq. (3.18) becomes

$$\left(\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \frac{(m - j_0)^2}{r^2} + \epsilon^2\right)\phi_{mj_0}(r) = 0 \quad (3.26)$$

and its solutions are

$$\begin{aligned}
&\lim_{v\to0} \phi_{mj_0}(r) \\
&= \lim_{v\to0} r^{m-j_0}(1 + vr^2)^{-m} \\
&\quad \times F(-j + m, j + 1 + m; 1 + m - j_0; vr^2/(1 + vr^2)) \\
&= \Gamma(1 + m - j_0)(2/\epsilon)^{m-j_0} J_{m-j_0}(\epsilon r). \qquad (3.27)
\end{aligned}$$

If ϵ^2 is real and positive, these functions yield infinite-dimensional unitary representation of the subgroup $E(2)$. For any other $\epsilon \neq 0$ values we get nonunitary representations. If $\epsilon = 0$, we either have a one-dimensional trivial unitary representation or finite-dimensional nonunitary ones. These latter will be discussed in the Appendix.

E. Anglelike Variables

In the literature, instead of our r and $vr^2/(1 + vr^2)$ variables, anglelike variables are more often used. Though in this paper we do not use them, we give the connection for the reader's convenience.

The introduction of the anglelike variables is based on the mapping of the complex plane onto an ellipsoid (sphere) or a hyperboloid, respectively. A suitable mapping is the following:

$$\begin{cases} \varphi = \varphi', & r = v^{-\frac{1}{2}} \tan(\frac{1}{2}v^{\frac{1}{2}}\vartheta), \\ \varphi = 2\pi - \varphi', & r = (-v)^{-\frac{1}{2}} \coth[\frac{1}{2}(-v)^{\frac{1}{2}}\vartheta] \\ & \quad \text{if } v < 0 \text{ and } -vr^2 > 1. \end{cases} \qquad (3.28)$$

This mapping maps the plane onto a unit sphere if $v = 1$, the region $r < 1$ is on the lower hemisphere, and the $r > 1$ is on the upper. As v decreases, the equator of the sphere grows as $v^{-\frac{1}{2}}$, and the north and south pole remain fixed. At $v = 0$ the ellipsoid degenerates into two parallel planes which curve to a hyperboloid if $v < 0$.

Inserting the anglelike variables into the solutions of Eq. (3.18), we very easily see that at the $v = 1$, 0, -1 points our results coincide with those known from the literature.[1,9]

F. Summary of BF's

Summarizing what we have obtained in this section, we collect the expressions for the BF's, $m \geq j_0 \geq 0$:

(a) For $v > 0$,

$$\begin{aligned}
\phi_{jm,\chi}(z, v) &= N_1 e^{i(m-j_0)\varphi} r^{m-j_0}(1 + vr^2)^{\sigma-1-m} \\
&\quad \times F(-j + m, j + 1 + m; 1 + m - j_0; vr^2/1 + vr^2).
\end{aligned} \qquad (3.29)$$

(b) For $v = 0$,

$$\begin{aligned}
\phi_{\epsilon m,\chi}(z, 0) &= N_2 e^{i(m-j_0)\varphi} J_{m-j_0}(\epsilon r) \\
&\quad \times (2/\epsilon)^{m-j_0}\Gamma(1 + m - j_0).
\end{aligned} \qquad (3.30)$$

(c) For $v < 0$,

$$\begin{aligned}
\phi_{jm,\chi}(z, v) &= \hat{\phi}^+_{mj_0}(r)\theta(1 + vr^2) \\
&\quad + \hat{\phi}^-_{mj_0}(r)\theta(-1 - vr^2),
\end{aligned} \qquad (3.31)$$

$$\begin{aligned}
\hat{\phi}^+_{mj_0}(r) &= N_3^+ e^{i(m-j_0)\varphi} r^{m-j_0}(1 + vr^2)^{-1-j} \\
&\quad \times F(-j + m, -j - j_0; 1 + m - j_0; -vr^2),
\end{aligned} \qquad (3.32)$$

$$\begin{aligned}
\hat{\phi}^-_{mj_0}(r) &= N_3^- e^{i(m-j_0)\varphi} r^{-m-j_0}(1 + 1/vr^2)^{-j}(1 + vr^2)^{\sigma-1} \\
&\quad \times F(-j + m, -j + j_0; 1 + m + j_0; -1/vr^2).
\end{aligned} \qquad (3.33)$$

The normalization factors will be calculated later.

As we discussed in Sec. 3D, for $\phi_{jm,\chi}(z, v)$ the form (3.31) is valid for any v. Formulas (3.29), (3.30), and (3.32) go into one another as v is varying at fixed r; the θ function standing before $\hat{\phi}^-$ is equal to zero for $v \geq 0$.

4. FURTHER NOTES ON THE IBF's

The subject of the further sections is the calculation of the IBF (2.3) with the BF's determined in the previous section.

As we mentioned in Sec. 2A, the scalar product, i.e., positive-definite, Hermitian IBF, exists only in such D_χ spaces, $\chi = (j_0, \sigma)$, for which (a) $\sigma = -\bar{\sigma}$, (b) $j_0 = 0$, σ real, $0 < |\sigma| < 1$, and (c) $j_0 = 0$, $\sigma = 1$. Since we want to consider other D_χ spaces as well when calculating matrix elements, we have to use the

IBF specified in Eq. (2.3). If we change in it not only the signs of j_0 and σ but the sign of m as well and if instead of j we put $-j-1$, then this $B(\varphi, \psi)$ coincides with the scalar product of the principal series.

Consequently, if $\varphi_{jm,\chi}(z, v)$ and $\psi_{kn,\chi}(z, u)$ are two elements of the same D_χ and we want to calculate the matrix element of an operator A between them, we shall write:

$$B(\varphi, A\psi) = \tfrac{1}{2}i \int dz\, d\bar{z} \varphi_{-j-1,-m,-\chi}(z, v) A \psi_{kn,\chi}(z, u). \tag{4.1}$$

As special cases the following basic quantities are important:

(i) The value of the normalization integral for the BF's, when both functions of Eq. (4.1) are taken at the same value of v, $A = 1$;

(ii) the elements of the transformation matrix between normalized BF's (so-called overlap functions), when φ and ψ belong to the same D_χ space, but one of them is an eigenfunction of $S^2(v)$, and the other of $S^2(u)$, $v \neq u$ [cf. Eq. (3.7)] and $A = 1$;

(iii) matrix elements of $SL(2, C)$ representations. Now both functions φ and ψ belong to the same value of v, $A = T(g)$, and $g \in SL(2, C)$. The effect of $T(g)$ onto a BF is given by Eq. (2.1). As special cases, the finite-boost representation matrix elements are the most important, i.e., when $g = \exp(i\alpha N_3)$. As a matter of fact, the boost matrix elements are special cases of (ii) too.

In this paper we discuss in detail only the first question, i.e., the normalization factors, which remained unknown in the previous section. Now we prescribe them as follows:

$$N_1 = \left(\frac{\Gamma(j+1+\alpha+\beta)\Gamma(j+1-\beta)}{\Gamma(j+1-\alpha-\beta)\Gamma(j+1+\beta)} \right)^{\frac{1}{2}} \times \frac{v^\alpha}{(\pi)^{\frac{1}{2}}\Gamma(1+\alpha)} \quad \text{for} \quad v > 0, \tag{4.2}$$

where $\alpha = |m - j_0|$ and $\beta = -\frac{1}{2}(|m - j_0| + |m + j_0|)$,

$$N_2 = \lim_{\substack{v \to 0 \\ vj(j+1) \to \frac{1}{4}\epsilon^2}} N_1 = \frac{2^\alpha}{\epsilon^\alpha \Gamma(1+\alpha)\pi^{\frac{1}{2}}} \quad \text{for} \quad v = 0, \tag{4.3}$$

$$N_3^+ = N_1 \quad \text{for} \quad v < 0, \quad 1 + vr^2 > 0,$$
$$N_3^- = N_1 v^{-\alpha} \quad \text{for} \quad v < 0, \quad 1 + vr^2 < 0. \tag{4.4}$$

For convenience, we introduce the following notation for the normalized BF's:

$$\phi_{j,m,\chi}(z, v) = |j, m; v\rangle = \begin{cases} |j, m; v > 0\rangle & \text{if } v \text{ is positive}, \\ |\epsilon, m;\rangle & \text{if } v \text{ is zero}, \\ |j, m; v < 0\rangle = |j, m; +\rangle + |j, m; -\rangle & \text{if } v \text{ is negative}. \end{cases}$$

The two terms in the case $v < 0$ are to be identified with the function inside ($|j, m; +\rangle$) and outside ($|j, m; -\rangle$) the circle $1 + vr^2 = 0$. For the IBF (4.1) we are going to use the bracket symbol like

$$\langle j, m; v \mid k, n; u \rangle$$

if the BF's in it are normalized.

5. THE NORMALIZATION INTEGRALS

For the $SU(2)$-like case, $v > 0$, we make use of the fact that our basis functions are essentially representation functions of the IG's. The application of the standard methods[9] gives for integer or half-integer values of j, $|m| \leq j$:

$$\langle j, m; v > 0 \mid k, n; v > 0 \rangle = \delta_{jk}\delta_{mn}/v(2_j + 1). \tag{5.1}$$

We turn to the $E(2)$ case, $v = 0$. This gives a simple example to demonstrate that, in general, the IBF (4.1) must be considered as a generalized function. Obviously, we have

$$\langle \epsilon, m \mid \epsilon', n \rangle = 2\delta_{mn} \int r\, dr J_{m-j_0}(\epsilon r) J_{m-j_0}(\epsilon' r). \tag{5.2}$$

We do not want to restrict ourselves to unitary representations, and thus we let ϵ and ϵ' be arbitrary complex numbers. The problem with the calculation of Eq. (5.2) is that the integral in it diverges. Nevertheless, it can be treated following the standard regularization technique of the generalized functions.[10] As a first step we define the Weber–Schaftlein-type integral[11]

$$I_\rho = \int r^\rho\, dr J_{m-j_0}(\epsilon r) J_{m-j_0}(\epsilon' r) \tag{5.3}$$

for Re $\rho < 1$ and for arbitrary complex values of ϵ and ϵ'. The above-mentioned regularization technique suggests defining the integral in Eq. (5.2) as the rhs of the formula [7.7.4.(29)] of Ref. 11, for complex values of $a = \epsilon$ and $b = \epsilon'$, after having performed the limit $\rho \to 1$. Straightforward manipulations yield

$$\lim_{\rho \to 1} I_\rho = 2(\epsilon/\epsilon')^{m-j_0} \left[\epsilon'^{-1} \left(\lim_{\rho \to 1} \frac{(1 - \epsilon^2/\epsilon'^2)^{-\rho}}{\Gamma((1-\rho)/2)} \right) \theta(|\epsilon| - |\epsilon'|) + \epsilon^{-2} \left(\lim_{\rho \to 1} \frac{(1 - \epsilon'^2/\epsilon^2)^{-\rho}}{\Gamma((1-\rho)/2)} \right) \theta(|\epsilon'| - |\epsilon|) \right]. \tag{5.4}$$

This expression shows that the basis functions with different ϵ and ϵ' are orthogonal but for $\epsilon = -\epsilon'$. Actually, such $E(2)$ representations are equivalent, and so we make the restriction $\operatorname{Re}\epsilon > 0$. Then, Eq. (5.4) gives, e.g., for real ϵ and ϵ',

$$\langle \epsilon, m \mid \epsilon', n \rangle = \delta_{mn}(2/\epsilon)\delta(\epsilon - \epsilon'). \tag{5.5}$$

Lastly, we discuss the $SU(1,1)$-like case, $v < 0$. First, we calculate $\langle j, m; + \mid k, n; + \rangle$. Here again the IBF is to be taken as a generalized function, since the integrals appearing diverge. We proceed similarly as above considering first a convergent integral

$$I_\lambda = \int dx\, x^\alpha (1-x)^{k-j-1+\lambda} \times F(-j-\beta+\lambda, -j+\alpha+\beta; 1+\alpha+\lambda; x) \times F(1+k+\alpha+\beta, 1+k-\beta; 1+\alpha; x) \tag{5.6}$$

in its domain of convergence; what we need is $\lim I_\lambda$ as $\lambda \to 0$. In this way, working in the spirit of the regularization technique,[10] we get

$$\langle j, m; + \mid k, n; + \rangle = \delta_{mn} \frac{\Gamma(-1-j-k)\Gamma(1+j+k) \lim_{\lambda\to 0} B(j-k, k-j+\lambda)}{\Gamma(j+1+\alpha+\beta)\Gamma(-j-\alpha-\beta)\Gamma(j+1+\beta)\Gamma(-j-\beta)} (-1)^{1+\alpha}, \tag{5.7}$$

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

If $j = -\frac{1}{2} + i\rho$ and $k = -\frac{1}{2} + i\tau$, then

$$\langle j, m; + \mid k, n; + \rangle = \delta_{mn}(-1)^{2\beta+2\epsilon} \frac{\tan(\pi j + \epsilon)}{2(2j+1)} \delta(\rho - \tau), \tag{5.8}$$

$$\epsilon = \begin{cases} 0 & \text{if } m, j_0 \text{ are integer,} \\ \frac{1}{2} & \text{if } m, j_0 \text{ are half-integer.} \end{cases}$$

If j and k are integer, then

$$\langle j, m; + \mid k, n; + \rangle = \begin{cases} 0 & \text{if } j \neq k, \\ \delta_{mn}/v(2j+1) & \text{if } j = k, \quad j \le \min(|m|, |j_0|), \\ & \qquad mj_0 \ge 0. \end{cases} \tag{5.9}$$

The results are the same for $\langle j, m; - \mid k, n; - \rangle$.

APPENDIX

In Sec. 3, we obtained the BF's, the elements of a D_χ space. If we omit the over-all $(1 + vr^2)^{\sigma-1}$ factor and arrange the elements in matrix form according to m and j_0, we get the matrix representation of two-parameter elements of the IG's, as we discussed in Sec. 3.

If the normalization is that of Eq. (4.2)–(4.4), the representation is unitary for the appropriate values of j, m, and j_0.

Let us change the normalization a bit by writing

$$N_1' = N_1 v^{\frac{1}{2}(m-j_0)}. \tag{A1}$$

As a consequence, an element of j_0 and m and one of $-j_0$ and $-m$ will now differ from each other by a factor v^{m-j_0}. This destroys the unitarity of the representations, as one can see by direct calculation, e.g., from Eq. (3.14). [For this purpose the anglelike variables (3.28) are preferred.] Now the $v \to 0$, $vj(j+1) \to \frac{1}{4}\epsilon^2$, $\epsilon \neq 0$ limit does not give generally regular representation. Taking j finite, we see that the $v \to 0$ limit yields

$$\lim_{v\to 0} \phi_{jm,\chi}(z, v) = 0 \quad \text{if} \quad m - j_0 > 0$$
$$= \text{const} \times \exp i(m - j_0)\varphi r^{m-j_0}. \tag{A2}$$

As can be proven by simple replacement, it is a solution of Eq. (3.26) with $\epsilon = 0$; so this is a nonunitary representation of $E(2)$.

As an IG representation, the matrix of Eq. (A2) is a lower-triangle matrix. Had we written $N_1' = N_1 v^{j_0 - m/2}$, the matrix would have been an upper-triangle one. An interesting feature is that, in this representation, one of the step operators, either S_+ or S_-, is identically zero. It means that this basis is not a cyclic one.

In this case it is possible to fix j, and, as v goes from 1 to -1, the representations change continuously from $SU(2)$ type to $SU(1,1)$ type via $E(2)$, through nonunitary, nontrivial ones, without the appearance of any singularity [cf. Eq. (3.14)]. If we take unitary representations and want to get regular representations (and nontrivial ones) for any v, we have to allow j to go to infinity [cf. Eq. (3.27)].

[1] N. Mukunda, J. Math. Phys. **9**, 50 (1968).
[2] A. O. Barut, in *Boulder Lectures in Theoretical Physics, Vol. 9-A: Mathematical Methods of Theoretical Physics*, edited by W. E. Brittin and A. O. Barut (Gordon and Breach, New York, 1967).
[3] K. Szegö and K. Tóth, Nuovo Cimento **66A**, 371 (1970).
[4] R. Delbourgo *et al.*, Nuovo Cimento **52A**, 1254 (1967).
[5] K. Szegö and K. Tóth, J. Math. Phys., **12**, 853 (1971).
[6] I. M. Gel'fand *et al.*, *Generalized Functions* (Academic, New York, 1968), Vol. 5.
[7] G. Domokos, *Proceedings of the Eighth Nobel Symposium* (Almquist and Wiksell, Stockholm, 1968).
[8] H. Bateman, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 1.
[9] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U.P., Princeton, N.J., 1957).
[10] I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic, New York, 1964), Vol. 1.
[11] Reference 8, Vol. 2.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

$SL(2, C)$ Representations in Explicitly "Energy-Dependent" Basis. II

K. SZEGÖ AND K. TÓTH
Central Research Institute for Physics, Budapest, Hungary

(Received 24 July 1970)

A systematic treatment is presented for the transformation matrix elements between different type of basis sets in the $SL(2, C)$ representation space and for the boost matrix elements. The treatment is based on the theory of invariant bilinear functionals and makes use of the knowledge of explicit basis functions. New expressions are given for the $SU(2) \leftrightarrow E(2)$ and $SU(1, 1) \leftrightarrow E(2)$ overlap functions, and it is shown that they can be considered as a Fourier transform of a function, which is also known. Some relations are proved between boost functions and transformation matrix elements. In the Appendix a new and relatively simple expression is derived for the boost matrix elements $d^{j_0\sigma}_{jmj'}(\xi)$ in $SU(2)$ basis.

1. INTRODUCTION

Developments in the last few years in physics have made important expansions with respect to $SL(2, C)$ representations, which include infinite-dimensional nonunitary representations as well. In connection with these expansions, it has also become clear that the exploration of $SL(2, C)$ representations in other bases than the conventional $SU(2)$ basis may lead to much new information. On the other hand, these expansions raise a lot of mathematical questions. Among them, we mention that, in general, physical applications favor certain types of basis sets, depending on the spacelike, timelike, or lightlike character of a 4-vector which plays a distinguished role in the physical problem. The squared length of this 4-vector used to be a fundamental variable, and its continuous change from positive to negative values is usually assumed. It is a natural desire to relax the description of the $SL(2, C)$ representations in such a way that the change of this variable be followed by the change of the character of the basis system chosen in the representation space. A detailed survey of basis set, in which the little group of the 4-vector

$$p_\mu = [\tfrac{1}{2}(1 + v), 0, 0, \tfrac{1}{2}(1 - v)]$$

is diagonal and depends explicitly on v, has been given in another paper.[1] (Hereafter referred to as I). The present paper is intended to make use of the knowledge of these basis functions for investigating further problems.

A question arising from the use of different basis systems is the calculation of the transformation matrix elements between them. This problem has already been studied by many authors[2–6] and most completely in Ref. 7. The method of induced representations is extensively used in most of these papers. In possession of the basis functions, we see that these matrix elements appear in a very natural manner. Nevertheless, our results may be preferred to those of Ref. 7 due to the following reasons:

(a) In many cases our formulas are simpler and more suitable for further calculations.

(b) The calculation of some auxiliary quantities is, to a certain extent, doubtful in the paper by Delbourgo *et al.*[7] These quantities, denoted by $\langle A, j_0 \,|\, j_0, \sigma\}$ are calculated in Ref. 7 for finite-dimensional representations and then extended to the unitary representations. The very exceptional position of the finite-dimensional representations in the theory of the $SL(2, C)$ representations makes questionable such a procedure. In our approach these quantities do not appear at all. A further nice feature of our method is that it gives, with a coherent normalization, all the transformation matrix elements and boost functions.

Another problem one meets is connected with the appearance of nonunitary representations. Physicists generally favor the use of the so-called bra–ket technique, and it makes necessary special care when the basis vectors are not elements of a Hilbert space. In this case, the usual scalar product is to be replaced by the more general notion of bilinear functional. As a matter of fact, the aim of this paper is the investigation of bilinear functionals, through use of the basis functions explicitly constructed in I.

In Sec. 2 the main points of Paper I are reviewed. In Secs. 3 and 4, bilinear functionals are calculated: in Sec. 3, the ones which are specifically called "overlap functions" and, in Sec. 4, the ones which are known as boost functions. In the Appendix a differential equation method is outlined, which gives useful ideas to the calculation of boost functions.

2. SUMMARY OF THE BASIS SETS

In Paper I the construction of explicit basis functions was studied; here we give a short review of the results.

As usual, we considered the space D_χ, $\chi = (j_0, \sigma)$, of the infinitely differentiable functions $\phi(z, \bar{z})$ [in the following, $\phi(z)$] of one complex variable $z = x + iy$ and $\bar{z} = x - iy$. The elements g of $SL(2, C)$ are represented by the operators $T_\chi(g)$ acting on the functions $\phi(z)$ as follows:

$$T_\chi(g)\phi(z) = (\beta z + \delta)^{-j_0+\sigma-1}(\bar{\beta}\bar{z} + \bar{\delta})^{j_0+\sigma-1}\phi\left(\frac{\alpha z + \gamma}{\beta z + \delta}\right), \tag{2.1}$$

where $\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$, $\alpha\delta - \beta\gamma = 1$, is the 2×2 matrix representation of g. On the other hand, we introduced the one-parameter elements of $SL(2, C)$ in the form $e^{-i\alpha M_k}$ and $e^{-i\alpha N_k}$, where M_k and N_k, $k = 1, 2, 3$, are the six generators of $SL(2, C)$ commuting like

$$\begin{aligned}
[M_i, M_j] &= -[N_i, N_j] = i\epsilon_{ijk}M_k,\\
[M_i, N_j] &= i\epsilon_{ijk}N_k.
\end{aligned} \tag{2.2}$$

Then, choosing the generators to be $M_k = \frac{1}{2}\sigma_k$ and $N_k = \frac{1}{2}i\sigma_k$ in the 2×2 matrix representation, where σ_k are the Pauli matrices, we find it easy to get M_k and N_k as differential operators acting on the functions $\phi(z)$. Namely, one gets

$$\begin{aligned}
M_1\phi(z) &= \frac{1}{2}\Big((\bar{z}^2 - 1)\frac{\partial}{\partial\bar{z}} - (z^2 - 1)\frac{\partial}{\partial z}\\
&\qquad + (\sigma - j_0 - 1)z - (\sigma + j_0 - 1)\bar{z}\Big)\phi(z),\\
M_2\phi(z) &= \frac{i}{2}\Big((\bar{z}^2 + 1)\frac{\partial}{\partial\bar{z}} + (z^2 + 1)\frac{\partial}{\partial z}\\
&\qquad - (\sigma - j_0 - 1)z - (\sigma + j_0 - 1)\bar{z}\Big)\phi(z),\\
M_3\phi(z) &= \Big(z\frac{\partial}{\partial z} - \bar{z}\frac{\partial}{\partial\bar{z}} + j_0\Big)\phi(z),\\
N_1\phi(z) &= \frac{i}{2}\Big((1 - \bar{z}^2)\frac{\partial}{\partial\bar{z}} + (1 - z^2)\frac{\partial}{\partial z}\\
&\qquad + (\sigma - j_0 - 1)z + (\sigma + j_0 - 1)\bar{z}\Big)\phi(z),\\
N_2\phi(z) &= \frac{1}{2}\Big((1 + \bar{z}^2)\frac{\partial}{\partial\bar{z}} - (1 + z^2)\frac{\partial}{\partial z}\\
&\qquad + (\sigma - j_0 - 1)z - (\sigma + j_0 - 1)\bar{z}\Big)\phi(z),\\
N_3\phi(z) &= i\Big(z\frac{\partial}{\partial z} + \bar{z}\frac{\partial}{\partial\bar{z}} - \sigma + 1\Big)\phi(z).
\end{aligned} \tag{2.3}$$

This realization of the generators serves as groundwork when we construct basis functions in the D_χ spaces. The Casimir operators $\mathbf{M}^2 - \mathbf{N}^2$ and $\mathbf{MN}$, as calculated from Eqs. (2.3), appear to be

$$\begin{aligned}
(\mathbf{M}^2 - \mathbf{N}^2)\phi(z) &= (\sigma^2 + j_0^2 - 1)\phi(z),\\
\mathbf{MN}\phi(z) &= -ij_0\sigma\phi(z)
\end{aligned} \tag{2.4}$$

for all the elements $\phi(z)$ of a given space D_χ, where j_0 is an integer or half-integer and σ is an arbitrary complex number. As is known, four independent operators, made from the generators M_k and N_k, can be diagonalized in a basis set, which one is going to use for representing $SL(2, C)$. In I, the two operators beyond $\mathbf{M}^2 - \mathbf{N}^2$ and $\mathbf{MN}$ were chosen to be $C = S_1^2(v) + S_2^2(v) + vS_3^2(v)$ and $S_3(v)$, where

$$\begin{aligned}
S_1(v) &= \tfrac{1}{2}(1 + v)M_2 + \tfrac{1}{2}(1 - v)N_1,\\
S_2(v) &= -\tfrac{1}{2}(1 + v)M_1 + \tfrac{1}{2}(1 - v)N_2,\\
S_3(v) &= M_3,
\end{aligned} \tag{2.5}$$

and v is an arbitrary real parameter. The $S_i(v)$ operators can be used for generating a subgroup of $SL(2, C)$ which, in the four-dimensional representation, leaves invariant the 4-vector $p_\mu = [\frac{1}{2}(1 + v), 0, 0, \frac{1}{2}(1 - v)]$, $p_\mu^2 = v$. This group is

(a) the $SU(2)$ group for $v = 1$ and isomorphic to $SU(2)$ for $v > 0$,

(b) the $E(2)$ group for $v = 0$,

(c) the $SU(1, 1)$ group for $v = -1$ and isomorphic to $SU(1, 1)$ for $v < 0$. Due to these properties in Paper I, this group was called the "interpolating group" (IG).

The basis vectors in the space D_χ can be determined as the eigenfunctions of the operators C and $S_3(v)$:

$$C\phi(z) = vj(j + 1)\phi(z), \tag{2.6a}$$

$$S_3\phi(z) = m\phi(z). \tag{2.6b}$$

When $v = 0$, a suitable definition for the eigenvalue of C is

$$\lim_{v\to 0} vj(j + 1) = \tfrac{1}{4}\epsilon^2. \tag{2.7}$$

In general, we are not going to restrict ourselves to unitary representations and allow arbitrary complex values for j and ϵ. An appropriate solution of Eqs. (2.6) is (see Paper I)

$$\begin{aligned}
\phi_{jm}^\chi(z, v) &= \phi_{jm}^{\chi(+)}(z, v) + \phi_{jm}^{\chi(-)}(z, v)\\
&= e^{i(m-j_0)\varphi}(1 + vr^2)^{\sigma-1}\\
&\quad\times \{\theta(1 + vr^2)\tilde{N}_1[(v)^{\frac{1}{2}}r]^\alpha(1 + vr^2)^{-j}\\
&\quad\times F(-j - \beta, -j + \alpha + \beta; 1 + \alpha; -vr^2)\\
&\quad+ \theta(-1 - vr^2)\tilde{N}_2[(v)^{\frac{1}{2}}r]^{-\alpha'}(1 + 1/vr^2)^{-j}\\
&\quad\times F(-j - \beta, -j + \alpha' + \beta; 1 + \alpha; -1/vr^2)\}.
\end{aligned} \tag{2.8}$$

In these formulas, polar variables $r = |z|$ and $\varphi = \arg z$ and the following notations are introduced:

$$\alpha = |m - j_0|, \quad \alpha' = |m + j_0|, \quad \beta = -\tfrac{1}{2}(\alpha + \alpha').$$

The normalization factors are prescribed as follows:

$$\tilde{N}_1 = \frac{1}{\pi^{\frac{1}{2}}} \frac{1}{\Gamma(1+\alpha)} \left(\frac{\Gamma(1+j-\beta)\Gamma(1+j+\alpha+\beta)}{\Gamma(1+j+\beta)\Gamma(1+j-\alpha-\beta)} \right)^{\frac{1}{2}}.$$

$\tilde{N}_2$ can be obtained from $\tilde{N}_1$ by changing α to α'. For any real value of v, the functions (2.8) can be used as basis states in the representation space D_χ. Specially, they are the functions of the $E(2)$ basis for $v = 0$, if we take $\lim_{v\to 0} vj(j+1) = \frac{1}{4}\epsilon^2$:

$$\lim_{v\to 0} \phi^\chi_{jm}(z, v) = \frac{1}{\pi^{\frac{1}{2}}} e^{i(m-j_0)\varphi} J_\alpha(\epsilon r). \tag{2.9}$$

In the rest of the paper we shall always take $m \geq j_0 \geq 0$.

3. BILINEAR FUNCTIONALS

It is well known that the representation spaces D_χ are generally not Hilbert spaces and an invariant, positive-definite scalar product does not always exist between the elements of a space D_χ. It exists only when:

(i) j_0 is integer or half-integer and σ is pure imaginary (principal series),

(ii) $j_0 = 0$, σ is real, and $0 < |\sigma| < 1$ (supplementary series),

(iii) $j_0 = 0$, $\sigma = 1$ (trivial representation).

In the other cases the substitute for the scalar product is the so-called invariant bilinear functional. A beautiful treatment of the theory of bilinear functionals, invariant under $SL(2, C)$ transformations, can be found in Ref. 8, and a short summary was given in I.

In Paper I we started with the calculation of the bilinear functional

$$B(\varphi, \psi) = \tfrac{1}{2}i \int dz\, d\bar{z} \varphi(z)\psi(z), \tag{3.1}$$

where $\varphi(z)$ and $\psi(z)$ are elements of the spaces D_χ and $D_{-\chi}$, respectively, $\chi = (j_0, \sigma)$, $-\chi = (-j_0, -\sigma)$. Once the basis functions in the D_χ spaces are explicitly known, we are able to calculate $B(\varphi, \psi)$ in many concrete cases.

We introduce, as we did also in I, the following notations:

(i) The ket symbols $|j, m; v\rangle$ or $|\epsilon, m\rangle$ will be used for the normalized basis functions in a space D_χ with given $\chi = (j_0, \sigma)$.

(ii) The bracket symbol $\langle j, m; v \,|\, j', m'; v'\rangle$ will be used for

$$\tfrac{1}{2}i \int dz\, d\bar{z} \varphi^{-\chi}_{-j-1,-m}(z, v) \psi^\chi_{j',m'}(z, v'). \tag{3.2}$$

It was shown in I that $\langle j, m; v \,|\, j', m'; v\rangle$, as defined by (3.2), coincides with the usual scalar product for the unitary principal series.

In the subsequent part of the paper we are going to calculate the integral (3.2) in such cases, when $v \neq v'$. Specially, in the remainder of this section the cases $v' = 0$, $v \neq 0$, and $v' > 0$, $v < 0$ will be considered.

A. The $SU(2) \leftrightarrow E(2)$ Overlap Functions $\langle \epsilon, m \,|\, j, m; v\rangle$

For positive values of v the second term is identically zero in the expression (2.8) of the basis function $|j, m; v\rangle$. After expanding into power series in the variable

$$x = vr^2/(1 + vr^2),$$

the hypergeometric function of the first term, and after using the binomial formula for

$$[vr^2/(1 + vr^2)]^n = [1 - 1/(1 + vr^2)]^n,$$

we integrate term by term to obtain

$$\begin{aligned}
\langle \epsilon, m \,|\, j, m'; v\rangle
&= \frac{2}{v} \frac{\delta_{mm'}}{\Gamma(m-j)\Gamma(m+j+1)} \\
&\times \left(\frac{\Gamma(1+j+m)\Gamma(1+j-j_0)}{\Gamma(1+j-m)\Gamma(1+j+j_0)} \right)^{\frac{1}{2}} \\
&\times \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{(-1)^k}{k!\,(n-k)!} \\
&\times \frac{\Gamma(m-j+n)\Gamma(m+j+1+n)}{\Gamma(1-j_0+m+n)\Gamma(1-\sigma+m+k)} \\
&\times \left(\frac{\epsilon}{2v^{\frac{1}{2}}} \right)^{k+m-\sigma} K_{\sigma-j_0-k}\left(\frac{\epsilon}{v^{\frac{1}{2}}} \right).
\end{aligned} \tag{3.3}$$

This formula is valid for arbitrary complex j and ϵ, $\epsilon \neq 0$, with the restriction that the summation for k must first be performed. In the case when $j - m$ is positive integer [unitary $SU(2)$ representations], the infinite sum with respect to n terminates, and (3.3) can be simplified. Indeed, for $j - m$ a positive integer, the hypergeometric function in $|j, m; v\rangle$ can be transformed in the following manner:

$$\begin{aligned}
&(1 + vr^2)^{-j+m} F(-j+m, -j-j_0; 1+m-j_0; -vr^2) \\
&= (-1)^{j-m} \frac{\Gamma(1+m-j_0)\Gamma(1+j+j_0)}{\Gamma(1+m+j_0)\Gamma(1+j-j_0)} \\
&\times F\left(-j+m, j+m+1; 1+m+j_0; \frac{1}{1+vr^2}\right).
\end{aligned}$$

The integration now yields:

$$\langle \epsilon, m \mid j, m'; v\rangle = \frac{2}{v}(-1)^{j-m}\delta_{mm'}\left(\frac{\Gamma(1+j-m)\Gamma(1+j+j_0)}{\Gamma(1+j+m)\Gamma(1+j-j_0)}\right)^{\frac{1}{2}} \sum_{k=0}^{j-m}(-1)^k \frac{1}{k!} \times \frac{\Gamma(1+j+m+k)}{\Gamma(1+j-m-k)\Gamma(1+j_0+m+k)\Gamma(1-\sigma+m+k)}\left(\frac{\epsilon}{2v^{\frac{1}{2}}}\right)^{k+m-\sigma} K_{\sigma-j_0-k}\left(\frac{\epsilon}{v^{\frac{1}{2}}}\right). \quad (3.4)$$

This formula was first derived in Ref. 6 by a different method.

B. The $SU(1,1) \leftrightarrow E(2)$ Overlap Functions $\langle \epsilon, m \mid j, m; \pm\rangle$

In this case there are two functions to be calculated correspondingly to the two terms in (2.8). We shall use for them the notations $\langle \epsilon, m \mid j, m'; +\rangle$ and $\langle \epsilon, m \mid j, m'; -\rangle$. We start with $\langle \epsilon, m \mid j, m'; +\rangle$. By similar trick as in the $SU(2)$ case it is easy to get

$$\langle \epsilon, m \mid j, m'; +\rangle = -\frac{1}{v}e^{\frac{1}{2}i\pi(m-j_0)}\frac{\delta_{mm'}}{\Gamma(-j+m)\Gamma(j+1+m)} \times \left(\frac{\Gamma(1+j+m)\Gamma(1+j-j_0)}{\Gamma(1+j-m)\Gamma(1+j+j_0)}\right)^{\frac{1}{2}} \times \sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{(-1)^k}{k!\,(n-k)!} \times \frac{\Gamma(-j+m+n)\Gamma(-j-j_0+n)\Gamma(\sigma-j+k)}{\Gamma(m-j_0+1+n)} \times \left(\frac{\epsilon}{2(-v)^{\frac{1}{2}}}\right)^{j-k-\sigma} J_{\sigma-j_0-j+m+k}\left(\frac{\epsilon}{(-v)^{\frac{1}{2}}}\right). \quad (3.5)$$

It is worth saying that an alternative form of $\langle \epsilon, m \mid j, m'; +\rangle$ could be obtained also by writing the $SU(1,1)$ representation function in $|j, m; +\rangle$ as a linear combination of second kind functions. This form would be less similar to (3.3) than (3.5) is, but simpler to such an extent that it does not contain a summation corresponding to the one over k in (3.3) and (3.5). The form (3.5) of the function $\langle \epsilon, m \mid j, m'; +\rangle$ does not appear to have been given before, and, together with the corresponding formulas in the $SU(2)$ case, may be preferred to the expressions of Ref. 7, because of the appearance of Bessel functions instead of Meijer functions.

Unfortunately, we have not been lucky enough to find a similar representation for the functions $\langle \epsilon, m \mid j, m'; -\rangle$. After expanding into power series the hypergeometric function in $|j, m'; -\rangle$ and integrating, we obtained:

$$\langle \epsilon, m \mid j, m'; -\rangle = -\frac{1}{v}e^{-\frac{1}{2}i\pi(m+j_0)}\delta_{mm'}\frac{\Gamma(\sigma-j)}{\Gamma(-j+m)\Gamma(-j+j_0)} \times \left(\frac{\Gamma(1+j+m)\Gamma(1+j+j_0)}{\Gamma(1+j-m)\Gamma(1+j-j_0)}\right)^{\frac{1}{2}} \times \sum_{n=0}^{\infty}\frac{\Gamma(-j+m+n)\Gamma(-j+j_0+n)}{n!\,\Gamma(m+j_0+1+n)} \times \left(\frac{\epsilon}{2(-v)^{\frac{1}{2}}}\right)^{-2j+m+j_0+2n} \times G_{13}^{20}\left(-\frac{\epsilon^2}{4v}\,\middle|\,\begin{matrix}0\\ j-\sigma, j-j_0-n, j-m-n\end{matrix}\right). \quad (3.6)$$

C. The $SU(2) \leftrightarrow SU(1,1)$ Overlap Functions

These functions appear as special case of (3.2), when the signs of v and v' are different. It is obvious that, putting the appropriate functions into the integral (3.2), we could reproduce the formulas first derived by the authors of Ref. 2 in the special case $v = 1$ and $v' = -1$. We just mention that it is easy to get alternative formulas by making use of the fact (see Paper I) that all the $E(2)$ basis functions, $|\epsilon, m\rangle$, with real, positive ϵ, form a complete orthogonal system, and we may write

$$\langle j, m; v \mid j', m; v'\rangle = \int_0^\infty \tfrac{1}{2}\epsilon\, d\epsilon \langle j, m; v \mid \epsilon, m\rangle\langle \epsilon, m \mid j', m; v'\rangle. \quad (3.7)$$

A term-by-term integration, applying the functions (3.3), (3.5), and (3.6), yields series in terms of hypergeometric functions, which are not remarkably simpler than the ones in Ref. 2.

We finish this section with a remark which may prove to be advantageous for many purposes and which, in our opinion, is interesting even in itself.

This remark is based on the following integral representation of the Bessel functions:

$$\int_0^\pi e^{ir\epsilon\cos\varphi}\sin^{n-2}\varphi\, d\varphi = \pi^{\frac{1}{2}}\Gamma[\tfrac{1}{2}(n-1)](\tfrac{1}{2}r\epsilon)^{-\frac{1}{2}n+1} J_{\frac{1}{2}n-1}(r\epsilon). \tag{3.8}$$

With this formula, a short calculation gives

$$\langle \epsilon, m \mid j, m; v\rangle = \frac{1}{\pi}\frac{1}{\Gamma(1-j_0+m)}\left(\frac{\Gamma(1+j+m)\Gamma(1+j-j_0)}{\Gamma(1+j-m)\Gamma(1+j+j_0)}\right)^{\frac{1}{2}}\left(\frac{v^{\frac{1}{2}}\epsilon}{2\pi}\right)^{m-j_0}$$
$$\times \int (1+v|x|^2)^{\sigma-1-m} F\left(-j+m, j+m+1; m-j_0+1; \frac{v|x|^2}{1+v|x|^2}\right) e^{i(x,s)}\, dx, \tag{3.9}$$

where $x = (x_1, \cdots, x_k)$, $k = 2(m - j_0 + 1)$, $|x|^2 = x_1^2 + \cdots + x_k^2 = r^2$, $s = (s_1, \cdots, s_k)$, $s_1 = \epsilon$, and $s_2 = \cdots = s_k = 0$. Obviously, (3.9) is a Fourier transform in a $2(m - j_0 + 1)$-dimensional space.

This remark makes us able to make a very simple rederivation of the normalization integral $\langle \epsilon, m \mid \epsilon', m\rangle$. First we notice that, when $\lim vj(j+1) = \frac{1}{2}\epsilon'^2$ as $v \to 0$, we have

$$\lim_{v\to 0} \langle \epsilon, m \mid j, m; v\rangle = \langle \epsilon, m \mid \epsilon', m\rangle.$$

Let us now consider, for simplicity, real ϵ and ϵ' values, and calculate the following quantity:

$$I = \int_0^\infty \tfrac{1}{2}\epsilon'\, d\epsilon' \left(\lim_{v\to 0} \langle \epsilon; m \mid j, m; v\rangle\right)\varphi(\epsilon'),$$

where $\varphi(\epsilon)$ is an infinitely differentiable function with finite support. It is straightforward to get the following by making use of (3.8) and (3.9):

$$I = \frac{1}{2\pi}\int_0^\infty d\epsilon' \left(\int_{-\infty}^\infty e^{i(\epsilon-\epsilon')y}\, dy\right)\varphi(\epsilon'),$$

that is,

$$\langle \epsilon, m \mid \epsilon', m\rangle = 2\epsilon^{-1}\delta(\epsilon - \epsilon').$$

It goes without saying that similar remarks hold good also for $\langle \epsilon, m \mid j, m; +\rangle$ and $\langle \epsilon, m \mid j, m; -\rangle$.

4. THE BOOST MATRIX ELEMENTS

The last type of bilinear functionals we want to deal with involves v and v' with the same sign. As a by-product we shall obtain also the boost representation matrix elements

$$d^{j_0\sigma}_{jmj'}(\xi) = \langle j, m; \pm 1 |\, e^{-i\xi N_3}\, | j', m; \pm 1\rangle.$$

Namely, the following relation is valid:

$$d^{j_0\sigma}_{jmj'}(\xi) = e^{j_0\xi}\langle j, m; \pm 1 \mid j', m; \pm v^{-1}\rangle, \tag{4.1}$$

where the notation $v = \exp(-2\xi)$ is introduced. This statement can be proved simply by referring to the formula (2.1). Let g be the boost transformation $\exp(-i\xi N_3)$. Then we have

$$T(e^{-i\xi N_3})\phi(z) = e^{j_0\xi}\phi(e^{2\xi}z),$$

and from this the validity of (4.1) is obvious.

The calculation of $\langle j, \bar{m}; v \mid j', m'; v'\rangle$ can be performed either by making use of Eq. (3.7) or by the method applied in Refs. 2 and 7. To save space, we write down explicit result only in the special case when $v = 1$, $v' > 0$, and both $j - m$ and $j' - m$ are nonnegative integers. For the calculation, Eq. (3.7) was used:

$$\langle j, m; 1 \mid j', m'; v\rangle = \delta_{mm'}(-1)^{j'-j}\Gamma(m - j_0 + 1)$$
$$\times \left(\frac{\Gamma(1+j-m)\Gamma(1+j+j_0)\Gamma(1+j'-m)\Gamma(1+j'+j_0)}{\Gamma(1+j+m)\Gamma(1+j-j_0)\Gamma(1+j'+m)\Gamma(1+j'-j_0)}\right)^{\frac{1}{2}} \sum_{k=0}^{j-m}\sum_{k'=0}^{j'-m}\frac{(-1)^{k+k'}}{k!\,k'!}$$
$$\times \frac{\Gamma(1+j+m+k)\Gamma(1+j'+m+k')\Gamma(1+j_0+m+k+k')}{\Gamma(1+j-m-k)\Gamma(1+j'-m-k')\Gamma(1+j_0+m+k)\Gamma(1+j_0+m+k')}$$
$$\times \frac{v^{-1-\frac{1}{2}(m-j_0)}}{\Gamma(2m+k+k'+2)} F(m-j_0+1, m-\sigma+1+k'; 2m+2+k+k'; 1-v^{-1}). \tag{4.2}$$

It is worth mentioning that (4.2) can be simplified by sophisticated partial integrations. These partial integrations lead to an expression with many less terms to sum. For this we refer the reader to the Appendix.

Finally, we call attention to a consequence of (4.1): If $\xi \to \infty$ together with $e^{-2\xi}j'(j'+1) \to \frac{1}{4}\epsilon^2$, then

$$\lim_{\xi\to\infty} d^{j_0\sigma}_{jmj'}(\xi)e^{-j_0\xi} = \langle j, m; \pm 1 \mid \epsilon, m\rangle.$$

APPENDIX

As we have seen, the bilinear functional approach presented in this paper for the calculation of transformation matrix elements and boost functions is extremely powerful. Nevertheless, it leaves undiscovered many properties of these functions, which can be used for simplifying the integrals to be calculated. The aim of this appendix is to outline a special differential equation method, which makes us able to transform the expression (4.2) into a much simpler form. This simplification would hardly be found out without the application of the results coming from this differential equation approach. The method first appeared in Ref. 5, where it was introduced in a different and less direct manner.

The construction is as follows. First of all, we observe the symmetric role of the variables ϵ and r in the functions of the $E(2)$ basis. Then, for practical reasons, we introduce the following operators:

$$E_\pm = \mp i(M_1 + iM_2) + (N_1 \pm iN_2),$$
$$F_\pm = \pm i(M_1 + iM_2) + (N_1 \pm iN_2).$$

Applying the differential operator realization (2.3) of the generators, the explicit form

$$|\epsilon, m\rangle = \frac{1}{(\pi)^{\frac{1}{2}}} e^{i(m-j_0)\varphi} J_{m-j_0}(\epsilon r)$$

of the functions of the $E(2)$ basis, and the differentiation formulas of the Bessel functions, one gets

$$E_\pm |\epsilon, m\rangle = \mp i\epsilon |\epsilon, m \pm 1\rangle, \quad \text{(A1)}$$

$$F_\pm |\epsilon, m\rangle = \mp i\left(\epsilon \frac{\partial^2}{\partial\epsilon^2} + (-2\sigma \pm 2m + 3)\frac{\partial}{\partial\epsilon} \mp \frac{(m - j_0 \pm 1)(2\sigma \mp j_0 \mp m - 1)}{\epsilon}\right) \times |\epsilon, m \pm 1\rangle, \quad \text{(A2)}$$

$$M_3 |\epsilon, m\rangle = m |\epsilon, m\rangle, \quad \text{(A3)}$$

$$N_3 |\epsilon, m\rangle = i\left(1 - \sigma + \epsilon \frac{\partial}{\partial\epsilon}\right) |\epsilon, m\rangle. \quad \text{(A4)}$$

This representation of the $SL(2, C)$ generators enables us to write down a differential equation for the overlap functions $\langle\epsilon, m \mid j, m; v\rangle$. One must only express the Casimir operator C, defined in Sec. 2, as a differential operator in the variable ϵ:

$$[\tfrac{1}{4}(E_+E_- - vF_+E_- - vE_+F_- + v^2F_+F_-) + vM_3^2 - vM_3] \times \langle\epsilon, m \mid j, m; v\rangle = vj(j + 1)\langle\epsilon, m \mid j, m; v\rangle. \quad \text{(A5)}$$

Unfortunately, this is a very complicated fourth-order differential equation. Nevertheless, it can be verified by direct calculation that the function $\langle\epsilon, m \mid j, m; v\rangle$, given as a Fourier transform by Eq. (3.9), is a solution of (A5).

Really useful information comes from the representation (A1)–(A4) in the case of $SU(2)$-like bases, when the representation of the IG is unitary. In this case Eq. (A5) can be replaced by a recursion relation.

Before going into the details, we introduce a new normalization, which makes the calculations in the variable ϵ much easier. We redefine the normalization factor $\tilde{N}_1$, given in Sec. 2, as follows:

$$\tilde{N}_1' = [v^{-\sigma}(\Gamma(1 - \sigma + j)/\Gamma(1 + \sigma + j))]^{\frac{1}{2}}\tilde{N}_1. \quad \text{(A6)}$$

Correspondingly, we take, in the $E(2)$ case,

$$\phi'_{\epsilon m}(z) = \pi^{-\frac{1}{2}} e^{i(m-j_0)\varphi}(\tfrac{1}{2}\epsilon)^{-\sigma} J_{m-j_0}(\epsilon r). \quad \text{(A7)}$$

The use of this normalization is that the realization (A2) of the operators $F_\pm$ simplifies. Using from now on the notation $|\epsilon, m\rangle$ for the function (A7) and $|j, m; v\rangle$ for the $SU(2)$-like basis functions normalized by (A6), we have

$$F_\pm |\epsilon, m\rangle = \mp i\left(\epsilon \frac{\partial^2}{\partial\epsilon^2} + (3 \pm 2m)\frac{\partial}{\partial\epsilon} + \frac{(m \pm 1)^2 - (j_0 \mp \sigma)^2}{\epsilon}\right) |\epsilon, m \pm 1\rangle,$$

$$N_3 |\epsilon, m\rangle = i\left(1 + \epsilon \frac{\partial}{\partial\epsilon}\right) |\epsilon, m\rangle, \quad \text{(A8)}$$

while the form of $E_\pm$ and M_3 remains unchanged. In the following we shall take $v = 1$ and denote the functions of the $SU(2)$ basis by $|j, m\rangle$. Introducing the raising operator $M_+ = M_1 + iM_2$, we see that

$$M_+ |j, j\rangle = 0.$$

On the other hand, the following relation is true:

$$M_+ = \tfrac{1}{2}i(E_+ - F_+),$$

which enables us to write

$$\left(\epsilon \frac{\partial^2}{\partial\epsilon^2} + (1 - 2j)\frac{\partial}{\partial\epsilon} + \frac{j^2 - (j_0 - \sigma)^2}{\epsilon} - \epsilon\right) \times \langle\epsilon, j \mid j, j\rangle = 0. \quad \text{(A9)}$$

Maneuvering similarly also with the lowering operator $M_- = M_1 - iM_2$, we can derive the following relation for the functions $\langle\epsilon, j \mid j, j\rangle$ and $\langle\epsilon, m \mid j, m\rangle$ (see Ref. 5):

$$\langle\epsilon, m \mid j, m\rangle = \left(-\frac{1}{2}\right)^{j-m}\left(\frac{(j + m)!}{(j - m)!\,(2j)!}\right)^{\frac{1}{2}} \times \epsilon^{-m} B^{j-m}_{j_0+\sigma}(\epsilon)\epsilon^j \langle\epsilon, j \mid j, j\rangle, \quad \text{(A10)}$$

where $B^{j-m}_{j_0+\sigma}(\epsilon)$ stands for the $(j - m)$th power of the

Bessel operator

$$B_{j_0+\sigma}(\epsilon) = \frac{d^2}{d\epsilon^2} + \frac{1}{\epsilon}\frac{d}{d\epsilon} - \frac{(j_0+\sigma)^2}{\epsilon^2} - 1. \quad \text{(A11)}$$

The differential equation system (A10), (A9) was first solved in Ref. 6 by a direct method. The same result is reproduced by the bilinear functional method in Sec. 3 of the present paper. For the reader's convenience we write down the counterpart of (3.4) corresponding to the normalization introduced by (A6) and (A7):

$$\langle \epsilon, m \mid j, m\rangle = (-1)^{j-m}2^{-m}\left(\frac{\Gamma(1+j-m)\Gamma(1+j-j_0)\Gamma(1+j-\sigma)}{\Gamma(1+j+m)\Gamma(1+j+j_0)\Gamma(1+j+\sigma)}\right)^{\frac{1}{2}}$$
$$\times \sum_{k=0}^{j-m}\left(-\frac{1}{2}\right)^k \frac{(j+m+k)!}{k!\,(j-m-k)!}\,\frac{1}{\Gamma(1+j_0+m+k)\Gamma(1-\sigma+m+k)}\,\epsilon^{k+m}K_{j_0-\sigma+k}(\epsilon). \quad \text{(A12)}$$

Now we turn to our main purpose, the calculation of the boost function $d^{j_0\sigma}_{jmj'}(\xi)$ in $SU(2)$ basis. We calculate the integral

$$d^{j_0\sigma}_{jmj'}(\xi) = e^{\xi}\int_0^\infty \tfrac{1}{2}\epsilon\, d\epsilon\langle j, m \mid \epsilon', m\rangle\langle \epsilon, m \mid j', m\rangle, \quad \epsilon' = \epsilon e^{\xi}, \quad \text{(A13)}$$

and proceed in the following manner: For definiteness, we assume $j \le j'$. For $\langle j, m \mid \epsilon', m\rangle$ we put into (A13) the expression (A12), and for $\langle \epsilon, m \mid j', m\rangle$ we use (A10). Then by partial integration we shift the Bessel operators from $\langle \epsilon, m \mid j', m\rangle$ to $\langle j, m \mid \epsilon, m\rangle$. This procedure leads to a sum the terms of which include a part of the type $B^k_c(y)[y^pK_q(y)]$, where c, p, and q are arbitrary complex numbers and k is a positive integer. It can be proved, by induction with respect to k and by using the relation

$$B_c(y)[y^pK_q(y)] = (p-q+c)(p-q-c)y^{p-2}K_q(y) - 2py^{p-1}K_{q-1}(y),$$

that the following relation is true:

$$B^k_c(y)[y^pK_q(y)] = 4^k\Gamma[1+\tfrac{1}{2}(p-q+c)]\Gamma[1+\tfrac{1}{2}(p-q-c)]$$
$$\times \sum_{r=0}^{k}\left(-\frac{1}{2}\right)^r \frac{k!}{r!\,(k-r)!}\,\frac{\Gamma(1+p-k+r)}{\Gamma[1+\frac{1}{2}(p-q+c)-(k-r)]\Gamma[1+\frac{1}{2}(p-q-c)-(k-r)]}$$
$$\times \frac{1}{\Gamma(1+p-k)}\,y^{p+r-2k}K_{q-r}(y).$$

Finally we get this result for $d^{j_0\sigma}_{jmj'}(\xi)$,

$$d^{j_0\sigma}_{jmj'}(\xi) = (-1)^{j'-j}\left(\frac{\Gamma(1+j-m)\Gamma(1+j+j_0)\Gamma(1+j+\sigma)}{\Gamma(1+j+m)\Gamma(1+j-j_0)\Gamma(1+j-\sigma)}\right)^{\frac{1}{2}}$$
$$\times \left(\frac{\Gamma(1+j'+m)\Gamma(1+j'-j_0)\Gamma(1+j'-\sigma)}{\Gamma(1+j'-m)\Gamma(1+j'+j_0)\Gamma(1+j'+\sigma)}\right)^{\frac{1}{2}}$$
$$\times \sum_{k=0}^{j-m}\sum_{n=0}^{k}(-1)^{k+n}\frac{\Gamma(1+j'-m)\Gamma(1+j+m+k)}{\Gamma(n+1)\Gamma(1+j'-m-n)\Gamma(1+j-m-n)\Gamma(1+k-n)}$$
$$\times \frac{\Gamma(1+j'+j_0+k-n)\Gamma(1+j'+\sigma+k-n)}{\Gamma(1+j_0+m+k)\Gamma(1+\sigma+m+k)\Gamma(2+2j'+k-n)}$$
$$\times (x^{-\frac{1}{2}})^{j_0-\sigma+m+2n+1}(x^{-1}-1)^{j'-m-n}F(1-j_0+j', 1+j'-\sigma; 2+2j'+k-n; 1-x), \quad \text{(A14)}$$

where $x = \exp(-2\xi)$. For practical purposes, it is useful to write the expression (A14) in another form which utilizes the fact that the hypergeometric function is of a special degenerate type:

$$\begin{aligned} d^{j_0\sigma}_{jmj'}(\xi) &= (-1)^{j'-j}\frac{\Gamma(1-j_0+\sigma)\Gamma(1+j_0-\sigma)}{j_0+\sigma} \\ &\times \left(\frac{\Gamma(1+j-m)\Gamma(1+j+j_0)\Gamma(1+j'+m)\Gamma(1+j'-j_0)}{\Gamma(1+j+m)\Gamma(1+j-j_0)\Gamma(1+j'-m)\Gamma(1+j'+j_0)}\right)^{\frac{1}{2}} \\ &\times [\Gamma(1+j-\sigma)\Gamma(1+j+\sigma)\Gamma(1+j'-\sigma)\Gamma(1+j'+\sigma)]^{-\frac{1}{2}} \\ &\times \sum_{k=0}^{j-m}\sum_{n=0}^{k}(-1)^{k+n}\frac{\Gamma(1+j'-m)\Gamma(1+j+m+k)}{\Gamma(n+1)\Gamma(1+j'-m-n)\Gamma(1+j-m-k)\Gamma(1+k-n)} \\ &\times \frac{\Gamma(1+j'+j_0+k-n)\Gamma(1+\sigma+j)}{\Gamma(1+j_0+m+k)\Gamma(1+\sigma+m+k)}\left[x^{\frac{1}{2}(m-j_0+\sigma+1)}(1-x)^{j_0-m-1-n}P^{(-j_0+\sigma;-j_0-\sigma-k+n)}_{j'+j_0+k-n}\left(\frac{1+x}{1-x}\right)\right. \\ &\left. - x^{\frac{1}{2}(m+j_0-\sigma+1)}(1-x)^{-j_0-m-1-k}P^{(j_0-\sigma;j_0+\sigma+k-n)}_{j'-j_0}\left(\frac{1+x}{1-x}\right)\right]. \end{aligned}$$

The $P^{(\alpha,\beta)}_n(y)$ symbols stand here for Jacobi polynomials. A similar formula has been known only in the special case $m = j$. A remarkable feature of these formulas is that the number of terms in the sum depends only on the lesser of j and j'.

[1] K.Szegö and K. Tóth, J. Maths, Phys. **12**, 846 (1971).
[2] A. Sciarrino and M. Toller, J. Math. Phys. **8**, 1252 (1967).
[3] S. Ström, Arkiv Fysik **34**, 215 (1967).
[4] M. A. Liberman and A. A. Makarov, Dubna Preprint P2-4067, 1968 (in Russian).
[5] S. J. Chang and L. O'Raifeartaigh, J. Math. Phys. **10**, 21 (1969).
[6] K. Tóth, ICTP, Internal Report No. IC/69/125, 1969.
[7] R. Delbourgo *et al.*, Nuovo Cimento **52A**, 1254 (1967).
[8] I. M. Gel'fand *et al.*, *Generalized Functions* (Academic, New York, 1968), Vol. 5.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Density of States

Platon C. Deliyannis
Department of Mathematics, Illinois Institute of Technology, Chicago, Illinois 60616

(Received 21 August 1970)

A characterization of the weak limits of states is given, leading to an extension of a theorem of Gudder and the verification of a conjecture by Mackey

In this paper we are concerned with the weak closure of the set of all states of a logic. Gudder has proved a related theorem[1] under the hypothesis that the logic is directed, i.e., that any two partitions of the unit element on the logic into pairwise disjoint elements admit a common refinement. This, however, implies that the logic is Boolean, and so the range of applicability of this theorem is reduced. We shall establish here under more general conditions that the states are dense in the set of all expectation functionals on the bounded observables, so that we shall obtain as a corollary the conclusion of Gudder's theorem.

We shall be working within Mackey's system. Our logic $\mathfrak{L}$ is a set with a partial order $\leq$, smallest and largest elements O and I respectively, and an involution $Q \rightarrow Q'$ such that $P \leq Q$ iff $Q' \leq P'$, $P \vee P' = I$, $P \wedge P' = O$, $O' = I$, and $I' = O$. Further, if $P \leq Q$, we assume the existence of a unique $R \leq P'$ such that $Q = P \vee R$, and finally assume that any sequence of pairwise disjoint Q_j (i.e., such that $Q_i \leq Q'_k$ for $i \neq k$) admits a least upper bound $\sum Q_j$. An observable is a σ-homomorphism A of the Borel sets in the reals to $\mathfrak{L}$, and for any Borel function f the observable fA is the map $A \circ f^{-1}$. The spectrum σA of A is the complement of the largest open set whose characteristic function acting on A produces O. The elements of $\mathfrak{L}$ are naturally identified to the idempotent observables. A state is a countably additive (on disjoint elements) probability measure on the logic. We shall define a quasistate as a finitely additive (on disjoint elements) probability measure on $\mathfrak{L}$. Clearly the weak limit of any net of states will be a quasistate. We shall write $\mathcal{M}$ for the set of all states on $\mathfrak{L}$, and assume the logic to be full.

utilizes the fact that the hypergeometric function is of a special degenerate type:

$$
\begin{aligned}
d^{j_0\sigma}_{jmj'}(\xi) = {} & (-1)^{j'-j}\frac{\Gamma(1-j_0+\sigma)\Gamma(1+j_0-\sigma)}{j_0+\sigma} \\
& \times \left(\frac{\Gamma(1+j-m)\Gamma(1+j+j_0)\Gamma(1+j'+m)\Gamma(1+j'-j_0)}{\Gamma(1+j+m)\Gamma(1+j-j_0)\Gamma(1+j'-m)\Gamma(1+j'+j_0)}\right)^{\frac{1}{2}} \\
& \times [\Gamma(1+j-\sigma)\Gamma(1+j+\sigma)\Gamma(1+j'-\sigma)\Gamma(1+j'+\sigma)]^{-\frac{1}{2}} \\
& \times \sum_{k=0}^{j-m}\sum_{n=0}^{k}(-1)^{k+n}\frac{\Gamma(1+j'-m)\Gamma(1+j+m+k)}{\Gamma(n+1)\Gamma(1+j'-m-n)\Gamma(1+j-m-k)\Gamma(1+k-n)} \\
& \times \frac{\Gamma(1+j'+j_0+k-n)\Gamma(1+\sigma+j)}{\Gamma(1+j_0+m+k)\Gamma(1+\sigma+m+k)}\Bigg[x^{\frac{1}{2}(m-j_0+\sigma+1)}(1-x)^{j_0-m-1-n}P^{(-j_0+\sigma;-j_0-\sigma-k+n)}_{j'+j_0+k-n}\left(\frac{1+x}{1-x}\right) \\
& - x^{\frac{1}{2}(m+j_0-\sigma+1)}(1-x)^{-j_0-m-1-k}P^{(j_0-\sigma;j_0+\sigma+k-n)}_{j'-j_0}\left(\frac{1+x}{1-x}\right)\Bigg].
\end{aligned}
$$

The $P_n^{(\alpha,\beta)}(y)$ symbols stand here for Jacobi polynomials. A similar formula has been known only in the special case $m = j$. A remarkable feature of these formulas is that the number of terms in the sum depends only on the lesser of j and j'.

[1] K.Szegö and K. Tóth, J. Maths, Phys. **12**, 846 (1971).
[2] A. Sciarrino and M. Toller, J. Math. Phys. **8**, 1252 (1967).
[3] S. Ström, Arkiv Fysik **34**, 215 (1967).
[4] M. A. Liberman and A. A. Makarov, Dubna Preprint P2-4067, 1968 (in Russian).
[5] S. J. Chang and L. O'Raifeartaigh, J. Math. Phys. **10**, 21 (1969).
[6] K. Tóth, ICTP, Internal Report No. IC/69/125, 1969.
[7] R. Delbourgo *et al.*, Nuovo Cimento **52A**, 1254 (1967).
[8] I. M. Gel'fand *et al.*, *Generalized Functions* (Academic, New York, 1968), Vol. 5.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Density of States

Platon C. Deliyannis

Department of Mathematics, Illinois Institute of Technology, Chicago, Illinois 60616

(Received 21 August 1970)

A characterization of the weak limits of states is given, leading to an extension of a theorem of Gudder and the verification of a conjecture by Mackey

In this paper we are concerned with the weak closure of the set of all states of a logic. Gudder has proved a related theorem[1] under the hypothesis that the logic is directed, i.e., that any two partitions of the unit element on the logic into pairwise disjoint elements admit a common refinement. This, however, implies that the logic is Boolean, and so the range of applicability of this theorem is reduced. We shall establish here under more general conditions that the states are dense in the set of all expectation functionals on the bounded observables, so that we shall obtain as a corollary the conclusion of Gudder's theorem.

We shall be working within Mackey's system. Our logic $\mathcal{L}$ is a set with a partial order $\leq$, smallest and largest elements O and I respectively, and an involution $Q \to Q'$ such that $P \leq Q$ iff $Q' \leq P'$, $P \vee P' = I$, $P \wedge P' = O$, $O' = I$, and $I' = O$. Further, if $P \leq Q$, we assume the existence of a unique $R \leq P'$ such that $Q = P \vee R$, and finally assume that any sequence of pairwise disjoint Q_j (i.e., such that $Q_i \leq Q_k'$ for $i \neq k$) admits a least upper bound $\sum Q_j$. An observable is a σ-homomorphism A of the Borel sets in the reals to $\mathcal{L}$, and for any Borel function f the observable fA is the map $A \circ f^{-1}$. The spectrum σA of A is the complement of the largest open set whose characteristic function acting on A produces O. The elements of $\mathcal{L}$ are naturally identified to the idempotent observables. A state is a countably additive (on disjoint elements) probability measure on the logic. We shall define a quasistate as a finitely additive (on disjoint elements) probability measure on $\mathcal{L}$. Clearly the weak limit of any net of states will be a quasistate. We shall write $\mathcal{M}$ for the set of all states on $\mathcal{L}$, and assume the logic to be full.

Lemma 1: For each $Q \in \mathfrak{L}$ let $\hat{Q}$ be the function $m \to mQ$ on $\mathcal{M}$. Then the partial order on $\mathfrak{L}$ is transferred to pointwise partial order by the map $\hat{}$, $\widehat{Q'} = 1 - \hat{Q}$ and $Q = \sum Q_j$ iff $\hat{Q}$ is the pointwise sum of the $\hat{Q}_j$.

Since the proof is straightforward, we shall omit it.

Now consider the sets $\{m \mid mQ \leq a\}$, for $Q \in \mathfrak{L}$ and a real, and the Boolean σ-algebra $\mathfrak{X}$ of sets that they generate in $\mathcal{M}$. Every function $\hat{Q}$ is measurable relative to $\mathfrak{X}$ and, if μ is a probability measure on $\mathfrak{X}$, then $Q \to \int \hat{Q}\, d\mu$ is a state of $\mathfrak{L}$, because of the last relation in Lemma 1.

One final remark before we prove our first theorem. Suppose that $\mathfrak{X}$ is a Boolean σ-algebra of sets and $\mathfrak{Y}$ a finite Boolean subalgebra. Given a (finitely additive) probability measure λ on $\mathfrak{Y}$, there exists a countably additive probability measure μ on $\mathfrak{X}$ which agrees with λ on $\mathfrak{Y}$. To construct μ, we consider all nonempty intersections $T = S_1 \cap S_2 \cap \cdots \cap S_n$ with $S_i \in \mathfrak{Y}$; they are pairwise disjoint, finitely many, and every element of $\mathfrak{Y}$ is the union of such intersections. Choose a point t in each T, and for $S \in \mathfrak{X}$ let μS be the sum of all λT for the various sets T for which $t \in S$; obviously μ is a countably additive probability measure on $\mathfrak{X}$, and for $S \in \mathfrak{Y}$ we see that $t \in S$ iff $T \subseteq S$, which implies $\mu S = \lambda S$.

Theorem 1: A quasistate q is the weak limit of states iff, for any real $c_1, c_2, \cdots, c_r$ and any $Q_1, Q_2, \cdots, Q_r \in \mathfrak{L}$, the condition $\sum_{i=1}^{r} c_i mQ_i \geq 0$ for all $m \in \mathcal{M}$ implies $\sum_{i=1}^{r} c_i qQ_i \geq 0$.

Proof: Necessity is obvious. Assume the condition to hold and consider $Q_1, Q_2, \cdots, Q_r \in \mathfrak{L}$ and an $\epsilon > 0$; we must produce a state m such that $|qQ_i - mQ_i| < \epsilon$ for $i = 1, 2, \cdots, r$. We may assume I to be one of the Q_i, and consider the vector space of functions on $\mathcal{M}$ spanned by the $\hat{Q}_i$. Our condition above implies that the map $\sum c_i\hat{Q}_i \to \sum c_i qQ_i$ is well-defined, linear, positive and normalized to 1 on the function $\hat{I}$. Thus[2] it extends to such a functional on the vector space of all bounded functions on $\mathcal{M}$, measurable relative to $\mathfrak{X}$. Therefore, there exists a finitely additive probability measure λ on $\mathfrak{X}$ which generates this functional by integration. Now consider the finite Boolean subalgebra $\mathfrak{X}_n$ of $\mathfrak{X}$ generated by the sets $\{m \mid (k-1)/n < mQ_i \leq k/n\}$, $k = 1, 2, \cdots, n$ and $i = 1, 2, \cdots, r$. We can find a countably additive measure μ_n on $\mathfrak{X}$ which agrees with λ on $\mathfrak{X}_n$, and we see that $|qQ_i - \int \hat{Q}_i\, d\mu_n| \leq n^{-1}$ since $qQ_i = \int \hat{Q}_i\, d\lambda$. As the map $Q \to \int \hat{Q}\, d\mu_n$ is a state, all we must do is to choose n so that $n^{-1} < \epsilon$.

For any quasistate q and any bounded observable A we can define the expectation value of A in q as follows. Consider first the case of a finite spectrum $\{a_1, a_2, \cdots, a_n\}$, so that $A = \sum a_i A(\{a_i\})$; then the expectation qA will be defined as $\sum a_i q(A(\{a_i\}))$, and it is not hard to show that if we write A as $\sum c_j Q_j$ with Q_j pairwise disjoint, then $qA = \sum c_j qQ_j$. We shall have $|qA| \leq \|A\|$, where for any observable we define $\|A\|$ to be $\sup \{|\lambda| \mid \lambda \in \sigma A\}$. Now, if the spectrum of A is bounded, we can find a sequence of observables A_n which commute with A and have finite spectrum, while $\|A - A_n\|$ tends to zero; then the limit of qA_n exists and is independent of the sequence chosen, so that we can define the expectation qA of A in q as this limit. It is not hard to see that this expectation functional is linear on commuting observables and positive on positive observables while $|qA| \leq \|A\|$ for all A.

Lemma 2: Let q be the weak limit of the states m_j, i.e., $\lim_j m_j Q = qQ$ for all $Q \in \mathfrak{L}$. Then for any bounded observable A we have $qA = \lim_j m_j A$.

Proof: Given $\epsilon > 0$, there exist reals a_i and pairwise disjoint $Q_i \in \mathfrak{L}$ such that A and $\sum_{i=1}^{r} a_i Q_i$ commute while $\|A - \sum_i a_i Q_i\| < \epsilon/3$. Also there exists a $j(\epsilon)$ such that for $j > j(\epsilon)$ we have $|qQ_i - m_j Q_i| < \epsilon/3 \sum_i |a_i|$ for all i. Thus we obtain

$$|\sum a_i qQ_i - \sum a_i m_j Q_i| < \epsilon/3,$$
$$|m_j A - \sum_i a_i m_j Q_i| < \epsilon/3$$

for all $j > j(\epsilon)$, and $|qA - \sum_i a_i qQ_i| < \epsilon/3$. So we have $|qA - m_j A| < \epsilon$ for $j > j(\epsilon)$.

Theorem 2: Suppose that every pair of bounded observables admits a unique sum, in the sense that given A and B there exists a unique C such that $mA + mB = mC$ for all $m \in \mathcal{M}$. Then the weak closure of $\mathcal{M}$ consists of those quasistates for which the associated expectation functional is additive on the bounded observables.

Proof: Since the states are additive, it is clear that their weak limits will also be, because of Lemma 2. Now suppose q is an additive expectation functional, so that it is actually linear, and that $\sum c_i mQ_i \geq 0$ for all $m \in \mathcal{M}$. Then $m(\sum c_i Q_i) \geq 0$ for all m which means that the observable $\sum c_i Q_i$ has nonnegative spectrum.[3] Thus $q(\sum c_i Q_i) \geq 0$ and, as q is linear, we have $\sum c_i qQ_i \geq 0$. By Theorem 1, q is the limit of states.

Since in the case of a Boolean σ-algebra sums are defined, we obtain at once Gudder's theorem[1]: The set of states is weakly closed iff every bounded positive functional on the bounded observables is a multiple of an expectation functional corresponding to a state.

We also observe that Theorem 2 supports a conjecture of Mackey[4] to the effect that expectation functionals in Segal's sense should be limits of states.

Finally we should remark that it is not clear under what *a priori* conditions on $\mathfrak{L}$ the expectation functional of a quasistate is additive on noncommuting observables.

[1] S. P. Gudder, Can. J. Math. **20**, 1276 (1968).
[2] J. L. Kelley and I. Namioka, *Linear Topological Spaces* (Van Nostrand, New York, 1963).
[3] S. P. Gudder, Trans. Am. Math. Soc. **119**, 428 (1965).
[4] G. W. Mackey, *The Mathematical Foundations of Quantum Mechanics* (Benjamin, New York, 1963), p. 70.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Strong Forces

J. T. Anderson*

Mathematics Department, University of Maryland, Heidelberg, Germany

(Received 26 May 1970)

A new analytic function method is utilized to study self-consistent vector interactions in the nonrelativistic case. A mass formula is obtained as a function of the current and charge densities and found to yield particle radii near the Compton lengths for a number of elementary particles. With spin-current interactions included, the mass formula gives the baryon masses within 10% at densities corresponding to the Compton lengths. The results lead to definite conclusions regarding the nature of strong forces.

INTRODUCTION

In earlier studies of the analytic properties of functions of several complex variables, the equivalence of analytic functions in complex variable representations with conserved functions in real variable representations was shown[1] and utilized to study the self-consistent scalar interactions of a particle gas at arbitrary coupling strength in the nonrelativistic case[2] and also relativistically for a spinor and a scalar field.[1] We consider now a nonrelativistic particle gas interacting by means of a spin-dependent vector field and a scalar field, utilizing the same method developed for scalar interactions. Relativistic self-consistent pseudoscalar and pseudovector interactions have been studied by Nambu,[3] who derived the condition for existence of massive nucleon states.

Again we make use of transformations of the form

$$\begin{aligned} H'(z_\alpha) &= e^{z_\alpha}H(z_\alpha)e^{-z_\alpha} \\ &= H(z_\alpha) + [z_\alpha, H] + \tfrac{1}{2}[z_\alpha, [z_\alpha, H]] + \cdots \\ &= H(z_\alpha) + \sigma_{z_\alpha}\frac{\partial H}{\partial z_\alpha^*} + \frac{\sigma_z^2}{2}\frac{\partial^2 H}{\partial z_\alpha^{*2}} + \cdots, \end{aligned} \tag{1}$$

where $z_\alpha = z_{1\alpha} + iz_{2\alpha}$, $z_1, z_2 > 0$, is a sum of normalized complex scalars and the σ_{z_α} are constants. The third line follows from the second after applying the fundamental quantum conditions to the classical brackets.[1] We look for eigenvalues z'_α for which $H' = H$, that is, z'_α satisfying

$$\begin{aligned} \lim_{z_\alpha \to z_\alpha'} [z_\alpha, H] &= \sigma_{z_\alpha'}\frac{\partial H}{\partial z_\alpha'^*} \\ &= 0, \end{aligned} \tag{2}$$

and it can be shown that if (2) holds, then all the higher-order brackets in (1) vanish as well. For real variables, the z'_α satisfying (2) are just the constants of motion of the conserved Hamiltonian $H(z'_\alpha)$. But Eq. (2) generates the Cauchy–Riemann equations in each variable so that, provided H satisfies sufficient conditions in each variable, the z'_α define the analytic regions of H and the analytic function for our complex variable representation corresponds to a conserved function for real variables. It is now a straightforward matter to impose momentum and energy conservation on H by writing down equations such as (2) for the eigenvalues of the momentum and energy.

However, the z'_α obtained from (2) are defined on the first quadrants so that $H(z'_\alpha)$ is conserved only on the first quadrants. From symmetry considerations corresponding to ordinary spatial reflections and time reversal in configuration space, it is necessary that the reflected function $H^\dagger(z'_\alpha)$ also be analytic and satisfy equations such as (2) on the other quadrants. With this requirement the Hamiltonian becomes a constant, or analytic on the entire plane provided that there are no divergences or oscillations at infinity and if H and $H^\dagger$ satisfy sufficient conditions for analyticity in each variable. It can be shown that the sufficient conditions are satisfied if H is independent of angle or, otherwise, if H satisfies additional constraints which require that H be a constant on the entire planes.

Our procedure now is to obtain equations such as (2) for all the conserved variables, that is, energy, momentum, and particle number in an interacting system, and solve them simultaneously for the new eigenvalues.

functional on the bounded observables is a multiple of an expectation functional corresponding to a state.

We also observe that Theorem 2 supports a conjecture of Mackey[4] to the effect that expectation functionals in Segal's sense should be limits of states.

Finally we should remark that it is not clear under what *a priori* conditions on $\mathfrak{L}$ the expectation functional of a quasistate is additive on noncommuting observables.

[1] S. P. Gudder, Can. J. Math. **20**, 1276 (1968).
[2] J. L. Kelley and I. Namioka, *Linear Topological Spaces* (Van Nostrand, New York, 1963).
[3] S. P. Gudder, Trans. Am. Math. Soc. **119**, 428 (1965).
[4] G. W. Mackey, *The Mathematical Foundations of Quantum Mechanics* (Benjamin, New York, 1963), p. 70.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Strong Forces

J. T. ANDERSON*
Mathematics Department, University of Maryland, Heidelberg, Germany

(Received 26 May 1970)

A new analytic function method is utilized to study self-consistent vector interactions in the nonrelativistic case. A mass formula is obtained as a function of the current and charge densities and found to yield particle radii near the Compton lengths for a number of elementary particles. With spin-current interactions included, the mass formula gives the baryon masses within 10% at densities corresponding to the Compton lengths. The results lead to definite conclusions regarding the nature of strong forces.

INTRODUCTION

In earlier studies of the analytic properties of functions of several complex variables, the equivalence of analytic functions in complex variable representations with conserved functions in real variable representations was shown[1] and utilized to study the self-consistent scalar interactions of a particle gas at arbitrary coupling strength in the nonrelativistic case[2] and also relativistically for a spinor and a scalar field.[1] We consider now a nonrelativistic particle gas interacting by means of a spin-dependent vector field and a scalar field, utilizing the same method developed for scalar interactions. Relativistic self-consistent pseudoscalar and pseudovector interactions have been studied by Nambu,[3] who derived the condition for existence of massive nucleon states.

Again we make use of transformations of the form

$$\begin{aligned} H'(z_\alpha) &= e^{z_\alpha} H(z_\alpha) e^{-z_\alpha} \\ &= H(z_\alpha) + [z_\alpha, H] + \tfrac{1}{2}[z_\alpha, [z_\alpha, H]] + \cdots \\ &= H(z_\alpha) + \sigma_{z_\alpha} \frac{\partial H}{\partial z_\alpha^*} + \frac{\sigma_z^2}{2} \frac{\partial^2 H}{\partial z_\alpha^{*2}} + \cdots, \end{aligned} \tag{1}$$

where $z_\alpha = z_{1\alpha} + i z_{2\alpha}$, $z_1, z_2 > 0$, is a sum of normalized complex scalars and the σ_{z_α} are constants. The third line follows from the second after applying the fundamental quantum conditions to the classical brackets.[1] We look for eigenvalues z'_α for which $H' = H$, that is, z'_α satisfying

$$\begin{aligned} \lim_{z_\alpha \to z_\alpha'} [z_\alpha, H] &= \sigma_{z_\alpha'} \frac{\partial H}{\partial z_\alpha'^*} \\ &= 0, \end{aligned} \tag{2}$$

and it can be shown that if (2) holds, then all the higher-order brackets in (1) vanish as well. For real variables, the z'_α satisfying (2) are just the constants of motion of the conserved Hamiltonian $H(z'_\alpha)$. But Eq. (2) generates the Cauchy-Riemann equations in each variable so that, provided H satisfies sufficient conditions in each variable, the z'_α define the analytic regions of H and the analytic function for our complex variable representation corresponds to a conserved function for real variables. It is now a straightforward matter to impose momentum and energy conservation on H by writing down equations such as (2) for the eigenvalues of the momentum and energy.

However, the z'_α obtained from (2) are defined on the first quadrants so that $H(z'_\alpha)$ is conserved only on the first quadrants. From symmetry considerations corresponding to ordinary spatial reflections and time reversal in configuration space, it is necessary that the reflected function $H^\dagger(z'_\alpha)$ also be analytic and satisfy equations such as (2) on the other quadrants. With this requirement the Hamiltonian becomes a constant, or analytic on the entire plane provided that there are no divergences or oscillations at infinity and if H and $H^\dagger$ satisfy sufficient conditions for analyticity in each variable. It can be shown that the sufficient conditions are satisfied if H is independent of angle or, otherwise, if H satisfies additional constraints which require that H be a constant on the entire planes.

Our procedure now is to obtain equations such as (2) for all the conserved variables, that is, energy, momentum, and particle number in an interacting system, and solve them simultaneously for the new eigenvalues.

1. DERIVATION OF THE EIGENVALUE EQUATIONS

We consider matrix elements of the Pauli Hamiltonian

$$H(k, E) = \langle k, E|\, H_0 - i\mu\boldsymbol{\sigma}\cdot\frac{\partial}{\partial \mathbf{x}} \times \mathbf{A}(\mathbf{x}, t) + \frac{r_0}{2}A^2(\mathbf{x}, t) - \lambda\phi(\mathbf{x}, t)\,|p\rangle \quad (3)$$

between initial momentum state $\mathbf{p}$ and final state E, $\mathbf{k} = \mathbf{p} + \mathbf{q}$, where k_α and E are complex variables, λ is the scalar coupling strength, and r_0 is the classical radius. The terms $(\partial/\partial\mathbf{x})\cdot\mathbf{A} + \mathbf{A}\cdot(\partial/\partial\mathbf{x})$ are eliminated from (3) by a suitable choice of boundary and initial conditions. The interactions are assumed to occur via scalar and vector fields defined by

$$\langle 0|\,\Box\phi(\mathbf{x}, t)\,|q, E\rangle = \rho(q, E) \quad (4)$$

and

$$\langle 0|\,\Box\mathbf{A}(\mathbf{x}, t)\,|q, E\rangle = \mathbf{j}(q, E), \quad (5)$$

where ρ and $\mathbf{j}$ are the charge density and current density, respectively. Noting that $\partial k/\partial k^* = 0$, we see that H satisfies Cauchy–Riemann equations in k_α provided that

$$\frac{d\rho}{dk_\alpha'^*} = 0 \quad\text{and}\quad \frac{dj_\beta}{dk_\alpha'^*} = 0, \quad (6)$$

and, since ρ and j_β are defined on the first quadrants, it is evident (6) are satisfied; there are similar conditions for E. The sufficient conditions require that H be independent of the angles defined by $\tan\theta_\alpha = \delta k_{2\alpha}/\delta k_{1\alpha}$, $\tan\phi = \delta E_2/\delta E_1$, and $\tan\gamma = \delta\rho_2/\delta\rho_1$ and that H be a constant on the entire planes, satisfying $dH/dk_\alpha' = dH/dE' = dH/d\rho' = 0$. Thus the sufficient conditions determine the eigenvalues k', E', and ρ' without "continuations," but the Hamiltonian so defined is not necessarily an invariant under symmetry transformations $k_\alpha \to -k_\alpha$ and $E \to -E$, corresponding to spatial and time reflections in configuration space and to rotations in the planes. The analytic "continuations" are associated with these transformations, and, in order that H be invariant on the entire planes, it is necessary that equations such as (2) be obtained on the entire planes. We therefore impose energy, momentum, and particle number conservation by means of the transformations

$$\begin{aligned} e^{E^*/E_0}H^\dagger e^{-E^*/E_0} &= H^\dagger + E_0^{-1}[E^*, H^\dagger] + \cdots \\ &= H^\dagger + \sigma_E^*\frac{dH^\dagger}{dE} + \cdots \\ &= H^\dagger(k_\alpha, E'), \end{aligned} \quad (7)$$

$$\begin{aligned} e^{k_\alpha^*/k_0}H^\dagger e^{-k_\alpha^*/k_0} &= H^\dagger + k_0^{-1}[k_\alpha^*, H^\dagger] + \cdots \\ &= H^\dagger + \sigma_{k_\alpha}^*\frac{dH^\dagger}{dk_\alpha} + \cdots \\ &= H^\dagger(k_\alpha', E), \end{aligned} \quad (8)$$

and

$$\begin{aligned} e^{\rho^*/\rho_0}H^\dagger e^{-\rho^*/\rho_0} &= H^\dagger + \rho_0^{-1}[\rho^*, H^\dagger] + \cdots \\ &= H^\dagger + \sigma_\rho^*\frac{dH^\dagger}{d\rho} \\ &= H^\dagger|_{\rho\to\rho'}. \end{aligned} \quad (9)$$

Here σ_E^*, $\sigma_{k_\alpha}^*$, and σ_ρ^* are constants, $H^\dagger$ is the "continuation" of H onto the third quadrants, and the third lines are the conserved functions in the limits $k_\alpha \to k_\alpha'$, $E \to E'$, and $\rho \to \rho'$, where k_α', E', and ρ' are the solutions of eigenvalue equations such as (2). Finally, we require that all conservation laws hold simultaneously so that

$$e^{\rho^*/\rho_0}e^{k_\alpha^*/k_0}e^{E^*/E_0}H^\dagger e^{-E^*/E_0}e^{-k_\alpha^*/k_0}e^{-\rho^*/\rho_0} = H^\dagger(k_\alpha', E')|_{\rho\to\rho'}, \quad (10)$$

which simply requires that (7)–(9) yield simultaneous eigenvalues in the limits $k_\alpha \to k_\alpha'$, $E \to E'$, and $\rho \to \rho'$. It can be shown that if (7)–(9) hold, then all brackets involving two or more different variables also vanish so that the eigenvalues k', E', and ρ' commute with each other as well as with $H^\dagger$.

It is evident from (2) that (7)–(9) are unitary transformations provided that there exist eigenvalues satisfying

$$\begin{aligned} [E^{*\prime}, H^\dagger] &= \sigma_{E'}^*\lim_{E\to E'}\frac{dH^\dagger}{dE} \\ &= 0, \end{aligned} \quad (11)$$

$$\begin{aligned} [k_\alpha'^*, H^\dagger] &= \sigma_{k_\alpha'}^*\lim_{k_\alpha\to k_\alpha'}\frac{dH^\dagger}{dk_\alpha} \\ &= 0, \end{aligned} \quad (12)$$

and

$$\begin{aligned} [\rho^{*\prime}, H^\dagger] &= \sigma_{\rho'}^*\lim_{\rho\to\rho'}\frac{dH^\dagger}{d\rho} \\ &= 0. \end{aligned} \quad (13)$$

The "continuations" giving $H \to H^\dagger$ are carried out by rotations through π on each plane carrying $k_\alpha \to -k_\alpha$ and $E \to -E$, and it is evident by inspection that $H^\dagger(-k, -E) \to H(k, E)$ provided that

$$\rho(-k, -E) \to \rho(k, E)$$

and

$$j_\alpha(-k, -E) \to -j_\alpha(k, E),$$

which are satisfied if the charge and current densities are functions of the kinetic energies of occupied states in the usual manner.

If the current and density are analytic on the entire planes, then

$$\frac{d\rho^\dagger}{dk_\alpha} = \frac{dj_\beta^\dagger}{dk_\alpha} = \frac{d\rho^\dagger}{dE} = \frac{dj_\beta^\dagger}{dE} = 0, \quad (14)$$

and with (5) we obtain

$$\frac{\partial\mathbf{A}}{\partial\mathbf{k}'} = 2\mathbf{q}'\mathbf{j}\Gamma^2, \quad \frac{\partial\mathbf{A}}{\partial E} = -\frac{2E}{c^2}\mathbf{j}\Gamma^2 \quad (15)$$

and with (4) also

$$\frac{\partial\phi}{\partial\mathbf{k}'} = 2\mathbf{q}'\rho\Gamma^2, \quad \frac{\partial\phi}{\partial E'} = -\frac{2E}{c^2}\rho\Gamma^2. \tag{16}$$

Combining equations (11), (12), (15), and (16) gives

$$\begin{aligned}(\sigma^*_{E'})^{-1}[E'^*, H^\dagger] &= -2E'\Gamma^2(\mu\boldsymbol{\sigma}\cdot\mathbf{k}'\times\mathbf{j} + r_0\mathbf{A}\cdot\mathbf{j} - \lambda\rho)\\ &= 0\end{aligned} \tag{17}$$

and

$$\begin{aligned}(\sigma^*_{k'})^{-1}[\mathbf{k}'^*, H^\dagger] &= (\mathbf{k}'/m) - \mu\boldsymbol{\sigma}\times\mathbf{A} + 2\Gamma^2\\ &\quad\times(\mu\boldsymbol{\sigma}\cdot\mathbf{k}'\times\mathbf{j}\mathbf{q} + r_0\mathbf{A}\cdot\mathbf{j}\mathbf{q} - \lambda\rho\mathbf{q})\\ &= 0.\end{aligned} \tag{18}$$

Now evaluating (18) in the limit $E \to E'$ gives for the momentum eigenvalues

$$\mathbf{k}' = (e/c)\boldsymbol{\sigma}\times\mathbf{A}, \tag{19}$$

and (17) gives the renormalized propagator

$$\Gamma = (r_0\mathbf{j}\cdot\mathbf{j})^{-1}(\lambda\rho - \mu\boldsymbol{\sigma}\cdot\mathbf{k}'\times\mathbf{j}), \tag{20}$$

where $\mathbf{j}\cdot\mathbf{j}$ represents the scalar product of three complex variables. With (20) we obtain for the excitation energies

$$\frac{E'^2}{mc^2} = \frac{1}{m}(k'_\alpha - p_\alpha)^2 + \frac{e^2}{mc^2}\frac{\mathbf{j}\cdot\mathbf{j}}{\lambda\rho - \mu\boldsymbol{\sigma}\cdot\mathbf{k}'\times\mathbf{j}} \tag{21}$$

and for the fields

$$\begin{aligned}\phi &= \phi_c + \phi_T\\ &= \frac{\lambda}{r_0}\frac{\rho^2}{\mathbf{j}\cdot\mathbf{j}} - \frac{e}{\alpha}\frac{\rho}{\mathbf{j}\cdot\mathbf{j}}\boldsymbol{\sigma}\cdot\mathbf{k}'\times\mathbf{j}\end{aligned} \tag{22}$$

and

$$\begin{aligned}\mathbf{A} &= \mathbf{A}_c + \mathbf{A}_T\\ &= \frac{\lambda}{r_0}\frac{\rho\mathbf{j}}{\mathbf{j}\cdot\mathbf{j}} - \frac{e}{\alpha}\frac{\mathbf{j}}{\mathbf{j}\cdot\mathbf{j}}\boldsymbol{\sigma}\cdot\mathbf{k}'\times\mathbf{j}\end{aligned} \tag{23}$$

so that

$$\rho\mathbf{A} = \mathbf{j}\phi \tag{24}$$

holds for both terms in (22) and (23).

In (21) the first term on the rhs is the new kinetic energy which reduces to the initial value in the absence of excited states, and the second term represents a mass eigenvalue introduced by the interactions in which the $e^2\mathbf{j}\cdot\mathbf{j}/\lambda\rho m^2c^2$ part is related to the terms ϕ_c and $\mathbf{A}_c$ in (22) and (23). The first terms in (22) and (23) evidently represent central fields proportional to λ/r_0 multiplied by kinematical factors, but the second terms represent spin and momentum-dependent fields which obviously predominate at sufficiently high momentum, thus at $\lambda = e$, at $N_j\boldsymbol{\sigma}\times\mathbf{k}\cdot\mathbf{v}/N_\rho > mc^2$, where N_j and N_ρ are the current number density and scalar number density, respectively. Note that the coupling of the noncentral field is e/α and that at $\rho/|\mathbf{j}| = 1$ the noncentral and central forces are equal at $|\mathbf{k}|/\alpha \simeq r_0^{-1}$ and the noncentral forces predominate below the Compton length.

The mass eigenvalues in (21) are related to the fields ϕ and $\mathbf{A}$ by

$$m'^2c^2 = \frac{r_0}{\lambda}\frac{\mathbf{j}\cdot\mathbf{j}}{\rho} \tag{25}$$

for the central fields and by

$$m''^2c^2 = \mathbf{j}\cdot\mathbf{j}/\mathbf{j}\cdot\mathbf{A} \tag{26}$$

for the total fields. Now comparing (25) and (26) and remembering that, at sufficiently high density, $\mu\boldsymbol{\sigma}\cdot\mathbf{k}'\times\mathbf{j}/\lambda\rho \gg 1$, we see that the central field contributes the larger part to the mass in this region. This suggests that for $|\mathbf{k}|/\alpha \geq r_0^{-1}$ the central force becomes the inertial force and the repulsive interaction is due to the noncentral force. If the signs of $\mathbf{k}$, $\mathbf{j}$, the spin, and angle dependences are such that ϕ_c and ϕ_T have opposite signs and ϕ_c is attractive [and it can be shown that the scalar potential energy must be attractive in the Hamiltonian (3)], then, as ϕ_c decreases relative to ϕ_T, the density increases so that ϕ_c yields an inertial force increasing the condensation. As the angular momentum forces do not increase absolutely with increasing density, they contribute small terms to the noncentral forces. We expect that, in a similar manner as for Coulomb forces, there exist equilibrium states which occur at values of m'' for which the total central and noncentral forces balance and that the (m'')'s occur in each density region at certain values of $\mathbf{k}$, $\mathbf{j}$, and the spin and for certain angle dependences; that is, the momentum, spin, and current have definite eigenvalues and selection rules corresponding to the observed masses. It remains, therefore, to establish the nature of the forces and the density dependence.

In Eq. (25) both the current and density are conserved, as is evident from the analyticity of m' in the j_α and ρ, provided, of course, that m' is differentiable in all variables. Writing $j_\alpha = \xi_\alpha + i\lambda_\alpha$ and $\rho = \rho_1 + i\rho_2$ gives

$$\begin{aligned}\frac{m'^2}{m^2} &= \frac{\mu^2}{\lambda mc^2}\frac{\mathbf{j}\cdot\mathbf{j}}{\rho}\\ &= \frac{\mu^2}{\lambda mc^2}\frac{(\xi_\alpha^2 - \lambda_\alpha^2)\rho_1 + (\xi_\alpha\lambda_\alpha + \lambda_\alpha\xi_\alpha)\rho_2 - i[(\xi_\alpha^2 - \lambda_\alpha^2)\rho_2 - (\xi_\alpha\lambda_\alpha + \lambda_\alpha\xi_\alpha)\rho_1]}{\rho_1^2 + \rho_2^2}.\end{aligned} \tag{27}$$

Now consider the stable masses determined by

$$\frac{\lambda mc^2}{\mu^2} \operatorname{Im} m'^2/m^2 = -(\xi_\alpha^2 - \lambda_\alpha^2)\rho_2 + (\xi_\alpha\lambda_\alpha + \lambda_\alpha\xi_\alpha)\rho_1 = 0, \tag{28}$$

which gives

$$\lambda = -\xi\left\{\frac{\rho_1}{\rho_2} \mp \left[\left(\frac{\rho_1}{\rho_2}\right)^2 + 1\right]^{\frac{1}{2}}\right\}, \tag{29}$$

with $\lambda^2 = \sum \lambda_\alpha^2$ and $\xi^2 = \sum \xi_\alpha^2$ so that $[\xi, \lambda] = 0$. Obviously, at $\lambda = 0$ we have $\rho_2 = 0$. Hence, as $\rho_2 \to 0$ the two singularities cancel on the rhs of (29) so that $\lambda \to 0$ uniformly as $\rho_2 \to 0$. Now consider

$$\lim_{\lambda\to 0} \lim_{\rho_2 \to 0} \frac{\rho_2}{\lambda} = \frac{2\rho_1}{\xi}, \tag{30}$$

which is obtained from (28). In the limits the rhs approaches a finite value and the lhs must approach a constant since ρ_2 and λ are not independent functions. With lim $\rho_2/\lambda = \epsilon$ as $\rho_2 \to 0$, $\lambda \to 0$, we obtain $\xi = (2/\epsilon)\rho_1$. Thus the stable masses given by (27) result always from a correlation of ρ_1 and ξ such that

$$\frac{\lambda mc^2}{\mu^2} \operatorname{Re} \frac{m'}{m} = \lim_{\rho_2 \to 0} \frac{\rho_1 \xi^2}{\rho_1^2 + \rho_2^2}\left(1 + \frac{3\rho_2^2}{4\rho_1^2}\right) \tag{31}$$

reduces to a constant. As has been indicated, this result can be derived from the requirement that m'^2/m^2 be analytic in the current, that is, that m'^2/m^2 conserve the current.

Since m'^2/m^2 satisfies the Cauchy–Riemann equations on the current planes, it is invariant under rotations on the current planes. Then m'^2/m^2 obviously has the symmetry of the current so that we identify $\xi = (2/\epsilon)\rho_1$ with a central potential. Noting that $\phi_c \geq mc^2/\lambda$, in the limit we have $\rho^2/\mathbf{j}\cdot\mathbf{j} = 1$, and from this we deduce that (30) is equal to 2. The value of ϵ can be obtained also from the condition

$$\operatorname{Re} \frac{\rho^2}{\mathbf{j}\cdot\mathbf{j}} = \operatorname{Re} \frac{N_\rho^2}{N_j^2}\frac{c^2}{v^2} \geq 1 \tag{32}$$

since $c/v \geq 1$ and $N_\rho/N_j \geq 1$ in order that the total number density exceed the current number density. At the lower bound one finds for stable masses $\xi = \rho_1$, which gives $\epsilon = 2$.

With these results we obtain the limiting value in (27), and with $\lambda = e$ we find

$$m'/m = \mu(N_\rho/mc^2)^{\frac{1}{2}} = 1 \cdot 10^{-17}\, \text{cm}^{\frac{3}{2}}/r'^{\frac{3}{2}}, \tag{33}$$

where r' is the particle radius corresponding to the mass m'. For any reasonable value of Re N_ρ in the nuclear region, one obtains, in units of the electron mass, $m'/m \simeq 10^3$ so that the mass eigenvalue is in the range of the nucleon in the correct density region. Moreover, the radii obtained from (33) for several particle masses are near, and apparently have simple relations with, the Compton lengths (see Table I). It appears, therefore, that the particle masses are simple functions of the charge density. For example, with (33) the exact nucleon mass gives $N_\rho = 8.3.10^{39}$ cm^{-3}, which gives a nucleon radius of $r' = 0.31$ F compared to the Compton length of 0.21 F. Evidently the Compton length λ_C corresponds to the bare nucleon with no meson whereas m'/m should include terms corresponding to the meson and binding energy. Hence, the mass difference corresponding to r' and λ_C should bound the meson plus binding energies. However, (33) gives $m'/m = 3300$ at the proton Compton length which leaves $\Delta m/m = 1460$, which is certainly too large for pion interactions; moreover, if one expects the extended radius to be of the order of the pion Compton length, one finds the pion length 1.41 F is completely at variance with r'_n obtained from (33). The charged pion radius obtained from (33) is 1.1 F, which is within 10% of the effective nuclear radius entering the formula for nuclear radii, and suggests that pion exchange predominates for nuclear interactions. In this connection it is probable that pions play the same role in nucleon–nucleon interactions as that of phonons in solids.

The ratios of the particle radii to the Compton lengths are $r'/\lambda_C = (\alpha m'/m)^{\frac{1}{3}}$, and the constant values

TABLE I. Comparison of particle radii calculated from (33) and Compton lengths.

Pseudoscalar mesons	$r'(F)$	$\lambda_C(F)$	$R = r'/\lambda_C$
$\pi^\pm$	1.10	1.41	0.78
π^0	1.13	1.46	0.77
$K^\pm$	0.47	0.40	1.19
K^0	0.47	0.39	1.19
η^0	0.44	0.36	1.23
Baryons			
p	0.31	0.21	1.47
n	0.31	0.21	1.47
Λ^0	0.27	0.18	1.56
Σ^+	0.26	0.16	1.60
Σ^0	0.26	0.16	1.60
Σ^-	0.26	0.16	1.60
Ξ^0	0.25	0.15	1.65
Ξ^-	0.25	0.15	1.65
Ω^-	0.21	0.11	1.97
Vector mesons			
ρ^0	0.35	0.26	1.38
ω	0.35	0.25	1.39
$4K^*$	0.32	0.20	1.61
ϕ	0.29	0.17	1.67

of these ratios within the observed multiplets (see Table I) suggests that the multiplet masses vary as $\alpha^{-1} \times$ const so that the splittings are due to electromagnetic interactions as expected.

In the electrodynamic region ($\sim 10^{33}\ \text{cm}^{-3}$) Eq. (33) predicts masses in the range of the electron mass; for example, for $m'/m = 2$ we find $N_\rho = 9.5.10^{33}\ \text{cm}^{-3}$ corresponding to a radius $r' = 0.29.10^{-11}$ cm, corresponding to the pair threshold. In the same way the density $N_\rho = 25.2.10^{39}\ \text{cm}^{-3}$ corresponds to the nucleon threshold, and so on. In this way one is led to the concept of the particle hierarchy as entirely a function of the density region (see Table I). More precisely, one expects the observed particle masses to occur as stable states of the renormalized fields (22) and (23), and it is evident from (25) and (26) that the mass is related to the fields in a simple manner. Equations (22), (23), and (33) lead to definite conclusions regarding the nature of strong forces; obviously, the current–current forces play an important role since $\mu^2 \mathbf{j}\cdot\mathbf{j}$ occurs in the numerator of the mass formula (33). A straightforward dimensional analysis shows that at sufficiently high density the current interactions predominate since the magnetic force $\mu\mathbf{j}$ varies as $\lambda_C r^{-3}$, whereas the electric force varies as r^{-2} and λ_C is the Compton length. Thus in a single-particle formulation for electrons, the ratio of magnetic to electric forces exceeds unity at a density corresponding to $r \simeq \lambda_C$. In the present formulation, which includes the many-body effects in ρ and $\mathbf{j}$, we see that the ratio exceeds unity at $\mu\,|\mathbf{j}|/\lambda\,|\mathbf{k}|\,\phi = \mu\,|\mathbf{j}|\,\mathbf{j}\cdot\mathbf{j}/\lambda\rho^2\,|k| > 1$, and it is evident this condition is satisfied at a sufficiently large value of the current density. It follows that, in the region of nuclear densities, m' must result predominantly from an inertial central field term opposed to the current interactions; this follows directly from the observation that m'^2 varies as $\mu^2\mathbf{j}\cdot\mathbf{j}/\rho$, where $\mu\mathbf{j}$ is the magnetic force $\sim \lambda_C r_j^{-3}$, where $(\frac{4}{3}\pi)r_j^3 = V_j$ is the current volume and λ_C is the *electron* Compton length. Hence, in the stationary limit [see Eq. (27) and discussion] we find $m'^2 = (3me^2/4\pi c^2)\lambda_C^2/r_\rho^3$. However, this gives a continuous spectrum so that the observed masses must be fixed by selection rules which give the proper sign for the noncentral field and which give the balance of central and noncentral forces.

From the discussion of the previous paragraph, one expects that at densities corresponding to the Compton lengths the current forces and central forces exactly balance and that the resulting particles are free and noninteracting. In this picture it is clear why the radii obtained from (33) are all (with the exception of the pions) larger than the Compton lengths, as only in this way can particles occur as bound states with finite decay times. Thus the mass differences corresponding to the $r' - \lambda_C$ differences are to be associated with binding potentials, that is, with mesons or with short-range collective excitations and their interactions with the bare particles.

In the calculations of the previous paragraphs we set $(\boldsymbol{\sigma}\times\mathbf{j})^2/\mathbf{j}\cdot\mathbf{j} = 0$ and $\rho_2 = 0 = \lambda$, thus neglecting the noncentral field and the oscillating terms in the current and density; the effects of the latter are expected to be small compared to the noncentral field corrections, which we now consider.

2. SOLUTION OF THE MOMENTUM EQUATION

The difference between the bare nucleon mass of $3300m$ obtained from (33) at the Compton length and the observed mass obtained at 0.31 F suggests that the extended radius is associated with the noncentral field which reduces the bare mass to approximately the observed value. We must therefore evaluate k' in the noncentral terms in order to obtain explicit expressions for ϕ_T and $\mathbf{A}_T$. With (23) the momentum becomes

$$\begin{aligned}
\mathbf{k}' &= \frac{\lambda\rho}{\mu}\frac{\boldsymbol{\sigma}\times\mathbf{j}}{\mathbf{j}\cdot\mathbf{j}}\left(1 - \frac{\mu}{\lambda\rho}\boldsymbol{\sigma}\cdot\mathbf{k}'\times\mathbf{j}\right) \\
&= \frac{\lambda}{\mu}\frac{\rho}{\mathbf{j}\cdot\mathbf{j}}\boldsymbol{\sigma}\times\mathbf{j}\left[1 + \frac{\mu}{\lambda\rho}\boldsymbol{\sigma}\times\mathbf{j}\right. \\
&\qquad \times\left.\left(\frac{\lambda}{\mu}\frac{\rho}{\mathbf{j}\cdot\mathbf{j}}\boldsymbol{\sigma}\times\mathbf{j}\right)\left(1 + \frac{\mu}{\lambda\rho}\cdots\right)\right] \\
&= \frac{\lambda}{\mu}\frac{\rho}{\mathbf{j}\cdot\mathbf{j}}\boldsymbol{\sigma}\times\mathbf{j}\left(1 + \frac{(\boldsymbol{\sigma}\times\mathbf{j})^2}{\mathbf{j}\cdot\mathbf{j}}\right. \\
&\qquad \left. + \frac{(\boldsymbol{\sigma}\times\mathbf{j})^4}{(\mathbf{j}\cdot\mathbf{j})^2} + \cdots\right)
\end{aligned} \tag{34}$$

and the fields and excitation energies are

$$\phi = \frac{\lambda}{r_0}\frac{\rho^2}{\mathbf{j}\cdot\mathbf{j}}\left[1 + \frac{(\boldsymbol{\sigma}\times\mathbf{j})^2}{\mathbf{j}\cdot\mathbf{j}} \times\left(1 + \frac{(\boldsymbol{\sigma}\times\mathbf{j})^2}{\mathbf{j}\cdot\mathbf{j}}\right)1 + \cdots\right], \tag{35}$$

$$\mathbf{A} = \frac{\lambda}{r_0}\frac{\rho\mathbf{j}}{\mathbf{j}\cdot\mathbf{j}}\left[1 + \frac{(\boldsymbol{\sigma}\times\mathbf{j})^2}{\mathbf{j}\cdot\mathbf{j}} \times\left(1 + \frac{(\boldsymbol{\sigma}\times\mathbf{j})^2}{\mathbf{j}\cdot\mathbf{j}}\right)1 + \cdots\right], \tag{36}$$

and

$$\frac{E'^2}{mc^2} = \frac{1}{m}(k_\alpha'^2 - p_\alpha^2) - \frac{r_0}{m\lambda}\frac{\mathbf{j}\cdot\mathbf{j}}{\rho}\left(1 + \frac{(\boldsymbol{\sigma}\times\mathbf{j})}{\mathbf{j}\cdot\mathbf{j}} + \cdots\right)^{-1}. \tag{37}$$

Noting that $\mathbf{j}\cdot\mathbf{j} = \rho^2\mathbf{A}\cdot\mathbf{A}/\phi^2$, we expect the series to terminate after a finite number of terms and at a value corresponding to one of the observed masses.

For $\lambda = Ze$ we find that $Ze\phi$ and r_0A^2 are independent of the coupling strength and are functions only of the mechanical degrees of freedom, but the m'^2 term is a function of the coupling in the central term and the noncentral contribution can be equated to functions of ϕ or A. Hence, we expect the mass eigenvalues to be fixed at certain ratios of the central field mass to the noncentral contribution. Thus we have

$$\frac{m''^2}{m^2} = \frac{\mu^2}{mc^2}\frac{\mathbf{j}\cdot\mathbf{j}}{\lambda\rho}\left(1 + \frac{(\boldsymbol{\sigma}\times\mathbf{j})^2}{\mathbf{j}\cdot\mathbf{j}} + \cdots\right)^{-1}, \quad (38)$$

and we expect the factor $1 + (\boldsymbol{\sigma}\times\mathbf{j})^2/\mathbf{j}\cdot\mathbf{j} + \cdots$ to converge at the observed masses. We find that (38) gives the observed nucleon mass within 10% when this factor equals 4, and the same value gives agreement for Λ^0, Σ^0, and Ξ^0 within 6%. Since (37) gives better than order of magnitude agreement with the observed baryon masses at densities corresponding to the Compton lengths, evidently the noncentral contribution can be treated as a perturbation since the reduction factor 4 corresponds to a correction of less than unity. Thus, for small noncentral contributions, (35)–(38) become

$$\mathbf{k}' = \frac{\lambda}{\mu}\frac{\rho}{\mathbf{j}\cdot\mathbf{j}}\boldsymbol{\sigma}\times\mathbf{j}\left(1 - \frac{(\boldsymbol{\sigma}\times\mathbf{j})^2}{\mathbf{j}\cdot\mathbf{j}}\right)^{-1}, \quad (39)$$

$$\phi = \frac{\lambda}{r_0}\frac{\rho^2}{\mathbf{j}\cdot\mathbf{j}}\left[1 + \frac{(\boldsymbol{\sigma}\times\mathbf{j})^2}{\mathbf{j}\cdot\mathbf{j}}\left(1 - \frac{(\boldsymbol{\sigma}\times\mathbf{j})^2}{\mathbf{j}\cdot\mathbf{j}}\right)^{-1}\right], \quad (40)$$

$$\mathbf{A} = \frac{\lambda}{r_0}\frac{\rho\mathbf{j}}{\mathbf{j}\cdot\mathbf{j}}\left[1 + \frac{(\boldsymbol{\sigma}\times\mathbf{j})^2}{\mathbf{j}\cdot\mathbf{j}}\left(1 - \frac{[\boldsymbol{\sigma}\times\mathbf{j}]^2}{\mathbf{j}\cdot\mathbf{j}}\right)^{-1}\right], \quad (41)$$

and

$$\frac{m''^2}{m^2} = \frac{\mu^2}{mc^2}\frac{\mathbf{j}\cdot\mathbf{j}}{\lambda\rho}\left(1 - \frac{(\boldsymbol{\sigma}\times\mathbf{j})^2}{\mathbf{j}\cdot\mathbf{j}}\right). \quad (42)$$

As pointed out in the preceding paragraphs, Eq. (42) gives good agreement with observed neutral baryon masses for $(\boldsymbol{\sigma}\times\mathbf{j})^2/\mathbf{j}\cdot\mathbf{j} = \frac{3}{4}$ at densities (see Table II) corresponding to the Compton lengths, which suggests that the spin–current interactions can be treated as the generators of the ordinary spin group with $s = \frac{1}{2}$ for baryons. For the vector mesons one obtains agreement within 17% for ρ^0 and within 2% for the decouplet K^* with $S = \frac{1}{2}$; for the pseudoscalar mesons, however, the calculated masses agree within about 40% with observed values. That the spin–current interactions generate the ordinary spin group is supported also by an examination of the T_3 and Y values for the baryon and meson multiplets, which show that the $(\boldsymbol{\sigma}\times\mathbf{j})^2/\mathbf{j}\cdot\mathbf{j}$ interaction cannot be assigned to either the isospin or hypercharge groups. Within each multiplet, however, the best results are found at strangeness $= \pm 1$, with quite good agreement within the baryon and vector meson multiplets. This suggests that the discrepancies at both ends of the baryon multiplet can be accounted for by the strangeness numbers.

TABLE II. Masses calculated from (49) with spin–current corrections and compared with observed masses.

Particle	s	m''/m	m''/m (observed)
π^0	0	181	264
K^0	0	1289	975
η^0	0	1490	1078
n	$\frac{1}{2}$	1661	1839
Λ^0	$\frac{1}{2}$	2137	2188
Σ^0	$\frac{1}{2}$	2347	2338
Ξ^0	$\frac{1}{2}$	2747	2578
ρ^0	$\frac{1}{2}$	1266	1535
$4K^*$	$\frac{1}{2}$	1779	1749

Table II gives the masses calculated at densities corresponding to the Compton lengths and the observed spin values for baryons and mesons. One expects the differences between the observed and calculated masses to be accounted for by the short-range collective effects associated with finite values of the current and density fluctuations λ_α and ρ_2. It is clear from the discussion that these are to be associated with isospin and strangeness conserving subgroups. According to this picture, with the spin–current interactions turned off, the central field contracts until the current–current forces are exactly balanced at densities corresponding to the Compton lengths. These bare particles occur with spins aligned along the currents, there are no current or density fluctuations, and no further interactions occur. With the spin–current interactions turned on, additional repulsive forces opposing the condensation result in extended radii and reduced masses which are given approximately by (42). The current and density fluctuations induced by the spin–current interactions then induce higher-order interactions and must be conserved with the total density and current.

* Permanent address: Box 94, La Jolla, California.

[1] J. T. Anderson, J. Math. Phys. **11**, 117 (1970).
[2] J. T. Anderson, J. Math. Phys. **11**, 2463 (1970).
[3] Y. Nambu and G. Jonas-Lasinio, Phys. Rev. **122**, 345 (1961); **124**, 246 (1961).

Continuous Representations and the Kernel Function

JAGANNATHAN GOMATAM* AND C. W. GARDINER†‡
Physics Department, Syracuse University, Syracuse, New York 13210

(Received 24 August 1970)

New examples of continuous representations are obtained utilizing the theory of kernel functions for a finite domain D in the complex z-plane. It is shown how, by starting with the continuous representation for a circular domain, the Riemann mapping theorem makes it possible to obtain the continuous representations with respect to any finite, simply connected *schlicht* domain with at least two boundary points. A method of constructing orthonormal polynomials with respect to a class of weight functions for both finite domains and infinite plane is given. The standard coherent states employed in quantum optics and quantum field theory emerge as a particular case of the present investigation.

1. INTRODUCTION

The relationship between reproducing kernels[1,2] and continuous representation[3] has been noted in general terms in the literature,[4,5] but not fully exploited in the construction of continuous representations. That the orthonormal, analytic functions of a complex variable provide a direct way of constructing reproducing kernel functions has been observed by Bergman[1] and used extensively in the study of differential equations. Bargmann[5] studied the Hilbert space of entire analytic functions defined by an inner product with respect to the measure $\rho = \exp(-\bar{z}z)$ and obtained coherent states[6] as a particular case.

The contribution of this investigation is threefold: first in showing how the theory of kernel functions defined on a finite domain D can be utilized in the construction of new continuous representations; second in the use of conformal mapping in conjunction with the Riemann mapping theorem to obtain continuous representations with respect to any finite, simply connected, *schlicht* domain having more than one boundary point; third in the explicit delineation of a method of construction of orthonormal polynomials with respect to a class of weight functions for both finite domains and infinite plane. The standard continuous representation emerge as a particular case of our treatment.

The general theory of reproducing kernels is summarized in Sec. 2, while the formalism for constructing continuous representations using kernel functions is outlined in Sec. 3. Two explicit examples for finite domains are treated in Sec. 4. The weighted kernel function is considered in Sec. 5; a method of constructing closed orthonormal system of functions with respect to weight function $\gamma(x, y)$ over a finite domain is presented. The extension of these ideas to the infinite plane calls for further restrictions on $\gamma(x, y)$. This is discussed in Sec. 6 with the aid of two examples. Further possible generalizations are indicated in Sec. 7. A discussion of the paper is presented in Sec. 8.

2. REPRODUCING KERNELS AND HILBERT SPACES OF ANALYTIC FUNCTIONS

In this section we briefly review the properties of reproducing kernels in Hilbert spaces of analytic functions.[1,2] Consider the class of all single-valued analytic functions $\{f\}$ defined over a finite domain D in the complex $z = x + iy$ plane. Introduce a norm by

$$\|f\|_D^2 = \int_D |f(z)|^2 \, d\omega, \tag{2.1}$$

where $d\omega = dx\, dy$, and the inner product

$$(f, g)_D = \int_D f(z)\overline{g(z)} \, d\omega. \tag{2.2}$$

Then Bergman has shown[1] that the class of all f such that $\|f\|_D$ exists is in fact complete in this norm and forms a Hilbert space with a countable base, which we shall call $\mathfrak{L}^2(D)$. He has also shown an analog of the Riesz–Fisher theorem for this Hilbert space.

The Riesz–Fisher Theorem

Let $\{\psi_n\}$ be an arbitrary orthonormal set in $\mathfrak{L}^2(D)$, and let the numbers $a_0, a_1, a_2 \cdots$ be such that

$$\sum_{n=0}^{\infty} |a_n|^2 < \infty. \tag{2.3}$$

Then there exists a function $f \in \mathfrak{L}^2(D)$ such that

$$a_n = (f, \psi_n) \tag{2.4}$$

and

$$\|f\| = \sum_{n=0}^{\infty} |a_n|^2. \tag{2.5}$$

It should be noted that the above theorem, while apparently identical in form with the more usual Riesz–Fisher theorem, is a much stronger theorem, since the Hilbert space is of analytic functions. It is this aspect which leads to the possibility of a continuous representation and a reproducing kernel.

Reproducing Kernel

The kernel function for the Hilbert space $\mathfrak{L}^2(D)$ is defined by

$$K_D(z, \overline{z'}) = \sum_{n=0}^{\infty} \psi_n(z)\overline{\psi_n(z')}, \tag{2.6}$$

where $\psi_n(z)$ is any complete orthonormal set in $\mathfrak{L}^2(D)$. It has been shown by Bergman[7] that (2.6) converges absolutely and uniformly with respect to $z(z')$ for any fixed $z'(z)$, when z and z' belong to any closed subdomain $D' \subset D$. The result of this is that the kernel function $K_D(z, \overline{z'})$ is an analytic function of the two complex variables z, $\overline{z'}$. Also,

$$K_D(z, \bar{z}) = \sum_{n=0}^{\infty} |\psi_n(z)|^2 < \infty. \tag{2.7}$$

The kernel function is a reproducing kernel

$$\begin{aligned}\int K_D(z, \overline{z'})f(z')\, d\omega' &= \int \sum_{n=0}^{\infty} \psi_n(z)\overline{\psi_n(z')}f(z')\, d\omega' \\ &= \sum_{n=0}^{\infty} a_n\psi_n(z) \\ &= f(z). \end{aligned} \tag{2.8}$$

From this reproducing property it can be seen that the kernel function can depend only on the domain D and not the set $\{\psi_n\}$ used to construct it.

One should note particularly the very great difference between (2.6) and the more usual type of relation—for example, true for complete orthonormal functions $\{\psi_n\}$ on the real line:

$$K_R(x, x') = \sum_{n=0}^{\infty} \psi_n(x)\overline{\psi_n(x')} = \delta(x - x'). \tag{2.9}$$

The delta function is a reproducing function, but is not continuous or analytic like the kernel (2.6). The basic difference is that $K(z, \overline{z'})$ is a reproducing function only for square integrable *analytic* functions.

The effect of conformal mapping of D_z to D_w on $\mathfrak{L}^2(D)$ has been studied.[1] Let

$$w = k(z) \tag{2.10}$$

define the conformal mapping of D_z to D_w with

$$z = h(w) \tag{2.11}$$

defining the inverse mapping. Then, if $f \in \mathfrak{L}^2(D)$ with $d\omega_w = du\, dv$,

$$\begin{aligned}\int_{D_w} |f(h(w))h'(w)|^2\, d\omega_w &= \int_{D_z} |f(z)|^2|h'(w)|^2\frac{\partial(u, v)}{\partial(x, y)}\, d\omega_z \\ &= \int_{D_z} |f(z)|^2\, d\omega_z, \end{aligned} \tag{2.12}$$

because the Jacobian

$$\frac{\partial(u, v)}{\partial(x, y)} = |h'(w)|^{-2}.$$

From (2.12) it follows that $f(h(w))h'(w)$ is a function of the class $\mathfrak{L}^2(D_w)$. Also it can be shown that the functions

$$\phi_n(w) = \psi_n(h(w))h'(w) \tag{2.13}$$

form a closed orthonormal system with respect to D_w, with

$$K_{D_w}(w, \overline{w'}) = K_{D_z}(h(w), \overline{h(w')})\frac{dz}{dw}\overline{\frac{dz'}{dw'}} \tag{2.14}$$

as the reproducing kernel. Thus, starting with a given set of orthonormal functions in a domain D, we can construct new orthonormal systems by studying the conformal mapping of the domain D. The Riemann mapping theorem[8] states that an arbitrary, finite, simply connected *schlicht* domain D with more than one boundary point can be conformally mapped into a unit circle. Thus starting with a complete orthonormal system in the unit circle [see later, (4.1)], we can construct a complete orthonormal system in any finite, simply connected, *schlicht* domain D with more than one boundary point by using (2.13), where (2.11) is the conformal mapping function of the unit circle into D_z. The completeness of the system in D_z follows from the completeness of the initial system in the unit circle.

3. CONTINUOUS REPRESENTATIONS USING THE ORTHONORMAL SYSTEM IN $\mathfrak{L}^2(D)$

The continuous representations in Hilbert space was introduced in its most general form by Klauder.[3] Roughly speaking, the continuous representation $|z)$ labeled by complex number z is such that the inner product $(z \mid z')$ is a continuous function of z as z' approaches z and that it constitutes an overcomplete family of states. The coherent states are a particular realization of these continuous representations. Our aim is to construct new realizations of the continuous representations with the help of the formalism set up in Sec. II.

We define a continuous representation $|z)_D$ as follows:

$$|z)_D \equiv \sum_{n=0}^{\infty} \psi_n(z)\, |n\rangle, \tag{3.1}$$

where the states $|n\rangle$ belong to the Fock space $\mathcal{F}$. They are characterized by

$$\langle n \mid m\rangle = \delta_{nm}, \tag{3.2}$$

$$\sum_{n=0}^{\infty} |n\rangle\langle n| = 1. \tag{3.3}$$

The states $|z)_D$ defined by (3.1) do indeed form a continuous representation. The norm of $|z)_D$ is given by

$$\| |z)_D\|^2 = (z \mid z)_D = \sum_{n=0}^{\infty} |\psi_n(z)|^2 = K_D(z, \bar{z}) < \infty. \tag{3.4}$$

The inner product $(z' \mid z)_D$ is given by

$$(z' \mid z)_D = \sum_{n=0}^{\infty} \psi_n(z)\overline{\psi_n(z')} = K_D(z, \bar{z}'), \tag{3.5}$$

where we have used (3.2). From the uniform convergence of (2.6) it follows that $(z' \mid z)_D$ is a continuous function of z at $z = z'$. Thus expression (3.1) defines a continuous representation. The complete set of classical orthonormal polynomials over a real line cannot be used because $K_R(x, x')$ is not a continuous function of x at $x = x'$, but rather a generalized function of x and x', for instance, $\delta(x - x')$. That $|z)_D$ form over a complete family of states is seen as follows:

$$\int_D |z)_D(z|_D \, d\omega = \int_D \sum_{\substack{n=0\\m=0}}^{\infty} \psi_n(z)\overline{\psi_m(z)} \, |n\rangle\langle m| \, d\omega$$
$$= \sum_{n=0}^{\infty} |n\rangle\langle n| = 1. \tag{3.6}$$

(3.6) is a reflection of the Riesz–Fisher theorem proved in Sec. 2 and, when the integral is interpreted to converge weakly, simply means that

$$\int_D g(z)\overline{h(z)} \, d\omega = (g, h)_D = \sum_{n=0}^{\infty} g_n \bar{h}_n,$$

which follows by linearity from (2.5).

4. EXAMPLES OF CONTINUOUS REPRESENTATIONS

A. Orthonormal Polynomials over a Circular Domain[1] $|z| < R$

The polynomials are given by

$$\psi_n(z) = [(n + 1)/\pi]^{\frac{1}{2}} z^n / R^{n+1}. \tag{4.1}$$

The inner product is

$$(z' \mid z) = \sum_{n=0}^{\infty} \frac{n + 1}{\pi} \frac{(\bar{z}')^n (z)^n}{R^{2(n+1)}} = \frac{R^2}{\pi(R^2 - \bar{z}'z)^2}. \tag{4.2}$$

Also

$$\| |z)\|^2 = R^2/\pi(R^2 - |z|^2)^2. \tag{4.3}$$

It is of interest to construct the operator A whose eigenvectors are $|z)$. To this end we assume that there exist annihilation and creation operators a and a^+, respectively, such that

$$[a, a^+] = 1, \tag{4.4}$$

$$a\,|n\rangle = \sqrt{n}\,|n - 1\rangle, \tag{4.5}$$

$$a^+\,|n\rangle = (n + 1)^{\frac{1}{2}}\,|n + 1\rangle, \tag{4.6}$$

$$N\,|n\rangle \equiv a^+a\,|n\rangle = n\,|n\rangle. \tag{4.7}$$

Using (4.5) and (4.7), we easily check that

$$A\,|z) = z\,|z), \tag{4.8}$$

where

$$A = \frac{R}{(N + 2)^{\frac{1}{2}}} a \tag{4.9}$$

and

$$|z) = \sum_{n=0}^{\infty} \left(\frac{n + 1}{\pi}\right)^{\frac{1}{2}} \frac{z^n}{R^{n+1}} |n\rangle. \tag{4.10}$$

Also

$$|z) = U\,|0\rangle, \tag{4.11}$$

where

$$U = e^{|z|^2/2R^2} \left[\frac{(N + 1)!}{\pi R}\right]^{\frac{1}{2}} e^{(1/R)(za^+ - \bar{z}a)}. \tag{4.12}$$

We note that U is not an unitary operator. We will now show that attempts to redefine $|z)$ so that the redefined states are generated from the vacuum by a unitary operator destroy the analyticity of the associated function space. Suppose we set

$$|z)) \equiv \exp(-|z|^2/2R^2)[\pi R/(N + 1)!]^{\frac{1}{2}}\,|z), \tag{4.13}$$

where we recall that N is the number operator, then $\| |z))\|^2 = 1$. But (4.13) together with (4.10) yields

$$|z)) = e^{(-|z|^2/2R^2)} \sum_{n=0}^{\infty} \frac{1}{\sqrt{n!}} \left(\frac{z}{R}\right)^n |n\rangle, \tag{4.14a}$$

which are just the standard continuous representations with the complex number z replaced by z/R. Note that the functions $\exp[(-|z|^2/2R^2)(1/\sqrt{n!})(z/R)^n]$ are no longer analytic in z. We also note that, though the vectors

$$\frac{|z)}{\| |z)\|} = \frac{(R^2 - |z|^2)}{R} \sum_{n=0}^{\infty} (n + 1)^{\frac{1}{2}} \frac{z^n}{R^{n+1}} |n\rangle \tag{4.14b}$$

have unit norm, the functions $[(n + 1)/R^{n+2}](R^2 - |z|^2)z^n$ are not analytic in z. Thus, in our presentation, we do not employ either of the definitions (4.13) or (4.14b). We will adhere to the definition (3.1).

B. Orthonormal Functions over an Elliptic Domain[9] $(x^2/\cosh^2 \sigma) + (y^2/\sinh^2 \sigma) < 1$

The functions are

$$\psi_n(z) = c_{n+1} \frac{e^{(n+1)\cosh^{-1} z}}{\sinh(\cosh^{-1} z)}, \tag{4.15}$$

where

$$c_{n+1} = \left(\frac{(n+1)}{\pi \sinh 2(n+1)\sigma}\right)^{\frac{1}{2}}. \quad (4.16)$$

The proof of orthonormality of (4.15) is outlined in the Appendix. The $|z)$ constructed with the aid of (4.15) can be shown to be eigenstates of

$$B = \left(\frac{\sinh [2(N+2)\sigma]}{(N+1)\sinh [2(N+1)\sigma]}\right)^{\frac{1}{2}} a, \quad (4.17)$$

with the eigenvalues $e^{\cosh^{-1} z} = z \pm (z^2 - 1)^{\frac{1}{2}}$. Note the degeneracy in the eigenvalue. Also

$$|z \pm (z^2 - 1)^{\frac{1}{2}}) = V_{\pm} |0\rangle, \quad (4.18)$$

where

$$V_{\pm} \equiv \exp (\tfrac{1}{2} |e^{\cosh^{-1} z}|^2) \times \frac{\exp (\cosh^{-1} z)}{\sinh (\cosh^{-1} z)} \left(\frac{(N+1)!}{\pi \sinh [(N+1)\sigma]}\right)^{\frac{1}{2}} \times \exp [\exp (\cosh^{-1} z) a^{\dagger} - \overline{\exp (\cosh^{-1} z)} a]. \quad (4.19)$$

5. WEIGHTED KERNELS[10] AND CONTINUOUS REPRESENTATIONS—FINITE DOMAIN

The theory of reproducing kernels and consequently that of the continuous representation can be further generalized by modifying (2.2) to

$$(f, g)_{\mu, D} = \int_D |\mu(z)|^2 f(z)\overline{g(z)}\, d\omega. \quad (5.1)$$

If $\mu(z)$ is an analytic weight function not equal to zero in D such that

$$\int |\mu(z)|^2\, d\omega < \infty, \quad (5.2)$$

then $\mu(z)$ can be absorbed into the definition of the function, and the inner product (5.1) will then be with respect to the unit weight function.

The orthonormal system with a general weight function has been sparsely studied in the literature.[5,10,11] Sufficient conditions have been imposed on the weight function $\gamma(x, y)$ to obtain closed orthonormal system. We will use the method of conformal mapping to construct weighted orthonormal functions. To the best of our knowledge this does not seem to have been carried out in the literature. We will also discuss the question of extending the domain D to cover the entire complex plane.

We first present a method of constructing orthonormal polynomials over any *schlicht* domain that is simply connected with more than one boundary point. We start with the orthonormal set

$$\int_a^b \alpha(\phi) R_n(\phi)\overline{R_m(\phi)}\, d\phi = \delta_{nm}, \quad (5.3)$$

where $\phi \in [a, b]$ and $\alpha(\phi)$ is a weight function. Multiplying (5.3) by

$$\int_0^R \rho\beta(\rho) G_n(\rho)\overline{G_m(\rho)}\, d\rho = c_{nm} < \infty, \quad (5.4)$$

where $\beta(\rho)$ and $G_n(\rho)$ are such that $c_{nm} < \infty$, to obtain

$$\int_{\rho=0}^R \int_{\phi=a}^b \rho\, d\rho\, d\phi \beta(\rho)\alpha(\phi) G_n(\rho) R_n(\phi) \overline{G_m(\rho)}\, \overline{R_m(\phi)} = c_{nn}\delta_{nm}, \quad (5.5)$$

we will now consider ρ and ϕ as the modulus and amplitudes of a complex variable $w = \rho e^{i\phi}$. The function

$$\psi_n(\rho, \phi) = G_n(\rho) R_n(\phi) e^{i\Lambda(\rho,\phi)}/(c_{nn})^{\frac{1}{2}}, \quad (5.6)$$

where $\overline{\Lambda(\rho, \phi)} = \Lambda(\rho, \phi)$, is orthonormal with respect to the weight function $\beta(\rho)\alpha(\phi)$ in the sector $S\{\rho, \phi;\ 0 \le \rho < R,\ \alpha \le \phi \le b\}$ in the w plane. Let

$$w = u + iv = F(z), \quad z = h(w), \quad (5.7)$$

$$du\, dv = |F'(z)|^2\, dx\, dy \quad (5.8)$$

define a conformal mapping of S to some domain D_z in the z plane. Since we are interested in analytic orthonormal polynomials, we want $\psi_n(\rho, \phi)$ to be an analytic function of z when ρ and ϕ are expressed in terms of z using (5.7). Since an analytic function of an analytic function is again an analytic function, we demand that $\psi_n(\rho, \phi)$ be an analytic function of $w = \rho e^{i\phi}$. The only way we can restrict (5.6) so as to be analytic function of w is to stipulate

$$G_n(\rho) = \rho^n, \quad (5.9)$$

$$R_n(\phi) = e^{in\phi}, \quad (5.10)$$

$$\Lambda(\rho, \phi) = \text{const}. \quad (5.11)$$

Without loss of generality we can set $\Lambda(\rho, \phi) = 0$. Then

$$\psi_n(w) = \frac{w^n}{\sqrt{c_{nn}}} = \frac{[F(z)]^n}{\sqrt{c_{nn}}}. \quad (5.12)$$

The choice (5.10) yields $a = -\pi$, $b = +\pi$, and $\alpha(\phi) = 2\pi$. From (5.9) it is clear that $\beta(\rho)$ should obey the restriction

$$\int_0^R \beta(\rho)\rho^{m+n+1}\, d\rho = c_{mn} < \infty. \quad (5.13)$$

Thus the polynomials

$$\phi_n(z) = [F(z)]^n/\sqrt{c_{nn}} \quad (5.14)$$

are orthonormal over D_z with respect to the weight function

$$\gamma(x, y) = \frac{\beta(|F(z)|)}{2\pi} |F'(z)|^2. \tag{5.15}$$

The crucial point in the construction outlined above is the separability of the area measure and the functions $\psi_n(\rho, \phi)$, in terms of ρ and ϕ, in the w plane. Thus, given a domain D_z, we first construct its conformal mapping, $w = F(z)$ to the circle C in the w plane. So long as the given domain is a simply connected *schlicht* domain with more than one boundary point, this can be done according to the Riemann mapping theorem.[2] Then we use this mapping function to construct the orthonormal set (5.14); the corresponding weight function is given by (5.15).

We use (5.14) and (5.15) to construct continuous representations. We consider the Hilbert space of analytic functions defined by

$$\|f\|^2_{\gamma,D} = \int_D \gamma(x, y)\, |f(z)|^2\, d\omega \tag{5.16}$$

and

$$(f, g)_{\gamma,D} = \int_D \gamma(x, y) f(z)\overline{g(z)}\, d\omega. \tag{5.17}$$

We denote this by $\mathfrak{L}^2_\gamma(D)$. Here $\gamma(x, y)$ is already restricted by (5.13) via (5.15). If $\gamma(x, y)$ is a positive, continuous, differentiable function of x and y in a domain that includes D, then there exists[11] an orthonormal system with respect to the inner product (5.16). The system also possesses an analytic kernel function. The continuous representations associated with (5.14) are

$$|z)_{\gamma,D} = [\gamma(x, y)]^{\frac{1}{2}} \sum_{n=0}^{\infty} \frac{[F(z)]^n}{\sqrt{c_{nn}}} |n\rangle \tag{5.18}$$

$$(z' \mid z)_{\gamma,D} = [\gamma(x, y)\gamma(x', y')]^{\frac{1}{2}} \sum_{n=0}^{\infty} \frac{[F(z)\overline{F(z')}]^n}{c_{nn}}. \tag{5.19}$$

The states $F(z)$ are eigenvalues of

$$C = [c_{N+1N+1}/(N+1)c_{NN}]^{\frac{1}{2}} a, \tag{5.20}$$

with $F(z)$ as eigenvalues. Again

$$|z)_{\gamma,D} = W\,|0\rangle, \tag{5.21}$$

where

$$W = [\gamma(x, y)]^{\frac{1}{2}} \exp\,[|F(z)|^2/2](N!/C_{NN})^{\frac{1}{2}} \times \exp\,[F(z)a^+ - \overline{F(z)}a]. \tag{5.22}$$

We note that W is not unitary, although the part $\exp\,[F(z)a^+ - \overline{F(z)}a]$ is. We will show later that if $\beta(\rho)$ is such as to yield c_{NN} proportional to $N!$, we retrieve the standard continuous representation.

6. WEIGHT FUNCTIONS FOR THE INFINITE PLANE

In the literature[6] the infinite z plane is used in the construction of continuous representations. A straightforward way of obtaining such representations here is studying the limit of the integral (5.13) under $R \to \infty$. Thus all $\beta(\rho)$ satisfying

$$\int_0^\infty \beta(\rho)\rho^{m+n+1}\, d\rho = c_{mn} < \infty \tag{6.1}$$

will yield orthonormal set (5.14) with $F(z) = z$ over the complex z plane. With $\beta(\rho) = e^{-\rho^2}/\pi$, we obtain $c_{nn} = n!/2\pi$ and

$$|z) = e^{-|z|^2/2} \sum_{n=0}^{\infty} z^n/\sqrt{n!}\, |n\rangle,$$

which is the standard continuous representation. New continuous representations can be obtained by making different choices of $\beta(\rho)$. For example, $\beta(\rho) = e^{-k_1\rho} \cosh k_2\rho$ with $k_1 > |k_2|$ leads to

$$c_{nn} = \tfrac{1}{2}(2n + 1)![1/(k_1 - k_2)^{2n+2} + 1/(k_1 + k_2)^{2n+2}]. \tag{6.2}$$

The corresponding continuous representation can be written down easily.

7. FURTHER GENERALIZATION

It is possible to generalize continuous representations further by setting

$$|z) = T(z) \sum_{n=0}^{\infty} \alpha_n \varphi_n(z)\, |n\rangle, \tag{7.1}$$

where $\{\alpha_n\}$ is a suitably restricted sequence of numbers and $T(z)$ is a function to be specified below. As before $\{\varphi_n(z)\}$ is an orthonormal system with respect to a domain D. Notice that $T(z)$ need not be the weight function for the domain D. Then the inner product is

$$(z' \mid z) = \overline{T(z')}T(z) \sum_{n=0}^{\infty} |\alpha_n|^2\, \varphi_n(z)\overline{\varphi_n(z')}. \tag{7.2}$$

The "completeness condition" reads

$$\int_D |z)(z|\, d\omega = \int_D \sum_{\substack{n=0\\m=0}}^{\infty} |T(z)|^2\, \varphi_n(z)\overline{\varphi_m(z)}\, |n\rangle\langle m|\, \alpha_n\bar{\alpha}_m\, d\omega. \tag{7.3}$$

If the domain is the unit circle, with $\varphi_n(z) = z^n$,

$$\alpha_n = \prod_{k=1}^{n} [f(k)]^{-1}, \quad n \geq 1, \quad \alpha_0 = 1,$$

and $T(z) = A(z)$, we obtain the representation discussed by Lerner, Huang, and Walters.[12] The quantities $A(z)$ and $f(k)$ are defined in Ref. 12. Then (5.7) reads

$$\int_{r=0}^{1}\int_{\theta=0}^{2\pi}|z)(z|\, r\,dr\,d\theta = 2\pi\sum_{n=0}^{\infty}\prod_{k=1}^{n}[f(k)]^{-2}\int r^{(2n+1)}|A(r)|^2\,dr, \quad (7.4)$$

where the assumption $|A(z)| = A(|z|)$ is made. The important point here is that the functions z^n used by Lerner *et al.* are orthogonal polynomials on a circular domain or a circular boundary.

8. DISCUSSION

The theory of kernel functions affords a mathematical setup for the construction of new examples of continuous representations. In our case, the representations $|z)$ are not obtainable from vacuum by a unitary transformation. However, we have explicitly constructed the nonunitary transformations that relate $|z)$ to $|0\rangle$. We also noted an example where an attempt to build normalized $|z)$ was not compatible with the analyticity of the associated function space.

In the case of infinite complex z plane, we have demonstrated the existence of further examples of continuous representations which contain the standard coherent states as a particular case. The physical applications for these new continuous representations are not evident at this time, though the boundedness[13] of A^+A might be of interest.

ACKNOWLEDGMENTS

We wish to thank Professor F. Rohrlich for useful discussions and encouragement during the course of this work. One of us (J. G.) is grateful to Professor L. Streit for critical comments.

APPENDIX: ORTHONORMALITY OF THE SYSTEM (4.15)

Consider the integral

$$(\psi_n, \psi_m) = c_{n+1}c_{m+1}\times\int_{\text{Ellipse}} dx\,dy\,\frac{e^{(n+1)\cosh^{-1}z}\,\overline{e^{(m+1)\cosh^{-1}z}}}{|\sinh\cosh^{-1}z|^2}. \quad (\text{A1})$$

Let us introduce the transformation

$$\cosh^{-1}z = w = u + iv \quad (\text{A2})$$

or, equivalently,

$$x = \cosh u\cos v, \quad (\text{A3a})$$

$$y = \sinh u\sin v. \quad (\text{A3b})$$

The area $dx\,dy$ is transformed into $du\,dv$ where

$$dx\,dy = \frac{\partial(x, y)}{\partial(u, v)}\,du\,dv. \quad (\text{A4})$$

Here $\partial(x, y)/\partial(u, v)$ is the Jacobian of the transformations (A3a) and (A3b):

$$\frac{\partial(x, y)}{\partial(u, v)} = (\sinh u\cos v)^2 + (\cosh u\sin v)^2 = |\sinh w|^2. \quad (\text{A5})$$

Thus:

$$(\psi_n, \psi_m) = \tfrac{1}{2}c_{n+1}c_{m+1}\int_{u=-\sigma}^{+\sigma}du\int_{v=-\pi}^{\pi}dv\,e^{u(n+m+2)}e^{iv(n-m)}. \quad (\text{A6})$$

When $n \neq m$, (6) vanishes because $\int_{-\pi}^{\pi} e^{iv(n-m)}\,dv$ does. When $n = m$, we have

$$(\psi_n, \psi_m) = \frac{(c_{n+1})^2}{2}\int_{-\sigma}^{+\sigma}du\,e^{2u(n+1)}\int_{-\pi}^{\pi}dv = \frac{(c_{n+1})^2}{2}\frac{\sinh 2(n+1)\sigma}{(n+1)}2\pi = 1. \quad (\text{A7})$$

We note that (A2) is the conformal mapping of twice the elliptic domain in the z plane over the rectangular domain, $-\sigma < u < \sigma$, $-\pi < v < \pi$ in the w plane.

* Work supported by the National Science Foundation.

† Work supported by the United States Atomic Energy Commission.

‡ Present address: Waikato University, Hamilton, New Zealand.

[1] S. Bergman, *The Kernel Function and Conformal Mapping* (Am. Math. Soc., New York, 1950).

[2] N. Aronszajn, Trans. Am. Math. Soc. **68**, 337 (1950).

[3] J. R. Klauder, J. Math. Phys. **4**, 1055 (1963).

[4] J. McKenna and J. R. Klauder, J. Math. Phys. **5**, 878 (1964).

[5] V. Bargmann, Commun. Pure Appl. Math. **14**, 187 (1961).

[6] J. R. Klauder and E. C. G. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1960).

[7] See Ref. 1, p. 9.

[8] Z. Nehari, *Conformal Mapping* (McGraw-Hill, New York, 1952).

[9] K. Friedrichs, Trans. Am. Math. Soc. **41**, 321 (1937).

[10] V. I. Smirnov and N. A. Lebedev, *Functions of a Complex Variable—Constructive Theory* (M.I.T. Press, Cambridge, Mass., 1960).

[11] P. R. Garabedian, Pacific J. Math. **1**, 485 (1951).

[12] E. C. Lerner, H. W. Huang, and G. E. Walters, J. Math. Phys. **11**, 1679 (1970).

[13] We are thankful to the referee for pointing this out to us.

Percolation Processes. II. The Pair Connectedness*

JOHN W. ESSAM†
Applied Mathematics Department, Brookhaven National Laboratory
Upton, L.I., New York 11973

(Received 14 August 1970)

The probability that two sites of a crystal lattice belong to the same cluster (the pair connectedness) is shown to play an important role in percolation theory. Use of the linked-graph method to obtain low-density series expansions leads to the discussion of topological invariants for rooted graphs. These are related to the k-weights which arose in a previous investigation of the mean number of clusters. The mean size of clusters is related to the pair connectedness via sum rules.

1. INTRODUCTION AND DEFINITIONS

A discussion of the relevance of percolation theory to physics has been given in the review by Frisch and Hammersley.[1] We shall be more interested here in the graph theoretic aspects of the problem.

The pair connectedness in percolation theory is the analog of the pair correlation function in statistical mechanics. Just as summation of the spin correlation function over all pairs of lattice sites gives the susceptibility of an Ising ferromagnet the same summation of the pair connectedness[2] leads to the mean size of clusters.[3] In the absence of exact results, high-temperature series expansions of the susceptibility provide the most accurate determination of the Curie temperature.[4] The mean size of clusters has a similar strong divergence at the critical probability p_c and this led Domb[5] to suggest the use of low density expansions for its location. A fairly extensive study of these expansions was made by Sykes and Essam[6] with the conclusion that useful information about the critical region can be derived by this method but that the initial terms of the series are less regular than those for the susceptibility. Our long term objective will therefore be to extend the mean size expansions via the pair connectedness and also to study the latter in its own right. In the meantime a graph theoretic interpretation of the coefficients in the expansion will be given together with a prescription for finding the graph weights.

We begin by defining the functions of interest for a finite linear graph with the idea of proceeding to the limit of uniform infinite lattice graphs[7] for which the polynomials become infinite series. Two cases will be distinguished, the "site problem" and the "bond problem."

In the site problem, particles are distributed over vertices of the graph (sites of the lattice) subject to the constraint of at most one particle per site. The particles are otherwise independently distributed, the probability p of finding an occupied site being given. If $G = (V, E)$ is a linear graph with vertex set V and edge set E, then there is a possible state of the system for every subset V' of V, namely the one in which the vertices of V' are occupied but those of $V - V'$ are unoccupied. The assumption of independence means that the probability of occurrence of the state corresponding to V' is

$$\pi(V') = p^{|V'|}(1-p)^{|V-V'|}. \qquad (1.1)$$

The expectation value of any function of state $A(V', G)$ is given by

$$\langle A; G\rangle = \sum_{V'\subseteq V} \pi(V')A(V', G); \qquad (1.2)$$

for example, the mean value of the occupation number

$$\nu_i = \begin{cases} 1 & \text{for } i \in V' \\ 0 & \text{otherwise} \end{cases} \qquad (1.3)$$

of the ith vertex is

$$\langle \nu_i; G\rangle = p. \qquad (1.4)$$

It is also found that

$$\langle \nu_i\nu_j; G\rangle = p^2 \quad \text{for} \quad i \neq j, \qquad (1.5)$$

which is consistent with the assumption of independence. There is therefore no correlation between particles on different vertices, as there would be if they were interacting.

In the bond problem it is the edges of the graph (bonds of the lattice) which are occupied by particles, and the above description may be taken over by replacing V by E. (Notice that occupied and unoccupied by a particle is just a way of thinking of the state of the vertex or edge, and we could equally well use spin up–spin down, black–white, open–closed, depending on the application.[1]) The site problem is more general than the bond problem, since the bond problem on any graph is isomorphic with the site problem on its covering graph.[7] It is nevertheless useful to retain the distinction since a covering graph is more complex than the graph from which it was derived (for example, the honeycomb covering lattice is the Kagomé lattice) and also since some theorems which are easily formulated for the bond problem are more obscure (or may not ever be true) when transcribed to the site problem.

We have seen that the positions of the particles are uncorrelated, but just by chance they form clusters as a result of the restricted space. In the case of a crystal lattice, the proximity of two particles may be measured in the geometrical sense, but in the general problem on a graph it is indicated by the interrelation between vertices and edges. We define the connectedness indicator $\gamma_{v,v'}$ for two vertices $v, v' \in V$ by

$$\gamma_{v,v'} = \begin{cases} 1 & \text{if } v, v' \text{ are connected by a chain of occupied edges,} \\ 0 & \text{otherwise.} \end{cases} \quad (1.6)$$

For the purpose of this definition, an edge is considered occupied in the site problem if both vertices it connects are occupied; in particular, v and v' must be occupied. The expectation value of $\gamma_{v,v'}$ will be called the vertex–vertex connectedness or, briefly, the *pair connectedness*. If v and v' are the same vertex, the expectation value is p for the site problem and unity for the bond problem. Two related quantities are $\gamma_{e,e'}$ and $\gamma_{v,e}$; these are defined as in (1.6), but $e, e' \in E$ are to be terminal links of the chain. Their mean values will be known as the edge–edge and vertex–edge connectedness, respectively, and are related to the pair connectedness by

$$\langle \gamma_{v,e'}; G \rangle = p \langle \gamma_{v,v'}; G_e^\gamma \rangle \quad (1.7)$$

and

$$\langle \gamma_{e,e'}; G \rangle = p^2 \langle \gamma_{v,v'}; G_{e,e'}^\gamma \rangle, \quad (1.8)$$

where G_e^γ and $G_{e,e'}^\gamma$ are the graphs obtained from G by contracting[7] the edges e and e' to yield the vertices v and v'. Thus G_e^γ and $G_{e,e'}^\gamma$ have one and two less vertices than G, respectively.

It is convenient to define the *pair connectedness of a two-rooted*[7] *graph* G^{ii}, obtained from G by designating the vertices v and v' as root points, by

$$D(p; G^{\text{ii}}) = \langle \gamma_{v,v'}; G \rangle. \quad (1.9)$$

It was shown in Ref. 3 that, for the site problem,

$$D(p; G^{\text{ii}}) = \sum_m [c_m^{\text{ii}}; G^{\text{ii}}] D(c_m^{\text{ii}}) p^{v_m}, \quad (1.10)$$

where

$$D(c_m^{\text{ii}}) = \sum_j (-1)^{v_m - v_j} [c_j^{\text{ii}}; c_m^{\text{ii}}]^F. \quad (1.11)$$

Essentially this shows that, to obtain the coefficient of p^n in the polynomial $D(p; G^{\text{ii}})$, it is necessary to enumerate all connected two-rooted section graphs[7] C^{ii} of G^{ii} with n vertices, each of which contributes $D(C^{\text{ii}})$, its *strong pair connectedness weight*, to the coefficient. The graph c_i^{ii} is the ith member of a list of connected two-rooted graphs no two of which are isomorphic. The list must be complete in the sense that every connected two-rooted section graph of G^{ii} is isomorphic to some graph in the list, and $[c_m^{\text{ii}}; G^{\text{ii}}]$ is the number of such section graphs isomorphic with c_m^{ii}. The grouping together of isomorphic graphs in this way is of more value when working with an infinite lattice, and for this reason the square bracket quantities are called strong lattice constants.[7] To complete the explanation of the above formulas, v_m is the number of vertices in c_m^{ii}, and F restricts c_j^{ii} to be a section graph of c_m^{ii} having full vertex perimeter.[7] Similarly, for the bond problem it was shown that[3] the pair connectedness, which is now distinguished by a bar, may be written

$$\bar{D}(p; G^{\text{ii}}) = \sum_m (c_m^{\text{ii}}; G^{\text{ii}}) \bar{d}(c_m^{\text{ii}}) p^{e_m}, \quad (1.12)$$

where

$$\bar{d}(c_m^{\text{ii}}) = \sum_j (-1)^{e_m - e_j} (c_j^{\text{ii}}; c_m^{\text{ii}})^F. \quad (1.13)$$

These formulas are similar to those for the site problem, but the round brackets denote weak lattice constants (i.e., subgraphs are enumerated rather than section graphs); e_m is the number of edges in c_m^{ii} and F denotes full edge perimeter.[7]

Since every subgraph of a graph corresponds to a unique section graph with the same number of vertices (the one with the same vertex set), we may rewrite (1.10) in the form

$$D(p; G^{\text{ii}}) = \sum_m (c_m^{\text{ii}}; G^{\text{ii}}) d(c_m^{\text{ii}}) p^{v_m}, \quad (1.14)$$

where $d(c_m^{\text{ii}})$ denotes the *weak pair connectedness weight* of c_m^{ii}. We notice that if C^{ii} is a connected two-rooted graph, then, when $p = 1$, the roots must be connected so that $D(1; C^{\text{ii}}) = 1$, which leads to a recursion formula for the d's,

$$d(C^{\text{ii}}) = 1 - \sum_m{}' (c_m^{\text{ii}}; C^{\text{ii}}) d(c_m^{\text{ii}}), \quad (1.15)$$

where the prime denotes omission of the term $c_m^{\text{ii}} = C^{\text{ii}}$. For the bond problem, $\bar{D}(1; C^{\text{ii}}) = 1$ so that from (1.12) we see that $\bar{d}$ also satisfies (1.15) and therefore

$$\bar{d}(c_m^{\text{ii}}) = d(c_m^{\text{ii}}), \quad (1.16)$$

which allows us to express the pair connectedness for both bond and site problems in terms of the weak weights.

The strong weights are more useful than the weak weights for computational purposes on an infinite lattice, but the weak weights are of more interest from a graph-theoretic point of view since they are topological invariants of the graph (i.e., homeomorphic graphs[7] have the same weak weight). The latter property was established for the weak weights in the mean number expansion (k-weights) in the first paper of this series[8] (subsequently referred to as I).

If $E' \subseteq E$ is the set of occupied edges in a given state for the bond problem on a graph $G = (V, E)$, then the mean number of clusters is defined by

$$\bar{K}_0(p; G) = \langle n; G\rangle, \tag{1.17}$$

where $n(E', G)$ is the number of components in the subgraph $G' = (V, E')$. The zero subscript serves to remind us that isolated vertices of G' are counted as clusters (see I). In the site problem the mean number of clusters $K(p; G)$ is defined similarly, but G' is the section graph of G defined by the subset $V' \subseteq V$ of occupied vertices. In I it was shown that both mean number functions were determined by the weak k-weights

$$\bar{K}_0(p; G) = \sum_m (c_m; G)k(c_m)p^{e_m} \tag{1.18}$$

and

$$K(p; G) = \sum_m (c_m; G)k(c_m)p^{v_m}, \tag{1.19}$$

although again for computational purposes it is more useful to use the strong lattice constant expansion

$$K(p; G) = \sum_m [c_m; G]K(c_m)p^{v_m} \tag{1.20}$$

in the case of the site problem. The properties of the mean number weights have been discussed in I, but subsequently the weak weights were investigated independently by Crapo[9] in the more general context of matroids. Two useful properties which were obtained by Crapo but which did not appear in I are

$$|k(G^{\mathrm{D}})| = |k(G)| \tag{1.21}$$

(where G^{D} is a dual of the planar graph G) and

$$k(G) = k(G_e^\gamma) - k(G_e^\delta), \tag{1.22}$$

where G_e^γ and G_e^δ are the graphs obtained from G by contracting and deleting the edge e of G. The deletion of an edge from a connected graph may separate it into one or more components. In I the k-weights were defined for connected graphs only, but for a general graph we adopt the definition[7]

$$\sum_{E' \subseteq E} k(G') = n(G),$$

where $G' = (V, E')$. This relation may be inverted to yield

$$k(G) = \sum_{E' \subseteq E} (-1)^{|E-E'|} n(G'). \tag{1.23}$$

If the number of components $n(G)$ in G is greater than one, the k-weight is zero unless all but one of the components are isolated vertices. When G is the trivial graph with n vertices but no edges, (1.23) shows that $k = n$. For the graph G with one arbitrary connected component C and $n - 1$ isolated vertices, we find $k(G) = k(C)$. Combining these results, we see that the definition of k for connected graphs may be written (in agreement with I) as

$$\sum_{C' \subseteq G}' k(C') = -r(G),$$

where C' is a connected subgraph of G, the rank $r(G) = |V| - n(G)$, and the prime on the sum indicates omission of subgraphs with one vertex. By using k rather than $\beta = |k|$, as in Crapo's work, we see that Eq. (1.22) holds for any edge including loops.

The main result of this paper is to relate the mean number and pair connectedness weights. Theorems are then developed which enable the pair connectedness weights of a general two-rooted graph, and hence the mean number weights, to be expressed in terms of the pair connectedness weights of the elementary[7] two-rooted graphs. This work parallels that of Van Leeuwen, Groeneveld, and de Boer[10] on the pair correlation in an imperfect classical gas. The technique used is to develop theorems for the mean number and pair connectedness functions and then use the following proposition.

Proposition 1: Consider the weight factors in the weak and strong lattice constant expansions of the function $\langle A; G\rangle$. The weak weight of a graph g is the coefficient of $p^{e(g)}$ in the polynomial $\langle A; g\rangle$ for the bond problem, and the strong weight is the coefficient of $p^{v(g)}$ in the same polynomial for the site problem.

This results from the fact that the only subgraph of g with $e(g)$ edges is g itself and the only section graph with $v(g)$ vertices is g itself. The proposition was used in I, and as a further example of its use we note that (1.22) follows from the result of Kasteleyn and Fortuin[11] for the bond problem

$$\langle n; G\rangle = p\langle n; G_e^\gamma\rangle + (1 - p)\langle n; G_e^\delta\rangle. \tag{1.24}$$

The above authors also note that (1.24) is valid for the bond problem pair connectedness, and so

$$d(G^{\mathrm{ii}}) = d(G_e^{\mathrm{ii}\gamma}) - d(G_e^{\mathrm{ii}\delta}). \tag{1.25}$$

The rest of the paper is broken down as follows. Section 2 is concerned with the site problem, Sec. 3 deals with modifications required for the bond problem, and Sec. 4 relates the mean size of clusters to the pair connectedness. In the latter section the mean-size weight factors which, even in the weak case, are not topological invariants are expressed as a sum over the pair connectedness weights for all possible two-rootings.

2. SITE PROBLEM

A. Pair Connectedness

In general, the pair connectedness $D(p; G^{ii})$ is a polynomial of degree $v(G^{ii})$, but under certain conditions some of the coefficients vanish. The most obvious case is when the roots are adjacent. In this case $D(p; G^{ii}) = p^2$, and, applying Proposition 1, we have our first theorem.

Theorem 1: If G^{ii} is a two-rooted graph with three or more vertices and adjacent roots, then the strong pair connectedness weight is zero [i.e., $D(G^{ii}) = 0$].

Clearly, if D_n is the coefficient of p^n, then

$$D_n = 0 \quad \text{for} \quad 0 \leq n < v_{\min}, \tag{2.1}$$

where $v_{\min}$ is the number of vertices in the shortest chain connecting the roots of G^{ii}. A more useful result for the computation of weight factors is the following:

$$D_n = 0 \quad \text{for} \quad v_{\max} < n \leq v(G^{ii}), \tag{2.2}$$

where $v_{\max}$ is the number of vertices in the maximal 1-irreducible[7] two-rooted subgraph S^{ii} of G^{ii}. This results from the fact that the connectedness of the roots of G^{ii} is unaltered by changing the state of occupation of vertices not in S^{ii}, and so $D(p; G^{ii}) = D(p; S^{ii})$.

If G^{ii} itself is 1-irreducible, then we learn nothing about the polynomial from (2.2), but it may still be that $D(p; S^{ii})$ has degree less than $v(S^{ii})$. The characteristic of a 1-reducible two-rooted graph is that either it is not connected or, if it is connected, then there must be at least one articulation point (vertex, the deletion of which separates the graph into two or more components at least one of which has no roots). A 1-irreducible graph may still have an articulation set V_x of higher order, and, if there is such a set, the corresponding section graph of which is complete, then (2.2) may be generalized. Suppose that S_x^{ii} is the maximal two-rooted section graph of S^{ii} having no such articulation set; then $D(p; S^{ii}) = D(p; S_x^{ii})$. If V_x contains both roots, the result is trivial, but otherwise it follows by the previous argument that the occupation of vertices not in S_x^{ii} is irrelevant to the connectedness of the roots. Thus (2.2) holds with $v_{\max}$ equal to the number of vertices in S_x^{ii}. Using Proposition 1 now leads to the following theorem.

Theorem 2: If G^{ii} is a two-rooted graph which either is disconnected or contains an articulation set the section graph of which is complete, then its strong pair connectedness weight is zero.

In calculating strong pair connectedness weights, we are thus led to consider only 1-irreducible graphs, and even some of their weights may vanish by the above theorems. In a case where the weight is not expected to be zero, it is still possible to simplify the calculation if the graph may be formed by series–parallel combination of smaller graphs.

If S^{ii} is the 1-irreducible two-rooted graph obtained by identifying the roots of the graphs S_1^{ii} and S_2^{ii} (i.e., connecting them in parallel), then

$$1 - p^{-2}D(p; S^{ii}) = [1 - p^{-2}D(p; S_1^{ii})][1 - p^{-2}D(p; S_2^{ii})]. \tag{2.3}$$

This is true since $p^{-2}D$ is the probability that if the roots are occupied, then there is a chain of occupied vertices between the roots. The probability that there is no such chain is the product of the probabilities of no chain for each of the parallel components. The equation is trivially satisfied when the roots of either component are adjacent. A *composite* graph[7] $\mathcal{C}$ has at least two parallel components, but the root points are not adjacent. By repeated factorization and use of Proposition 1, it is possible to express the weight of a composite graph as a product over the weights of its *simple*[7] components. If $\mathcal{C}$ is the parallel combination of n simple two-rooted graphs $\mathcal{S}_1, \cdots, \mathcal{S}_n$, then

$$D(\mathcal{C}) = (-1)^{n+1} \prod_{t=1}^{n} D(\mathcal{S}_t). \tag{2.4}$$

The pair connectedness of a simple graph may be further factorized if it has a vertex through which all paths between the roots must pass. Such a vertex is called a nodal point; a simple two-rooted graph with one or more nodal points is called *nodal*.[7] Nodal graphs can thus be formed by series combination of smaller graphs. If S^{ii} is a 1-irreducible two-rooted graph obtained by series combination of S_1^{ii} and S_2^{ii}, then

$$D(p; S^{ii}) = p^{-1}D(p; S_1^{ii})D(p; S_2^{ii}). \tag{2.5}$$

The strong pair connectedness weight of a nodal graph $\mathcal{N}$, which is the series combination of n nonnodal graphs $S_1^{ii}, \cdots, S_n^{ii}$, may be written, by repeated use of (2.5) and then Proposition 1, as

$$D(\mathcal{N}) = \prod_{t=1}^{n} D(S_t^{ii}). \tag{2.6}$$

If the nonnodal graph S_t^{ii} is simple and therefore *elementary*,[7] no further reduction is possible, but, if it is composite, (2.4) may be applied again. Finally,

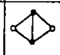
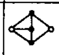
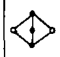
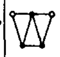
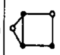
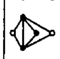
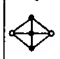
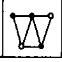
n	e_n	$D(e_n)$	$d(e_n)$	n	e_n	$D(e_n)$	$d(e_n)$
1		-1	+2	6		0	-3
2		0	+2	7		-1	-3
3		-1	+2	8		0	+4
4		0	-2	9		+1	+4
5		0	-2	10		+1	-6

FIG. 1. Pair connectedness weights for the elementary graphs with four and five vertices.

if it has a root-connecting edge or multi-edge, it contributes a factor p^2 to the pair connectedness and hence a factor of zero to the strong weight unless it is just a two-rooted multi-edge, in which case it contributes a factor of one to the weight. By repeated application of (2.4) and (2.6) the weight of any graph may be written, apart from a sign, as the product of the weights of its elementary constituents. These weights for graphs with five or less vertices are given in Fig. 1.

B. Relationship between the Pair Connectedness and Mean Number

Consider graphs G^{ii} and G_n, where G^{ii} is obtained by rooting two vertices of a graph G, and G_n is obtained by connecting the same vertices of G by a chain of n edges. It will be shown that

$$K(p; G_n) = K(p; G) + (n-1)p - np^2 + p^{n-1}D(p; G^{ii}) \quad \text{for} \quad n \geq 1. \tag{2.7}$$

As a first step we establish the result for $n = 1$ which involves finding the change in mean number when two vertices of G are connected by an additional edge. The only case in which a change occurs is when both roots are occupied but do not already belong to the same cluster. This happens with probability $p^2 - D(p; G^{ii})$ and reduces the number of clusters by one. Hence (2.7) is true with $n = 1$. Suppose now we wish to go from $n = 1$ to $n = 2$ by insertion of an additional vertex. The number of clusters is increased by one if either

(i) the inserted vertex is occupied and both roots are unoccupied [probability $p(1-p)^2$]

or

(ii) the inserted vertex is unoccupied, the roots are both occupied and do not belong to the same cluster on G {probability $(1-p)[p^2 - D(p; G^{ii})]$},

but is otherwise unchanged. Thus

$$K(p; G_2) = K(p; G_1) + p(1-p) - (1-p)D(p; G^{ii}) \tag{2.8}$$

and applying this result taking G to be G_{n-1} with one edge deleted from the chain

$$K(p; G_n) = K(p; G_{n-1}) + p(1-p) - (1-p)p^{n-2}D(p; G^{ii}) \quad \text{for} \quad n \geq 2. \tag{2.9}$$

Iterating (2.9), we find

$$K(p; G_n) = K(p; G_1) + (n-1)p(1-p) - (1-p^{n-1})D(p; G^{ii}) \quad \text{for} \quad n \geq 1, \tag{2.10}$$

which together with the result for $n = 1$ establishes (2.7) for $n > 1$. In fact, (2.7) holds for $n = 0$, but the proof is omitted.

Using Proposition 1, we find, from (2.7) the following relationships involving the strong K-weights:

$$K(G_1) = K(G) + D(G^{ii}) \quad \text{for} \quad v(G) > 2, \tag{2.11}$$

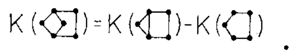
$$K(G_n) = D(G^{ii}) \quad \text{for} \quad n \geq 2 \quad \text{and} \quad v(G) > 2. \tag{2.12}$$

From (2.12) it follows that the strong mean number weights are independent of the number of edges in a given bridge[7] provided that there are at least two.

Combining (2.11) and (2.12) results in the following relation between K-weights:

$$K(G_n) = K(G_1) - K(G) \quad \text{for} \quad n \geq 2 \quad \text{and} \quad v(G) > 2 \tag{2.13}$$

e.g.,



$$-1 = 0 - +1$$

In practice, strong pair connectedness weights are determined by breaking down the graph into its elementary constituents and then using (2.11) or (2.12) to determine the weights of the elementary graphs from a list of K-weights. Such a list will eventually be much shorter than a list of pair connectedness weights since there are far more two-rooted graphs. Equations (2.11) and (2.12) are sometimes useful in hand computation of K-weights as is (2.13).

3. BOND PROBLEM

This section runs parallel to the previous section, and for this reason detailed arguments will not always be given. Formulas for weak pair connectedness weights will be obtained via the bond problem functions and Proposition 1, but, as we saw in the Introduction, they also determine the site problem pair connectedness.

A. Pair Connectedness

Theorem 1 is no longer valid for weak weights, and Theorem 2 holds only for disconnected graphs and graphs with articulation sets of order one. This is because even when G^{ii} has an articulation set V_x of order greater than two corresponding to a complete section graph, the vertices of V_x can be connected either directly by an occupied edge or indirectly by a chain of occupied edges which lies outside S_x^{ii}. The latter edges are therefore relevant to the connectedness of the roots in the case of the bond problem.

Equation (2.3) takes on a simpler form for the bond problem since the occupation of the roots need not be considered:

$$1 - \bar{D}(p; S^{\text{ii}}) = [1 - \bar{D}(p; S_1^{\text{ii}})][1 - \bar{D}(p; S_2^{\text{ii}})]. \tag{3.1}$$

Adjacent roots no longer give a trivial result so that we now consider the factorization of the weak weight of a ladder graph[7] $\mathfrak{L}$, which is similar to a composite graph, but the roots may be connected by a multi-edge. If e is the multiplicity of the multi-edge, it contributes a factor $(1 - p)^e$ to $1 - \bar{D}$ and so

$$d(\mathfrak{L}) = (-1)^{e+n+1} \prod_{t=1}^{n} d(\mathcal{S}_t), \tag{3.2}$$

where $\mathfrak{L}$ is the parallel combination of n simple graphs and e edges. In particular a two-rooted multi-edge of multiplicity e has weak weight $(-1)^{e+1}$.

The rule for series combination is also very simple, that is,

$$\bar{D}(p; S^{\text{ii}}) = \bar{D}(p; S_1^{\text{ii}})\bar{D}(p; S_2^{\text{ii}}). \tag{3.3}$$

Using Proposition 1, we may write the weak weight of a nodal graph $\mathcal{N}$ as the product of the weights for its nonnodal constituents:

$$d(\mathcal{N}) = \prod_{t=1}^{n} d(S_t^{\text{ii}}). \tag{3.4}$$

The nonnodal graphs this time fall into three classes, elementary, ladder (includes multi-edge of multiplicity two or more), and single edge with two roots (contributes a factor of unity). The ladder graph contributions may be further factorized using (3.2) and, as before, repeated use of (3.2) and (3.4) enables the weight of any graph to be expressed as a product of the weights of elementary graphs, the first few of which are listed in Fig. 1 beside the strong weights.

B. Relationship between the Pair Connectedness and Mean Number

The basic formula is almost the same as (2.7):

$$\bar{K}_0(p; G_n) = \bar{K}_0(p; G) + (n - 1) - np + p^n \bar{D}(p; G^{\text{ii}}) \quad \text{for} \quad n \geq 0. \tag{3.5}$$

The case $n = 0$ is easily proven since G_0 is formed by identifying the roots of G^{ii} and a cluster is lost whenever the roots of G^{ii} are not already connected by a chain of occupied edges [probability $1 - \bar{D}(p; G^{\text{ii}})$]. The result

$$\bar{K}_0(p; G_n) = \bar{K}_0(p; G_{n-1}) + (1 - p) \times [1 - p^{n-1}\bar{D}(p; G^{\text{ii}})], \tag{3.6}$$

which allows the length of the chain to be increased, follows from the fact that insertion of a new link only changes the mean number if it is unoccupied and then only if its incident vertices are not already connected through the rest of the graph [probability $1 - p^{n-1}\bar{D}(p; G^{\text{ii}})$]. Iteration of (3.6) yields

$$\bar{K}_0(p; G_n) = \bar{K}_0(p; G_0) + n(1 - p) - (1 - p^n)\bar{D}(p; G^{\text{ii}}) \quad \text{for} \quad n \geq 0, \tag{3.7}$$

which together with the result for $n = 0$ establishes (3.5) for general n.

Using Proposition 1 on (3.5) gives the basic relation between the weak mean number and pair connectedness weights

$$d(G^{\text{ii}}) = k(G_n) \quad \text{for} \quad e(G^{\text{ii}}) \geq 1 \quad \text{and} \quad n \geq 1. \tag{3.8}$$

Since (3.8) is true for all n, the k-weights of homeomorphic graphs are equal (see also I), and consequently the same equation implies that $d(G^{\text{ii}})$ is also a topological invariant of G^{ii}. Theorem IV of I also follows from (3.8) and the previous result that $d(G^{\text{ii}})$ is zero unless G^{ii} is 1-irreducible. As an example of (3.8), we have

d() = k() = 2.

C. Generalization of the Contraction-Deletion Rule (1.23) and the Effect of Edge Substitution on Weak Weights

Consider the graph

$$G_{e \to G_1{}^{\text{ii}}}$$

obtained from G by replacing the edge e of G by the two-rooted graph G_1^{ii}. A simple, but nevertheless useful, extension of (1.24) is

$$\bar{K}(p; G_{e \to G_1{}^{\text{ii}}}) = \bar{D}(p; G_1^{\text{ii}})\bar{K}(p; G_e^{\gamma}) + [1 - \bar{D}(p; G_1^{\text{ii}})]\bar{K}(p; G_e^{\delta}). \tag{3.9}$$

Using Proposition 1, we find

$$k(G_{e \to G_1{}^{\text{ii}}}) = d(G_1^{\text{ii}})[k(G_e^{\gamma}) - k(G_e^{\delta})] = d(G_1^{\text{ii}})k(G), \tag{3.10}$$

where we have used (1.22) to introduce $k(G)$. Thus the effect of replacing an edge of G by G_1^{ii} is to multiply the weak k-weight by the weak pair-connectivity weight of G^{ii}. This together with (3.8) enables the k-weight of any 2-reducible graphs to be expressed in terms of the k-weights of graphs with fewer vertices. A similar result is valid for the pair connectivity, namely

$$\bar{D}(p; G^{\text{ii}}_{e\to G_1}{}^{\text{ii}}) = \bar{D}(p; G_1^{\text{ii}})\bar{D}(p; G_e^{\text{ii}\gamma}) + [1 - \bar{D}(p; G_1^{\text{ii}})]\bar{D}(p; G_e^{\text{ii}\delta}), \tag{3.11}$$

which leads to

$$d(G^{\text{ii}}_{e\to G_1}{}^{\text{ii}}) = d(G_1^{\text{ii}})d(G^{\text{ii}}). \tag{3.12}$$

For example, graphs 2, 3, 4, and 5 in Fig. 1 are 2-reducible. The d-weights of 2 and 3 may be obtained from 1 by replacing an edge by a chain of length two. Similarly replacing the edges of 1 by a two-rooted triangle, for which $d = -1$, yields the d-weights of 4 and 5. Finally, by taking G_1^{ii} in (3.10) and (3.12) to be a two-rooted double bond,[7] for which $d = -1$, we find that doubling an edge changes the sign of the weak weights.

4. THE MEAN SIZE OF CLUSTERS AND HIGHER MOMENTS

A. Definitions

The mean number of clusters was defined in (1.17) as the mean number of components in the graph G', which is either the subgraph defined by the occupied edges or the section graph defined by the occupied vertices depending on the problem. We now define moments of the cluster size distribution by

$$M_{rs}(p; G) = \left\langle \sum_{i=1}^{n} e_i^r v_i^s; G \right\rangle, \tag{4.1}$$

where e_i and v_i are the numbers of edges and vertices in the ith component of G'. The bond problem moments are denoted by $\bar{M}_{rs}$. Clearly M_{00} is the mean number of clusters previously discussed. M_{01} and M_{10} are the mean number of vertices and edges respectively, and are simple functions

$$M_{01} = |V|\,p, \quad M_{10} = |E|\,p^2,$$
$$\bar{M}_{01} = |V|, \quad \bar{M}_{10} = |E|\,p. \tag{4.2}$$

The second moments yield various measures of the mean size of clusters. If the size of a cluster is defined by vertex content, then we may either choose a vertex, calculate the mean size of clusters containing that vertex, and average over all vertices, which gives

$$|V|^{-1} M_{02}(p; G), \tag{4.3}$$

or choose an edge, calculate the mean size of clusters containing that edge, and average over all edges, which gives

$$|E|^{-1} M_{11}(p; G). \tag{4.4}$$

Alternatively, it may be useful to measure size by edge content, in which case the two methods of computation yield

$$|V|^{-1} M_{11}(p; G) \quad \text{and} \quad |E|^{-1} M_{20}(p; G), \tag{4.5}$$

respectively. To establish contact with previous notation,[6]

$$S(p)^S = M_{02}(p; G)/M_{10}(p; G) \tag{4.6}$$

and

$$S(p)^B = \bar{M}_{20}(p; G)/\bar{M}_{10}(p; G). \tag{4.7}$$

B. Relationship between Mean Size and Pair Connectedness

It was shown in Ref. 3 that $S(p)^S$ could be expressed as a sum over the pair connectedness. This sum rule may be generalized to all second moments for site and bond problems. The essential observation is that

$$\sum_{\substack{v,v'\in V\\ v\neq v'}} \gamma_{v,v'} = \sum_i v_i(v_i - 1), \tag{4.8}$$

which when averaged gives

$$M_{02}(p; G) = M_{01}(p; G) + 2 \sum_t D(p; G_t^{\text{ii}}), \tag{4.9}$$

where the sum is over all two rootings of G. Similarly,

$$\sum_{\substack{v\in V\\ e\in E}} \gamma_{v,e} = \sum_i v_i e_i, \tag{4.10}$$

so that

$$M_{11}(p; G) = \sum_{\substack{v\in V\\ e\in E}} \langle \gamma_{v,e}; G\rangle \tag{4.11}$$

$$= p \sum_{v,e} \langle \gamma_{v,v'}; G_e^{\gamma}\rangle, \tag{4.12}$$

which is again a sum over the pair connectedness, and finally

$$M_{20}(p; G) = M_{00}(p; G) + \sum_{\substack{e,e'\in E\\ e\neq e'}} \langle \gamma_{e,e'}; G\rangle \tag{4.13}$$

$$= M_{10}(p; G) + p^2 \sum_{\substack{e,e'\in E\\ e\neq e'}} \langle \gamma_{v,v'}, G^{\gamma}_{e,e'}\rangle. \tag{4.14}$$

All the above equations hold for the corresponding bond problem quantities.

The equations may be generalized to higher moments; for example,

$$M_{03}(p; G) = M_{01}(p; G) + 6 \sum D(p; G^{\text{ii}}) + 6 \sum D(p; G^{\text{iii}}) \tag{4.15}$$

and

$$M_{04}(p; G) = M_{01}(p; G) + 14 \sum D(p; G^{\text{ii}}) + 36 \sum D(p; G^{\text{iii}}) + 24 \sum D(p; G^{\text{iv}}), \tag{4.16}$$

where $D(p; G^{\text{iii}})$ and $D(p; G^{\text{iv}})$ are the connectedness functions appropriate to three- and four-rooted graphs and the sums are over all possible rootings.

C. Graphical Expansions

1. Strong Lattice Constant Expansion for the Site Problem

Substitution of (1.10) in (4.9) provides a graphical expansion for $M_{02}(p; G)$:

$$M_{00}(p; G) = |V|\, p + 2 \sum_t \sum_m [c_m^{\text{ii}}; G_t^{\text{ii}}] D(c_m^{\text{ii}}) p^{v_m} \tag{4.17}$$

$$= \sum_r [c_r; G] M_{02}(c_r) p^{v_r}, \tag{4.18}$$

where in going from (4.17) to (4.18) we have grouped together contributions from c_m^{ii} which are isomorphic with c_r when the roots are ignored. By applying Proposition 1, the weight functions may be related

$$M_{02}(G) = \begin{cases} 2 \sum_t D(G_t^{\text{ii}}) & \text{for } |V| > 1 \\ 1 & \text{for } |V| = 1 \end{cases}. \tag{4.19}$$

The sum over rootings in (4.19) is more conveniently replaced by a sum over 1-irreducible two-rooted graphs no two of which are isomorphic:

$$M_{02}(G) = 2 \sum_t ((s_t^{\text{ii}}; G)) D(s_t^{\text{ii}}) \tag{4.20}$$

where $((s_t^{\text{ii}}; G))$ is the number of rootings[7] of G isomorphic with s_t^{ii}. A similar analysis of M_{11} and M_{20} via Eqs. (4.12) and (4.14) gives the following expressions for the strong weights:

$$M_{11}(G) = \sum_t D(H_t^{\text{ii}}) \quad \text{for} \quad |V| > 2, \tag{4.21}$$

where H_t^{ii} is obtained by rooting any vertex and any nonincident edge of G and then contracting the edge. We have

$$M_{20}(G) = 2 \sum_t D(H_t^{\text{ii}}) \quad \text{for} \quad |V| > 3, \tag{4.22}$$

where H_t^{ii} is obtained by choosing a pair of nonincident edges of G and contracting. Each pair is counted once only.

2. Weak Lattice Constant Expansions for Site and Bond Problems

Following the arguments of I, we may expand the moments for site and bond problems in terms of weak lattice constants. Thus

$$M_{rs}(p; G) = \sum_t (c_t; G) m_{rs}(c_t) p^{v_t} \tag{4.23}$$

and

$$\bar{M}_{rs}(p; G) = \sum_t (c_t; G) \bar{m}_{rs}(c_t) p^{e_t}. \tag{4.24}$$

Now for a connected graph C we have

$$M_{rs}(1; G) = \bar{M}_{rs}(1; G) = |E|^r |V|^s, \tag{4.25}$$

and so in analogy with (1.15) we obtain the recursive definition of the site-problem weak weights

$$m_{rs}(C) = |E|^r |V|^s - \sum_t{}' (c_t; C) m_{rs}(c_t). \tag{4.26}$$

Clearly, $\bar{m}_{rs}(c_t)$ satisfies the same equation, and so

$$\bar{m}_{rs}(c_t) = m_{rs}(c_t). \tag{4.27}$$

There is therefore only one set of weak weights to be determined, and these will be derived from bond problem formulas together with Proposition 1. Substitution of (1.12) into the barred versions of (4.9), (4.12), and (4.14) enables the relations

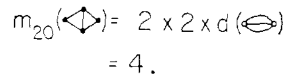
$$m_{02}(G) = 2 \sum d(G_t^{\text{ii}}) \quad \text{for} \quad |E| > 2 \tag{4.28}$$

$$m_{11}(G) = \sum_t d(H_t^{\text{ii}}) \quad \text{for} \quad |E| > 1, \tag{4.29}$$

and

$$m_{20}(G) = 2 \sum_t d(H_t^{\text{ii}}) \quad \text{for} \quad |E| > 2 \tag{4.30}$$

to be obtained; e.g.,



The sums in (4.29) and (4.30) are over the same graphs as in (4.21) and (4.22).

5. CONCLUSION

The pair connectedness, which is an interesting concept in its own right, has been shown to provide an important link between the previously discussed concepts of mean number and mean size. It is hoped that the formulas of Sec. 4, which relate the mean size weights to the pair connectedness weights and hence to the mean number weights, will enable a significant extension of the mean size expansions to be made.

* Work performed under the auspices of the U.S. Atomic Energy Commission.

† Present address: Westfield College, University of London, Hampstead, London N.W.3, England.

[1] H. L. Frisch and J. M. Hammersley, J. Soc. Ind. Appl. Math. **11**, 894 (1963).

[2] In Ref. 2 the same function was known as the "pair connectivity." Professor F. Harary pointed out that connectivity is widely used by graph theorists in a different sense and suggested the use of "connectedness" in the present context.

[3] J. W. Essam, Proc. Cambridge Phil. Soc. **67**, 523 (1970).

[4] C. Domb and M. F. Sykes, J. Math. Phys. 2, 63 (1961).
[5] C. Domb, Nature 184, 509 (1959).
[6] M. F. Sykes and J. W. Essam, Phys. Rev. A133, 310 (1964).
[7] J. W. Essam and M. E. Fisher, Rev. Mod. Phys. 42, 271 (1970). This reference is a list of definitions to which we shall constantly refer.
[8] J. W. Essam and M. F. Sykes, J. Math. Phys. 7, 1573 (1966).
[9] H. H. Crapo, J. Combinatorial Theory 2, 406 (1967).
[10] J. M. J. Van Leeuwen, J. Groeneveld, and J. de Boer, Physica 25, 792 (1959).
[11] P. W. Kasteleyn and C. M. Fortuin, J. Phys. Soc. Japan, Suppl. 26, 11 (1969).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Discrete Lorentz and Singleton Representations of the Universal Covering Group of the 3 + 2 de Sitter Group

Leon Jaffe
Center for Particle Theory, Department of Physics, University of Texas, Austin, Texas 78712

(Received 26 May 1970)

We consider the relationship between the discrete reduction of $\bar{S}O(3, 2)$ with respect to $\bar{S}O(3, 1)$ and the singleton reduction of $\bar{S}O(3, 2)$ with respect to $\bar{S}O(3) \otimes \bar{S}O(2)$.

INTRODUCTION

The unitary irreducible representations (UIR's) of $S = \bar{S}O(3, 2)$ [$\bar{S}O(p, q)$ denotes the universal covering group of $SO(p, q)$] have been considered in Ref. 1 for those UIR's which have a singleton reduction with respect to its maximal pseudocompact subgroup $K = \bar{S}O(3) \otimes \bar{S}O(2)$. [A singleton reduction of a representation of group with respect to a subgroup means that each irreducible representation of the subgroup occurs once only in the reduction.] These will simply be referred to as singleton UIR's.

However, for many physical applications, we are interested in S, for example, because it contains L_+, the covering group of the proper Lorentz group [$L_+ \cong \bar{S}O(3, 1)$], as a subgroup. It therefore becomes useful to know the reduction of representations of S with respect to L_+. This has been done by the author[2] for unitary and nonunitary irreducible representations of S with a discrete singleton reduction with respect to L_+, which we shall refer to as the "discrete Lorentz" representations of S.

We arrive at the remarkable result that all the discrete Lorentz UIR's are nothing but singleton representations [i.e., with a singleton reduction with respect to K]. There is no *a priori* reason for supposing that this should be so.

We thus independently arrive at many of the singleton UIR's obtained by Ehrman. In fact, we obtain all those singleton UIR's in which for each value of the angular momentum the eigenvalues of the generator of $\bar{S}O(2)$, Γ_0, are bounded.

There is reason to believe that a discrete reduction of UIR's of S with respect to L_+ must of necessity be a singleton reduction.[3] If this is correct, then we have considered all UIR's of S with a discrete reduction with respect to L_+. It follows that all other UIR's of S have a nondiscrete reduction with respect to L_+. Hence, all the nonsingleton UIR's of S do not have a discrete reduction with respect to $\bar{S}O(3, 1)$. This applies also to those singleton representations in which Γ_0 has an unbounded spectrum of eigenvalues within each angular momentum subspace.

The discrete reduction of the UIR's of S with respect to L_+ is quite simple. We give a brief review of this. (The complete analysis can be found in Ref. 2, which includes nonunitary representations as well.) The bases of the representation spaces of the restricted class of irreducible representations we thus obtain are diagonal with respect to the Casimir operators of L_+. We effect a similarity transformation which takes this "Lorentz" basis into a "maximal compact" basis that is diagonal with respect to the Casimir operators of K, and determine the spectrum of eigenvalues of Γ_0 in each subspace with definite angular momentum. This determines the reduction with respect to K. This procedure is in many ways simpler than a direct determination of the reduction as in Ref. 1.

Class II(c) and V representations (see text) are the Majorana representations.

Class IV(a) and IV(b) representations (see text) provide a nautral generalization of the finite nonunitary Dirac representation to unitary representation

[4] C. Domb and M. F. Sykes, J. Math. Phys. **2**, 63 (1961).
[5] C. Domb, Nature **184**, 509 (1959).
[6] M. F. Sykes and J. W. Essam, Phys. Rev. **A133**, 310 (1964).
[7] J. W. Essam and M. E. Fisher, Rev. Mod. Phys. **42**, 271 (1970). This reference is a list of definitions to which we shall constantly refer.
[8] J. W. Essam and M. F. Sykes, J. Math. Phys. **7**, 1573 (1966).
[9] H. H. Crapo, J. Combinatorial Theory **2**, 406 (1967).
[10] J. M. J. Van Leeuwen, J. Groeneveld, and J. de Boer, Physica **25**, 792 (1959).
[11] P. W. Kasteleyn and C. M. Fortuin, J. Phys. Soc. Japan, Suppl. **26**, 11 (1969).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Discrete Lorentz and Singleton Representations of the Universal Covering Group of the 3 + 2 de Sitter Group

LEON JAFFE
Center for Particle Theory, Department of Physics, University of Texas, Austin, Texas 78712

(Received 26 May 1970)

We consider the relationship between the discrete reduction of $\bar{S}O(3, 2)$ with respect to $\bar{S}O(3, 1)$ and the singleton reduction of $\bar{S}O(3, 2)$ with respect to $\bar{S}O(3) \otimes \bar{S}O(2)$.

INTRODUCTION

The unitary irreducible representations (UIR's) of $S = \bar{S}O(3, 2)$ [$\bar{S}O(p, q)$ denotes the universal covering group of $SO(p, q)$] have been considered in Ref. 1 for those UIR's which have a singleton reduction with respect to its maximal pseudocompact subgroup $K = \bar{S}O(3) \otimes \bar{S}O(2)$. [A singleton reduction of a representation of group with respect to a subgroup means that each irreducible representation of the subgroup occurs once only in the reduction.] These will simply be referred to as singleton UIR's.

However, for many physical applications, we are interested in S, for example, because it contains L_+, the covering group of the proper Lorentz group [$L_+ \cong \bar{S}O(3, 1)$], as a subgroup. It therefore becomes useful to know the reduction of representations of S with respect to L_+. This has been done by the author[2] for unitary and nonunitary irreducible representations of S with a discrete singleton reduction with respect to L_+, which we shall refer to as the "discrete Lorentz" representations of S.

We arrive at the remarkable result that all the discrete Lorentz UIR's are nothing but singleton representations [i.e., with a singleton reduction with respect to K]. There is no *a priori* reason for supposing that this should be so.

We thus independently arrive at many of the singleton UIR's obtained by Ehrman. In fact, we obtain all those singleton UIR's in which for each value of the angular momentum the eigenvalues of the generator of $\bar{S}O(2)$, Γ_0, are bounded.

There is reason to believe that a discrete reduction of UIR's of S with respect to L_+ must of necessity be a singleton reduction.[3] If this is correct, then we have considered all UIR's of S with a discrete reduction with respect to L_+. It follows that all other UIR's of S have a nondiscrete reduction with respect to L_+. Hence, all the nonsingleton UIR's of S do not have a discrete reduction with respect to $\bar{S}O(3, 1)$. This applies also to those singleton representations in which Γ_0 has an unbounded spectrum of eigenvalues within each angular momentum subspace.

The discrete reduction of the UIR's of S with respect to L_+ is quite simple. We give a brief review of this. (The complete analysis can be found in Ref. 2, which includes nonunitary representations as well.) The bases of the representation spaces of the restricted class of irreducible representations we thus obtain are diagonal with respect to the Casimir operators of L_+. We effect a similarity transformation which takes this "Lorentz" basis into a "maximal compact" basis that is diagonal with respect to the Casimir operators of K, and determine the spectrum of eigenvalues of Γ_0 in each subspace with definite angular momentum. This determines the reduction with respect to K. This procedure is in many ways simpler than a direct determination of the reduction as in Ref. 1.

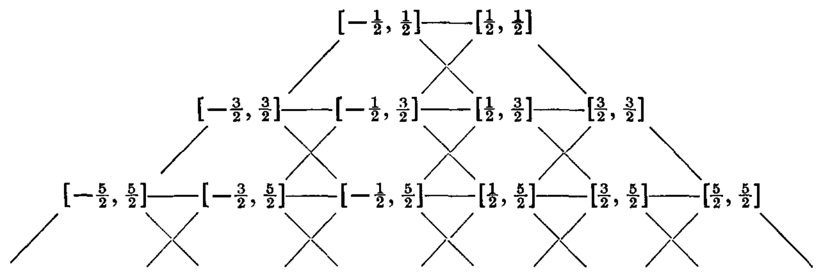
Class II(c) and V representations (see text) are the Majorana representations.

Class IV(a) and IV(b) representations (see text) provide a nautral generalization of the finite nonunitary Dirac representation to unitary representation for half-integral and integral spin, respectively. The Dirac representation has the following reduction with respect to K:

$$[-\tfrac{1}{2},\tfrac{1}{2}] \text{——} [\tfrac{1}{2},\tfrac{1}{2}],$$

where $[j,\mu]$ indicates a representation of $\bar{S}O(3)\otimes\bar{S}O(2)$, j being the angular momentum [i.e., $\bar{S}O(3)$ label], and μ is the value of Γ_0 [i.e., the $\bar{S}O(2)$ label]. $[j,\mu]$ —— $[j',\mu']$ indicates that the representations $[j,\mu]$ and $[j',\mu']$ are "interlocked"—i.e., J_{0i} has nonzero matrix elements connecting these representations.

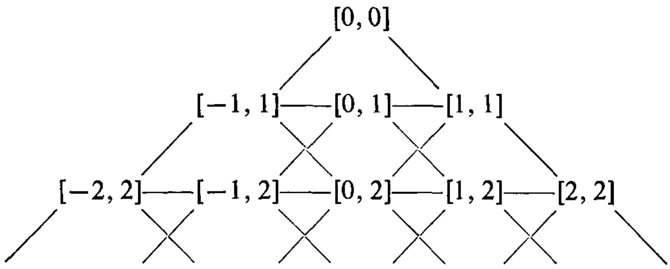
Class IV(a) representations have the following reduction:



etc.

Unlike the Dirac representation this contains all half-integral spins.

Class IV(b) representations have the following reduction:



etc.

This representation is a natural counterpart for integral spin of the above representation.

It may be useful to investigate these as representing infinite towers of particles with the same internal quantum numbers, but different spin.[4]

For these two classes of representations, we determine the expansion coefficients of the similarity transformation connecting the Lorentz and maximal compact bases. We also determine explicitly the matrix elements of all the generators on both these bases.

1. COMMUTATION RELATIONS

The Lie algebra of $S=\bar{S}O(3,2)$ is given by the following commutation relations:

$$[J_{\alpha\beta},J_{\gamma\delta}]=i(g_{\alpha\delta}J_{\beta\gamma}+g_{\beta\gamma}J_{\alpha\delta}-g_{\alpha\gamma}J_{\beta\delta}-g_{\beta\delta}J_{\alpha\gamma}),$$

where the metric $g_{\alpha\beta}$ is given by

$$\begin{aligned} g_{\alpha\beta} &= 1 && \text{for } \alpha=\beta=0 \text{ or } 4,\\ &= -1 && \text{for } \alpha=\beta=1,2,3,\\ &= 0 && \text{otherwise}, \end{aligned}$$

and $J_{\alpha\beta}=-J_{\beta\alpha}$, $\alpha,\beta=0,1,2,3,4$, are the generators.

If we make the identification

$$\Gamma_\mu = J_{\mu 4},\quad \mu=0,1,2,3,$$

we obtain

$$[J_{\mu\nu},J_{\rho\sigma}]=i(g_{\mu\sigma}J_{\nu\rho}+g_{\nu\rho}J_{\mu\sigma}-g_{\mu\rho}J_{\nu\sigma}-g_{\nu\sigma}J_{\mu\rho}), \tag{1a}$$

$$[\Gamma_\mu,J_{\rho\sigma}]=i(g_{\mu\rho}\Gamma_\sigma-g_{\mu\sigma}\Gamma_\rho), \tag{1b}$$

$$[\Gamma_\mu,\Gamma_\nu]=-iJ_{\mu\nu}, \tag{1c}$$

where the range of the indices μ, ν, ρ, and σ is 0, 1, 2, 3.

$J_{\mu\nu}$ are the generators of the proper Lorentz group and Γ_μ is a Lorentz vector, its vector properties being defined by (b). Relation (c) is one of the simplest to close the Lie algebra with only $J_{\mu\nu}$ and Γ_μ as basis elements.

2. LORENTZ BASIS

The Casimir operators of $\bar{S}O(3,2)$ are

$$\tilde{D}_1=-\tfrac{1}{2}J_{\alpha\beta}J^{\alpha\beta},\quad \tilde{D}_2=+W_\alpha W^\alpha,$$

where $W_\alpha = \frac{1}{8}\epsilon_{\alpha\beta\gamma\delta\epsilon}J^{\beta\gamma}J^{\delta\epsilon}$. The range of the indices is 0, 1, 2, 3, 4, and $\epsilon_{\alpha\beta\gamma\delta\epsilon}$ is the antisymmetric tensor with $\epsilon_{01234} = +1$. Hence we obtain

$$\tilde{D}_1 = -(\Gamma_\mu\Gamma^\mu + \tilde{C}_1), \quad (2a)$$

$$\tilde{D}_2 = W_\mu W^\mu + \tilde{C}_2^{\,2}, \quad (2b)$$

where

$$\tilde{C}_1 = \tfrac{1}{2}J_{\mu\nu}J^{\mu\nu}, \quad (3a)$$

$$\tilde{C}_2 = \tfrac{1}{8}\epsilon_{\mu\nu\rho\sigma}J^{\mu\nu}J^{\rho\sigma}, \quad (3b)$$

$$W_\mu = \tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma}J^{\nu\rho}\Gamma^\sigma, \quad (4)$$

and $\epsilon_{\mu\nu\rho\sigma}$ is the Levi-Cevita tensor with $\epsilon_{0123} = +1$. $\tilde{C}_1$ and $\tilde{C}_2$ are the Casimir operators of L_+, and an irreducible representation of L_+ is characterized by (k, c), where

$$C_1 = k^2 + c^2 - 1; \quad C_2 = ikc$$

are the eigenvalues of $\tilde{C}_1$ and $\tilde{C}_2$ respectively. $|k|$ specifies the minimum value of the angular momentum in the reduction of (k, c) with respect to $\bar{S}O(3)$. Note that (k, c) and $(-k, -c)$ specify the same irreducible representation.

The eigenvalues D_1 and D_2 of $\tilde{D}_1$ and $\tilde{D}_2$ are not sufficient in general to label an irreducible representation of S. We need additional labels, D_0 say, and label the representation by $\mathsf{D} = (D_0, D_1, D_2)$.

An irreducible representation D of S generates *ipso facto* a representation $\tau(\mathsf{D})$ of L_+, obtained by restricting ourselves to those operators which correspond to L_+ transformations. We consider those representations of S in which $\tau(\mathsf{D})$ can be decomposed into a direct sum of irreducible representations τ of L_+, each τ occuring once only in the reduction (such a reduction will be called a singleton reduction with respect to L_+). If τ is a component of the decomposition of $\tau(\mathsf{D})$, we shall write $\tau \in \mathsf{D}$.

We decompose τ in the usual fashion[5–7] with respect to the rotation group and obtain $|\mathsf{D}\,\tau jm\rangle$ as a basis of the representation space of τ, where $j(j+1)$ and m are the eigenvalues of $\mathrm{J}^2 = \frac{1}{2}J_{ij}J^{ij}$, $i, j = 1, 2, 3$, and $J_3 = J_{12}$, respectively. We put $\tau = (k, c)$ in accordance with the labeling defined above.

The basis $|\mathsf{D}\,\tau jm\rangle$ extends to a basis of D. If $\tau(\mathsf{D})$ has a nondiscrete reduction into irreducible representations of L_+, then the action of Γ_μ on the Lorentz basis $|\mathsf{D}\,\tau jm\rangle$ will be singular, and we cannot use the usual infinitesimal approach used below. (See the introduction of Ref. 2.) We, therefore, restrict ourselves to those UIR's of S with a discrete singleton reduction with respect to L_+.

We can, hence, write the action of Γ_0 in a well-defined fashion as

$$\Gamma_0\,|\,\mathsf{D}\,\tau jm\rangle = \sum_{\tau'\in\mathsf{D}} |\,\mathsf{D}\,\tau' jm\rangle\langle\tau'\,|\,\tau\rangle_j, \quad (5)$$

where we have used the fact that Γ_0 is a rotation scalar operator. The coefficients $\langle\tau'\,|\,\tau\rangle_j$ depend on D, but for notational convenience we omit this dependence.

From the Lorentz vector properties (1b) of Γ_μ, we obtain[5,7] for $(k, c) \in \mathsf{D}$

$$\langle k'c'\,|\,kc\rangle_j = 0 \quad \text{unless} \quad (k', c') \equiv (k \pm 1, c) \quad \text{or} \quad (k', c') \equiv (k, c \pm 1),$$

where

$$(k', c') \equiv (k, c) \quad \text{iff} \quad (k', c') = \pm(k, c).$$

The j dependence of $\langle\tau'\,|\,\tau\rangle_j$ is given by

$$\langle k \pm 1, c\,|\,kc\rangle_j = \langle k \pm 1, c\,|\,kc\rangle[(j \mp k)(j \pm k + 1)]^{\frac{1}{2}}, \quad (6a)$$

$$\langle k, c \pm 1\,|\,kc\rangle_j = \langle k, c \pm 1\,|\,kc\rangle[(j \mp c)(j \pm c + 1)]^{\frac{1}{2}}. \quad (6b)$$

We introduce a notation for $\langle\tau'\,|\,\tau\rangle$ so that we can write

$$\begin{aligned}\Gamma_0\,|\mathsf{D}\,kc\,jm\rangle &= a_+(c, k + \tfrac{1}{2})[(j - k)(j + k + 1)]^{\frac{1}{2}}\,|\mathsf{D}\,k+1\,c\,jm\rangle \\ &+ a_-(c, k - \tfrac{1}{2})[(j + k)(j - k + 1)]^{\frac{1}{2}}\,|\mathsf{D}\,k-1\,c\,jm\rangle \\ &+ b_+(k, c + \tfrac{1}{2})[(j - c)(j + c + 1)]^{\frac{1}{2}}\,|\mathsf{D}\,k\,c+1\,jm\rangle \\ &+ b_-(k, c - \tfrac{1}{2})[(j + c)(j - c + 1)]^{\frac{1}{2}}\,|\mathsf{D}\,k\,c-1\,jm\rangle. \end{aligned} \quad (7)$$

Since we are considering unitary representations, we have that

$$a_+(c, k \pm \tfrac{1}{2}) = a_-^*(c, k \pm \tfrac{1}{2}),$$

$$b_+(k, c \pm \tfrac{1}{2}) = b_-^*(k, c \pm \tfrac{1}{2})$$

for Γ_μ to be self-adjoint.

If $\tau \in \mathsf{D}$ and $\langle\tau'\,|\,\tau\rangle_j \neq 0$ for some j, then $\tau' \in \mathsf{D}$. For a unitary representation we also have $\langle\tau\,|\,\tau'\rangle_j \neq 0$. We shall say that τ and τ' interlock and depict this by

$$\tau \text{ --- } \tau'.$$

Furthermore, (see Ref. 2, Theorem 4) we can choose a phase convention for the basis $|\mathsf{D}\,kc\,jm\rangle$ so that

$$a_\pm(c, \gamma) = a(c, \gamma), \quad b_\pm(k, \gamma) = b(k, \gamma).$$

Hence, $a(c, \gamma)$ and $b(k, \gamma)$ from the unitarity condition are real, and we can choose the phase for the basis so that they are nonnegative unless (k, c) is $(\frac{1}{2}, 0)$ or $(0, \frac{1}{2})$. In the latter case $a(0, 0) < 0$ and $a(0, 0) > 0$ define inequivalent representations in which $(\frac{1}{2}, 0)$ is self-interlocking, and $b(0, 0) < 0$ and $b(0, 0) > 0$ define inequivalent representations in which $(0, \frac{1}{2})$ is self-interlocking.

3. UNITARITY

If $(k, c) \in \mathsf{D}$, then, since (k, c) is unitary, we have the following two possibilities:

(i) (k, c) is a main series representation; i.e., c is pure imaginary and k is any integral or half-integral number;

(ii) (k, c) is a supplementary series representation; i.e., $k = 0$ and c is real with $|c| \leq 1$.

We see from (7) that we cannot have the trivial representation of L_+, i.e., $(0, 1)$, in D. Since $j = 0$ is the only value of j in this representation, $(0, 1)$ can only interlock with $(0, 2)$. However, $(0, 2)$ is nonunitary, and hence $(0, 1)$ does not interlock with any other representation for a unitary representation of $\bar{S}O(3, 2)$. Since D is irreducible, $(0, 1)$ is the only component in the reduction, and we hence have $\Gamma_\mu = 0$ which violates relation (1c).

Case (i): If $(k, c) \in$ D with c pure imaginary and $(k, c) \neq (0, 0)$, we can only have $(k \pm 1, c)$ interlocking with (k, c).

If $(0, 0) \in$ D, since $(0, 1) \equiv (0, -1)$ cannot belong to D, we can only have this interlocking with $(1, 0) \equiv (-1, 0)$.

Hence, for these cases we can write

$$\begin{aligned}\Gamma_0 &|\mathrm{D}\, k\, c\, jm\rangle \\ &= a(c, k + \tfrac{1}{2})[(j - k)(j + k + 1)]^{\frac{1}{2}} |\mathrm{D}\, k + 1\, c\, jm\rangle \\ &\quad + a(c, k - \tfrac{1}{2})[(j + k)(j - k + 1)]^{\frac{1}{2}} |\mathrm{D}\, k - 1\, c\, jm\rangle.\end{aligned} \tag{8}$$

Case (ii): If $(0, c) \in$ D and c is real with $c \neq 0$ and $|c| \leq 1$, then we can, without loss of generality, take $1 > c > 0$ ($c = 1$ gives the trivial representation). So the only interlocking we have is

$$(0, c) —— (0, c - 1).$$

We see that we can write this as

$$(0, \tfrac{1}{2} + \mu) —— (0, \tfrac{1}{2} - \mu)$$

with $\frac{1}{2} > \mu \geq 0$ and

$$\begin{aligned}\Gamma_0 &|\mathrm{D}\, 0\, \tfrac{1}{2} \pm \mu\, jm\rangle \\ &= b(0, \pm\mu)[(j + \tfrac{1}{2} + \mu)(j + \tfrac{1}{2} - \mu)]^{\frac{1}{2}} \\ &\qquad \times |\mathrm{D}\, 0\, \tfrac{1}{2} \mp \mu\, jm\rangle.\end{aligned} \tag{9}$$

4. REDUCED MATRIX ELEMENTS

In order to determine the reduced matrix elements $a(c, \gamma)$ and $b(k, \gamma)$, we use the relation (1c). However, the calculations are much simplified by using commutation relations involving the Casimir operators $\tilde{C}_1$ and $\tilde{C}_2$:

$$[\Gamma_0, [\Gamma_0, \tilde{C}_1]] = 2(\Gamma_0^2 - \mathrm{J}^2 + 2\tilde{C}_1 + \tilde{D}_1), \tag{10}$$

$$[\Gamma_0, [\Gamma_0, \tilde{C}_2]] = \tilde{C}_2. \tag{11}$$

In case (i) we obtain

$$\begin{aligned}&[(k + 1)a(c, k + \tfrac{1}{2})^2 - (k - 1)a(c, k - \tfrac{1}{2})^2 - \tfrac{1}{2}] \\ &\quad \times j(j + 1) = k(k + 1)^2 a(c, k + \tfrac{1}{2})^2 \\ &\quad - k(k - 1)^2 a(c, k - \tfrac{1}{2})^2 - (k^2 + c^2 - 1) - \tfrac{1}{2}D_1,\end{aligned} \tag{12}$$

$$\begin{aligned}j(j + 1)c[a(c, k + \tfrac{1}{2})^2 - a(c, k - \tfrac{1}{2})^2] &\\ = kc[(k + 1)a(c, k + \tfrac{1}{2})^2 &\\ - (k - 1)a(c, k - \tfrac{1}{2}) - \tfrac{1}{2}].&\end{aligned} \tag{13}$$

(12) and (13) follow from (10) and (11), respectively.

In case (ii) we obtain

$$\begin{aligned}j(j + 1)&[(1 \pm \mu)b(0, \mu)b(0, -\mu) - 1] \\ &= (\mu^2 - \tfrac{1}{4})(1 \pm \mu)b(0, \mu)b(0, -\mu) \\ &\qquad - 2[\mu^2 \mp \mu - \tfrac{3}{4}] - D_1.\end{aligned} \tag{14}$$

Relation (11) reduces to a tautology here.

Since we are dealing with unitary representations and the trivial case $(0, 1) \in$ D is excluded, j assumes more than one value (indeed, infinitely many) in (k, c), and in case (i) we must have

$$(k + 1)a(c, k + \tfrac{1}{2})^2 - (k - 1)a(c, k - \tfrac{1}{2})^2 = \tfrac{1}{2}, \tag{15}$$

$$c[a(c, k + \tfrac{1}{2})^2 - a(c, k - \tfrac{1}{2})^2] = 0, \tag{16}$$

$$\begin{aligned}\tfrac{1}{2}D_1 = (1 - k^2 - c^2) + k[(k + 1)^2 a(c, k + \tfrac{1}{2})^2 &\\ - (k - 1)^2 a(c, k - \tfrac{1}{2})^2].&\end{aligned} \tag{17}$$

In case (ii) we deduce that

$$\mu = 0, \quad b(0, 0)^2 = 1, \quad \text{and} \quad D_1 = \tfrac{5}{4}.$$

We determine D_2. Note that if we have any 4-vector A_μ given by

$$\begin{aligned}A_0 &|\mathrm{D}\, k\, c\, jm\rangle \\ &= a_+(c, k + \tfrac{1}{2})[(j - k)(j + k + 1)]^{\frac{1}{2}} |\mathrm{D}\, k + 1\, c\, jm\rangle \\ &\quad + a_-(c, k - \tfrac{1}{2})[(j + k)(j - k + 1)]^{\frac{1}{2}} |\mathrm{D}\, k - 1\, c\, jm\rangle,\end{aligned} \tag{18}$$

then from the Lorentz invariance of $A_\mu A^\mu$ we have that $A_\mu A^\mu |\mathrm{D}\, k\, c\, jm\rangle = \langle A_\mu A^\mu\rangle |\mathrm{D}\, k\, c\, jm\rangle$ and

$$\begin{aligned}\langle A_\mu A^\mu\rangle = f(k, c)a_-(c, k + \tfrac{1}{2})a_+(c, k + \tfrac{1}{2}) &\\ + g(k, c)a_+(c, k - \tfrac{1}{2})a_-(c, k - \tfrac{1}{2}).&\end{aligned}$$

Now

$$W_\mu = i[\tilde{C}_2, \Gamma_\mu], \tag{19}$$

and we see that if Γ_μ is given by (18), W_μ will be given by (18) with $a_\pm(c, k \pm \frac{1}{2}) \to \mp c a_\pm(c, k + \frac{1}{2})$. We see immediately therefore that

$$\langle W_\mu W^\mu\rangle = -c^2\langle\Gamma_\mu\Gamma^\mu\rangle.$$

Hence from (2) we have

$$D_2 = c^2(D_1 + c^2 - 1). \tag{20}$$

This applies for case (i) representations.

In case (ii) we see that $\tilde{C}_2 = 0$, and hence from (19) and (2b) we see $D_2 = 0$.

5. L_+ INTERLOCKING DIAGRAMS

We consider the irreducible representations determined by (15) and (16) for case (i), Sec. 4. Consider all $(k, c) \in \mathsf{D}$. Let $l = \min |k|$. Then we have the following possibilities.

1. $c = 0, l = 0$

Since $(1, 0) \equiv (-1, 0)$, we may take $a(0, -\frac{1}{2}) = 0$ without loss of generality. Hence, we obtain

$$a(0, \tfrac{1}{2}) = \sqrt{\tfrac{1}{2}},$$
$$a(0, k + \tfrac{1}{2}) = \tfrac{1}{2}, \quad \text{for} \quad k \geq 1.$$

2. $c = 0, l = \frac{1}{2}$

Put $a(0, 0) = M$. Then we obtain

$$a(0, n) = \left(\frac{(n + M)(n - M)}{4(n + \frac{1}{2})(n - \frac{1}{2})}\right)^{\frac{1}{2}}, \quad \text{integral } n \geq 1.$$

Since $a(0, 1)^2 \geq 0$, we must have $1 \geq |M|$. If $|M| = 1$, then $a(0, 1) = 0$, and we only have $(\frac{1}{2}, 0) \in \mathsf{D}$.

3. $c = 0, l \geq 1$

Hence, $a(0\ l - \frac{1}{2}) = 0$, and we obtain

$$a(0, k + \tfrac{1}{2}) = \left(\frac{(k + l)(k - l + 1)}{4k(k + 1)}\right)^{\frac{1}{2}}, \quad \text{for } k \geq l.$$

4. $c \neq 0$

From (15) and (16) we see that

$$a(c, k \pm \tfrac{1}{2}) = \tfrac{1}{2}.$$

N.B.: Throughout we have used the fact that we can choose the phase of the basis so that $a(c, \gamma) \geq 0$ unless $c = \gamma = 0$.

Hence, we obtain irreducible representations of $\bar{S}O(3, 2)$ depicted by the following interlocking systems.

1. $D_1 = +2, D_2 = 0$

$$[(0, 0) —— (1, 0) —— (2, 0) —— \text{etc.},$$
$$a(0, \tfrac{1}{2}) = 1/\sqrt{2},$$
$$a(0, k + \tfrac{1}{2}) = \tfrac{1}{2}, \quad k \geq 1.$$

2. $D_1 = \frac{9}{4} - M^2, D_2 = 0, 1 \geq M \geq -1$

(a) $|M| \neq 1$ and $M \neq 0$:

$$[(\tfrac{1}{2}, 0) —— (\tfrac{3}{2}, 0) —— (\tfrac{5}{2}, 0) —— \text{etc.},$$

$a(0, 0) = M,$

$$a(0, n) = \left(\frac{(n + M)(n - M)}{4(n + \frac{1}{2})(n - \frac{1}{2})}\right)^{\frac{1}{2}}, \quad \text{for } n \geq 1.$$

N.B.: $[(\frac{1}{2}, 0)$ indicates that $(\frac{1}{2}, 0)$ is self-interlocking.

(b) $M = 0$:

$$(\tfrac{1}{2}, 0) —— (\tfrac{3}{2}, 0) —— (\tfrac{5}{2}, 0) —— \text{etc.},$$
$$a(0, n) = n/[2(n^2 - \tfrac{1}{4})^{\frac{1}{2}}], \quad n \geq 1.$$

(c) $|M| = 1$:

$$[(\tfrac{1}{2}, 0),$$

(i) $M = 1, \quad a(0, 0) = +1,$
(ii) $M = -1, \quad a(0, 0) = -1.$

3. $D_1 = 2 - l(l - 1)$, $D_2 = 0$, *Integral or Half-Integral* $l \geq 1$.

$$(l, 0) —— (l + 1, 0) —— (l + 2, 0) —— \text{etc.},$$
$$a(0, k + \tfrac{1}{2}) = \left(\frac{(k + l)(k - l + 1)}{4k(k + 1)}\right)^{\frac{1}{2}}, \quad k \geq l.$$

4. $D_1 = 2(1 - c^2)$, $D_2 = c^2(1 - c^2)$, *c Pure Imaginary and* $c \neq 0$

(a) Integral angular momentum:

$$\text{etc.} —(-2, c)—(-1, c)—(0, c)—(1, c)—(2, c)— \text{etc.}$$
$$a(c, n) = \tfrac{1}{2} \text{ for all integral } n.$$

(b) half-integral angular momentum:

$$—— (-\tfrac{3}{2}, c) —— (-\tfrac{1}{2}, c) —— (\tfrac{1}{2}, c) —— (\tfrac{3}{2}, c) ——,$$
$$a(c, n) = \tfrac{1}{2} \text{ for all half-integral } n.$$

5. $D_1 = \frac{5}{4}, D_2 = 0$

$$(0, \tfrac{1}{2}),$$

(i) $b(0, 0) = +1,$
(ii) $b(0, 0) = -1.$

V follows from case (ii).

We shall refer to these as the class of "discrete Lorentz" UIR's of $\bar{S}O(3, 2)$.

6. MAXIMAL COMPACT BASIS

We now decompose these UIR's with respect to K and obtain a basis $|\mathsf{D}\ \lambda\ \mu\ jm\rangle$, where $\mu, j(j + 1)$, and m are the eigenvalues of Γ_0, $\mathbf{J}^2$, and J_3, respectively, and λ is any additional label necessary to specify the multiplicity of the reduction.

We show below that the multiplicity of this reduction is one, i.e., that it is a singleton reduction, i.e., the discrete Lorentz UIR's are singleton UIR's.

Consider the subspace $R(\mathsf{D}, j)$ of UIR, D, of $\bar{S}O(3, 2)$ with definite angular momentum j. For D, one of the discrete Lorentz UIR's, this subspace is finite dimensional, and we have a finite-dimensional transformation connecting the basis $|\mathsf{D}\ kc\ jm\rangle$ and $\|\mathsf{D}\ \lambda\ \mu\ jm\rangle$ within this subspace. For class I–IV UIR's we have

$$\|\mathsf{D}\ \lambda\ \mu\ jm\rangle = \sum_{k \in R(\mathsf{D}, j)} |\mathsf{D}\ kc\ jm\rangle \langle \mathsf{D}\ kc \mid \lambda\ \mu\ j\rangle, \quad (21)$$

where we write $k \in R(\mathsf{D}, j)$ if $|\mathsf{D}\ kc\ jm\rangle \in R(\mathsf{D}, j)$. For class V UIR's we have

$$\|\mathsf{D}\ \lambda\ \mu\ jm\rangle = |\mathsf{D}\ 0\ \tfrac{1}{2} jm\rangle \langle \mathsf{D}\ 0\ \tfrac{1}{2} \mid \lambda\ \mu\ j\rangle. \quad (22)$$

We write $[\mu, j]$ to indicate a UIR of K where $j(j+1)$ and μ are the eigenvalues of J^2 and Γ_0, respectively. If $[\mu, j]$ is a component in the reduction of a UIR D of S, we shall write $[\mu, j] \in \mathrm{D}$.

Note that the coefficients $\langle \mathrm{D}\, kc \mid \lambda\, \mu j\rangle$ are nontrivial [i.e., $\exists\, k \in R(\mathrm{D}, j)$ such that $\langle \mathrm{D}\, kc \mid \lambda\, \mu j\rangle \neq 0$)] iff $[\mu, j] \in \mathrm{D}$.

For class II(c) and V UIR's we see immediately that these are singleton UIR's, and we can write $\|\mathrm{D}\, \lambda\, \mu\, jm\rangle$ as

$$|\mathrm{D}\, \mu\, jm\rangle = |\mathrm{D}\, \tfrac{1}{2}\, 0\, jm\rangle \qquad (23)$$

for D, a class II(c) UIR, and

$$|\mathrm{D}\, \mu\, jm\rangle = |\mathrm{D}\, 0\, \tfrac{1}{2}\, jm\rangle \qquad (24)$$

for D, a class V UIR.

We consider the remaining discrete Lorentz UIR's. From (21) and (8) we obtain, for $k \in R(\mathrm{D}, j)$,

$$\begin{aligned}\mu\langle \mathrm{D}\, kc \mid \lambda\, \mu j\rangle &= a(k - \tfrac{1}{2})[(j - k + 1)(j + k)]^{\frac{1}{2}}\langle \mathrm{D}\, k - 1\, c \mid \lambda\, \mu j\rangle \\ &\quad + a(k + \tfrac{1}{2})[(j + k + 1)(j - k)]^{\frac{1}{2}}\langle \mathrm{D}\, k + 1\, c \mid \lambda \mu\, j\rangle,\end{aligned} \qquad (25)$$

where we have put $a(\gamma)$ for $a(c, \gamma)$ as it is c independent. Now $k = j \in R(\mathrm{D}, j)$, and we see that $\langle \mathrm{D}\, jc \mid \lambda \mu\, j\rangle \neq 0$ for $[\mu, j] \in \mathrm{D}$. Furthermore, $\langle \mathrm{D}\, kc \mid \lambda \mu j\rangle / \langle \mathrm{D}\, jc \mid \lambda \mu j\rangle$ is uniquely determined by (25) as a function of D, k, μ, and j only. Hence, if $\|\mathrm{D}\, \lambda'\, \mu\, jm\rangle$ and $\|\mathrm{D}\, \lambda\, \mu\, jm\rangle$ are two vectors in $R(\mathrm{D}, j)$, then

$$\frac{\langle \mathrm{D}\, jc \mid \lambda'\, \mu j\rangle}{\langle \mathrm{D}\, jc \mid \lambda\, \mu j\rangle} \|\mathrm{D}\, \lambda\, \mu\, jm\rangle = \|\mathrm{D}\, \lambda'\, \mu\, jm\rangle.$$

It follows that each UIR of K occurs once only in the reduction of these UIR's of $\bar{S}O(3, 2)$. Hence, these are "singleton" UIR's.

Hence, we may omit the label λ and write $\|\mathrm{D}\, \lambda\, \mu\, jm\rangle$ as $|\mathrm{D}\, \mu\, jm\rangle$. From (25) we see that the coefficients $\langle \mathrm{D}\, kc \mid \lambda \mu j\rangle$ are independent of λ and c and write these as $\langle k \mid \mu j\rangle$, omitting the dependence on D for notational convenience. Hence we have

$$|\mathrm{D}\, \mu\, jm\rangle = \sum_{k \in R(\mathrm{D}, j)} |\mathrm{D}\, k\, c\, jm\rangle \langle k \mid \mu j\rangle \qquad (26)$$

and, for $k \in R(\mathrm{D}, j)$,

$$\begin{aligned}\mu\langle k \mid \mu j\rangle &= a(k - \tfrac{1}{2})[(j - k + 1)(j + k)]^{\frac{1}{2}}\langle k - 1 \mid \mu j\rangle \\ &\quad + a(k + \tfrac{1}{2})[(j + k + 1)(j - k)]^{\frac{1}{2}}\langle k + 1 \mid \mu j\rangle.\end{aligned} \qquad (27)$$

For $|\mathrm{D}\, \mu\, jm\rangle$ not a null vector, i.e., $[\mu, j] \in \mathrm{D}$, we can choose

$$\sum_{k \in R(\mathrm{D}, j)} \langle k \mid \mu j\rangle^* \langle k \mid \mu j\rangle = 1,$$

and hence from (27) we can easily show that

$$\sum_{k \in R(\mathrm{D}, j)} \langle k \mid \mu' j\rangle \langle k \mid \mu j\rangle = \delta_{\mu'\mu}. \qquad (28)$$

7. K INTERLOCKING DIAGRAMS

From the commutation relation

$$[\Gamma(q), \Gamma_0] = i\, K(q),$$

we have that for $[\mu, j] \in \mathrm{D}$, $K^{\mu'\mu}_{j'j} = 0$ unless $\mu' = \mu \pm 1$ (for the notation, see Appendix C).

For class II(c) UIR's we have from (8) and (23) that (j half-integral)

II(c)(i) $\Gamma_0\, |\mathrm{D}\, \mu\, jm\rangle = (j + \tfrac{1}{2})\, |\mathrm{D}\, \mu\, jm\rangle$.

Hence, only $[j + \tfrac{1}{2}, j] \in \mathrm{D}$ and $K^{j+\frac{3}{2}}_{j+1}{}^{j+\frac{1}{2}}_{j} \neq 0$ by the infinite dimensionality of D.

II(c)(ii) Similarly here, only $[-j - \tfrac{1}{2}, j] \in \mathrm{D}$ and $K^{-j-\frac{3}{2}}_{j+1}{}^{-j-\frac{1}{2}}_{j} \neq 0$.

For class V UIR's we have from (9) and (24) that (j integral)

V(i) only $[j + \tfrac{1}{2}, j] \in \mathrm{D}$ and $K^{j+\frac{3}{2}}_{j+1}{}^{j+\frac{1}{2}}_{j} \neq 0$,
V(ii) only $[-j - \tfrac{1}{2}, j] \in \mathrm{D}$ and $K^{-j-\frac{3}{2}}_{j+1}{}^{-j-\frac{1}{2}}_{j} \neq 0$.

We consider the remaining discrete Lorentz UIR's.

From the action of $K(q)$ we obtain the following recurrence relation for these UIR's, for $[\mu, j] \in \mathrm{D}$:

$$\begin{aligned}K_{j'j}\langle(k, c)\langle k \mid \mu j\rangle &= \langle k \mid \mu - 1\, j'\rangle K^{\mu-1}_{j'}{}^{\mu}_{j}(\mathrm{D}) \\ &\quad + \langle k \mid \mu + 1\, j'\rangle K^{\mu+1}_{j'}{}^{\mu}_{j}(\mathrm{D}).\end{aligned} \qquad (29)$$

We use (29) to show that if $[\mu, j] \in \mathrm{D}$, then, for the UIR's $\neq$ II(c), V,

1. $K^{\mu'}_{j+1}{}^{\mu}_{j}(\mathrm{D}) \neq 0$ for $\mu' = \mu \pm 1$,
2. if $D_2 = 0$, then $K_j{}^{\mu'}{}^{\mu}_{j} = 0$,
3. for class IV UIR's $K^{\mu'}_{j}{}^{\mu}_{j} \neq 0$ for $\mu' = \mu \pm 1$ unless $(\mu', \mu) = \pm(j + 1, j)$,
4. From the finite dimensionality of $R(\mathrm{D}, j)$ we see immediately, therefore, that, for class IV UIR's, $[\mu, j] \in \mathrm{D}$ for $\mu = -j, -j + 1, \cdots, j - 1, j$ and that $K^{\mu'}_{j}{}^{\mu}_{j} = 0$ for $(\mu', \mu) = \pm(j + 1, j)$.

We prove proposition 1 by contradiction: Suppose $[\mu, j] \in \mathrm{D}$ and $K^{\mu+1}_{j+1}{}^{\mu}_{j} = 0$. Then

$$K_{j+1\, j}(k, c)\langle k \mid \mu j\rangle = \langle k \mid \mu - 1\, j + 1\rangle K^{\mu-1}_{j+1}{}^{\mu}_{j}\,.$$

Now $[\mu, j] \in \mathrm{D}$ iff $\langle j \mid \mu j\rangle \neq 0$ [this follows from (27)] and hence $K^{\mu-1}_{j+1}{}^{\mu}_{j} \neq 0$ since $K_{j+1\, j}(j, c) \neq 0$. Hence, $[\mu - 1, j + 1] \in \mathrm{D}$ and $\langle j + 1 \mid \mu - 1\, j + 1\rangle \neq 0$, contradicting $K_{j+1\, j}(j + 1, c) = 0$. Hence, $K^{\mu+1}_{j+1}{}^{\mu}_{j} \neq 0$. Similarly, $K^{\mu-1}_{j+1}{}^{\mu}_{j} \neq 0$.

For proposition 2 we note that if $D_2 = 0$ for a discrete Lorentz UIR, then $C_2 = 0$ and hence

$$K_{j\, j}\,(k, c) = 0.$$

Assume for example that $K^{\mu+1}_{j}{}^{\mu}_{j} \neq 0$. Then

$$[\mu + 1, j] \in \mathrm{D}$$

and, from (28) and (29), we have

$$K_j^{\mu+1}{}^{\mu}_j = \sum_{k\in R(\mathrm{D},j)} K_{jj}(k,c)\langle k \mid \mu j\rangle\langle k \mid \mu+1\, j\rangle^* = 0.$$

Hence $K_j^{\mu+1}{}^{\mu}_j = 0$, and similarly $K_j^{\mu-1}{}^{\mu}_j = 0$.

For proposition 3 assume, for example, that $K_j^{\mu+1}{}^{\mu}_j = 0$. Then

$$K_{jj}(k,c)\langle k \mid \mu j\rangle = \langle k \mid \mu-1\, j\rangle K_j^{\mu-1}{}^{\mu}_j.$$

For class IV UIR's, $K_{jj}(j,c) \neq 0$ unless $j=0$. If $j=0$, we see that $\mu=0$ from (27), and there is nothing to prove. If $j>0$, then $K_j^{\mu-1}{}^{\mu}_j \neq 0$, and

$$\langle k \mid \mu - 1\, j\rangle = C\, k\, \langle k \mid \mu j\rangle,$$

where $C \neq 0$ is a constant independent of k. From (25) we therefore have

$$\frac{\langle j-1 \mid \mu j\rangle}{\langle j \mid \mu j\rangle} = \frac{2\mu}{(2j)^{\frac{1}{2}}} = \frac{2(\mu-1)}{(2j)^{\frac{1}{2}}}\frac{j}{j-1}.$$

Hence $\mu = j$.

Similarly, $K_j^{\mu-1}{}^{\mu}_j \neq 0, j \neq 0$, unless $\mu = -j$.

If $[\mu, j] \in \mathrm{D}$ and $K_{j'}^{\mu'}{}^{\mu}_j \neq 0$, then $[u', j'] \in \mathrm{D}$, and we say $[j,\mu]$ and $[j',\mu']$ are interlocked, depicting this by

$$[\mu, j] \text{——} [\mu', j'].$$

Using propositions 1, 2, and 3, we now deduce the interlocking diagrams depicting the reduction of the UIR's $\neq$ II(c), V, with respect to K, by examining the spectrum of the eigenvalues of Γ_0 in $R(\mathrm{D}, l)$, where $l = \min j$.

I: $l=0$ and, for $j=0$, only $[0,0]\in\mathrm{D}$ since (27) gives

$$\mu\langle 0 \mid \mu 0\rangle = 0;$$

II(a), (b): $l=\frac{1}{2}$ and, for $j=\frac{1}{2}$, only $[M,\frac{1}{2}]\in\mathrm{D}$ since

$$\mu\langle\tfrac{1}{2} \mid \mu\tfrac{1}{2}\rangle = M\langle\tfrac{1}{2} \mid \mu\tfrac{1}{2}\rangle;$$

III: $l \geq 1$ and, for $j=l$, only $[0,l]\in\mathrm{D}$ since

$$\mu\langle l \mid \mu l\rangle = 0;$$

IV(a): for $j=0$, only $[0,0]\in\mathrm{D}$ since (27) gives

$$\mu\langle 0 \mid \mu 0\rangle = 0;$$

IV(b): for $j=\frac{1}{2}$, only $[\pm\frac{1}{2},\frac{1}{2}]\in\mathrm{D}$ since (27) gives

$$\mu\langle\tfrac{1}{2} \mid \mu\tfrac{1}{2}\rangle = \tfrac{1}{2}\langle -\tfrac{1}{2} \mid \mu\tfrac{1}{2}\rangle,$$
$$\mu\langle -\tfrac{1}{2} \mid \mu\tfrac{1}{2}\rangle = \tfrac{1}{2}\langle\tfrac{1}{2} \mid \mu\tfrac{1}{2}\rangle.$$

We can now write down the interlocking diagrams depicting the reduction of the discrete Lorentz UIR's with respect to K. These are tabulated in Appendix A.

8. EXPANSION COEFFICIENTS

For class IV UIR's, we determine the expansion coefficients in

$$|\mathrm{D}\,\mu\, jm\rangle = \sum_{|k|\leq j} |\mathrm{D}\, kc\, jm\rangle\langle k \mid \mu j\rangle.$$

Put

$$\langle k \mid \mu j\rangle = \langle j \mid \mu j\rangle \frac{[(2j)!]^{\frac{1}{2}}}{[(j+k)!\,(j-k)!]^{\frac{1}{2}}} P^j(k,\mu), \qquad |k| \leq j. \quad (30)$$

Then $P^j(k,\mu)$ is determined by

$$P^j(j,\mu) = 1, \quad (31a)$$

$$2\mu P^j(k,\mu) = (j+k)P^j(k-1,\mu) + (j-k)P^j(k+1,\mu). \quad (31b)$$

$\langle j \mid \mu j\rangle$ is determined up to a phase by

$$(2j)!\,|\langle j \mid \mu j\rangle|^2 \sum_{|k|\leq j} \frac{[P^j(k,\mu)]^2}{[(j+k)!\,(j-k)!]} = 1. \quad (32)$$

We use the method of Laplace to solve (31) and obtain[8]

$$P^j(k,\mu) = \frac{N^j(\mu)}{2\pi i}\int_C \frac{t^{j+k}}{(1-t)^{j+1-\mu}(1+t)^{j+1+\mu}}\,dt, \quad (33a)$$

where the contour C satisfies

$$\left(\frac{t^{j+k}}{(1-t)^{j-\mu}(1+t)^{j+\mu}}\right)_C = 0. \quad (33b)$$

Substituting $u = (1-t)/(1+t)$ in (33a), we obtain

$$P^j(k,\mu) = \frac{M^j(\mu)}{2\pi i}\int_{C'} \frac{(1-u)^{j+k}(1+u)^{j-k}}{u^{j-\mu+1}}\,du = M^j(\mu)\sum_{n=0}^{j-\mu}(-1)^n\, {}^{j+k}C_n\, {}^{j-k}C_{j-\mu-n}$$

(for the contour C' a circle around $u=0$), ${}^nC_r = n!/r!\,(n-r)!$. From $P^j(j,\mu)=1$, we finally obtain

$$P^j(k,\mu) = \frac{(j+\mu)!\,(j-\mu)!}{(2j)!}\sum_{n=0}^{j-\mu} {}^{j+k}C_n\, {}^{j-k}C_{j-\mu-n}(-1)^{j-\mu-n}. \quad (34)$$

Substituting $u = -t$ in (33a), we see that

$$P^j(k,-\mu) = (-1)^{j-k}P^j(k,\mu). \quad (35)$$

Substituting $n' = n - k + \mu$ in the summation (34), we see that

$$P^j(k,\mu) = P^j(\mu,k). \quad (36)$$

Put

$$\Sigma^j(\mu',\mu) = \sum_{|k|\leq j} \frac{P^j(k,\mu')P^j(k,\mu)}{[(j+k)!\,(j-k)!]}.$$

From (28), $\Sigma^j(\mu', \mu) = 0$ for $\mu' \neq \mu$. We determine $\Sigma^j(\mu, \mu) = \Sigma^j(\mu)$: From (36) and (31) we have

$$2kP^j(k, \mu) = (j + \mu)P^j(k, \mu - 1) + (j - \mu)P^j(k, \mu + 1).$$

Hence,

$$(j + \mu)\Sigma^j(\mu - 1) = \sum_{|k|\leq j} 2k \frac{P^j(k, \mu)P^j(k, \mu - 1)}{(j + k)!\,(j - k)!} = (j - \mu + 1)\Sigma^j(\mu);$$

$$\therefore\quad \Sigma^j(\mu) = \frac{(j + \mu)!\,(j - \mu)!}{(2j)!}\Sigma^j(j),$$

$$\Sigma^j(j) = \sum_{|k|\leq j} \frac{1}{(j + k)!\,(j - k)!} = \frac{1}{(2j)!}\sum_{v=0}^{2j} {}^{2j}C_v = \frac{2^{2j}}{(2j)!}.$$

From (32) we have

$$\langle j \mid \mu j\rangle = \phi(j, \mu)\frac{1}{2^j}\left(\frac{(2j)!}{(j + \mu)!\,(j - \mu)!}\right)^{\frac{1}{2}}, \quad (37)$$

where $\phi(j, \mu)$ is an arbitrary phase.

From (34), (35), (36), and (37) we obtain the expressions in Appendix B for $\langle k \mid \mu j\rangle$ [substituting $r = j - \mu - n$ in (34)].

9. MATRIX ELEMENTS

The matrix elements of the action of the generators on the Lorentz basis follow from the expressions for the reduced matrix elements $a(c, \gamma)$ and $b(0, 0)$.

For the action on the maximal compact basis we have to determine the $K^{j'\ j}_{\mu'\ \mu}$. We do this for class IV UIR's, using (B7) and (B8) of Appendix B and (29) for $k = j', j' - 1$. This gives two simultaneous equations for $K^{j'\ j}_{\mu-1\ \mu}$ and $K^{j'\ j}_{\mu+1\ \mu}$ and finally the result in Appendix C.

10. $\bar{S}O(3, 1)$ SUBGROUPS

$S = \bar{S}O(3, 2)$ has two $\bar{S}O(3, 1)$ subgroups, L_+ generated by J and K and M_+ generated by J and $\mathbf{\Gamma}$. Under the correspondence $J_{\mu\nu} \to J'_{\mu\nu}$ defined by

$$\mathrm{J}' = \mathrm{J}, \quad (38a)$$
$$\Gamma'_0 = \Gamma_0, \quad (38b)$$
$$\mathrm{K}' = \mathbf{\Gamma}, \quad (38c)$$
$$\mathbf{\Gamma}' = -\mathrm{K}, \quad (38d)$$

we see that $J'_{\mu\nu}$ satisfies exactly the same commutation relations as $J_{\mu\nu}$. Hence, S has exactly the same reduction with respect to M_+ as it has with respect to L_+, and we obtain in an analogous fashion an "M_+-basis" $|\mathrm{D}\, kc\, jm\rangle'$ on which the action of $J_{\mu\nu}$ is the same as in Appendix C [(C4)–(C7)] with

$$|\mathrm{D}\, kc\, jm\rangle \to |\mathrm{D}\, kc\, jm\rangle',$$
$$J_{\mu\nu} \to J'_{\mu\nu}.$$

Also, for class IV UIR's,

$$|\mathrm{D}\, \mu\, jm\rangle = \sum_{|k|\leq j} |\mathrm{D}\, kc\, jm\rangle'\langle k \mid \mu j\rangle', \quad (39)$$

where

$$\langle k \mid \mu j\rangle' = \psi(j, \mu)C^j(k, \mu), \quad (40)$$

$\psi(j, \mu)$ being a phase factor dependent on the choice for $\phi(j, \mu)$. Putting $\psi(j, \mu) = \sigma^*(j, \mu)\phi(j, \mu)$, we see that

$$\sigma(j, \mu)\,|\mathrm{D}\, \mu\, jm\rangle = \sum_{|k|\leq j} |\mathrm{D}\, kc\, jm\rangle'\langle k \mid \mu j\rangle. \quad (41)$$

Considering the action of $\Gamma(q)$, we obtain

$$\sigma(j, \mu)\sum_{\mu'} |\mathrm{D}\, \mu' j'm'\rangle\langle j'm' \mid 1q; jm\rangle[2j + 1]^{\frac{1}{2}}\Gamma^{\mu'\ \mu}_{j'\ j}$$
$$= \sum_{|k|\leq j} |\mathrm{D}\, kc\, j'm'\rangle'\langle j'm' \mid 1q; jm\rangle[2j + 1]^{\frac{1}{2}} \times K_{j'j}(k, c)\langle k \mid \mu j\rangle$$
$$= \sum_{\mu'} \sigma(j', \mu')\,|\mathrm{D}\, \mu' j'm'\rangle\langle j'm' \mid 1q; jm\rangle[2j + 1]^{\frac{1}{2}}K^{\mu'\ \mu}_{j'\ j},$$

the latter equation via (29) and (41), and hence, from (C15), $\sigma(j', \mu') = i(\mu - \mu')\sigma(j, \mu)$, giving

$$\psi(j, \mu) = (i)^\mu\phi(j, \mu). \quad (42)$$

Finally, we have that

$$|\mathrm{D}\, kc\, jm\rangle' = \sum_{|k'|\leq j} |\mathrm{D}\, k'c\, jm\rangle\langle k' \mid k\rangle^j, \quad (43)$$

and from the action of $K(q)$ we obtain the following recurrence relation for $\langle k' \mid k\rangle^j$:

$$2k'\langle k' \mid k\rangle^j = -i\langle k' \mid k + 1\rangle^j[(j - k)(j + k + 1)]^{\frac{1}{2}} + i\langle k' \mid k - 1\rangle^j[(j + k)(j - k + 1)]^{\frac{1}{2}}. \quad (44)$$

Hence, from (27) we have that

$$\langle k' \mid k\rangle^j = (i)^k C^j(k', k)f^j(k'). \quad (45)$$

But we also have from (43), (42), (40), (39), (26), (B1), and (B11) that

$$\langle k' \mid k\rangle^j = \sum_{|\mu|\leq j} C^j(k', \mu)i^{-\mu}C^j(k, \mu) \quad (46)$$

and hence

$$\langle k' \mid k\rangle^j = \langle k \mid k'\rangle^j = (i)^{k+k'}C^j(k', k)f^j.$$

Using (46) with $k' = k = j$ and (B7), we obtain $f^j = (-1)^{-j}$.

The results are tabulated in Appendix D.

ACKNOWLEDGMENT

It is a pleasure to thank Professor A. Böhm for suggesting this problem and providing constant encouragement.

APPENDIX A

	D_1	D_2	L_+ Reduction	K Reduction
I	2	0	(0, 0) — (1, 0) — (2, 0) — etc.	[0,0]; [-1,1] [1,1]; [-2,2] [2,0] [2,2]; etc.
II	$[\frac{9}{4}-M^2]$, $-1\le M\le 1$		(a) / (b)	
(a)	$M\neq 0,1$	0	(a) $(\frac{1}{2},0)$ — $(\frac{3}{2},0)$ — $(\frac{5}{2},0)$ — etc.	$[M,\frac{1}{2}]$; $[M-1,\frac{3}{2}]$ $[M+1,\frac{3}{2}]$; $[M-2,\frac{5}{2}]$ $[M,\frac{5}{2}]$ $[M+2,\frac{5}{2}]$; etc.
(b)	$M=0$	0	(b) $(\frac{1}{2},0)$ — $(\frac{3}{2},0)$ — $(\frac{5}{2},0)$ — etc.	
II (c)(i)	$M=1$	0	$(\frac{1}{2},0)$	(i) $[1,\frac{1}{2}]$; $[2,\frac{3}{2}]$; $[3,\frac{5}{2}]$; etc.
(c)(ii)	$M=-1$	0		(ii) $[-1,\frac{1}{2}]$; $[-2,\frac{3}{2}]$; $[-3,\frac{5}{2}]$; etc.
III	$2-\ell(\ell-1)$, $\ell\ge 1$, Integral 2ℓ	0	$(\ell,0)$ — $(\ell+1,0)$ — $(\ell+2,0)$ — etc.	$[0,\ell]$; $[-1,\ell+1]$ $[1,\ell+1]$; $[-2,\ell+2]$ $[0,\ell+2]$ $[2,\ell+2]$; etc.
IV	$2(1-c^2)$ Pure Imaginary $c\neq 0$	$c^2(1-c^2)$		
(a)			$(0, c)$; $(-1, c)$ $(1, c)$; $(-2, c)$ $(2, c)$; etc.	[0,0]; [-1,1] [0,1] [1,1]; [-2,2] [-1,2] [0,2] [1,2] [2,2]; etc.
IV (b)		–	$(-\frac{1}{2}, c)$ — $(\frac{1}{2}, c)$; $(-\frac{3}{2}, c)$ $(\frac{3}{2}, c)$; $(-\frac{5}{2}, c)$ $(\frac{5}{2}, c)$; etc.	$[-\frac{1}{2},\frac{1}{2}]$ $[\frac{1}{2},\frac{1}{2}]$; $[-\frac{3}{2},\frac{3}{2}]$ $[-\frac{1}{2},\frac{3}{2}]$ $[\frac{1}{2},\frac{3}{2}]$ $[\frac{3}{2},\frac{3}{2}]$; $[-\frac{5}{2},\frac{5}{2}]$ $[-\frac{3}{2},\frac{5}{2}]$ $[-\frac{1}{2},\frac{5}{2}]$ $[\frac{1}{2},\frac{5}{2}]$ $[\frac{3}{2},\frac{5}{2}]$ $[\frac{5}{2},\frac{5}{2}]$; etc.
V	$\frac{5}{4}$	0	$(0, \frac{1}{2})$	(i) $[\frac{1}{2},0]$; $[\frac{3}{2},1]$; $[\frac{5}{2},2]$; etc. (ii) $[-\frac{1}{2},0]$; $[-\frac{3}{2},1]$; $[-\frac{5}{2},2]$; etc.

$$\tilde{D}_1 = -\tfrac{1}{2} J_{\alpha\beta} J^{\alpha\beta}$$
$$\tilde{D}_2 = W_\alpha W^\alpha$$
$$W = \tfrac{1}{8}\, \epsilon_{\alpha\beta\gamma\delta\epsilon} J^{\beta\gamma} J^{\delta\epsilon}$$

APPENDIX B

$$|D\,\mu\, jm\rangle = \sum_{|k|\le j} |D\, kc\, jm\rangle\langle k \mid \mu j\rangle, \quad \text{D class IV UIR,}$$

with

$$\langle k \mid \mu j\rangle = \phi(j,\mu) C^j(k,\mu), \tag{B1}$$

where $\phi(j, \mu)$ is an arbitrary phase.

$$C^j(k,\mu) = \frac{(2j)!}{2^j}[(j+\mu)!\,(j-\mu)!\,(j+k)!\,(j-k)!]^{-\frac{1}{2}} P^j(k,\mu), \tag{B2}$$

$$P^j(k,\mu) = \sum_{r=m}^{n} \frac{(-1)^r (j+\mu)!\,(j-\mu)!\,(j+k);(j-k)!}{r!\,(r+k+\mu)!\,(j-r-k)!\,(j-r-\mu)!}, \tag{B3}$$

where

$$n = \min\{j-k, j-\mu\},$$
$$m = \max\{0, -(k+\mu)\}.$$

$C^j(k, \mu)$ has the following symmetries:

$$C^j(k,\mu) = C^j(\mu, k), \tag{B4}$$
$$C^j(k,\mu) = (-1)^{j-k} C^j(k, -\mu), \tag{B5}$$
$$C^j(k,\mu) = (-1)^{k+\mu} C^j(-k, -\mu). \tag{B6}$$

We tabulate the results for some simple cases:

$$C^j[\pm j, t] = C^j[t, \pm j] = (\pm 1)^{j-t}\left(\frac{1}{2}\right)^j \left(\frac{(2j)!}{(j+t)!\,(j-t)!}\right)^{\frac{1}{2}}, \tag{B7}$$

$$C^j[\pm(j-1), t] = C(t, \pm(j-1)] = 2t(\pm 1)^{j-t}\left(\frac{1}{2}\right)^j \left(\frac{(2j-1)!}{(j+t)!\,(j-t)^{\frac{1}{2}}}\right)^{\frac{1}{2}}, \tag{B8}$$

$$C^j[\pm(j-2), t] = C[t, \pm(j-2)] = [2\mu^2 - j](\pm 1)^{j-t}\left(\frac{1}{2}\right)^j \left(\frac{2(2j-2)!}{(j+t)!\,(j-t)^{\frac{1}{2}}}\right)^{\frac{1}{2}}. \tag{B9}$$

The orthonormality condition is

$$\sum_{|k|\le j} C^j(k,\mu')C^j(k,\mu) = \delta_{\mu'\mu}, \tag{B10}$$
$$\sum_{|\mu|\le j} C^j(k',\mu)C^j(k,\mu) = \delta_{k'k}. \tag{B11}$$

The recurrence relations for $P^j(k, \mu)$ are, $|k| \le j$, $|\mu| \le j$,

$$2\mu P^j(k,\mu) = (j+k)P^j(k-1,\mu) + (j-k)P^j(k+1,\mu), \tag{B12}$$
$$2kP^j(k,\mu) = (j+\mu)P^j(k,\mu-1) + (j-\mu)P^j(k,\mu+1), \tag{B13}$$
$$P^j(j,\mu) = P^j(k,j) = 1. \tag{B14}$$

From (29) we also obtain

$$2[(j+1)^2 - k^2]P^j(k,\mu) = (j+1)(2j+1)[P^{j+1}(k,\mu+1) - P^{j+1}(k,\mu-1)], \tag{B15}$$

$$2(j+1)(2j+1)P^{j+1}(k,\mu) = (j+\mu+1)(j+\mu)P^j(k,\mu-1) - (j-\mu+1)(j-\mu)P^j(k,\mu+1). \tag{B16}$$

APPENDIX C

We explicitly determine the action of the generators on the Lorentz basis and on the maximal compact basis.

We use the Wigner–Eckart theorem and define

$$J(0) = J_{12}, \tag{C1a}$$

$$J(\pm 1) = \mp 2^{-\frac{1}{2}}(J_{23} \pm iJ_{31}), \tag{C1b}$$

$$K(0) = J_{03}, \tag{C2a}$$

$$K(\pm 1) = \mp 2^{-\frac{1}{2}}(J_{01} \pm iJ_{02}), \tag{C2b}$$

$$\Gamma(0) = \Gamma_3, \tag{C3a}$$

$$\Gamma(\pm 1) = \mp 2^{-\frac{1}{2}}(\Gamma_1 \pm i\Gamma_2). \tag{C3b}$$

Then on the Lorentz basis $|\mathsf{D}\,\tau\, jm\rangle$ we have

$$J(q)\,|\mathsf{D}\,\tau\, jm\rangle = \sum_{m'} |\mathsf{D}\,\tau\, jm'\rangle\langle jm' \mid 1q; jm\rangle J_j, \tag{C4}$$

$$K(q)\,|\mathsf{D}\,\tau\, jm\rangle = \sum_{j'm'} |\mathsf{D}\,\tau\, j'm'\rangle\langle j'm' \mid 1q; jm\rangle \times [2j+1]^{\frac{1}{2}} K_{j'j}(\tau), \tag{C5}$$

$$\Gamma_0\,|\mathsf{D}\,\tau\, jm\rangle = \sum_{\tau'} |\mathsf{D}\,\tau\, jm\rangle\, \Gamma_j^{\tau'\tau}(\mathsf{D}), \tag{C6}$$

$$\Gamma(q)\,|\mathsf{D}\,\tau\, jm\rangle = \sum_{\tau' j'm'} |\mathsf{D}\,\tau'\, j'm'\rangle\langle j'm' \mid 1q; jm\rangle \times [2j+1]^{\frac{1}{2}} \Gamma_{j'j}^{\tau'\tau}(\mathsf{D}). \tag{C7}$$

On the maximal compact basis $|\mathsf{D}\,\mu\, jm\rangle$ we have

$$J(q)\,|\mathsf{D}\,\mu\, jm\rangle = \sum_{m'} |\mathsf{D}\,\mu\, jm'\rangle\langle jm' \mid 1q; jm\rangle J_j, \tag{C8}$$

$$\Gamma_0\,|\mathsf{D}\,\mu\, jm\rangle = \mu\,|\mathsf{D}\,\mu\, jm\rangle, \tag{C9}$$

$$K(q)\,|\mathsf{D}\,\mu\, jm\rangle = \sum_{\mu' j'm'} |\mathsf{D}\,\mu'\, j'm'\rangle\langle j'm' \mid 1q; jm\rangle \times [2j+1]^{\frac{1}{2}} K_{j'j}^{\mu'\mu}(\mathsf{D}), \tag{C10}$$

$$\Gamma(q)\,|\mathsf{D}\,\mu\, jm\rangle = \sum_{\mu' j'm'} |\mathsf{D}\,\mu'\, j'm'\rangle\langle j'm' \mid 1q; jm\rangle \times [2j+1]^{\frac{1}{2}} \Gamma_{j'j}^{\mu'\mu}(\mathsf{D}). \tag{C11}$$

$\langle j'm' \mid 1q; jm\rangle$ is the Clebsch–Gordan coefficient, with the phase convention as in Rose,[9] $[\langle jm \mid j_1m_1; j_2m_2\rangle = C(j_1j_2j;\, m_1m_2m)]$:

$$J_j = -[j(j+1)]^{\frac{1}{2}}. \tag{C12}$$

From the commutation relation

$$\Gamma_i = i[J_{0i}, \Gamma_0], \tag{C13}$$

we have

$$\Gamma_{j'j}^{\tau'\tau} = +i[K_{j'j}(\tau')\Gamma_j^{\tau'\tau} - \Gamma_{j'}^{\tau'\tau}K_{j'j}(\tau)], \tag{C14}$$

$$\Gamma_{j'j}^{\mu'\mu} = i(\mu - \mu')K_{j'j}^{\mu'\mu}. \tag{C15}$$

Hence it suffices to determine $\Gamma_j^{\tau'\tau}$ and $K_{j'j}^{\mu'\mu}$.

For the Lorentz basis:
In general

$$K_{j-1,j}(k, c) = -i\,\frac{[(j^2 - k^2)(j^2 - c^2)]^{\frac{1}{2}}}{[j(2j-1)(2j+1)]^{\frac{1}{2}}}, \quad \text{for} \quad j > |k|, \tag{C16a}$$

$$= 0, \quad \text{for} \quad j = |k|, \tag{C16b}$$

$$K_{j,j}(k, c) = i\,\frac{kc}{[(2j+1)j(j+1)]^{\frac{1}{2}}}, \quad \text{for} \quad j > 0. \tag{C16c}$$

N.B.: $K_{0,0}(0, c)$ is superfluous as $\langle 00 \mid 1q; 00\rangle = 0$ in any event:

$$K_{j+1,j}(k, c) = K_{j,j+1}(k, c). \tag{C16d}$$

For class IV UIR's:

$$\Gamma_j^{(k\pm1,c)(k,c)} = \tfrac{1}{2}[(j \mp k)(j \pm k + 1)]^{\frac{1}{2}}, \tag{C17}$$

$$K_{j-1\ j}^{\mu\pm1\ \mu} = \pm\,\frac{\phi(j,\mu)}{\phi(j-1,\mu\pm1)}\,\frac{i[j^2 - c^2]^{\frac{1}{2}}}{2[j(2j-1)(2j+1)]^{\frac{1}{2}}} \times [(j \mp \mu)(j \mp \mu - 1)]^{\frac{1}{2}}, \quad \text{for} \quad j > 0, \tag{C18a}$$

$$= 0, \quad \text{for} \quad j = 0, \tag{C18b}$$

$$K_{j\ j}^{\mu\pm1\ \mu} = \frac{\phi(j,\mu)}{\phi(j,\mu\pm1)}\,\frac{ic}{2[(2j+1)j(j+1)]^{\frac{1}{2}}} \times [(j \mp \mu)(j \pm \mu + 1)]^{\frac{1}{2}}, \quad \text{for} \quad j > 0, \tag{C18c}$$

$$K_{j+1\ j}^{\mu\pm1\ \mu} = \mp\,\frac{\phi(j,\mu)}{\phi(j+1,\mu\pm1)} \times \frac{i[(j+1)^2 - c^2]^{\frac{1}{2}}}{2[(j+1)(2j+1)(2j+3)]^{\frac{1}{2}}} \times [(j \pm \mu + 1)(j \pm \mu + 2)]^{\frac{1}{2}}. \tag{C18d}$$

APPENDIX D

$$|\mathsf{D}\,\mu\, jm\rangle = \sum_{|k|\le j} |\mathsf{D}\,kc\, jm\rangle'\langle k \mid \mu j\rangle', \tag{D1}$$

where

$$\langle k \mid \mu j\rangle' = (i)^{\mu}\phi(j,\mu)C^j(k,\mu), \tag{D2}$$

$$|\mathsf{D}\,k'c\, jm\rangle' = \sum_{|k|\le j} |\mathsf{D}\,kc\, jm\rangle(i)^{k+k'-2j}C^j(k', k). \tag{D3}$$

[1] J. B. Ehrman, Proc. Cambridge Phil. Soc. **53**, 290 (1957); thesis, Princeton University, 1954.
[2] L. Jaffe, University of Texas, Preprint CPT-28, 1969.
[3] A discrete reduction of $\bar{S}O(3, 2)$ with respect to $\bar{S}O(3, 1)$ smooths out the difficulties due to these being noncompact, and there is reason to believe that the reduction must be a singleton reduction in the same way that the reduction of $\bar{S}O(5)$ with respect to $\bar{S}O(4)$ is.
[4] A. Bohm, Phys. Rev. Letters **23**, 436 (1969).
[5] I. M. Gel'fand, R. A. Minlos, and Z. J. Shapiro, *Representations of the Rotation and Lorentz Groups and Their Applications* (Pergamon, Oxford, 1963).
[6] H. Joos, Fortschr. Physik **10**, 65 (1962).
[7] M. A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon, New York, 1964).
[8] P. M. Batchelder, *Linear Difference Equations* (Harvard U.P., Cambridge, Mass., 1927), Chap. III, Sec. 7.
[9] M. E. Rose, *Elementary Theory of Angular Momentum* (Wiley, New York, 1957).

Lattice Green's Function. Introduction

SHIGETOSHI KATSURA,* TOHRU MORITA,† SAKARI INAWASHIRO,* TSUYOSHI HORIGUCHI,‡ AND YOSHIHIKO ABE‡
Department of Applied Physics, Department of Applied Science,† Department of Physics‡*
Tohoku University, Sendai, Japan

(Received 20 July 1970)

Physical, analytical, and numerical properties of the lattice Green's functions for the various lattices are described. Various methods of evaluating the Green's functions, which will be developed in the subsequent papers, are mentioned.

This paper is intended to be a general introduction to a series of papers[1–4] which are presented by a group of authors to investigate the lattice Green's functions in several new methods.

The Helmholtz equation in the continuous space is given by

$$\tfrac{1}{2}\Delta\psi + E\psi = 0 \tag{1}$$

for the wavefunction $\psi(\mathbf{r})$. The Green's function $g(E, \mathbf{r})$ is defined to be a solution of

$$\tfrac{1}{2}\Delta g + Eg = \delta(\mathbf{r}). \tag{2}$$

The quantity corresponding to d^2f/dx^2 in the discrete lattice space is

$$\frac{\Delta}{\Delta x}\left(\frac{\Delta}{\Delta x}f\right) = f_{n+1} - 2f_n + f_{n-1}, \tag{3}$$

where Δx, the lattice spacing, is taken as a unit of the length. Generalizing Eq. (3) to a lattice of specified type, we have

$$\tfrac{1}{2}\sum_{\Delta}^{z} \phi(\mathbf{r} + \mathbf{\Delta}) + (E - \tfrac{1}{2}z)\phi(\mathbf{r}) = 0, \tag{4}$$

for the wavefunction $\phi(\mathbf{r})$, and

$$\tfrac{1}{2}\sum_{\Delta}^{z} G(E, \mathbf{r} + \mathbf{\Delta}) + (E - \tfrac{1}{2}z)G(E, \mathbf{r}) = \delta_{\mathbf{r},0}, \tag{5}$$

for the Green's function $G(E, \mathbf{r})$ as the Helmholtz equations in lattice spaces. Here $\mathbf{r} = (l, m, n)$ and $\mathbf{\Delta}$ is the nearest neighbor vector, i.e.,

$$\begin{aligned}
\mathbf{\Delta}_{\mathrm{sc}} &= (\pm1, 0, 0), (0, \pm1, 0), (0, 0, \pm1),\\
\mathbf{\Delta}_{\mathrm{bcc}} &= (\pm1, \pm1, \pm1),\\
\mathbf{\Delta}_{\mathrm{fcc}} &= (0, \pm1, \pm1), (\pm1, 0, \pm1), (\pm1, \pm1, 0),
\end{aligned} \tag{6}$$

where sc stands for the simple cubic lattice, bcc for the body-centered cubic lattice, and fcc for the face-centered cubic lattice, respectively.

Substituting

$$\phi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \tag{7}$$

into Eq. (4) and taking Eqs. (6) into consideration, we have the eigenvalue of Eq. (4),

$$\begin{aligned}
E_{\mathbf{k}} &= 3 - \cos k_x - \cos k_y - \cos k_z, \quad \text{for sc},\\
E_{\mathbf{k}} &= 4(1 - \cos k_x \cos k_y \cos k_z), \quad \text{for bcc},\\
E_{\mathbf{k}} &= 2(3 - \cos k_x \cos k_y - \cos k_y \cos k_z\\
&\qquad - \cos k_z \cos k_x), \quad \text{for fcc}.
\end{aligned} \tag{8}$$

The periodic boundary condition $\phi(\mathbf{r} + L\mathbf{\Delta}) \equiv \phi(\mathbf{r})$ gives the quantum number

$$\mathbf{k} = 2\pi\boldsymbol{\lambda}/L, \quad \lambda_x, \lambda_y, \lambda_z = 0, 1, 2, \cdots, L.$$

Putting

$$G(E, \mathbf{r}) = \frac{1}{N}\sum_{\mathbf{k}} A_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}, \quad N \equiv L^3, \tag{9}$$

into Eq. (5) and equating the coefficients of $\exp(i\mathbf{k}\cdot\mathbf{r})$ on both sides, we obtain $A_{\mathbf{k}} = 1/(E - E_{\mathbf{k}})$. Hence

$$G(E, \mathbf{r}) = \frac{1}{N}\sum_{\mathbf{k}} \frac{1}{E - E_{\mathbf{k}}} e^{i\mathbf{k}\cdot\mathbf{r}} = \frac{1}{E - \mathbf{H}_0}. \tag{10}$$

The lattice Green's functions $G(E \pm i\epsilon, 0)$, $\epsilon \to 0$, are real for $E > E_{\mathbf{k}\,\max}$ and $E < E_{\mathbf{k}\,\min}$ (outside the energy band). They are complex for $E_{\mathbf{k}\,\min} < E < E_{\mathbf{k}\,\max}$. They have singularities at $E = E_{\mathbf{k}\,\min}$, $E = E_{\mathbf{k}\,\max}$ and may have them at some value of E within the band. Equation (10) gives the lattice Green's function in a general form.

By the formula $1/(x - i\epsilon) = P(1/x) + i\pi\delta(x)$, we have

$$\operatorname{Im} G(E - i\epsilon, 0) = \pi\frac{1}{N}\sum_{\mathbf{k}} \delta(E - E_{\mathbf{k}}). \tag{11}$$

On integrating both sides from $-\infty$ to E, the right-hand side gives the number of states whose energy is less than E. Hence $(1/\pi)\operatorname{Im} G(E - i\epsilon, 0)$ represents the density of states.

From Eqs. (6) and (10) we encounter the following nontrivial integrals:

$$I_{sc}(a; l, m, n) = I_{ortho}(a; l, m, n; 1, 1, 1) = I_{tetra}(a; l, m, n; 1),$$

$$I_{tetra}(a; l, m, n; \gamma) = I_{ortho}(a; l, m, n; \gamma, 1, 1),$$

$$I_{ortho}(a; l, m, n; \gamma_1, \gamma_2, \gamma_3) = \frac{1}{\pi^3}\iiint_0^\pi \frac{\cos lx \cos my \cos nz}{a - \gamma_1 \cos x - \gamma_2 \cos y - \gamma_3 \cos z}\, dx\, dy\, dz,$$

$$I_{bcc}(a; l, m, n) = \frac{1}{\pi^3}\iiint_0^\pi \frac{\cos lx \cos my \cos nz}{a - \cos x \cos y \cos z}\, dx\, dy\, dz, \qquad (12)$$

$$I_{fcc}(a; l, m, n) = \frac{1}{\pi^3}\iiint_0^\pi \frac{\cos lx \cos my \cos nz}{a - \cos x \cos y - \cos y \cos z - \cos z \cos x} \times dx\, dy\, dz,$$

where

$$I_{sc}(a; l, m, n) = -G(E, \mathbf{r}), \quad a - 3 = -E + i\epsilon,$$
$$I_{bcc}(a; l, m, n) = -4G(E, \mathbf{r}), \quad a - 1 = \tfrac{1}{4}(-E + i\epsilon),$$
$$I_{fcc}(a; l, m, n) = -2G(E, \mathbf{r}), \quad a - 3 = \tfrac{1}{2}(-E + i\epsilon),$$
$$\mathbf{r} = (l, m, n), \qquad (13)$$

for the sc, bcc, and fcc lattices. The integrals $I_{tetra}(a; l, m, n; \gamma)$ and $I_{ortho}(a; l, m, n; \gamma_1, \gamma_2, \gamma_3)$ for the tetragonal and orthorhombic lattices, respectively, are introduced for the convenience of the following discussions. It is to be noted that $\mathrm{Re}\, I_{bcc}(a; l, m, n)$ is an odd function of a and $\mathrm{Im}\, I_{bcc}(a; l, m, n)$ is an even function of a, and also that $\mathrm{Re}\, I_{ortho}(a; l, m, n; \gamma_1, \gamma_2, \gamma_3)$, $\mathrm{Re}\, I_{tetra}(a; l, m, n; \gamma)$ and $\mathrm{Re}\, I_{sc}(a; l, m, n)$ are odd or even functions of a and $\mathrm{Im}\, I_{ortho}(a; l, m, n; \gamma_1, \gamma_2, \gamma_3)$, $\mathrm{Im}\, I_{tetra}(a; l, m, n; \gamma)$, and $\mathrm{Im}\, I_{sc}(a; l, m, n)$ are even or odd functions of a, according as the sum of l, m, and n is even or odd.

The lattice Green's functions (12) appear in the problems of lattice vibrations, of spin wave theory of Heisenberg model of magnetism, of localized modes of oscillation at the lattice defect, and of others. They are integrals which have many applications in solid state physics.[5,6] The one-body Green's function of the orthorhombic lattice $I_{ortho}(a; l, m, n; \gamma_1, \gamma_2, \gamma_3)$ has appeared as the two-body Green's function of the simple cubic lattice.[7] In that case $\gamma_1 = \cos \frac{1}{2}K_x$, $\gamma_2 = \cos \frac{1}{2}K_y$, and $\gamma_3 = \cos \frac{1}{2}K_z$, where K_x, K_y, and K_z are the x, y, and z components of the total momentum of the two-body system.

The integrals $I_{sc}(3; 0, 0, 0)$, $I_{bcc}(1; 0, 0, 0)$, and $I_{fcc}(3; 0, 0, 0)$ (at the bottom of the band) were evaluated by Watson[8] in a closed form in terms of products of elliptic integrals.

The imaginary part of I_{sc} in Eqs. (12) for $l = m = n = 0$, the negative of the density of states in the ideal spin waves, was calculated by Bowers and Rosenstock,[9] by Montroll,[10] and by others.[11] The nature of the singularity was discussed by Van Hove,[12] in general, and by Montroll.[10]

The integral $I_{sc}(a; 0, 0, 0)$ is real for $a \geq 3$, complex for $0 \leq a < 3$, and it has singularities at $a = 1$ and $a = 3$. The expansion coefficients of I_{sc} $(a \geq 3; 0, 0, 0)$ in powers of $1/a^2$ and a short table of it were given by Tickson.[13] Maradudin *et al.*[5] obtained the first few coefficients of the expansion in terms of $a - 3$, and showed that the leading term at $a = 3$ is $I_{sc}(3) + O((a - 3)^{\frac{1}{2}})$. They also gave a table of I_{tetra} $(a \geq 2 + \gamma; l, m, n; \gamma)$ for $l^2 + m^2 + n^2 \leq 15$, $\gamma = 1, 2, 4, 8, 16$. Mannari and Kawabata[6] expressed I_{sc} $(a \geq 3; 0, 0, 0)$, I_{bcc} $(a \geq 1; 0, 0, 0)$, and I_{fcc} $(a \geq 3; 0, 0, 0)$ in terms of the definite single integral of the elliptic functions and gave extensive tables by carrying out the numerical calculations of these integrals.

Short tables of the real and the imaginary parts of $I_{sc}(a; l, m, n)$ were given by Koster and Slater[14] $(0 \leq a \leq 3, 000; 3 \leq a \leq 5, 000, 100, 200, 300, 400; a = 3.5$, up to 333), by Wolfram and Callaway[15] $(0 \leq a \leq 6, 000, 100, 011, 200)$, and by Hone, Callen, and Walker[16] $(0 \leq a \leq 9, 000, 011, 200, 022)$ by numerical integration of

$$I_{sc}(a \geq 3; l, m, n) = \int_0^\infty dt e^{-at} I_l(t) I_m(t) I_n(t),$$
$$I_{sc}(a < 3; l, m, n) = i^{l+m+n+1}\int_0^\infty dt e^{-iat} J_l(t) J_m(t) J_n(t). \qquad (14)$$

Yussouff and Mahanty[17] also gave tables of $I_{sc}(a; l, m, n)$ $(0 \leq a \leq 7$, up to 200). Vashishta and Yussouff[18] gave rough tables of $I_{sc}(a; l, m, n)$ $(0 \leq a \leq 9$, up to 200), $I_{bcc}(a; l, m, n)$ $(0 \leq a \leq 3$, up to 220), and $I_{fcc}(a; l, m, n)$ $(-5 \leq a \leq 7$, up to 220) by the Fourier expansion method. Values of $I_{bcc}(a; l, m, n)$ were also given by Yussouff and Mahanty[19] $(0 \leq a \leq 3$, up to 222), by Walker, Celtin, and Hone[20] $(0 \leq a \leq 2$, up to 444).

Other available sources of the lattice Green's functions for other crystal structures are the following: The lattice Green's functions for the square and rectangular lattices are defined by

$$I_{sq}(a; l, m) = I_{rect}(a; l, m; 1, 1), \tag{15}$$

$$I_{rect}(a; l, m; \gamma_1, \gamma_2) = \frac{1}{\pi^2} \iint_0^\pi \frac{\cos lx \cos my}{a - \gamma_1 \cos x - \gamma_2 \cos y} \, dx \, dy. \tag{15'}$$

The real parts of these functions are odd or even functions of a and the imaginary parts are even or odd functions of a, according as the sum of l and m is even or odd. In a similar way, $I_{rect}(a; l, m; \gamma_1, \gamma_2)$ can be regarded as the two-body Green's function of the square lattice.

For the square lattice, the density of states $\text{Im}\, I_{sq}\ (0 \le a \le 2; 0, 0)$ was given by Bowers and Rosenstock[9] and by Montroll,[10,21] the real part of $I_{sq}\ (a \ge 2; 0, 0)$ outside the band by Mannari and Kageyama.[22] The density of states of the orthorhombic lattice at the band edge was given by Montroll.[10] In Ref. 10, the imaginary part for the rectangular lattice $\text{Im}\, I_{rect}(a; 0, 0; \gamma_1, \gamma_2)$ was also expressed in terms of the complete elliptic integral of the first kind.

The function $I_{bcc}\ (a \ge 1; 0, 0, 0)$ was expressed as a square of the complete elliptic integral of the first kind by Maradudin.[23] The functions $I_{fcc}\ (a \ge 3; 0, 0, 0)$ and $I_{fcc}\ (a \le -1; 0, 0, 0)$ were expressed as products of the complete elliptic integrals (with different moduli) by Iwata.[24]

In spite of great interest in the physical applications, little is known about the analytical nature and the accurate numerical values. For example, the exact value of $I_{sc}(1; 0, 0, 0)$ is not yet known. The authors are presenting several new methods of calculating the lattice Green's functions in a series of papers. In the second paper of this series,[1] $I_{sc}\ (a \ge 3; 0, 0, 0)$ is transformed into a Mellin–Barnes type integral, and it is shown that its analytic continuation gives the real and the imaginary parts of $I_{sc}\ (0 \le a \le 1; 0, 0, 0)$ and $I_{sc}\ (1 \le a \le 3; 0, 0, 0)$ in the form of a double series, which gives simple and rapid subroutines for numerical calculations. The exact values of $\text{Re}\, I_{sc}(1; 0, 0, 0)$, $\text{Im}\, I_{sc}\ (1; 0, 0, 0)$, $\text{Im}\, I_{sc}(0; 0, 0, 0)$, $\text{Re}\, I_{sc}(\sqrt{5}; 0, 0, 0)$, and $\text{Im}\, I_{sc}(\sqrt{5}; 0, 0, 0)$ are also obtained. This method has the advantages that the analytic properties of the function are easily discussed and that the generalization to cases $I_{sc}(a; l, m, n)$ is straightforward.

In the third paper,[2] the expressions of $I_{tetra}\ (a \ge 2 + \gamma; 0, 0, 0; \gamma)$, $I_{bcc}\ (a \ge 1; 0, 0, 0)$, and $I_{fcc}\ (a \ge 3; 0, 0, 0)$ given in the form of the definite integral of the complete elliptic integral of the first kind are analytically continued to the whole range of $-\infty < a < \infty$. The results are expressed as a sum of definite integrals of the complete elliptic integrals and have been used for the numerical calculation of $I_{sc}\ (0 \le a < \infty; 0, 0, 0)$, $I_{bcc}\ (0 \le a < \infty; 0, 0, 0)$, and $I_{fcc}\ (-\infty < a < \infty; 0, 0, 0)$.

A divergence of the Green's function[3,4,25] is found for the bcc and fcc lattices. It has been discussed in connection with Van Hove's discussion of the lattice spectrum.[25,4]

The analytic continuations of the closed expressions of $I_{bcc}\ (a > 1; 0, 0, 0)$ due to Maradudin[5] and $I_{fcc}\ (a \ge 3; 0, 0, 0)$ and $I_{fcc}\ (a \le -1; 0, 0, 0)$ due to Iwata[24] are discussed in two different standpoints. In one paper,[3] $\text{Re}\, I_{bcc}\ (0 \le a \le 1; 0, 0, 0)$ and $\text{Im}\, I_{bcc}\ (0 \le a \le 1; 0, 0, 0)$ are expressed in terms of the hypergeometric function of real variables. The asymptotic behaviors at $a \gtrsim 0$, $a \lesssim 1$, and $a \gtrsim 1$ and the nature of the divergence at $a = 0$ are easily derived. These expressions provide a simple and rapid subroutine for numerical calculations.

In another paper,[26] it is pointed out that the arithmetic–geometric means method is powerful in calculating the elliptic integrals (and the Jacobian elliptic functions) also when the modulus is complex. The method is quite convenient in evaluating the analytic continuations for the Green's function which are expressed in terms of the complete elliptic integral with complex modulus.[4] Such an expression is provided for $I_{sq}(a; 0, 0)$, $I_{rect}(a; 0, 0; \gamma_1, \gamma_2)$, $I_{bcc}(a; 0, 0, 0)$, and $I_{fcc}(a; 0, 0, 0)$.[4] The numerical calculations of these quantities for an arbitrary complex variable a can be performed as easily as for the real variable a, $-\infty < a < +\infty$. These expressions are also used in discussing the analytic behaviors. These methods[3,4] are more convenient in evaluating $I_{bcc}\ (-\infty < a < \infty; 0, 0, 0)$ and $I_{fcc}\ (-\infty < a < \infty; 0, 0, 0)$ than the one given in the third paper[2] of this series.

We notice here that the expressions derived in the third paper[2] involves one of the coordinates x as an integration variable over which the final integration is still to be performed. Hence, when one has an additional function of x in the integrand of sc in Eqs. (12), the integrals in the final expressions involve the same factor. For instance, if the function is $\cos lx$, one obtains an expression useful for the calculation of $I_{tetra}(a; l, 0, 0; \gamma)$. This result is known to give the values of $I_{tetra}(a; l, m, n; \gamma)$ for some sets of l, m, and n.[5] The method of the third paper is expected to be generalized to $I_{ortho}(a; l, m, n; \gamma_1, \gamma_2, \gamma_3)$.

The calculations of $I_{sc}(a; l, m, n)$ by the methods mentioned in this paper are now in progress and will be reported as parts of this series.

[1] S. Katsura, S. Inawashiro, and Y. Abe, J. Math. Phys. **12**, 895 (1971).
[2] T. Morita and T. Horiguchi, "Lattice Green's Functions of the Cubic Lattices in Terms of the Complete Elliptic Integral," J. Math. Phys. (1971) (to be published).
[3] S. Katsura and T. Horiguchi, J. Math. Phys. **12**, 230 (1971).
[4] T. Morita and T. Horiguchi, "Calculation of the Lattice Green's Functions for the bcc, fcc, and Rectangular Lattices," J. Math. Phys. (1971) (to be published).
[5] A. A. Maradudin, E. W. Montroll, G. H. Weiss, R. Herman, and H. W. Milnes, Acad. Roy. Belg., Classe Sci., Mem: Collection in 4°, **14**, (1960).
[6] I. Mannari and C. Kawabata, Okayama Univ., Physics Research Notes No. 15, 1964.
[7] F. J. Dyson, Phys. Rev. **102**, 1217 (1956); T. Morita, Progr. Theoret. Phys. (Kyoto) **20**, 728 (1958); J. Hanus, Phys. Rev. Letters **11**, 336 (1963); N. Fukuda and M. Wortis, J. Phys. Chem. Solids **24**, 1675 (1963); R. G. Boyd and J. Callaway, Phys. Rev. **138**, A1621 (1965).
[8] G. N. Watson, Quart. J. Math. (Oxford) **10**, 266 (1939).
[9] W. A. Bowers and H. B. Rosenstock, J. Chem. Phys. **18**, 1056 (1950).
[10] E. W. Montroll, in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman, Ed. (U. of California Press, Berkeley, Calif., 1956), Vol. 3, p. 209.
[11] G. F. Newell, J. Chem. Phys. **21**, 1877 (1953).
[12] L. Van Hove, Phys. Rev. **89**, 1189 (1953).
[13] M. Tickson, J. Res. Natl. Bur. Std. **B50**, 177 (1953).
[14] G. F. Koster and J. S. Clater, Phys. Rev. **96**, 1208 (1954).
[15] T. Wolfram and J. Callaway, Phys. Rev. **130**, 2207 (1963).
[16] D. Hone, H. Callen, and L. R. Walker, Phys. Rev. **144**, 283 (1966).
[17] M. Yussouff and J. Mahanty, Proc. Phys. Soc. (London) **85**, 1223 (1965).
[18] P. Vashishta and M. Yussouff, Department of Physics, Indian Institute of Technology, Kanpur, India, Technical Report No. 12/66, 1966.
[19] M. Yussouff and J. Mahanty, Proc. Phys. Soc. (London) **87**, 689 (1966).
[20] L. R. Walker, B. B. Celtin, and D. Hone, J. Phys. Chem. Solids **30**, 923 (1969) The tables for bcc's and fcc's in Refs. 18–20 were calculated for even *l*, *m*, and *n*. The values for odd *l*, *m*, and *n* in bcc's can be obtained from those for even *l*, *m*, and *n*.
[21] E. W. Montroll, J. Chem. Phys. **15**, 575 (1947).
[22] I. Mannari and H. Kageyama, Progr. Theoret. Phys. (Kyoto), Suppl. Extra Number, 269 (1968).
[23] A. A. Maradudin, cited in Ref. 5.
[24] G. Iwata, Natural Sci. Rept., Ochanomizu Univ., **20**, 13 (1969).
[25] T. Horiguchi and T. Morita, Phys. Letters **32A**, 191 (1970).
[26] T. Morita and T. Horiguchi, submitted to Numerische Math.

Note added in Proof: In the course of publication the authors found that the following references should be mentioned: S. Takeno, Progr. Theoret. Phys. (Kyoto) **28**, 33 (1962); **30**, 731 (1963); Y. Mitani and S. Takeno, *ibid.* **33**, 779 (1965); M. Okazaki, M. Inoue, Y. Toyozawa, T. Inui, and E. Hanamura, J. Phys. Soc. Japan **22**, 1349 (1967); R. J. Jelitto, J. Phys. Chem. Solids **30**, 609 (1969); T. Oguchi, J. Phys. Soc. Japan **26**, 32 (1969); I. Mannari and T. Kawabe, Progr. Theoret. Phys. (Kyoto) **44**, 359 (1970). Also the author's group submitted following articles as extensions of the present methods: S. Katsura and S. Inawashiro, "Lattice Green's Function for the Rectangular and the Square Lattices at Arbitrary Points," J. Math. Phys. (to be published); T. Morita and T. Horiguchi, "Formulas for the Lattice Green's Functions for the Cubic Lattices in Terms of the Complete Elliptic Integral," J. Phys. Soc. Japan **30**, No. 4 (1971) (to be published); T. Morita, "Useful Procedure for Computing the Lattice Green's Function—Square, Tetragonal, and bcc Lattices," J. Math. Phys. (to be published); T. Horiguchi, "Lattice Green's Function for the Simple Cubic Lattice," J. Phys. Soc. Japan **30**, No. 5 (1971) (to be published); T. Horiguchi, Y. Yamazaki, and T. Morita, "Lattice Green's Function for the Orthorhombic Lattice in Terms of Complete Elliptic Integral," J. Math. Phys. (to be published).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Lattice Green's Function for the Simple Cubic Lattice in Terms of a Mellin–Barnes Type Integral

SHIGETOSHI KATSURA,* SAKARI INAWASHIRO,* AND YOSHIHIKO ABE†
Department of Applied Physics and Department of Physics,† Tohoku University, Sendai, Japan*

(Received 20 July 1970)

The lattice Green's function for the simple cubic lattice

$$I(a) = \frac{1}{\pi^3}\iiint_0^\pi \frac{dx\,dy\,dz}{a + i\epsilon - \cos x - \cos y - \cos z}$$

for $a > 3$ is expressed as a Mellin–Barnes type integral. The analytic continuation gives simple and useful expressions in series for the numerical calculation of the real part $I_R(a)$ and the imaginary part $I_I(a)$ of the integral for $0 < a < 1$ and $1 < a < 3$. The values at $a = 1$, $a = 0$, and $a = \sqrt{5}$ are obtained exactly: $I_R(1) = (\pi/2)[\Gamma(5/8)\Gamma(7/8)]^{-2}$, $I_I(1)/(-i) = (\sqrt{2})I_R(1)$, and $I_I(0) = 3 \cdot 2^{-11/3}\pi^{-4}[(\Gamma(1/3)]^6$.

1. INTRODUCTION

The lattice Green's function for the simple cubic lattice

$$I(a) = \frac{1}{\pi^3}\iiint_0^\pi \frac{dx\,dy\,dz}{a + i\epsilon - \cos x - \cos y - \cos z} \quad (1)$$

is considered.[1] For $\epsilon \to 0$, the integral (1) is real for $a \geq 3$ and complex for $0 < a < 3$. The integral (1) has singularities at $a = 3$ and $a = 1$. The value $I(3)$ is given by Watson[2]:

$$I(3) = (18 + 12\sqrt{2} - 10\sqrt{3} - 7\sqrt{6})(2/\pi)^2 \times \{\mathbf{K}[(2 - \sqrt{3})(\sqrt{3} - \sqrt{2})]\}^2. \quad (2)$$

[1] S. Katsura, S. Inawashiro, and Y. Abe, J. Math. Phys. **12**, 895 (1971).
[2] T. Morita and T. Horiguchi, "Lattice Green's Functions of the Cubic Lattices in Terms of the Complete Elliptic Integral," J. Math. Phys. (1971) (to be published).
[3] S. Katsura and T. Horiguchi, J. Math. Phys. **12**, 230 (1971).
[4] T. Morita and T. Horiguchi, "Calculation of the Lattice Green's Functions for the bcc, fcc, and Rectangular Lattices," J. Math. Phys. (1971) (to be published).
[5] A. A. Maradudin, E. W. Montroll, G. H. Weiss, R. Herman, and H. W. Milnes, Acad. Roy. Belg., Classe Sci., Mem: Collection in 4°, **14**, (1960).
[6] I. Mannari and C. Kawabata, Okayama Univ., Physics Research Notes No. 15, 1964.
[7] F. J. Dyson, Phys. Rev. **102**, 1217 (1956); T. Morita, Progr. Theoret. Phys. (Kyoto) **20**, 728 (1958); J. Hanus, Phys. Rev. Letters **11**, 336 (1963); N. Fukuda and M. Wortis, J. Phys. Chem. Solids **24**, 1675 (1963); R. G. Boyd and J. Callaway, Phys. Rev. **138**, A1621 (1965).
[8] G. N. Watson, Quart. J. Math. (Oxford) **10**, 266 (1939).
[9] W. A. Bowers and H. B. Rosenstock, J. Chem. Phys. **18**, 1056 (1950).
[10] E. W. Montroll, in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman, Ed. (U. of California Press, Berkeley, Calif., 1956), Vol. 3, p. 209.
[11] G. F. Newell, J. Chem. Phys. **21**, 1877 (1953).
[12] L. Van Hove, Phys. Rev. **89**, 1189 (1953).
[13] M. Tickson, J. Res. Natl. Bur. Std. **B50**, 177 (1953).
[14] G. F. Koster and J. S. Clater, Phys. Rev. **96**, 1208 (1954).
[15] T. Wolfram and J. Callaway, Phys. Rev. **130**, 2207 (1963).
[16] D. Hone, H. Callen, and L. R. Walker, Phys. Rev. **144**, 283 (1966).
[17] M. Yussouff and J. Mahanty, Proc. Phys. Soc. (London) **85**, 1223 (1965).
[18] P. Vashishta and M. Yussouff, Department of Physics, Indian Institute of Technology, Kanpur, India, Technical Report No. 12/66, 1966.
[19] M. Yussouff and J. Mahanty, Proc. Phys. Soc. (London) **87**, 689 (1966).
[20] L. R. Walker, B. B. Celtin, and D. Hone, J. Phys. Chem. Solids **30**, 923 (1969) The tables for bcc's and fcc's in Refs. 18–20 were calculated for even *l*, *m*, and *n*. The values for odd *l*, *m*, and *n* in bcc's can be obtained from those for even *l*, *m*, and *n*.
[21] E. W. Montroll, J. Chem. Phys. **15**, 575 (1947).
[22] I. Mannari and H. Kageyama, Progr. Theoret. Phys. (Kyoto), Suppl. Extra Number, 269 (1968).
[23] A. A. Maradudin, cited in Ref. 5.
[24] G. Iwata, Natural Sci. Rept., Ochanomizu Univ., **20**, 13 (1969).
[25] T. Horiguchi and T. Morita, Phys. Letters **32A**, 191 (1970).
[26] T. Morita and T. Horiguchi, submitted to Numerische Math.

Note added in Proof: In the course of publication the authors found that the following references should be mentioned: S. Takeno, Progr. Theoret. Phys. (Kyoto) **28**, 33 (1962); **30**, 731 (1963); Y. Mitani and S. Takeno, *ibid.* **33**, 779 (1965); M. Okazaki, M. Inoue, Y. Toyozawa, T. Inui, and E. Hanamura, J. Phys. Soc. Japan **22**, 1349 (1967); R. J. Jelitto, J. Phys. Chem. Solids **30**, 609 (1969); T. Oguchi, J. Phys. Soc. Japan **26**, 32 (1969); I. Mannari and T. Kawabe, Progr. Theoret. Phys. (Kyoto) **44**, 359 (1970). Also the author's group submitted following articles as extensions of the present methods: S. Katsura and S. Inawashiro, "Lattice Green's Function for the Rectangular and the Square Lattices at Arbitrary Points," J. Math. Phys. (to be published); T. Morita and T. Horiguchi, "Formulas for the Lattice Green's Functions for the Cubic Lattices in Terms of the Complete Elliptic Integral," J. Phys. Soc. Japan **30**, No. 4 (1971) (to be published); T. Morita, "Useful Procedure for Computing the Lattice Green's Function—Square, Tetragonal, and bcc Lattices," J. Math. Phys. (to be published); T. Horiguchi, "Lattice Green's Function for the Simple Cubic Lattice," J. Phys. Soc. Japan **30**, No. 5 (1971) (to be published); T. Horiguchi, Y. Yamazaki, and T. Morita, "Lattice Green's Function for the Orthorhombic Lattice in Terms of Complete Elliptic Integral," J. Math. Phys. (to be published).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 5 MAY 1971

Lattice Green's Function for the Simple Cubic Lattice in Terms of a Mellin–Barnes Type Integral

SHIGETOSHI KATSURA,* SAKARI INAWASHIRO,* AND YOSHIHIKO ABE†
Department of Applied Physics and Department of Physics,† Tohoku University, Sendai, Japan*

(Received 20 July 1970)

The lattice Green's function for the simple cubic lattice

$$I(a)=\frac{1}{\pi^3}\iiint_0^\pi \frac{dx\,dy\,dz}{a+i\epsilon-\cos x-\cos y-\cos z}$$

for $a>3$ is expressed as a Mellin–Barnes type integral. The analytic continuation gives simple and useful expressions in series for the numerical calculation of the real part $I_R(a)$ and the imaginary part $I_I(a)$ of the integral for $0<a<1$ and $1<a<3$. The values at $a=1$, $a=0$, and $a=\sqrt{5}$ are obtained exactly: $I_R(1)=(\pi/2)[\Gamma(5/8)\Gamma(7/8)]^{-2}$, $I_I(1)/(-i)=(\sqrt{2})I_R(1)$, and $I_I(0)=3\cdot 2^{-11/3}\pi^{-4}[(\Gamma(1/3))]^6$.

1. INTRODUCTION

The lattice Green's function for the simple cubic lattice

$$I(a)=\frac{1}{\pi^3}\iiint_0^\pi \frac{dx\,dy\,dz}{a+i\epsilon-\cos x-\cos y-\cos z} \quad (1)$$

is considered.[1] For $\epsilon\to 0$, the integral (1) is real for $a\geq 3$ and complex for $0<a<3$. The integral (1) has singularities at $a=3$ and $a=1$. The value $I(3)$ is given by Watson[2]:

$$I(3)=(18+12\sqrt{2}-10\sqrt{3}-7\sqrt{6})(2/\pi)^2 \times\{\mathbf{K}[(2-\sqrt{3})(\sqrt{3}-\sqrt{2})]\}^2. \quad (2)$$

In this paper, $I(a)$ for $a > 3$ is expressed as a Mellin–Barnes type integral giving series representations for $a > 3$. The analytic continuation of the integral gives series representations of the real parts and the imaginary parts for $0 < a < 1$ and $1 < a < 3$. These expressions provide simple and rapid subroutines for the numerical calculations. The values at $a = 0$, $a = 1$, and $a = \sqrt{5}$ are obtained exactly.

2. $a > 3$

Equation (1) can be written

$$I(a) = -\frac{i}{\pi^3}\int_0^\infty dt\int_0^\pi dx\int_0^\pi dy\int_0^\pi dz \times \exp\{i[(a+i\epsilon) - \cos x - \cos y - \cos z]t\}.$$

When we carry out the integration $dx\,dy\,dz$, then

$$I(a) = -i\int_0^\infty e^{i(a+i\epsilon)t}[J_0(t)]^3\,dt. \tag{3}$$

Transforming the series representation of $[J_0(t)]^2$ into the Mellin–Barnes type integral representation, we have

$$[J_0(t)]^2 = \sum_{m=0}^{\infty}\frac{(-)^m(\frac{1}{2}t)^{2m}(m+1)_m}{m!\,[\Gamma(m+1)]^2} = \frac{1}{\sqrt{\pi}}\frac{1}{2\pi i}\int_C ds\,\frac{\Gamma(-s)\Gamma(s+\frac{1}{2})t^{2s}}{[\Gamma(s+1)]^2}, \tag{4}$$

where C is a straight line parallel to the imaginary axis crossing at the point $s = -\Delta(\Delta \to +0)$. Substituting Eq. (4) into Eq. (3) and changing the order of the integrations ds and dt, we obtain

$$I(a) = \frac{-i}{\sqrt{\pi}}\frac{1}{2\pi i}\int_C ds\,\frac{\Gamma(-s)\Gamma(s+\frac{1}{2})}{[\Gamma(s+1)]^2} \times \int_0^\infty dt\, e^{i(a+i\epsilon)t}J_0(t)t^{2s}.$$

For $\epsilon > 0$ and $2s + 1 > 0$, the integral $\int_0^\infty dt$ converges:

$$\int_0^\infty dt\cdots = \frac{\Gamma(2s+1)}{(-ia)^{2s+1}}\,{}_2F_1(s+\tfrac{1}{2}, s+1; 1; a^{-2}).$$

Hence

$$I(a) = \frac{1}{\pi a}\frac{1}{2\pi i}\int ds\,\frac{\Gamma(-s)[\Gamma(s+\frac{1}{2})]^2}{\Gamma(s+1)}\left(-\frac{4}{a^2}\right)^s \times {}_2F_1(s+\tfrac{1}{2}, s+1; 1; a^{-2}). \tag{5}$$

For $|a| > 1$, expanding ${}_2F_1$ and changing the order of the integration and the summation, we obtain

$$I(a) = \frac{1}{\pi a}\sum_{n=0}^{\infty}\frac{1}{n!\,\Gamma(1+n)}\left(\frac{1}{a^2}\right)^n\frac{1}{2\pi i}\int ds\left(-\frac{4}{a^2}\right)^s \times \frac{\Gamma(-s)\Gamma(s+\frac{1}{2})\Gamma(s+\frac{1}{2}+n)\Gamma(s+1+n)}{[\Gamma(s+1)]^2}. \tag{6}$$

For $|a| > 2$, the path of the integration is closed with a semicircle in the right half-plane. The pole of the integrand in this region is $s = 0, 1, 2, \cdots, m, \cdots$. By calculating the residues we have

$$I(a) = \frac{1}{\pi a}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \times \frac{\Gamma(m+\frac{1}{2})\Gamma(m+n+\frac{1}{2})\Gamma(m+n+1)4^m}{n!\,\Gamma(1+n)m!\,[\Gamma(1+m)]^2}\left(\frac{1}{a^2}\right)^{m+n}. \tag{7}$$

We investigate the region of the convergence of the double series[3]

$$I(x, y) = \frac{1}{\pi a}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \times \frac{\Gamma(m+\frac{1}{2})\Gamma(m+n+\frac{1}{2})\Gamma(m+n+1)}{n!\,\Gamma(n+1)m!\,[\Gamma(1+m)]^2}4^m x^m y^n \equiv \sum A_{mn}x^m y^n, \tag{8}$$

which is generalized from Eq. (7). Put $m = t\mu$ and $n = t\nu$; then, from

$$\frac{1}{r} \equiv \lim_{t\to\infty}\frac{A_{m+1,n}}{A_{m,n}} = \frac{4(\mu+\nu)^2}{\mu^2},$$

$$\frac{1}{s} \equiv \lim_{t\to\infty}\frac{A_{m,n+1}}{A_{m,n}} = \frac{(\mu+\nu)^2}{\nu^2},$$

we have

$$4r + s + (16rs)^{\frac{1}{2}} = 1. \tag{9}$$

For $r = s$, we have $r = \frac{1}{9}$. That is, Eq. (7) converges for $|a| > 3$.

Putting $m + n = p$, we transform Eq. (7) into

$$I(a) = \frac{1}{\pi a}\sum_{p=0}^{\infty}\Gamma(p+\tfrac{1}{2})\Gamma(p+1)\left(\frac{1}{a^2}\right)^p \times \sum_{m=0}^{p}\frac{\Gamma(m+\frac{1}{2})4^m}{(m!)^3[(p-m)!]^2}. \tag{10}$$

Equation (10) is a power series of $1/a^2$, when the coefficients are calculated by a finite series. Equation (10) is convenient for the numerical calculation. The imaginary part of $I(a)$ vanishes for $|a| > 3$.

3. $0 < a < 1$

The expansion of $I(a)$ for small a can be obtained by the analytic continuation of the integral representation (5). Transforming ${}_2F_1(\ ;\ ; 1/a^2)$ into ${}_2F_1(\ ;\ ; a^2)$ with use of the Kummer's relation, we obtain

$$I(a) = \frac{1}{\pi a}\frac{1}{2\pi i}\int ds \frac{\Gamma(-s)[\Gamma(s+\frac{1}{2})]^2}{\Gamma(s+1)}\left(-\frac{4}{a^2}\right)^s \times \left[\frac{\Gamma(\frac{1}{2})}{\Gamma(s+\frac{1}{2})\Gamma(-s+\frac{1}{2})}\left(-\frac{1}{a^2}\right)^{-s-\frac{1}{2}} \times {}_2F_1(s+\tfrac{1}{2}, s+\tfrac{1}{2}; \tfrac{1}{2}; a^2) + \frac{\Gamma(-\frac{1}{2})}{\Gamma(s+\frac{1}{2})\Gamma(-s)}\left(-\frac{1}{a^2}\right)^{-s-1} \times {}_2F_1(s+1, s+1; \tfrac{3}{2}; a^2)\right]. \quad (11)$$

The first and the second terms in the square brackets give the imaginary and the real parts, respectively. We denote $I(a) = I_R(a) + I_I(a)$.

Real part: Expanding ${}_2F_1$ in the second term and changing the order of the summation and the integration, we have

$$I_R(a) = a\sum_{n=0}^{\infty}\frac{a^{2n}}{n!\,\Gamma(n+\frac{3}{2})} \times \frac{1}{2\pi i}\int ds \frac{\Gamma(s+\frac{1}{2})[\Gamma(s+1+n)]^2 4^s}{[\Gamma(s+1)]^3}. \quad (12)$$

The path of the integration is closed with a semicircle in the left half-plane. The pole of the integrand is given by $s + \frac{1}{2} = 0, -1, -2, \cdots$. By calculating the residue, we have

$$I_R(a) = \frac{a}{2\pi}\sum_n^{\infty}\sum_m^{\infty}\frac{[\Gamma(\frac{1}{2}+m)]^3(\frac{1}{4})^m a^{2n}}{n!\,m!\,\Gamma(\frac{3}{2}+n)[\Gamma(\frac{1}{2}-n+m)]^2}. \quad (13)$$

The region of the convergence of the double series $\sum A_{mn}x^m y^n$ generalized from Eq. (13) is given by

$$r^{-1} = \mu^2/(\mu-\nu)^2, \quad s^{-1} = (\mu-\nu)^2/\nu^2.$$

Hence

$$r^{-\frac{1}{2}} - s^{\frac{1}{2}} = 1. \quad (14)$$

For $x = \frac{1}{4}$, we have $y = 1$. Hence Eq. (13) converges for $|a| < 1$.

Imaginary part: The contribution of the first term in the bracket in Eq. (11) gives

$$I_I(a) = \frac{-i}{\sqrt{\pi}}\frac{1}{2\pi i}\int ds 4^s \frac{\Gamma(-s)[\Gamma(s+\frac{1}{2})]^2}{[\Gamma(s+1)]^2\Gamma(-s+\frac{1}{2})} \times {}_2F_1(s+\tfrac{1}{2}, s+\tfrac{1}{2}; \tfrac{1}{2}; a^2). \quad (15)$$

Expanding ${}_2F_1$ in powers of a^2, we obtain

$$I_I(a) = -i\sum_{n=0}^{\infty}\frac{a^{2n}}{n!\,\Gamma(\frac{1}{2}+n)} \times \frac{1}{2\pi i}\int ds \frac{\Gamma(-s)[\Gamma(s+n+\frac{1}{2})]^2}{[\Gamma(s+1)]^2\Gamma(-s+\frac{1}{2})}4^s. \quad (16)$$

The path of the integration is closed with a semicircle in the left-hand plane, in which double poles $s = -\frac{1}{2} - n - m$, $m = 0, 1, 2, \cdots$, are found. By calculating the residues of the double poles, we have

$$\frac{1}{2\pi i}\int ds 4^s \frac{\Gamma(-s)[\Gamma(s+n+\frac{1}{2})]^2}{[\Gamma(s+1)]^2\Gamma(-s+\frac{1}{2})}$$
$$= \sum_{m=0}^{\infty}\left[\frac{d}{ds}(s+\tfrac{1}{2}+n+m)^2 \times \frac{[\Gamma(s+n+\frac{1}{2})]^2\Gamma(-s)4^s}{[\Gamma(s+1)]^2\Gamma(-s+\frac{1}{2})}\right]_{s=-\frac{1}{2}-n-m}$$
$$= \sum_{m=0}^{\infty}\left(\frac{d}{ds}\frac{\pi^2(s+\frac{1}{2}+n+m)^2}{\sin^2[\pi(-s-n-\frac{1}{2})]} \times \frac{\Gamma(-s)4^s}{[\Gamma(1-s-n-\frac{1}{2})]^2[\Gamma(s+1)]^2\Gamma(-s+\frac{1}{2})}\right)_{s=-\frac{1}{2}-n-m}$$
$$= \sum_{m=0}^{\infty}\frac{\Gamma(\frac{1}{2}+m+n)4^{-\frac{1}{2}-n-m}}{[\Gamma(1+m)]^2[\Gamma(\frac{1}{2}-m-n)]^2\Gamma(1+n+m)} \times [-\psi(\tfrac{1}{2}+m+n) + 2\psi(1+m) - 2\psi(\tfrac{1}{2}-m-n) + \psi(1+n+m) + \log 4]. \quad (17)$$

Hence

$$I_I(a) = -\frac{i}{2\pi^2}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \times \frac{(a^2/2^2)^n(\frac{1}{4})^m[\Gamma(\frac{1}{2}+m+n)]^3}{n!\,\Gamma(\frac{1}{2}+n)[\Gamma(1+m)]^2\Gamma(1+m+n)} \times [-3\psi(\tfrac{1}{2}+m+n) + 2\psi(1+m) + \psi(1+n+m) + \log 4]. \quad (18)$$

The radius of the convergence of Eq. (18) is also $a = 1$.

4. $1 < a < 3$

The expansion of $I(a)$ at $a^2 = 5$, which is convergent for $|a^2 - 5| < 4$, i.e., $1 < a < 3$, is derived by the analytic continuation from the expansion (7).

For large value of a^2, Eq. (7) is convergent and is equal to

$$I(a) = \frac{1}{\pi a}\sum_{m=0}^{\infty}\frac{[\Gamma(\frac{1}{2}+m)]^2}{(m!)^2}\left(\frac{4}{a^2}\right)^m \times {}_2F_1(m+1, m+\tfrac{1}{2}, -1; a^{-2}). \quad (19)$$

Transforming ${}_2F_1(\ ;\ ; 1/a^2)$ into ${}_2F_1(\ ;\ ; 1/(1-a^2))$ with use of Kummer's relation

$${}_2F_1(a, b, c; z) = (1-z)^{-a}\,{}_2F_1(a, c-b, c; z/(z-1)), \quad (20)$$

we have

$$I(a)=\frac{1}{\pi}\left(\frac{1}{a^2-1}\right)^{\frac{1}{2}}\sum_{m=0}^{\infty}\left(\frac{\Gamma(\frac{1}{2}+m)}{m!}\right)^2\left(\frac{4}{a^2-1}\right)^m \times {}_2F_1\left(m+\tfrac{1}{2},-m,1;\frac{1}{1-a^2}\right), \tag{21}$$

which is valid for large a^2.

Expanding the hypergeometric function and changing the order of the summation, we have

$$I(a)=\frac{1}{\pi}\left(\frac{1}{a^2-1}\right)^{\frac{1}{2}}\sum_{n=0}^{\infty}\frac{\Gamma(\frac{1}{2}+n)\Gamma(\frac{1}{2}+2n)}{(n!)^3}\left(\frac{2}{a^2-1}\right)^{2n} \times {}_2F_1\left(\tfrac{1}{2}+n,\tfrac{1}{2}+2n;1+n;\frac{4}{a^2-1}\right). \tag{22}$$

Again applying the relation (20), we transform ${}_2F_1(\ ;\ ;4/(a^2-1))$ into $F(\ ;\ ;4/(5-a^2))$ and we have

$$\begin{aligned}I(a)&=\frac{1}{\pi}\sum_{n=0}^{\infty}\frac{\Gamma(\frac{1}{2}+n)\Gamma(\frac{1}{2}+2n)}{(n!)^3}4^n\left(\frac{1}{a^2-5}\right)^{2n+\frac{1}{2}} \times {}_2F_1\left(\tfrac{1}{2},\tfrac{1}{2}+2n;1+n;\frac{4}{5-a^2}\right)\\&=\frac{1}{\pi}\left(\frac{1}{a^2-5}\right)^{\frac{1}{2}}\frac{1}{2\pi i}\int_{\mathrm{Re}\,t=-\Delta}\left\{dt\left[-4\left(\frac{1}{a^2-5}\right)^2\right]^t \times\frac{\Gamma(-t)\Gamma(\frac{1}{2}+t)\Gamma(\frac{1}{2}+2t)}{[\Gamma(1+t)]^2}\right\}\\&\quad\times F\left(\tfrac{1}{2},\tfrac{1}{2}+2t;1+t;\frac{4}{5-a^2}\right).\end{aligned} \tag{23}$$

Transformation of $F(\ ;\ ;4/(5-a^2))$ into $F(\ ;\ ;(5-a^2)/4)$ and the expansion of the hypergeometric function leads to

$$\begin{aligned}I(a)&=\frac{1}{\pi}\left(\frac{1}{a^2-5}\right)^{\frac{1}{2}}\frac{1}{2\pi i}\int_{\mathrm{Re}\,t=-\Delta}\left\{dt\left[-4\left(\frac{1}{a^2-5}\right)^2\right]^t \times\frac{\Gamma(-t)\Gamma(\frac{1}{2}+t)\Gamma(\frac{1}{2}+2t)}{[\Gamma(1+t)]^2}\right\}\\&\quad\times\left[\sum_{n=0}^{\infty}\frac{(-1)^n\Gamma(\frac{1}{2}+n)\Gamma(2t-n)}{n!\,\Gamma(\frac{1}{2}-n+t)}\left(\frac{a^2-5}{4}\right)^{n+\frac{1}{2}}\right.\\&\quad+\sum_{n=0}^{\infty}\frac{(-1)^n\Gamma(\frac{1}{2}+2t+n)\Gamma(-n-2t)}{n!\,\Gamma(\frac{1}{2}-n-t)}\\&\quad\left.\times\left(\frac{a^2-5}{4}\right)^{n+2t+\frac{1}{2}}\right].\end{aligned} \tag{24}$$

Changing the order of the integration and summation and closing the path of integration with a semicircle in the left half-plane, we finally obtain

$$I(a)=I_{\mathrm{R}}(a)+I_{\mathrm{I}}(a), \tag{25}$$

where

$$\begin{aligned}I_{\mathrm{R}}(a)&=\frac{1}{4\pi^{\frac{5}{2}}}\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\left(\frac{a^2-5}{4}\right)^n\\&\quad\times\sum_{m=0}^{\infty}\frac{(-1)^m\Gamma(\frac{1}{2}+m)\Gamma(\frac{1}{2}+n+m)\Gamma(\frac{1}{2}+n+2m)}{(m!)^2(n+2m)!}\left(\frac{1}{4}\right)^m\\&\quad\times[-\psi(\tfrac{1}{2}+m)-\psi(\tfrac{1}{2}+n+m)\\&\quad-2\psi(\tfrac{1}{2}+n+2m)+2\psi(1+m)\\&\quad+2\psi(1+n+2m)+\log 4]\end{aligned} \tag{26}$$

and

$$\begin{aligned}I_{\mathrm{I}}(a)&=\frac{-i}{4\pi^2}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{\Gamma(\frac{1}{2}+m)[\Gamma(\frac{1}{2}+n)]^2(m+n)!}{m!\,n!\,\Gamma(n+\frac{1}{2}m+1)\Gamma(n+\frac{1}{2}m+\frac{3}{2})}\\&\quad\times\frac{1}{2^m}\left(\frac{5-a^2}{4}\right)^{m+2n+1}\\&\quad-\frac{i}{4\pi^{\frac{3}{2}}}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{\Gamma(\frac{1}{2}+m)\Gamma(\frac{1}{2}+n+m)\Gamma(\frac{1}{2}+n+2m)}{(m!)^2n!(n+2m)!}\\&\quad\times\left(\frac{-1}{4}\right)^m\left(\frac{5-a^2}{4}\right)^n.\end{aligned} \tag{27}$$

The first term in Eq. (27) can also be written as

$$-\frac{i}{2\pi^{\frac{5}{2}}}\sum_{p=0}^{\infty}\frac{1}{(p+1)!}\left(\frac{5-a^2}{4}\right)^{p+1} \times\sum_{n=0}^{\leq p/2}\frac{4^n[\Gamma(\frac{1}{2}+n)]^2\Gamma(p-2n+\frac{1}{2})\Gamma(p-n+1)}{n!\,\Gamma(p-2n+1)}. \tag{27'}$$

Equations (26) and (27) are power series of $(a^2-5)/4$, which is convergent for $|a^2-5|<4$, i.e., $1<a<3$.

5. $I_{\mathrm{R}}(1)$, $I_{\mathrm{I}}(1)$, $I_{\mathrm{I}}(0)$, $I_{\mathrm{R}}(\sqrt{5})$, AND $I_{\mathrm{I}}(\sqrt{5})$

The value of $I_{\mathrm{R}}(1)$ can be obtained exactly. Carrying out $\sum_n$ in Eq. (13), we obtain

$$I_{\mathrm{R}}(a)=\frac{a}{2\pi^{\frac{3}{2}}}\sum_{m=0}^{\infty}\frac{\Gamma(m+\frac{1}{2})}{m!}\left(\frac{1}{4}\right)^m \times {}_2F_1(\tfrac{1}{2}-m,\tfrac{1}{2}-m;\tfrac{3}{2};a^2). \tag{28}$$

Putting $a=1$, and using the formula of ${}_2F_1(\ ;\ ;1)$,

$$I_{\mathrm{R}}(1)=\frac{1}{2\pi}\sum_{m=0}^{\infty}\frac{(\frac{1}{4})^m\Gamma(\frac{1}{2}+m)\Gamma(\frac{1}{2}+2m)}{m!\,[\Gamma(1+m)]^2}. \tag{29}$$

The summation $\sum_m$ can be expressed in terms of ${}_3F_2$. By using the formula for ${}_3F_2(\ ;\ ;1)$, $I_{\mathrm{R}}(1)$ is given by

$$I_R(1) = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{4})\Gamma(\frac{3}{4})}{2^{\frac{3}{2}}\pi^{\frac{3}{2}}} {}_3F_2(\tfrac{1}{4}, \tfrac{1}{2}, \tfrac{3}{4}; 1, 1; 1)$$

$$= (1 - 2^{-\frac{1}{2}})\left(\frac{2}{\pi}\right)^2 [\mathbf{K}((2\sqrt{2} - 2)^{\frac{1}{2}})]^2 \tag{30}$$

$$= \tfrac{1}{2}\pi[\Gamma(\tfrac{5}{8})]^{-2}[\Gamma(\tfrac{7}{8})]^{-2} \tag{30'}$$

$$= 0.642882248294458 \cdots. \tag{30''}$$

The value of $I_I(0)$ can also be obtained exactly. From Eq. (18) we have

$$\frac{I_I(0)}{-i} = \frac{1}{2\pi} \sum_{m=0}^{\infty} \left(\frac{\Gamma(m + \frac{1}{2})}{\pi^{\frac{1}{2}}\,\Gamma(m + 1)}\right)^3 \left(\frac{1}{2}\right)^{2m} \times [3\psi(1 + m) - 3\psi(m + \tfrac{1}{2}) + 2 \log 2]. \tag{31}$$

The right-hand side is expressed in terms of product of elliptic integrals[2] $(2/\pi^2)\mathbf{K}(k)\mathbf{K}(k')$, where $kk' = \frac{1}{4}$. Hence

$$\frac{I_I(0)}{-i} = \frac{2}{\pi^2} \mathbf{K}\left(\left(\frac{2 + \sqrt{3}}{4}\right)^{\frac{1}{2}}\right) \mathbf{K}\left(\left(\frac{2 - \sqrt{3}}{4}\right)^{\frac{1}{2}}\right) \tag{32}$$

$$= \frac{3[\Gamma(\frac{1}{3})]^6}{2^{\frac{11}{3}}\pi^4} \tag{32'}$$

$$= 0.896440788776763 \cdots. \tag{32''}$$

Transformation of ${}_2F_1(\ ;\ ; a^2)$ in Eq. (11) into ${}_2F_1(\ ;\ ; 1 - a^2)$ or ${}_2F_1(\ ;\ ; 1 - 1/a^2)$ gives the values of the real and the imaginary parts at $a = 1$:

$$\left.\begin{matrix} I_R(1) \\ I_I(1)/(-i) \end{matrix}\right\} = \frac{1}{4}\frac{[\Gamma(\frac{1}{4})]^2}{[\Gamma(\frac{5}{8})]^4} \mp \frac{1}{8}\frac{[\Gamma(\frac{3}{4})]^2}{[\Gamma(\frac{7}{8})]^4}$$

$$= \begin{cases} \frac{1}{2}\pi[\Gamma(\frac{5}{8})\Gamma(\frac{7}{8})]^{-2}, \\ 2^{-\frac{1}{2}}\pi[\Gamma(\frac{5}{8})\Gamma(\frac{7}{8})]^{-2} = 2^{\frac{1}{2}} I_R(1), \end{cases} \tag{33}$$

$$= \begin{cases} 0.642882248294458 \cdots, \\ 0.909172794546930 \cdots, \end{cases} \tag{33'}$$

respectively.

The values of $I_R(\sqrt{5})$ and $I_I(\sqrt{5})$ can be obtained in a similar way. The results are

$$I_R(\sqrt{5}) = \tfrac{1}{2}(\sqrt{5} - 2)^{\frac{1}{2}}(2/\pi)^2\mathbf{K}(k)\mathbf{K}(k')$$

$$= 0.547093244340170$$

$$= \sqrt{5}\, I_I(\sqrt{5})/(-i), \tag{34}$$

$$I_I(\sqrt{5})/(-i) = \tfrac{1}{2}(\sqrt{5} - 2)^{\frac{1}{2}}(2/\pi)^2[\mathbf{K}(k)]^2$$

$$= 0.244667536875105 \cdots, \tag{35}$$

where

$$k^2 = \tfrac{1}{2} - (\sqrt{5} - 2)^{\frac{1}{2}}. \tag{36}$$

6. CONCLUSION

Equations (10), (13), (18), (26), and (27) are the desired series representations of $I_R(a > 3; 0, 0, 0)$, $I_R(0 < a < 1;\ 0,\ 0,\ 0)$, $I_I(0 < a < 1;\ 0,\ 0,\ 0)$, $I_R(1 < a < 3; 0, 0, 0)$, and $I_I(1 < a < 3; 0, 0, 0)$, respectively. They are useful for analyzing the nature of analytical properties and supply simple and rapid subroutines for carrying out numerical calculations. Though these expressions seem to be double series, they become single series of independent variables after the coefficients have once been calculated.

The numerical values obtained from Eq. (10) reproduces Mannari and Kawabata's table.[4] Those from Eqs. (13), (18), (26), and (27) agree with values obtained by the method presented in the next paper by Morita and Horiguchi.[5] The exact values of $I_R(1)$, $I_I(1)$, $I_I(0)$, $I_R(\sqrt{5})$, and $I_I(\sqrt{5})$ are given in Eqs. (30), (32), (33), (34), and (35). It is to be noted that all these values and $I_R(3)$ are expressed in the product of the complete elliptic integrals of the first kind. The general expression for arbitrary a may perhaps exist in a generalized form.

ACKNOWLEDGMENTS

The authors acknowledge helpful discussions with Dr. T. Morita and Mr. T. Horiguchi.

[1] As a general introduction, see S. Katsura, T. Morita, S. Inawashiro, T. Horiguchi, and Y. Abe, J. Math. Phys. **12**, 892 (1971) (preceding paper).

[2] G. N. Watson, Quart. J. Math. **10**, 266 (1939).

[3] Bateman Project, *Higher Transcendental Functions*, A. Erdélyi, Ed. (McGraw-Hill, New York, 1953), Vol. 1.

[4] I. Mannari and C. Kawabata, Department of Physics, Okayama Univ., Research Notes, No. 15, 1964.

[5] T. Morita and T. Horiguchi, "Lattice Green's Functions of the Cubic Lattices in Terms of the Complete Elliptic Integral," J. Math. Phys. (to be published).