JOURNAL OF MATHEMATICAL PHYSICS VOLUME 13, NUMBER 1 JANUARY 1972

The Jost Function of a Momentum-Dependent Potential

Chindhu S. Warke*
Physics Department, McMaster University, Hamilton, Ontario
(Received 19 July 1971)

It is proved that the Jost function of a central, momentum-dependent, S-state interaction is equal to the ratio of the Fredholm determinant of the kernel of the integral equation for the outgoing scattering solution in the momentum space representation to that of the regular solution in the same representation. This proof of the expected result is more general than that given by Warke and Bhaduri for a nonlocal interaction.

I. INTRODUCTION

Jost and Pais[1] proved that the "Jost function" of a central, local potential is equal to the Fredholm determinant of the kernel of the integral equation for the outgoing scattering solution ψ^+ of the Schrödinger equation. This result, valid only for S-state interactions, was later generalized by Newton[2,3] to the higher partial waves as well as for local, noncentral potentials. Recently, these derivations have been extended for the nonlocal, central, S-state interaction by Warke and Bhaduri[4] and to the higher partial waves and for the nonlocal tensor forces by Singh and Warke.[5] It is found that the Jost function for a nonlocal potential differs from the Fredholm determinant of the ψ^+ solution by a real, k^2-dependent factor. This factor gives rise to the redundant zeros of the Jost function for positive imaginary values of k which do not correspond to bound states.

The momentum-dependent one-boson exchange nucleon–nucleon potentials and phenomenological momentum-dependent potentials are being used in the problems of nuclear physics. These potentials reproduce experimental phase shifts up to 350 MeV. Therefore, it would be useful to investigate the general form of the Jost function for a momentum-dependent potential. Contrary to local and nonlocal interactions, in this case, the Wronskian of two independent solutions of the Schrödinger equation at $r = 0$ is not equal to that at $r = \infty$. Moreover, the momentum-dependent interaction U does not commute with the radial wavefunction. Because of these two reasons, the usual derivations quoted above are not applicable. One has to investigate a new approach to find out the general form of the Jost function in this case. In this paper we prove that the Jost function for a central, momentum-dependent, S-state interaction is equal to the ratio of the Fredholm determinant of the kernel of the integral equation in the momentum space representation of the ψ^+ solution to that of the regular solution ϕ in the same representation.

II. DERIVATION OF THE JOST FUNCTION AND DISCUSSION

For reasons of simplicity, the proof given here is restricted only for the S-state Jost function. Let the central momentum-dependent nucleon–nucleon interaction be γU, where

$$U = [p^2\omega(r) + \omega(r)p^2]/m \equiv \hbar^2 V(r)/m, \tag{1}$$

where

$$V(r) = -\left(\frac{d^2\omega(r)}{dr^2} + 2\omega(r)\frac{d^2}{dr^2} + 2\frac{d\omega(r)}{dr}\frac{d}{dr}\right),$$

m being the nucleon mass, p the relative momentum, and the radial part $\omega(r)$; its first and second derivatives go to zero as $r \to \infty$. The notations used in this paper are close to that in Newton's book.[5] Let us introduce the various solutions of the Schrödinger equation

$$\left(k^2 + \frac{d^2}{dr^2} - \gamma V(r)\right)\psi = 0. \tag{2}$$

The outgoing scattering solution is

$$\psi^+(r) = \psi_0(r) + \gamma\int G_+(r,r')V(r')\psi^+(r')dr', \tag{3}$$

where

$$\psi_0(r) = \sin kr$$

and

$$G_+(r,r') = -\sin kr_< \exp(ikr_>)/k. \tag{4}$$

The regular solution which obeys the boundary conditions

$$\phi(r=0) = 0 \quad \text{and} \quad \frac{d\phi}{dr}\Big|_{r=0} = 1 \tag{5}$$

is given by

$$\phi(r) = \phi_0(r) + \gamma\int G(r,r')V(r')\phi(r')dr', \tag{6}$$

where

$$\phi_0(r) = \sin kr/k$$

$$G(r,r') = \begin{cases} \sin k(r-r')/k & \text{for } r' \le r \\ 0 & \text{otherwise} \end{cases}. \tag{7}$$

Finally, the Jost solutions of (2) have the form

$$f^\pm(r) = f_0^\pm(r) + \gamma\int \tilde{G}(r,r')V(r')f^\pm(r')dr', \tag{8}$$

where

$$f_0^\pm(r) = e^{\pm ikr} \quad \text{and} \quad \tilde{G}(r,r') = G(r',r). \tag{9}$$

These solutions satisfy the boundary conditions

$$e^{\mp ikr}f^\pm(r) \to 1 \quad \text{as} \quad r \to \infty. \tag{10}$$

Note that, in Eqs. (3)–(10) and what follows in this section, we suppress the k subscript to the various solutions and their Green's functions. The Jost functions $f^\pm(k)$ are defined to be the values of the Jost solutions $f^\pm(r)$ at $r = 0$.

Let the Wronskian of two independent solutions f and g of (2) be

$$W[f,g] = f(r)\frac{dg(r)}{dr} - g(r)\frac{df(r)}{dr}. \tag{11}$$

From (1), (2), and (11) it is easy to show that

$$[1 + 2\gamma\omega(r)]W[f,g] = [1 + 2\gamma\omega(r=0)]W[f,g]_{r=0}. \tag{12}$$

If we let $r \to \infty$ in (12), we obtain a relation

$$W[f,g]_{r\to\infty} = [1 + 2\gamma\omega(0)]W[f,g]_{r=0},$$

from which it follows that

$$W[f^+,f^-]_{r\to\infty} = [1 + 2\gamma\omega(0)]W[f^+,f^-]_{r=0} = -2ik. \tag{13}$$

The physical interpretation of Eq. (13) is that the flux at $r = \infty$, described by the Jost solution $f^+(r)$, is $[1 + 2\gamma\omega(0)]$ times its flux at $r = 0$. In the case of a repulsive, momentum-dependent interaction, $\gamma\omega(0) > 0$, the flux increases as the scattered particles are separated apart, and reaches its maximum value as $r \to \infty$. Before we proceed to the derivation of the Jost function, let us find out relations between the various solutions. The regular solution $\phi(r)$, being the solution of the same second-order differential equation whose two independent solutions are $f^\pm(r)$, can in general be written as a linear combination of $f^\pm(r)$. Its coefficients can easily be found by using (12) and (13). The expression for $\phi(r)$ thus obtained is

$$\phi(r) = [1 + 2\gamma\omega(0)][f^+(r)f^-(k) - f^-(r)f^+(k)]/2ik. \quad (14)$$

Comparing the asymptotic form of the integral equation (5) of $\phi(r)$ and that of (14), one obtains

$$[1 + 2\gamma\omega(0)]f^\pm(k) = 1 + \gamma\int_0^\infty e^{\pm ikr} V(r)\phi(r)dr. \quad (15)$$

Similarly, from the comparison of the asymptotic forms of (14) and (3) it follows that

$$\psi^+(r) = k\phi(r)[1 + 2\gamma\omega(0)]^{-1}/f^+(k), \quad (16)$$

and that the scattering matrix $S(k) = f^-(k)/f^+(k)$. The interacting Green's functions are the solutions of the following differential equation:

$$\left(\frac{d^2}{dr^2} + k^2 - \gamma V(r)\right) \mathcal{G}(r, r') = \delta(r - r'). \quad (17)$$

The Green's function for outgoing boundary condition $\mathcal{G}^+(r, r')$, satisfies the following conditions:

(a) $\mathcal{G}^+(r, r') = 0$ at $r = 0$; and $\to e^{ikr}g(r')$ as $r \to \infty$.

(b) It is continuous at $r = r'$ and satisfies

$$[1 + 2\gamma\omega(r')]\frac{d\mathcal{G}(r, r')}{dr}\Big|_{r' - \epsilon}^{r' + \epsilon} = 1 \text{ as } \epsilon \to 0.$$

Using the known solutions of Eq. (2) along with their Wronskians (12) and (13), it is straightforward to show that the required solution of (17) is

$$\mathcal{G}^+(r, r') = -[1 + 2\gamma\omega(0)]^{-1}\phi(r_<)f^+(r_>)/f^+(k). \quad (18)$$

The Green's function $\mathcal{G}(r, r')$ corresponding to the noninteracting $G(r, r')$ given by (7) also satisfies condition (b). However, instead of (a), it is $\mathcal{G}(r, r') = 0$ for $r' \geq r$. Knowing $\mathcal{G}^+(r, r')$ in (18) and that $\phi(r)$ and $f^\pm(r)$ are the solutions of (2), it is easy to see that

$$\mathcal{G}(r, r') = \begin{cases} [1 + 2\gamma\omega(0)]^{-1}[\phi(r)f^+(r') - \phi(r')f^+(r)] & \text{for } r \geq r' \\ 0 \quad \text{for } r < r'. \end{cases} \quad (19)$$

Via (14), it can also be written in a suitable form

$$\mathcal{G}(r, r') = \begin{cases} [f^+(r)f^-(r') - f^+(r')f^-(r)]/2ik & \text{for } r \geq r' \\ 0 \quad \text{for } r < r'. \end{cases} \quad (20)$$

Another formal solution of (17) for $\mathcal{G}(r, r')$ is

$$\begin{aligned}\mathcal{G}(r, r') &= G(r, r') + \gamma\int G(r, r'')V(r'')\mathcal{G}(r'', r')dr'' \\ &= G(r, r') + \gamma\int G(r, r'')V(r'')G(r'', r')dr'' \\ &+ \gamma^2\int G(r, r'')V(r'')G(r'', r''')V(r''')G(r''', r')dr''dr''' \\ &+ \cdots. \end{aligned} \quad (21)$$

It is to be observed that

$$\begin{aligned}&\int G(r, r'')V(r'')G(r'', r') \\ &= \int G(r'', r)V(r'')G(r', r'')dr'' \\ &+ \left\{2\omega(r'')W[G(r'', r), G(r', r'')]\right\}_{r''=0}^{r''=\infty}. \end{aligned} \quad (22)$$

The second term on the right-hand side of (22) vanishes because $\omega(r'') = 0$ as $r'' \to \infty$ and $G(r'' = 0, r) = 0$. Substituting (22) in (21), one obtains

$$\mathcal{G}(r'r) = \tilde{G}(r, r') + \gamma\int \tilde{G}(r, r'')V(r'')\mathcal{G}(r', r'')dr'' \quad (23)$$

In deriving (23), r and r' in (21) are interchanged and Eq. (9) is used. Thus, from (23), the interacting Green's function $\tilde{\mathcal{G}}(r, r')$, corresponding to $\tilde{G}(r, r')$ is related to $\mathcal{G}(r, r')$ in the following way:

$$\tilde{\mathcal{G}}(r, r') = \mathcal{G}(r', r). \quad (24)$$

Differentiating both the sides of (15) with respect to γ, one gets

$$\begin{aligned}&[1 + 2\gamma\omega(0)]\frac{df^+(k)}{d\gamma} + 2\omega(0)f^+(k) \\ &= \int_0^\infty e^{ikr}V(r)\phi(r)dr + \gamma\int_0^\infty e^{ikr}V(r)\frac{d\phi(r)}{d\gamma}dr. \end{aligned} \quad (25)$$

From (5),

$$\begin{aligned}\phi(r) &= \langle r|(1 - \gamma GV)^{-1}\phi_0\rangle, \\ \frac{d\phi(r)}{d\gamma} &= \langle r|(1 - \gamma GV)^{-2}GV\phi_0\rangle \\ &= \langle r|(1 - \gamma GV)^{-1}GV\phi\rangle \\ &= \langle r|\mathcal{G}V\phi\rangle. \end{aligned} \quad (26)$$

In the last step of the derivation of (26), Eq. (21) is used. Let us simplify the second term of (25):

$$\begin{aligned}&\int_0^\infty e^{ikr}V(r)\frac{d\phi(r)}{d\gamma}dr \\ &= \int_0^\infty e^{ikr}V(r)\mathcal{G}(r, r')V(r')\phi(r')drdr' \\ &= \int_0^\infty \phi(r')V(r')\mathcal{G}(r, r')V(r)e^{ikr}drdr' \\ &+ 2\omega(0)[1 + 2\gamma\omega(0)]^{-1}\int_0^\infty \phi(r)V(r)e^{ikr}dr. \end{aligned} \quad (27)$$

In deriving (27), the following relations are used which can easily be checked using Eqs. (12)–(14), (19), and (20):

$$\int_0^\infty e^{ikr}V(r)\mathcal{G}(r, r')dr = \int_0^\infty \mathcal{G}(r, r')V(r)e^{ikr}dr, \quad (28)$$

and

$$\begin{aligned}\int_0^\infty \mathcal{G}(r, r')V(r')\phi(r')dr' &= \int_0^\infty \phi(r')V(r')\mathcal{G}(r, r')dr' \\ &+ 2\omega(0)[1 + 2\gamma\omega(0)]^{-1}\phi(r). \end{aligned} \quad (29)$$

Similarly, it can be seen that

$$\int_0^\infty e^{ikr} V(r)\phi(r)dr = \int_0^\infty \phi(r)\, V(r) e^{ikr} dr + 2\omega(0). \quad (30)$$

Substitution of (27) and (30) in Eq. (25) gives

$$[1 + 2\gamma\omega(0)]\frac{df^+(k)}{d\gamma} + 2\omega(0) f^+(k)$$
$$= \int_0^\infty \phi(r)\, V(r)\left(e^{ikr} + \gamma\int_0^\infty \mathcal{G}(r'r)\, V(r') e^{ikr'} dr'\right)dr$$
$$+ 2\omega(0) + 2\gamma\omega(0)[1 + 2\gamma\omega(0)]^{-1}$$
$$\times \int_0^\infty \phi(r)\, V(r) e^{ikr} dr. \quad (31)$$

From (8), (23), and (24) it follows that

$$f^+(r) = e^{ikr} + \gamma\int_0^\infty \mathcal{G}(r, r')\, V(r') e^{ikr'} dr'. \quad (32)$$

One could also rewrite:

$$\int_0^\infty \phi(r) V(r) e^{ikr} dr = \int_0^\infty e^{ikr} V(r)\phi(r) dr - 2\omega(0). \quad (33)$$

Via (15), (32), and (33), Eq. (31) could be put into its simpler form

$$[1 + 2\gamma\omega(0)]\frac{df^+(k)}{d\gamma} = \int_0^\infty \phi(r)\, V(r) f^+(r) dr. \quad (34)$$

In order to solve (34) for $f^+(k)$, let us convert it from r-space to the momentum space representation. A complete set of momentum states $\psi_p(r)$ to be used are

$$\psi_p(r) = \sqrt{2/\pi}\, \sin pr. \quad (35)$$

In terms of these basis states, the following relation between $\langle p|\mathcal{G}^+|p'\rangle$ and $\langle p|\mathcal{G}|p'\rangle$ could be derived from (18) and (19):

$$\langle p|\phi\rangle\langle p'|f^+\rangle = f^+(k)[1 + 2\gamma\omega(0)](\langle p|\mathcal{G}|p'\rangle - \langle p|\mathcal{G}^+|p'\rangle). \quad (36)$$

Inserting the set of states ψ_p in (34) and using (36), one obtains

$$\frac{1}{f^+(k)}\frac{df^+(k)}{d\gamma} = \int_0^\infty\int_0^\infty dp dp' (\langle p|\mathcal{G}|p'\rangle - \langle p|\mathcal{G}^+|p'\rangle)\langle p'|V|p\rangle. \quad (37)$$

In deriving (37), use is made of the fact that

$$\langle p|V|p'\rangle = \langle p'|V|p\rangle. \quad (38)$$

Henceforth, the traces (denoted by Tr) are to be taken in the momentum space representation of the operators. Integration of (37) yields

$$\log f^+(k) = \int_0^\gamma [\mathrm{Tr}\mathcal{G}V - \mathrm{Tr}\mathcal{G}^+V] d\gamma'$$
$$= \mathrm{Tr}\ \log(1 - \gamma G_+ V) - \mathrm{Tr}\ \log(1 - \gamma G V).$$

Thus, the final expression for the Jost function is

$$f^+(k) = \exp[\mathrm{Tr}\ \log(1 - \gamma G_+ V)]/\exp[\mathrm{Tr}\ \log(1 - \gamma G V)]$$
$$= \det(1 - \gamma G_+ V)/\det(1 - \gamma G V). \quad (39)$$

The numerator in (39) is the Fredholm determinant of the kernel of the integral equation for the outgoing scattering solution (3) in the momentum space representation, while the denominator is the determinant of the kernal of that of the regular solution (5) in the momentum space representation. The integral equations of ψ^+ and ϕ in the momentum space representation have the same form for a nonlocal interaction or for a momentum-dependent interaction. From this fact and from the result in Ref. 3, $f^+(k)$ in (39) has the expected form. In the evaluation of the determinants in momentum space, the following momentum representation of G_+ and G would be useful:

$$\langle p|G_+|p'\rangle = \delta(p - p')/(k^2 - p^2 + i\epsilon) \quad \text{and}$$
$$\langle p|G|p'\rangle = \delta(p - p')/(k^2 - p^2).$$

The $p = k$ point is to be avoided in the evaluation of the determinant by choosing a proper mesh for p. As observed in the nonlocal potentials, the denominator in (39) introduces "redundant zeros" of $f^+(k)$ for positive imaginary values of k which do not correspond to bound states. Via the approach given in this paper, it can easily be shown that the Jost function $f^+(k)$ given by (39) is the most general one, valid for the superposition of a local, nonlocal, and momentum-dependent potential. It is also straightforward to generalize it to the higher partial waves and for the non-central forces.

ACKNOWLEDGMENTS

It is the pleasure to acknowledge the useful discussions with Professor R. K. Bhaduri. The author would like to thank the National Research Council of Canada for the award of a Research Associateship and to Professor M. W. Johns for the hospitality of the Physics Department at McMaster.

---

* On leave from Tata Institute of Fundamental Research, Colaba, Bombay 5.

[1] R. Jost and A. Pais, Phys. Rev. **82**, 840 (1951).
[2] R. G. Newton, J. Math. Phys. **2**, 188 (1961).
[3] R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966).
[4] C. S. Warke and R. K. Bhaduri, Nucl. Phys. **A162**, 289 (1970).
[5] Y. Singh and C. S. Warke, Can. J. Phys. **49**, 1029 (1971).

Green's Dyadics for Elastic and Electromagnetic Waves in a Random Medium

Sun Ho Chung
Korea Institute of Science and Technology, Seoul, Korea
and
Yung Ming Chen
Department of Applied Mathematics, State University of New York, Stony Brook, New York 11790
(Received 9 June 1969)

A method for obtaining Green's dyadics of renormalized stochastic vector wave operators is presented schematically and is applied to propagation of elastic and electromagnetic waves in statistically homogeneous and isotropic media. Formula for effective propagation constants and renormalized Green's dyadics are derived and then computed explicitly for various special cases of elastic and electromagnetic waves with a special type of correlation function. Furthermore, parameters which affect the validity of such a renormalization approximation are also derived by estimating the second nonvanishing term of the renormalized series for average waves.

1. INTRODUCTION

Let L be a stochastic linear operator and u denote a solution of the equation

$$Lu = f, \tag{1.1}$$

where f represents a nonrandom source function. The random part of the operator L, represented by L', is a statistically homogeneous random operator and is smaller than the nonrandom part, $L_0 = L - L'$. Renormalization methods[1-5] have been used to obtain an equation for $\langle u\rangle$, the average solution of (1.1), and for the effective propagation constant of a propagating wave.

The Nth-order renormalized operator L_N can be written [3, 4, 6] as

$$L_N = L_0 + \sum_{n=0}^{N} \langle L'T^n\rangle \tag{1.2}$$

where T is the so called "smoothing operator"[3] defined by

$$Tv = -L_0^{-1}\,(L'v - \langle L'v\rangle) \tag{1.3}$$

for a suitable random function v. Then the expansion for $\langle u\rangle$ in terms of[6] L_N and T is

$$\langle u\rangle = \sum_{i=0}^{\infty}\left(-L_N^{-1}\left\langle L'T^N \sum_{j=0}^{\infty}(T^N)^j \sum_{l=1}^{N} T^l\right\rangle\right)^i L_N^{-1} f. \tag{1.4}$$

Let the random part of the solution be denoted by u'. Then

$$u' \equiv u - \langle u\rangle = \sum_{i=0}^{\infty}(T^N)^i\left(\sum_{j=1}^{N} T^j\langle u\rangle\right). \tag{1.5}$$

To carry out even a few terms of (1.4) and (1.5) involves very complicated computations. If the random field is not Gaussian, higher-order moments of the random operator L' must also be known. Similar difficulties are also encountered when one attempts to use L_N given by (1.2) for $N > 1$. Therefore, from a practical point of view, approximations of the form

$$\langle u\rangle \cong L_1^{-1} f \tag{1.6}$$

and

$$u' \cong T\langle u\rangle = TL_1^{-1} f \tag{1.7}$$

are most useful. In fact, all of the work reported so far is confined to first-order renormalization approximations of the form (1.6) and (1.7).

For example, Tatarskii and Gertenshtein[4] have computed the Green's function of L_1 for scalar waves and later Tatarskii[5] has also computed asymptotic expressions for the Green's dyadics of L_1 for electromagnetic waves. Karal and Keller,[1] with a different approach, have derived dispersion relations for scalar, electromagnetic, and elastic waves. There they have assumed a plane wave type solution of the equation

$$L_1\langle u\rangle = 0 \tag{1.8}$$

to derive dispersion relations. From these dispersion relations, they have obtained formulas for propagation constants and carried out computations explicitly for some of the special cases.

In this paper a method for obtaining Green's dyadics of the vector renormalized operator is presented schematically, and it is applied to elastic and electromagnetic waves. Explicit results for effective propagation constants and renormalized Green's dyadics are obtained for several special cases of elastic and electromagnetic waves under the assumption of a homogeneous, isotropic random medium characterized by exponential correlation functions. From homogeneity, it follows that L_N is translationally invariant. At the end of each example, far field expressions for the average waves are presented. Our result on the effective propagation constant agrees with that of Karal and Keller[1] in the electromagnetic case. However, for elastic waves, their computation contains some errors which we shall point out later.

The limit of applicability of the renormalization approximation is investigated by estimating the second nonvanishing term on the right side of (1.4). This limit has been discussed briefly by Frisch,[3] Tatarskii and Gertenshtein,[4] and Tatarskii.[5] In particular, for the case of scalar waves with a Gaussian refrative index, Frisch has employed a simple nondimensionalizing technique to discuss the convergence of the expansions involved in the renormalization and to derive parameters which affect the convergence.

2. GREEN'S DYADICS OF N-FOLD RENORMALIZED TIME HARMONIC WAVE OPERATORS

The Green's dyadic of the N-fold renormalized time harmonic wave operator, denoted by $G_N(\bar{x}, \bar{x}')$, satisfies the equation

$$L_N G_N(\bar{x}, \bar{x}') = \bar{\bar{I}}\,\delta(\bar{x} - \bar{x}') \tag{2.1}$$

where $\bar{x}$ and $\bar{x}'$ are three-dimensional space variables, $\bar{\bar{I}}$ is an identity dyadic and $\delta(\bar{x} - \bar{x}')$ is the delta function. To obtain $G_N(\bar{x}, \bar{x}')$ explicitly, we take the threedimensional Fourier transformation

of (2.1) and get

$$\int [L_N G_N(\bar{x}, \bar{x}')] \cdot (\bar{\bar{I}} e^{-i\bar{s}\cdot\bar{x}}) d\bar{x} = \bar{\bar{I}} e^{-i\bar{s}\cdot\bar{x}'}. \quad (2.2)$$

By means of Green's vector identities and integration by parts we derive L_N^+, which is defined by the equation

$$\int [L_N^+ \cdot (\bar{\bar{I}} e^{-i\bar{s}\cdot\bar{x}})] \cdot G_N(\bar{x}, \bar{x}') d\bar{x} = \int [L_N G_N(\bar{x}, \bar{x}')] \cdot (\bar{\bar{I}} e^{-i\bar{s}\cdot\bar{x}}) d\bar{x}. \quad (2.3)$$

Then, in view of the translation invariance of L_N^+, (2.2) can be written as

$$\mathcal{L}_N^{(\bar{s})}\left(\frac{1}{\sqrt{8\pi^3}} \int G_N(\bar{x}, \bar{x}') e^{-i\bar{s}\cdot\bar{x}} d\bar{x}\right) = \frac{1}{\sqrt{8\pi^3}} \bar{\bar{I}} e^{-i\bar{s}\cdot\bar{x}'}, \quad (2.4)$$

where

$$\mathcal{L}_N(\bar{s}) = [L_N^{+\bullet}(\bar{\bar{I}} e^{-i\bar{s}\cdot\bar{x}})] e^{i\bar{s}\cdot\bar{x}} \quad (2.5)$$

and the integral inside the parenthesis is nothing but the Fourier transform of $G_N(\bar{x}, \bar{x}')$.

In the case that $\mathcal{L}_N(\bar{s})$ can be expressed in a symmetric form

$$\mathcal{L}_N(\bar{s}) = \mathcal{L}_{N1}(\bar{s})\bar{\bar{I}} + \mathcal{L}_{N2}(\bar{s})\hat{s}\hat{s}, \quad (2.6)$$

where $s = |\bar{s}|$ and $\hat{s} = \bar{s}/|\bar{s}|$, the inverse of $\mathcal{L}_N(\bar{s})$ becomes

$$\mathcal{L}_N^{-1}(\bar{s}) = \mathcal{L}_{N1}^{-1}(s)\bar{\bar{I}} - \mathcal{L}_{N1}^{-1}(\mathcal{L}_{N1}(s) + \mathcal{L}_{N2}(s))^{-1}\mathcal{L}_{N2}(s)\hat{s}\hat{s}. \quad (2.7)$$

Applying (2.7) to both sides of (2.4) and taking the inverse Fourier transformation yields the useful formula for the Green's dyadic $G_N(\bar{x}, \bar{x}')$,

$$G_N(\bar{x}, \bar{x}') = \frac{\bar{\bar{I}}}{2\pi^2 r} \times \int_0^\infty \left[\frac{s \sin(sr)}{\mathcal{L}_{N1}(s)} - \frac{\partial}{\partial r}\left(\frac{\mathcal{L}_{N2}(s)\sin(sr)}{rs\mathcal{L}_{N1}(s)[\mathcal{L}_{N1}(s) + \mathcal{L}_{N2}(s)]}\right)\right] ds - \frac{r}{2\pi^2}\hat{r}\hat{r} \times \int_0^\infty \frac{\partial}{\partial r}\left[\frac{1}{r}\frac{\partial}{\partial r}\left(\frac{\mathcal{L}_{N2}(s)\sin(sr)}{rs\mathcal{L}_{N1}(s)[\mathcal{L}_{N1}(s) + \mathcal{L}_{N2}(s)]}\right)\right] ds \quad (2.8)$$

where $r = |\bar{x} - \bar{x}'|$ and $\hat{r} = (\bar{x} - \bar{x}')/|\bar{x} - \bar{x}'|$.

3. GREEN'S DYADICS FOR ELASTIC AND ELECTROMAGNETIC WAVES

In the following subsections, elastic and electromagnetic waves in statistically homogeneous random media are considered and the explicit expressions for their Green's dyadics are computed. Since the problems treated here are exactly the same as those treated by Karal and Keller,[1] who studied plane wave propagation, many of our discussions and expressions are the same as theirs. Therefore, we will try not to repeat them here again but simply refer to theirs.

A. Elastic Waves

Following Karal and Keller's[1] formulation of the problem of solving for the displacement vector $\bar{u}$ in a randomly inhomogeneous and isotropic medium generated by a monochromatic source $\bar{f}$, we have

$$L_0\bar{u} = (\lambda_0 + \mu_0)\nabla(\nabla\cdot\bar{u}) + \mu_0\nabla^2\bar{u} + \omega^2\rho_0\bar{u}, \quad (3.1)$$

$$L'\bar{u} = (\lambda_1 + \mu_1)\nabla(\nabla\cdot\bar{u}) + \mu_1\nabla^2\bar{u} + \nabla\lambda(\nabla\cdot\bar{u}) + \nabla\mu_1 \times (\nabla\times\bar{u}) + 2(\nabla\mu_1\cdot\nabla)\bar{u} + \omega^2\rho_1\bar{u}. \quad (3.2)$$

The Green's dyadic for the nonrandom operator L_0 denoted by[1,7] $G_0(\bar{x} - \bar{x}')$ can be decomposed into the following forms:

$$G_0(x - x') = G_{01}\bar{\bar{I}} + G_{02}\hat{r}\hat{r} \quad (3.3)$$

$$= G_\delta^c(\bar{x} - \bar{x}') + G_\delta^s(\bar{x} - \bar{x}') \quad (3.4)$$

$$= (G_{\delta 1} + G_{\delta 1})\bar{\bar{I}} + (G_{\delta 2} + G_{\delta 2})\hat{r}\hat{r}. \quad (3.5)$$

Here G_{01} and G_{02}, coefficients of the dyadics $\bar{\bar{I}}$ and $\hat{r}\hat{r}$ respectively, are given by Eqs. (37) of Ref. 1 except that the delta function in G_{01} should be deleted. Superscripts "c" and "s" in (3.4) and (3.5) denote dilatational and rotational components of the dyadic, respectively.

From (1.2), L_1 acting on a nonrandom function $\bar{w}(\bar{x})$ can be written as

$$L_1\bar{w}(\bar{x}) = L_0\bar{w}(\bar{x}) - \int \langle L'(\bar{x})G_0(\bar{x} - \bar{x}')L'\rangle\bar{w}(\bar{x}')d\bar{x}'. \quad (3.6)$$

From (3.2) and (3.3), we obtain

$$L'G_0 = \left[(\lambda_1 + \mu_1)\left(\frac{G_{01}'}{r} + \frac{G_{02}'}{r} + \frac{2G_{02}}{r^2}\right) + \mu_1\left(\nabla^2 G_{01} + \frac{2G_{02}}{r^2}\right) + (\nabla\mu_1\cdot\hat{r})\left(G_{01}' + \frac{G_{02}}{r}\right) + \omega^2\rho_1 G_{01}\right]\bar{\bar{I}} + \left[2\left(G_{02}' - \frac{2G_{02}}{r}\right)(\nabla\mu_1\cdot\hat{r}) + \mu_1\left(\nabla^2 G_{02} - \frac{6G_{02}}{r}\right) + (\lambda_1 + \mu_1)\left(G_{01}'' - \frac{G_{01}'}{r} + G_{02}'' + \frac{G_{02}'}{r} - \frac{4G_{02}}{r^2}\right) + \omega^2\rho_1 G_{02}\right]\hat{r}\hat{r} + \left(G_{01}' + G_{02}' + \frac{2G_{02}}{r}\right)\nabla\lambda_1\hat{r} + \frac{2G_{02}}{r}\hat{r}\nabla\mu_1 + \left(G_{01}' + \frac{G_{02}}{r}\right) \times \nabla\mu_1\hat{r}, \quad (3.7)$$

where primes denote partial derivatives of the function with respect to r. Karal and Keller's[1] Eq. (42) differs from (3.7) and appears to be incorrect.

Assuming that the random process involved is homogeneous and isotropic, we have

$$\langle L'(\bar{x})G_0(\bar{x} - \bar{x}')L'(\bar{x}')\rangle = \bar{\bar{B}}_1\cdot\nabla'(\nabla'\cdot\,) + \bar{\bar{B}}_2\cdot\nabla'^2 + \bar{\bar{B}}_3\cdot\hat{r}(\nabla'\,\bullet\,) + \bar{\bar{B}}_4\cdot(\hat{r}\times\nabla'\times + 2\hat{r}\nabla') + \bar{\bar{B}}_5\cdot\bar{\bar{I}}, \quad (3.8)$$

where the prime denotes that the ∇ operator acts on $\bar{x}'$

$$\bar{\bar{B}}_1 = -k_c^2(\mathcal{R}_{\lambda\lambda} + 2\mathcal{R}_{\lambda\mu} + \mathcal{R}_{\mu\mu})G_0^c - (\mathcal{R}_{\lambda\mu} + \mathcal{R}_{\mu\mu}) \times (k_c^2 G_0^c + k_s^2 G_0^s) + (\mathcal{R}'_{\lambda\mu} + \mathcal{R}'_{\mu\mu})\,\mathbb{E} + \omega^2(\mathcal{R}_{\lambda\rho} + \mathcal{R}_{\mu\rho})G_0 + k_c^2(\mathcal{R}'_{\lambda\lambda} + \mathcal{R}'_{\lambda\mu})rG_{01}^c\hat{r}\hat{r}, \tag{3.9}$$

$$\bar{\bar{B}}_2 = -k_c^2(\mathcal{R}_{\lambda\mu} + \mathcal{R}_{\mu\mu})G_0^c + \mathcal{R}'_{\mu\mu}\,\mathbb{E} - (k_c^2 G_0^c + k_s^2 G_0^s)\mathcal{R}_{\mu\mu} + \omega^2\mathcal{R}_{\mu\rho}G_0 - k_c^2\mathcal{R}'_{\lambda\mu}rG_{01}^c\hat{r}\hat{r}, \tag{3.10}$$

$$\bar{\bar{B}}_3 = k_c^2(\mathcal{R}'_{\lambda\lambda} + \mathcal{R}'_{\lambda\mu})G_0^c + \mathcal{R}'_{\lambda\mu}(k_0^2 G_0^c + k_s^2 G_0^s) - \mathcal{R}''_{\lambda\mu}\,\mathbb{E} - \omega^2\mathcal{R}'_{\lambda\rho}G_0 + k_c^2\mathcal{R}''_{\lambda\lambda}rG_{01}^c\hat{r}\hat{r}, \tag{3.11}$$

$$\bar{\bar{B}}_4 = k_c^2(\mathcal{R}'_{\lambda\mu} + \mathcal{R}'_{\mu\mu})G_0^c + \mathcal{R}'_{\mu\mu}(k_c^2 G_0^c + k_s^2 G_0^s) - \mathcal{R}''_{\mu\mu}\,\mathbb{E} - \omega^2\mathcal{R}'_{\mu\rho}G_0 + k_c^2\mathcal{R}''_{\lambda\mu}rG_{01}^c\hat{r}\hat{r}, \tag{3.12}$$

$$\bar{\bar{B}}_5 = \omega^2[-k_c^2(\mathcal{R}_{\lambda\rho} + \mathcal{R}_{\mu\rho})G_0^c + \mathcal{R}'_{\mu\rho}\mathbb{E} - \mathcal{R}_{\mu\rho}(k_c^2 G_0^c + k_s^2 G_0^s) + \omega^2\mathcal{R}_{\rho\rho}G_0 - k_c^2\mathcal{R}'_{\lambda\rho}rG_{01}^c\hat{r}\hat{r}], \tag{3.13}$$

$$\mathbb{E} = (4\pi^2\omega^2\rho_0 r^4)^{-1}\{[(6 - 6ik_c r - 2k_c^2r^2)e^{ik_c r} - (6 - 6ik_s r - 3k_s^2r^2 + ik_s^3r^3)e^{ik_s r}]\bar{\bar{I}} - [(18 - 18ik_c r - 8k_c^2r^2 + 2ik_c^3r^3)e^{ik_c r} - (18 - 18ik_s r - 7k_s^2r^2 + ik_s^3r^3)e^{ik_s r}]\hat{r}\hat{r}\}. \tag{3.14}$$

Now upon substituting (3.6) into (2.3) with the help of (3.8), Green's vector identities and integration by parts, we get

$$L_1^{\dagger}\bar{w}(\bar{x}) = L_0\bar{w}(\bar{x}) - \int\{\bar{\bar{B}}_1\cdot\nabla'[\nabla'\cdot\bar{w}(\bar{x}')] + \bar{\bar{B}}_2\cdot\nabla'^2\bar{w}(\bar{x}') - \bar{\bar{B}}_3\cdot\hat{r}\times\nabla'\bar{w}(\bar{x}') + \bar{\bar{B}}_4\times\hat{r}\cdot[\nabla'\times\bar{w}(\bar{x}')] - 2\bar{\bar{B}}_4\cdot\hat{r}\cdot\nabla'\cdot\bar{w}(\bar{x}') + \bar{\bar{B}}_5\cdot\bar{w}(\bar{x}')\}d\bar{x}'. \tag{3.15}$$

We insert $\bar{\bar{I}}e^{-i\bar{s}\cdot\bar{x}}$ for $\bar{w}(\bar{x})$ into (3.15) and then multiply (3.15) by $e^{i\bar{s}\cdot\bar{x}}$ to obtain

$$\mathcal{L}_1(\bar{s}) = (\omega^2\rho_0 - \mu_0 s^2)\bar{\bar{I}} - (\lambda_0 + \mu_0)s^2\hat{s}\hat{s} + \int s^2(\bar{\bar{B}}_1\cdot\hat{s}\hat{s} + \bar{\bar{B}}_2) - is\bar{\bar{B}}_3\cdot\hat{r}\hat{s} - is\bar{\bar{B}}_4\cdot[\hat{r}\hat{s} + (\hat{r}\hat{s})\bar{\bar{I}}] - \bar{\bar{B}}_5\}e^{-i\bar{s}\cdot\bar{r}}d\bar{r}. \tag{3.16}$$

Since the $\bar{\bar{B}}_i$, $i = 1, 2, \ldots, 4$, are dyadic functions involving symmetric dyadics I and $\hat{r}\hat{r}$ only, we can apply mean value theorems[8] to the integral in (3.16). After some simplification, $\mathcal{L}_1(\bar{s})$ is given by (2.6) with

$$\mathcal{L}_{11}(s) = \omega^2\rho_0 - \mu_0 s^2 - D_1(s), \tag{3.17}$$

$$\mathcal{L}_{12}(s) = -(\lambda_0 + \mu_0)s^2 - D_2(s), \tag{3.18}$$

where

$$D_1(s) = \int_0^{\infty} -s^2\left(B_2^{\mathrm{I}}h - B_2^{rr}\frac{1}{sr^2}\frac{\partial h}{\partial s}\right) - s\left[B_4^{\mathrm{I}}\frac{1}{r}\times\frac{\partial h}{\partial s} - B_4^{rr}\frac{1}{r^3}\left(\frac{\partial^2 h}{\partial s^2} - \frac{1}{s}\frac{\partial h}{\partial s}\right)\right] + \left(B_5^{\mathrm{I}}h - B_5^{rr}\frac{1}{sr^2}\times\frac{\partial h}{\partial s}\right)\Bigg\} dr, \tag{3.19}$$

$$D_2(s) = \int_0^{\infty}\left[s^2\left(B_1^{rr}\frac{1}{r^2}\frac{\partial^2 h}{\partial s^2} - B_1^{\mathrm{I}}h\right) + B_2^{rr}\frac{s^2}{r^2}\left(\frac{\partial^2 h}{\partial s^2} - \frac{1}{s}\frac{\partial h}{\partial s}\right) - (B_3^{\mathrm{I}} + B_4^{\mathrm{I}} + B_3^{rr} + B_4^{rr})\frac{s}{r}\frac{\partial h}{\partial s} + B_4^{rr}\frac{s}{r^4}\left(\frac{\partial^3 h}{\partial s^3} - \frac{1}{s}\frac{\partial^2 h}{\partial s^2} + \frac{1}{s^2}\frac{\partial h}{\partial s}\right) - B_5^{rr}\frac{1}{r^2}\left(\frac{\partial^2 h}{\partial s^2} - \frac{1}{s}\frac{\partial h}{\partial s}\right)\right] dr. \tag{3.20}$$

Here superscripts "I" and "rr" denote the components of the symmetric dyadics $\bar{\bar{I}}$ and $\hat{r}\hat{r}$, respectively, and

$$h = (4\pi r/s)\sin(sr). \tag{3.21}$$

Equations (3.17)–(3.20) are very similar to those obtained by Karal and Keller[1] through a different approach. In fact, if they had not made mistakes in computing $L'G_0$, they would have been identical.

When correlation functions of fluctuating parts of the medium constants are given, $D_1(s)$ and $D_2(s)$ can be calculated. Once roots of equations

$$\mathcal{L}_{11}(s) = 0 \tag{3.22}$$

and

$$\mathcal{L}_{11}(s) + \mathcal{L}_{12}(s) = 0 \tag{3.23}$$

are determined, the inverse transform (2.8) can be computed by employing residue technique. To proceed further, let us assume that $\pm K_{sl}$ and $\pm K_{cj}$, for $l = 1, 2, \ldots$ and $j = 1, 2, \ldots$, are the roots of (3.22) and (3.23) respectively. Then the Green's dyadic may be written as

$$G_1(\bar{x} - \bar{x}') = \frac{1}{2\pi\omega^2\rho_0 r^3}\left[\sum_l\left(\frac{k_s^2 e^{iK_{sl}r}}{W_{sl}}[(1 - iK_{sl}r - K_{sl}^2r^2)\bar{\bar{I}} - (3 - 3iK_{sl}r - K_{sl}^2r^2)\hat{r}\hat{r}]\right) + \sum_j\left(\frac{k_c^2 e^{iK_{cj}r}}{W_{cj}}[-(1 - iK_{cj}r)\bar{\bar{I}} + (3 - 3iK_{cj}r - K_{cj}^2r^2)\hat{r}\hat{r}]\right)\right], \tag{3.24}$$

where

$$W_{sl} = 2K_{sl}^2 + \frac{k_s^2 k_{sl}^2}{\omega^2\rho_0}\frac{\partial}{\partial K_{sl}}D_1(K_{sl}), \tag{3.25}$$

$$W_{cj} = 2K_{cj}^2 + \frac{k_c^2 K_{cj}^2}{\omega^2\rho_0}\frac{\partial}{\partial K_{cj}}[D_1(K_{cj}) + D_2(K_{cj})]. \tag{3.26}$$

From (3.24) we observe that the Green's dyadic for L_1 is in a very similar form to that of G_0 except that the coefficients of the dilatational and rotational parts are k_c^2/W_{cj} and k_s^2/W_{sl}, respectively.

When all the correlation functions except $\mathcal{R}_{\rho\rho}$ are zero, (3.19) and (3.20) can be simplified to

$$D_1(s) = \omega^4\int_0^{\infty}\left(G_{01}h - G_{02}\frac{1}{sr^2}\frac{\partial h}{\partial s}\right)\mathcal{R}_{\rho\rho}(r)dr, \tag{3.27}$$

$$D_2(s) = \omega^4 \int_0^\infty G_{02} \frac{1}{r^2}\left(\frac{1}{s}\frac{\partial h}{\partial s} - \frac{\partial^2 h}{\partial s^2}\right) \mathcal{R}_{\rho\rho}(r)dr. \tag{3.28}$$

Equations (3.27) and (3.28) have been previously derived by Karal and Keller.[1] To obtain explicit results, we choose for $\mathcal{R}_{\rho\rho}$ form

$$\mathcal{R}_{\rho\rho}(r) = \langle \rho_1^2 \rangle e^{-r/\alpha}, \tag{3.29}$$

where $\langle \rho_1^2 \rangle$ and α are respectively mean square value and the correlation length of the random function $\rho_1(\bar{x})$. Since we are interested primarily in the case of when $k_{c,s}\alpha \ll 1$, we insert (3.29) into (3.27) and (3.28) and carry out the integration to obtain

$$D_1(s) = -\frac{\omega^2\langle \rho_1^2 \rangle}{\rho_0}\left(\frac{k_s^2\alpha^2}{s^2\alpha^2 + (1 - ik_c\alpha)^2} - X(k_c) + X(k_s)\right), \tag{3.30}$$

$$D_1(s) + D_2(s) = -\frac{\omega^2\langle \rho_1^2 \rangle}{\rho_0}\left(\frac{k_c^2\alpha^2}{s^2\alpha^2 + (1 - ik_s\alpha)^2} + 2[X(k_c) + X(k_s)]\right), \tag{3.31}$$

where

$$X(k) = \frac{1}{(1 - ik\alpha)^2} \sum_{n=0}^{\infty} (-1)^n \left(\frac{s\alpha}{1 - ik\alpha}\right)^{2n} \times \left(\frac{1 - (2n + 2)\, ik\alpha}{(2n + 3)(2n + 1)}\right). \tag{3.32}$$

To obtain the roots of equations (3.22) and (3.23) in the neighborhood of $\pm k_s$ and $\pm k_c$, respectively, we write

$$K_s^2/k_s^2 = 1 - (1/\omega^2\rho_0) D_1(K_s), \tag{3.33}$$

$$K_c^2/k_c^2 = 1 - (1/\omega^2\rho_0)[D_1(K_c) + D_2(K_c)]. \tag{3.34}$$

Upon substituting (3.30)–(3.32) into (3.33) and (3.34) and then iterating (3.33) and (3.34) with assumptions $k_c\alpha \ll 1$ and $k_s\alpha \ll 1$, we get

$$K_s/k_s = 1 + (\langle \rho_1^2 \rangle/\rho_0^2)[\tfrac{1}{2} k_s^2\alpha^2 + ik_s^3\alpha^3 + O(k_s^4\alpha^4)], \tag{3.35}$$

$$K_c/k_c = 1 + (\langle \rho_1^2 \rangle/\rho_0)[\tfrac{1}{2} k_c^2\alpha^2 + ik_c^3\alpha^3 + O(k_c^4\alpha^4)]. \tag{3.36}$$

Then with the insertion of (3.35) and (3.36), (3.25) and (3.26) become

$$W_c = 2k_c^2\{1 + (\langle \rho_1^2 \rangle/\rho_0^2)[k_c^2\alpha^2 + i2k_c^2\alpha^3 + O(k_c^4\alpha^4)]\}, \tag{3.37}$$

$$W_s = 2k_s^2\{1 + (\langle \rho_1^2 \rangle/\rho_0^2)[k_s^2\alpha^2 + i2k_s^3\alpha^3 + O(k_s^4\alpha^4)]\}. \tag{3.38}$$

Next, when all the correlation functions except

$$\mathcal{R}_{\lambda\lambda}(r) = \langle \lambda_1^2 \rangle e^{-r/\alpha} \tag{3.39}$$

are zero, we have

$$D_1(s) = 0, \tag{3.40}$$

$$D_2(s) = k_c^2 \int_0^\infty \left\{ s^2 \mathcal{R}_{\lambda\lambda}(r)\left[G_{01}^c h - G_{02}^c \frac{1}{r^2}\frac{\partial^2 h}{\partial s^2}\right] + \mathcal{R}'_{\lambda\lambda}(r)\left[G_{02}^c \frac{s}{r}\frac{\partial h}{\partial s} - G_{01}^c\left(\frac{s^2}{r^2}\frac{\partial^2 h}{\partial s^2} + \frac{s}{r}\frac{\partial h}{\partial s}\right) + G_{02}^c \frac{s}{r}\frac{\partial h}{\partial s}\right] + \mathcal{R}''_{\lambda\lambda} G_{01} s \frac{\partial h}{\partial s}\right\} dr. \tag{3.41}$$

Substitution of (3.39) into (3.41) yields

$$D_2(s) = \frac{-\langle \lambda_1^2 \rangle}{\omega^2\rho_0} k_c^2 s^2 \left(\frac{2}{3} + \frac{k_c^2\alpha^2}{s^2\alpha^2 + (1 - ik_c\alpha)^2} + \frac{2ik_c\alpha}{(1 - ik_c\alpha)^2} \sum_{n=0}^{\infty} (-1)^n \frac{[s\alpha/(1 - ik_c\alpha)]^{2n}}{(2n + 3)(2n + 1)}\right). \tag{3.42}$$

Similarly, by iteration we obtain

$$K_c/k_c = 1 + [\langle \lambda_1^2 \rangle/(\lambda_0 + 2\mu_0)^2]\left[(\tfrac{1}{3} - \tfrac{1}{6} k_c^2\alpha^2) + \frac{ik_c\alpha}{3} + O(k_c^3\alpha^3)\right] \tag{3.43}$$

under assumptions

$$\langle \lambda_1^2 \rangle/(\lambda_0 + 2\mu_0)^2 \ll 1 \tag{3.44}$$

and

$$k_c\alpha \ll 1.$$

Hence

$$W_c = 2k_c^2\{1 + [\langle \lambda_1^2 \rangle(1 - k_c)/(\lambda_0 + 2\mu_0)^2] \times [(\tfrac{2}{3} - \tfrac{1}{3} k_c^2\alpha^2) + \tfrac{2}{3} ik_c\alpha + O(k_c^3\alpha^3)]\}. \tag{3.45}$$

Since $D_1(s) = 0$ when $\lambda_1(x)$ is the only random variable, the rotational part of the wave remains unchanged.

Finally, when all the correlation functions except

$$\mathcal{R}_{\mu\mu}(r) = \langle \mu_1^2 \rangle e^{-r/\alpha} \tag{3.46}$$

are zero, we proceed as above to obtain D_1 and D_2. Then, by assuming that $\langle \mu_1^2 \rangle/\mu_0^2 \ll 1$, $k_s\alpha \ll 1$, and $k_c\alpha \ll 1$, we obtain

$$K_s/k_s = \left[1 + \frac{\langle \mu_1^2 \rangle}{\mu_0^2}\left(\frac{1}{3} + \frac{k_c^2}{15k_s^2}\right) + i\frac{\langle \mu_1^2 \rangle}{\mu_0^2} k_s\alpha\left(\frac{4}{35} + \frac{373}{840}\frac{k_c^3}{k_s^3}\right)\right] + O(k_s^3), \tag{3.47}$$

$$K_c/k_c = \left[1 + \frac{\langle \mu_1^2 \rangle}{\mu_0^2}\left(\frac{35}{24} + \frac{119k_c^2}{60k_s^2}\right) + i\frac{\langle \mu_1^2 \rangle}{\mu_0^2} k_s\alpha\left(\frac{83}{10} + \frac{53k_c^3}{20k_s^3}\right)\right] + O(k_s^3\alpha^2), \tag{3.48}$$

and then

$$W_s = 2k_s^2\left\{1 + \frac{\langle \mu_1^2 \rangle}{\mu_0^2}\left[\frac{13}{10}\left(1 - \frac{k_c^2}{k_s^2}\right) + \frac{297}{350} ik_s\alpha\left(1 - \frac{k_c^3}{k_s^3}\right) + O(k_s^4\alpha^2)\right]\right\}, \tag{3.49}$$

$$W_c = 2k_c^2\left\{1 + \frac{\langle \mu_1^2 \rangle}{\mu_0^2}\frac{273}{10}\left(\frac{k_c^2}{k_s^2} - 1\right) + \frac{465}{14} ik_s\alpha\left(\frac{k_c^3}{k_s^3} - 1\right) + O(k_s^4\alpha^2)\right]\right\}. \tag{3.50}$$

It must be pointed out that the results obtained above for three special cases may be used to obtain Green's dyadics of L_1 for other special cases with all the correlation functions being zero except combinations of $\mathcal{R}_{\rho\rho}$, $\mathcal{R}_{\mu\mu}$, and $\mathcal{R}_{\lambda\lambda}$. For example, if all correlation functions are zero except $\mathcal{R}_{\rho\rho}$ and $\mathcal{R}_{\mu\mu}$, $D_1(s)$ will be simply the sum of the results given by (3.30) and the $D_1(s)$ for $\mathcal{R}_{\mu\mu} \neq 0$. Likewise $D_2(s)$ for each special case can be easily obtained.

Therefore, we shall not consider such cases in detail.

To obtain Green's dyadic of L_1 for each special case, we simply insert the previously derived K_s, K_c, W_c, and W_S into (3.24). Then the average wave may be approximated by

$$\langle \bar{u} \rangle \cong \int G_1(\bar{x} - \bar{x}') f(\bar{x}') d\bar{x}'. \tag{3.51}$$

In the region where $|\bar{x} - \bar{x}'|$ is approximately equal to $|\bar{x}| - \bar{x} \cdot \bar{x}'/|\bar{x}|$, (3.51) may be expressed as

$$\langle \bar{u} \rangle \cong [G_{11}^c(|\bar{x}|)\bar{\bar{I}} + G_{12}^c(|\bar{x}|)\hat{r}\hat{r}] \cdot \bar{S}_c + [G_{11}^s(|\bar{x}|)\bar{\bar{I}} + G_{12}^s(|x|)\hat{r}\hat{r}] \cdot \bar{S}_s, \tag{3.52}$$

where

$$\bar{S}_c \equiv \int e^{-iK_c \bar{x} \cdot \bar{x}'/|\bar{x}|} f(\bar{x}') d\bar{x}', \tag{3.53}$$

$$\bar{S}_s \equiv \int e^{-iK_s \bar{x} \cdot \bar{x}'/|\bar{x}|} f(\bar{x}') d\bar{x}', \tag{3.54}$$

$$G_1(|\bar{x} - \bar{x}'|) = [G_{11}^c(|\bar{x} - \bar{x}'|) + G_{11}^s(|\bar{x} - \bar{x}'|)]\bar{\bar{I}} + [G_{12}^c(|\bar{x} - \bar{x}'|) + G_{12}^s(|\bar{x} - \bar{x}'|)]\hat{r}\hat{r}. \tag{3.55}$$

From (3.24) and (3.51), we see that the approximate average wave has the same form as that of the wave in the absence of the random fluctuation except that its propagation constants and the coefficients of the Green's dyadic $1/4\pi\omega^2\rho_0$ are replaced by corresponding effective propagation constants and $k_{c,s}^2/2\pi\omega^2\rho_0 W_{c,s}$, respectively. Therefore, the effect of the random inhomogeneity of the medium on waves can be seen by comparing $k_{c,s}$ with $K_{c,s}$ and $2k_{c,s}^2$ with $W_{c,s}$. From (3.55), (3.36), (3.44), (3.47), and (3.48), we see that $\mathrm{Re}(K_{c,s} - k_{c,s}) > 0$ and $\mathrm{Im}\,(K_{c,s}) > 0$. Since the real part of a propagation constant is inversely proportional to the velocity of the wave and the imaginary part determines the damping ratio of the wave, the results obtained show that the random inhomogeneity causes delay of the wave propagation as well as attenuation.

B. Electromagnetic Waves

$$\nabla \times \nabla \times \bar{E} - \omega^2 \mu\epsilon \bar{E} - (\nabla\mu/\mu) \times \nabla \times \bar{E} = -i\omega\mu\bar{J}, \tag{3.56}$$

which is derived from Maxwell's equations, assuming that permeability μ and permitivity ϵ are random functions of position. We again follow Karal and Keller's[1] notations for the average and random parts of dielectric constant, permeability and conductivity. Then

$$L_0 \bar{E} = \nabla \times \nabla \times \bar{E} - \omega^2 \mu_0 \epsilon_0 [1 + \langle \mu_1(\bar{x}) \epsilon_1(\bar{x}) \rangle] \bar{E}, \tag{3.57}$$

$$L' \bar{E} = -\omega^2 \mu_0 \epsilon_0 [\mu_1 + \epsilon_1 + \mu_1 \epsilon_1 - \langle \mu_1(\bar{x}) \epsilon_1(\bar{x}) \rangle] \bar{E} - [\nabla\mu_1/(1 + \mu_1)] \times \nabla \times \bar{E}. \tag{3.58}$$

It must be pointed out that the operator L_0 is different from the operator in the absence of fluctuations, for (3.57) has an additional term involving $\langle \mu_1(\bar{x}) \epsilon_1(\bar{x}) \rangle$ which in general is not negligible. Since the random process is assumed to be statistically isotropic, if we denote the correlation function $\langle \mu_1(\bar{x}) \epsilon_1(\bar{x}') \rangle$ by $\mathcal{R}_{\mu\epsilon}(|\bar{x} - \bar{x}'|)$, then $\langle \mu_1(\bar{x}) \epsilon_1(\bar{x}) \rangle$ is equal to $\mathcal{R}_{\mu\epsilon}(0)$. Taking the point of observation $\bar{x}$ away from the source region, the Green's dyadic for L_0 can be readily[1,7] written down as

$$G_0(\bar{x} - \bar{x}') = G_{01}(|\bar{x} - \bar{x}'|)\bar{\bar{I}} + G_{02}(|\bar{x} - \bar{x}'|)\hat{r}\hat{r}, \tag{3.59}$$

where

$$G_{01}(r) = -(1 - ikr - k^2 r^2) e^{ikr}/4\pi k^2 r^3, \tag{3.60}$$

$$G_{02}(r) = (3 - 3ikr - k^2 r^2) e^{ikr}/4\pi k^2 r^3, \tag{3.61}$$

$$k^2 = \omega^2 \mu_0 \epsilon_0 [1 + \mathcal{R}_{\mu\epsilon}(0)]. \tag{3.62}$$

Applying L' defined by (3.58) on (3.59), we obtain

$$L' G_0(\bar{x} - \bar{x}') = -\omega^2 \mu_0 \epsilon_0 [\epsilon_1 + \mu_1 + \epsilon_1 \mu_1 - \mathcal{R}_{\mu\epsilon}(0)] G_0(\bar{x} - \bar{x}') + [H(r)/(1 + \mu_1)][\nabla\mu_1 \cdot \hat{r} - (\nabla\mu_1 \cdot \hat{r})\bar{\bar{I}}], \tag{3.63}$$

where

$$H(r) = (k^2/r^2)(1 - ikr) e^{ikr}. \tag{3.64}$$

Using similar notations for correlation functions, we have

$$\langle L'(x) G_0(\bar{x} - \bar{x}') L'(\bar{x}') \rangle = \bar{\bar{C}}_1 \cdot \bar{\bar{I}} + \bar{\bar{\bar{C}}}_2 \cdot \hat{r} \times \nabla' \times, \tag{3.65}$$

where

$$\bar{\bar{C}}_1 = \omega^4 \mu_0^2 \epsilon_0^2 (\mathcal{R}_{\mu\mu} + 2\mathcal{R}_{\mu\epsilon} + \mathcal{R}_{\epsilon\epsilon} + \mathcal{R}_{\mu\mu}\mathcal{R}_{\epsilon\epsilon} + \mathcal{R}_{\mu\epsilon}^2) G_0 + \omega^2 \mu_0 \epsilon_0^2 \left[\left(\sum_{n=0}^{\infty} \langle \mu_1^{2n} \rangle \right) (\mathcal{R}'_{\mu\mu} + \mathcal{R}'_{\mu\epsilon}) + \left(\sum_{n=0}^{\infty} (2n + 1) \langle \mu_1^{2n} \rangle \right) (\mathcal{R}_{\mu\mu}\mathcal{R}'_{\mu\epsilon} + \mathcal{R}'_{\mu\mu}\mathcal{R}_{\mu\epsilon}) \right] \times H(r)(\hat{r}\hat{r}) - \bar{\bar{I}}), \tag{3.66}$$

$$\bar{\bar{\bar{C}}}_2 = -\omega^2 \mu_0 \epsilon_0 \sum_{n=0}^{\infty} \langle \mu_1^{2n} \rangle [(\mathcal{R}'_{\mu\mu} + \mathcal{R}'_{\mu\epsilon}) + (2n + 1) \times (\mathcal{R}_{\mu\mu}\mathcal{R}'_{\mu\epsilon} + \mathcal{R}'_{\mu\mu}\mathcal{R}_{\mu\epsilon})] G_0 - H(r) \frac{\partial^2}{\partial r^2} \left[\sum_{n=0}^{\infty} \left(\sum_{m=0}^{\infty} \times \frac{\langle \mu_1^{2n+1}(\bar{x}) \mu_1^{2m+1}(\bar{x}') \rangle}{(2n + 1)(2m + 1)} \right) \right] (\hat{r}\hat{r} - \bar{\bar{I}}), \tag{3.67}$$

$$\frac{1}{1 + \mu_1} = \sum_{n=0}^{\infty} (-\mu_1)^n. \tag{3.68}$$

Following the same process which led (3.8) to (3.16), we obtain

$$\mathcal{L}_{11}(s) = s^2 - k^2 - D_1(s), \tag{3.69}$$

$$\mathcal{L}_{12}(s) = -s^2 - D_2(s), \tag{3.70}$$

where

$$D_1(s) = \int_0^\infty \left[C_1^I h - C_1^{rr} \frac{1}{sr^2} \frac{\partial h}{\partial s} + C_2^I \frac{s}{r} \frac{\partial h}{\partial s} + C_2^{rr} \frac{1}{r^3} \left(\frac{\partial^2 h}{\partial s^2} - \frac{1}{s} \frac{\partial h}{\partial s} \right) \right] dr, \quad (3.71)$$

$$D_2(s) = \int_0^\infty \left[-C_2^{rr} \frac{s}{r^3} \left(\frac{\partial^3 h}{\partial s^3} - \frac{1}{s} \frac{\partial^2 h}{\partial s^2} + \frac{1}{s} \frac{\partial h}{\partial s^2} \right) + (C_2^I + C_2^{rr}) \frac{s}{r} \frac{\partial h}{\partial s} - C_1^{rr} \frac{1}{r^2} \left(\frac{\partial^2 h}{\partial s^2} - \frac{1}{s} \frac{\partial h}{\partial s} \right) \right] dr. \quad (3.72)$$

The superscripts "I" and "rr" denote components of the dyadics $\bar{\bar{I}}$ and $\hat{r}\hat{r}$, respectively. Here the Green's dyadics for L_1 can also be evaluated upon the determination of the roots of the equations $\mathcal{L}_{11}(s) = 0$ and $\mathcal{L}_{11}(s) + \mathcal{L}_{12}(s) = 0$. It has the form

$$G_1(\bar{x}, \bar{x}') = \frac{\bar{\bar{I}}}{2\pi r} \left\{ \sum_{l=1} \frac{1}{W_{sl}} \left[K_{sl}^2 e^{iK_{sl}r} - \frac{\partial}{\partial r} \left(\frac{1}{r} e^{iK_{sl}r} \right) \right] - \sum_{j=1} \frac{1}{W_{cj}} \frac{\partial}{\partial r} \left(\frac{1}{r} e^{iK_{cj}r} \right) \right\} - \frac{r}{2\pi} \hat{r}\hat{r} \frac{\partial}{\partial r} \times \left\{ \frac{1}{r} \frac{\partial}{\partial r} \left[\frac{1}{r} \left(\sum_{l=1} e^{iK_{sl}r} + \sum_{j=1} e^{iK_{cj}r} \right) \right] \right\}, \quad (3.73)$$

where

$$W_{sl} = 2K_{sl}^2 - K_{sl} \frac{\partial}{\partial K_{sl}} D_1(K_{sl}), \quad (3.74)$$

$$W_{cj} = K_{cj} \left(\frac{\partial}{\partial K_{cj}} D_1(K_{cj}) + \frac{\partial}{\partial K_{cj}} D_2(K_{cj}) \right). \quad (3.75)$$

The most common and important case is when the dielectric constant ϵ' is the only random variable. Then, from the definition of permitivity[1] ϵ, we have

$$\mathcal{R}_{\epsilon\epsilon}(r) = \left(\frac{\epsilon_0'}{\epsilon_0' + i\sigma_0/\omega} \right)^2 \mathcal{R}_{\epsilon'\epsilon'}(r), \quad (3.76)$$

$$\mathcal{R}_{\mu\epsilon}(r) = 0. \quad (3.77)$$

Then from (3.66) and (3.67)

$$\bar{\bar{C}}_1(r) = \omega^4 \mu_0^2 \epsilon_0^2 \mathcal{R}_{\epsilon\epsilon}(r) G_0, \quad (3.78)$$

$$\bar{\bar{C}}_2(r) = 0. \quad (3.79)$$

From (3.71) and (3.72)

$$D_1(s) = \omega^4 \mu_0^2 \epsilon_0^2 \int_0^\infty \mathcal{R}_{\epsilon\epsilon}(r) \left(G_{01} h - G_{02} \frac{1}{r^2 s} \frac{\partial h}{\partial s} \right) dr, \quad (3.80)$$

$$D_2(s) = -\omega^4 \mu_0^2 \epsilon_0^2 \int_0^\infty G_{02}(r) \mathcal{R}_{\epsilon\epsilon}(r) \frac{1}{r^2} \left(\frac{\partial^2 h}{\partial s^2} - \frac{1}{s} \frac{\partial h}{\partial s} \right) dr. \quad (3.81)$$

Let

$$\mathcal{R}_{\epsilon\epsilon}(r) = \left(\frac{\epsilon_0'}{\epsilon_0' + i\sigma_0/\omega} \right)^2 \langle \epsilon_1'^2 \rangle e^{-r/\alpha}, \quad (3.82)$$

where α is the correlation length of the random inhomogeneity. Then we can carry out the integrations in (3.80) and (3.81) with the help of (3.82). The results are

$$D_1(s) = \left(\frac{\epsilon_0'}{\epsilon_0' + i\sigma_0/\omega} \right)^2 \langle \epsilon_1'^2 \rangle k^2 \left\{ \frac{k_s^2 \alpha^2}{s^2\alpha^2 + (1 - ik_s\alpha)^2} - \frac{1}{(1 - ik_s\alpha)^2} \left[\frac{k^2\alpha^2}{3} + \sum_{n=1}^\infty (-1)^n \left(\frac{s\alpha}{1 - ik\alpha} \right)^{2n} \times \frac{1 - (2n+2)ik\alpha}{(2n+3)(2n+1)} \right] \right\}, \quad (3.83)$$

$$D_1(s) + D_2(s) = 2 \left(\frac{\epsilon_0'}{\epsilon_0' + i\sigma_0/\omega} \right)^2 \frac{\langle \epsilon_1'^2 \rangle k^2}{(1 - ik\alpha)^2} \left[\frac{k^2\alpha^2}{3} + \sum_{n=1}^\infty (-1)^n \left(\frac{s\alpha}{1 - ik\alpha} \right)^{2n} \frac{1 - (2n+2)ik\alpha}{(2n+3)(2n+1)} \right]. \quad (3.84)$$

Now to obtain roots of $\mathcal{L}_{11}(K) = 0$ in the neighborhood of $K^2 = k^2$, upon inserting (3.83) into (3.69) and performing the iteration under the assumption of small $\langle \epsilon_1^2 \rangle k^2\alpha^2$, we obtain

$$K^2 = k^2 \left[1 + \left(\frac{\epsilon_0^2}{\epsilon_0' + i\sigma_0/\omega} \right)^2 \langle \epsilon_1'^2 \rangle k^2 \left[\tfrac{11}{15} k^2\alpha^2 + i\tfrac{4}{3} k^3\alpha^3 + O(k^4\alpha^4) \right] \right]. \quad (3.85)$$

Consider

$$\mathcal{L}_{11}(s) + \mathcal{L}_{12}(s) = -[k^2 + D_1(s) + D_2(s)] = 0. \quad (3.86)$$

It can be easily shown that if $k\alpha$ is either very small or very large compared to unity, (3.86) does not have a zero. This implies that the Green' s dyadic will not have a longitudinal component. Finally with the insertion of (3.83), (3.25) yields

$$W_s = 2k^2 \left[1 - \left(\frac{\epsilon_0^2}{\epsilon_0' + i\sigma_0/\omega} \right)^2 \langle \epsilon_1'^2 \rangle k^2 \left[\tfrac{1}{15} \alpha^2 (1 - 11k^2) + i\tfrac{4}{3} k\alpha^3 (\tfrac{1}{5} - k^2) + O(k^4\alpha^4) \right] \right]. \quad (3.87)$$

When (3.86) does not have a zero in the complex s plane, the Green's dyadic $G_1(\bar{r})$ is given by the equation

$$G_1(\bar{x} - \bar{x}') = (e^{iK_s r}/2\pi r^3 W_s)[(1 - iK_s r - K_s^2 r^2)\bar{\bar{I}} - (3 - 3iK_s r - K_s^2 r^2)\hat{r}\hat{r}]. \quad (3.88)$$

Let us consider the following special cases:
(1) when σ is the only random variable,
(2) when ϵ is the only random variable,
(3) when μ is the only random variable.

Recalling the relation

$$\epsilon = \epsilon' + i\sigma/\omega, \quad (3.89)$$

We can easily observe that

$$\mathcal{R}_{\epsilon\epsilon}(r) = -[\omega^{-2}\sigma_0^2/(\epsilon_0' + i\omega^{-1}\sigma_0)^2] \mathcal{R}_{\sigma\sigma}(r). \quad (3.90)$$

Therefore, the form of equations for $D_1(s)$ and $D_2(s)$ must be the same as (3.80) and (3.81), respectively, but different by a constant which relates $\mathcal{R}_{\epsilon'\epsilon'}$ to $\mathcal{R}_{\sigma\sigma}$ or $\mathcal{R}_{\epsilon\epsilon}$ when cases (1) or (2) are considered.

In the case (3), $\bar{\bar{C}}_1$ and $\bar{\bar{C}}_2$ do not reduce to simple forms and $\bar{\bar{C}}_2$ involves higher-order moments of the random variable $\mu_1(\bar{x})$. Even when the random process is Gaussian, such series of higher-order

moments is very difficult to handle. However, one can avoid this difficulty by considering the equation for $\bar{H}$ field,

$$\nabla\times\nabla\times\bar{H}-\omega^2\mu\epsilon\bar{H}-(\nabla\epsilon\times\nabla\times\bar{H})/\epsilon = \nabla\times\bar{J}-(\nabla\epsilon\times\bar{J})/\epsilon, \quad (3.91)$$

which is of the same form as that of the $\bar{E}$ field (3.57). When ϵ is a constant, (3.91) reduces to

$$\nabla\times\nabla\times\bar{H}-\omega^2\mu\epsilon\bar{H}=\nabla\times\bar{J}. \quad (3.92)$$

Equation (3.92) is similar to the above mentioned special case (2) in which the effective propagation constant and $\langle\bar{H}\rangle$ can be easily computed. And then $\langle\bar{E}\rangle$ can be obtained from $\langle\bar{E}\rangle = (j\omega\epsilon)^{-1}\nabla\times\langle\bar{H}\rangle$.

Finally, the average electric field $\langle\bar{E}\rangle$ can be obtained from (3.51)-(3.55) by setting $\bar{S}_c$, G_{11}^c, and G_{12}^c to zero. In the limits $k\alpha\ll 1$ and $r\gg\alpha$, our result given by (3.88) agrees with that obtained by Tatarski.[5] Our $D_1(s)$ and $D_2(s)$ given by (3.80)-(3.84) have been previously obtained by Karal and Keller[1] through a different method.

The effects of the random inhomogeneity in medium parameters can be seen from the differences between K and k and between W_s and $2k^2$. It has been shown that K is a complex number with its real part larger than k. Therefore, the velocity of the average wave becomes lower than that of the wave in the absence of the fluctuation and the wave attenuates as it propagates away from the source, with attenuation coefficient proportional to $k^3\alpha^3$. $\mathrm{Re}(K-k)$ is proportional to $k^2\alpha^2$. These results agree with those obtained by Tatarski.[5]

4. APPLICABILITY OF THE RENORMALIZATION APPROXIMATION

When we rewrite (1.4) in a form

$$\langle\bar{u}\rangle = \int G(\bar{x},\bar{x}')\bar{f}(\bar{x}')d\bar{x}', \quad (4.1)$$

where

$$\begin{aligned}G(\bar{x},\bar{x}') = {} & G_1(\bar{x}-\bar{x}')-\int\int\int\int G_1(\bar{x}-\bar{x}')[\langle L'(\bar{x}^{V})\\ & \times G_0(\bar{x}^{V}-\bar{x}^{II})L'(\bar{x}^{II})G_0(\bar{x}^{II}-\bar{x}^{III})\rangle\\ & \cdot\langle L'(\bar{x}^{III})G_0(\bar{x}^{III}-\bar{x}^{IV})L'(\bar{x}^{IV})\\ & \times G_1(\bar{x}^{IV}-\bar{x}^{I})\rangle+\langle L'(\bar{x}^{II})G_0(\bar{x}^{II}-\bar{x}^{III})\\ & \times L'(\bar{x}^{III})G_0(\bar{x}^{III}-\bar{x}^{IV})\rangle\cdot\langle L'(\bar{x}^{V})\\ & \times G_0(\bar{x}^{V}-\bar{x}^{II})L'(\bar{x}^{IV})G_1(\bar{x}^{IV}-\bar{x}^{I})\rangle]\\ & \times d\bar{x}^{II}d\bar{x}^{III}d\bar{x}^{IV}d\bar{x}^{V}+\cdots. \end{aligned} \quad (4.2)$$

We apply $L_1(\bar{x})$ to (4.2) and integrate over the $\bar{x}$-space. After the proper change of variables, we obtain

$$\begin{aligned}\int L_1 G d\bar{x} = {} & \bar{\bar{I}}-\int\int\int\int[\langle L'(\bar{r}_1)G_0(\bar{r}_1-\bar{r}_2)L'(\bar{r}_3)\\ & \times G_0(\bar{r}_3-\bar{r}_4)\rangle\langle L'(\bar{r}_2)G_0(\bar{r}_2-\bar{r}_3)\\ & \times L'(\bar{r}_4)G_1(\bar{r}_4)\rangle+\langle L'(\bar{r}_2)G_0(\bar{r}_2-\bar{r}_3)\\ & \times L'(\bar{r}_3)G_0(\bar{r}_3-\bar{r}_4)\rangle\cdot\langle L'(\bar{r}_1)G_0(\bar{r}_1-\bar{r}_2)\\ & \times L'(\bar{r}_4)G_1(\bar{r}_4)\rangle]d\bar{r}_1d\bar{r}_2d\bar{r}_3d\bar{r}_4+\cdots. \end{aligned} \quad (4.3)$$

Recalling (3.8)-(3.14) and (3.65)-(3.67), we see that $\langle L'G_0L'G_0\rangle$ can be written in the form

$$\begin{aligned}&\langle L'(\bar{r}_1)G_0(\bar{r}_1-\bar{r}_2)L'(\bar{r}_2)G_0(\bar{r}_2-\bar{r}_3)\rangle\\ &\quad=\sum_{m,n}[\mathcal{R}_{mn}(|\bar{r}_1-\bar{r}_2|)\mathbf{F}_m^0(\bar{r}_1-\bar{r}_2)\\ &\qquad\cdot\mathbf{F}_n^0(\bar{r}_2-\bar{r}_3)], \end{aligned} \quad (4.4)$$

where superscript "0" denotes that the tensor function contains components of G_0 only. Since in the previous section we have shown that the Green's dyadic G_1 has a form very similar to that of G_0, we simply put a superscript "1" on the tensor function $\mathbf{F}_n^1$ generated by applying $L'(\bar{x})$ on $G_1(\bar{x}-\bar{x}')$.

We now rewrite (4.3) in terms of the new expression given by (4.4). After introducing the proper change of variables, we can write the integral term in (4.3) in the form

$$\begin{aligned}\mathcal{I} \equiv {} & \int\int\int\int\Big[\Big(\sum_{m,n}\mathcal{R}_{mn}(|\bar{t}_1+\bar{t}_3|)\mathbf{F}_m^0(\bar{t}_1)\mathbf{F}_n^0(\bar{t}_3)\Big)\\ & \times\Big(\sum_{m,n}\mathcal{R}_{mn}(|\bar{t}_2+\bar{t}_3|)\mathbf{F}_m^1(\bar{t}_2)\mathbf{F}_n^1(\bar{t}_4)\Big)\\ & +\Big(\sum_{m,n}\mathcal{R}_{mn}(|\bar{t}_2|)\mathbf{F}_m^0(\bar{t}_2)\mathbf{F}_n^0(\bar{t}_3)\Big)\\ & \times\Big(\sum_{m,n}\mathcal{R}_{mn}(|\bar{t}_1+\bar{t}_2+\bar{t}_3|)\mathbf{F}_m^0(\bar{t}_1)\,\mathbf{F}_n^1(\bar{t}_4)\Big)\Big]\\ & \times d\bar{t}_1d\bar{t}_2d\bar{t}_3d\bar{t}_4. \end{aligned} \quad (4.5)$$

Let us consider a correlation function

$$\mathcal{R}_{mn}(|t_1+t_2|)=C_{mn}e^{-|\bar{t}_1+\bar{t}_2|/\alpha}. \quad (4.6)$$

Since $|\bar{t}_1+\bar{t}_2|\geq|t_1-t_2|$, where $t_1\equiv|\bar{t}_1|$ and $t_2\equiv|\bar{t}_2|$, it follows that

$$e^{-|\bar{t}_1+\bar{t}_2|/\alpha}\leq e^{-|t_1-t_2|/\alpha}\quad\text{for all }\bar{t}_1\text{ and }\bar{t}_2. \quad (4.7)$$

We can estimate the integral $\mathcal{I}$ by inserting correlation functions of the form

$$\mathcal{R}_{mn}^e(\bar{t}_1,\bar{t}_2)=C_{mn}e^{-|t_1-t_2|/\alpha} \quad (4.8)$$

into (4.5) for $\mathcal{R}_{mn}$. Then we have

$$\begin{aligned}|\mathcal{I}|\leq {} & \int_0^\infty\int\int\int\Big[\Big(\sum_{m,n}\mathcal{R}_{mn}^e(\bar{t}_1,\bar{t}_3)|I_m^0(\bar{t}_1)||I_n^0(\bar{t}_3)|\Big)\\ & \times\Big(\sum_{m,n}\mathcal{R}_{mn}^e(\bar{t}_2,\bar{t}_3)|I_m^0(\bar{t}_2)||I_n^1(\bar{t}_4)|\Big)\\ & \times\Big(\sum_{m,n}\mathcal{R}_{mn}^e(\bar{t}_2)|I_m^0(\bar{t}_2)||I_n^0(\bar{t}_3)|\Big)\\ & \times\Big(\sum_{m,n}\mathcal{R}_{mn}^e(\bar{t}_1,\bar{t}_2,\bar{t}_3)|I_m^0(\bar{t}_1)||I_n^1(\bar{t}_4)|\Big)\Big]\\ & \times dt_1dt_2dt_3dt_4, \end{aligned} \quad (4.9)$$

where

$$I_m^i(t)\equiv\oint\!\!\!\!\!\!\int_S\mathbf{F}_m^i(\bar{t})dS,\qquad i=0,1, \quad (4.10)$$

S denoting the surface of a sphere with radius t and the center at the origin.

The dyadic functions $\mathbf{F}_m^i$ can be reduced to very simple forms when they are integrated over the

angular variables. For elastic waves, if $\rho_1(\bar{x})$ is the only random variable,

$$I_1^0(t) = -(\omega^2/3\rho_0)(k_c^2 t e^{ik_c t} + 2k_s^2 t e^{ik_s t}), \quad (4.11)$$

if $\lambda_1(\bar{x})$ is the only random variable,

$$I_1^0(t) = I_2^0(t) = (k_c^4/3\rho_0) t e^{ik_c t}, \quad (4.12)$$

$$I_3^0(t) = (k_c^2/3\rho_0)(1 - ik_c t)e^{ik_c t}, \quad (4.13)$$

and, if $\mu_1(\bar{x})$ is the only random variable,

$$I_1^0(t) = I_2^0(t) = (1/3\rho_0)(2k_c^4 t e^{ik_c t} + k_s^2 t e^{ik_s t}), \quad (4.14)$$

$$I_3^0(t) = (2/3\rho_0)[k_c^2(1 - ik_c t)e^{ik_c t} + k_s^2(1 - ik_s t)e^{ik_s t}]. \quad (4.15)$$

For the representative case of electromagnetic waves,

$$I_1^0(t) = -\tfrac{2}{3} k^2 t e^{ikt}. \quad (4.16)$$

To find the quantity Y such that we may write

$$\langle u \rangle = \int G_1(\bar{x} - \bar{x}')\bar{f}(\bar{x}')d\bar{x}'\{1 + O[Y]\} \quad (4.17)$$

we insert (4.8) and (4.11)–(4.16) into (4.9) and carry out the integration.

When $\rho_1(x)$ is the only random variable,

$$Y = (\langle \rho_1^2 \rangle^2/\rho_0^4)\omega^8(k_s^2\alpha^2 + k_s^4\alpha^4), \quad (4.18)$$

when $\lambda_1(\bar{x})$ is the only random variable,

$$Y = [\langle \lambda_1^2 \rangle^2 \omega^8/(\lambda_0 + 2\mu_0)^4](k_c^4\alpha^4 + 1/k_c^2\alpha^2), \quad (4.19)$$

and, when $\mu_1(\bar{x})$ is the only random variable,

$$Y = (\langle \mu_1^2 \rangle^2 \omega^8/\mu_0^4)(k_s^4\alpha^4 + 1/k_s^2\alpha^2). \quad (4.20)$$

For the representative special case of electromagnetic waves, we have

$$Y = \langle \epsilon_1^2 \rangle^2 k^2\alpha^2(1 + k^2\alpha^2). \quad (4.21)$$

When $k_{c,s}\alpha \ll 1$, from (4.19) the second nonvanishing term of $\langle u \rangle$ of the special elastic wave, when $\rho_1(\bar{x})$ is the only random variable, is of the order $Y = \langle \rho_1^2 \rangle^2 \omega^8 k_s^2\alpha^2/\rho_0^4$. This means that if $k_s^2\alpha^2$ is small enough, the first-order renormalization approximation is reasonably good even when $\langle \rho_1^2 \rangle$ is comparable to ρ_0^2 (so-called "strong fluctuation"). For the other two cases, from (4.19) and (4.20) we obtain $\langle \lambda_1^2 \rangle^2 \omega^8/(\lambda_0 + 2\mu_0)^4 k_c^2\alpha^2 \ll 1$ and $\langle \mu_1^2 \rangle^2 \omega^8/\mu_0^4 k_s^2\alpha^2 \ll 1$, respectively. These imply that either the frequency or the mean square fluctuation of the elastic wave must be small enough for the first-order renormalization approximation to be applicable. These implications are to be expected, because the random operator L' of the first case contains no derivatives of random variables and the second two cases do.

ACKNOWLEDGMENTS

The authors wish to express their appreciation to National Science Foundation for their support of this research through Grant No. GK-2756. Both the authors thank Professor J. B. Keller of New York University for many helpful discussions. One of the authors (S. H. C.) also wishes to thank Professor J. B. Keller for the invitation to Courant Institute of Mathematical Sciences, where the last stage of this research was completed.

[1] F. C. Karel, Jr., and J. B. Keller, J. Math. Phys. 5, 537 (1964).
[2] J. B. Keller, in *Symposium on Turbulence of Fluids and Plasmas* (Brooklyn Polytechnic Institute, Brooklyn, N.Y., 1968), pp. 131–42.
[3] U. Frisch, "Wave Propagation in Random Media," in *Probabilistic Methods in Applied Mathematics, Vol. 1*, edited by A. T. Barucha-Reid (Academic, New York, 1968).
[4] V. I. Tatarskii and M. E. Gertenshtein, Zh. Eksp. Teor. Fiz. **44**, 676 (1963) [Sov. Phys. JETP 17, 458 (1963)].
[5] V. I. Tatarskii, Zh. Eksp. Teor. Fiz. 46, 1399 (1964) [Sov. Phys. JETP 19, 946 (1964)].
[6] S. H. Chung, thesis, State University of New York at Stony Brook, 1968.
[7] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. I, II.
[8] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1966), Vol. II.

Weaker Versions of Zeeman's Conjectures on Topologies for Minkowski Space

S. Nanda
Mathematics Department, R.E. College, Rourkela-8, Orissa, India
(Received 8 October 1970)

In a paper [Topology **6**, 161 (1967)] Zeeman conjectured that (a) the finest topology on Minkowski space that induces the one-dimensional Euclidean topology on every timelike line and (b) the finest topology on Minkowski space that induces the three-dimensional Euclidean topology on every spacelike hyperplane have the same group of homeomorphisms G which is generated by the inhomogeneous Lorentz group and the dilatations. This paper deals with two topologies on Minkowski space which are weaker than those in (a) and (b), respectively, and have the property that they induce the one-dimensional and the three-dimensional Euclidean topology on timelike lines and spacelike hyperplanes, respectively. It is then shown that both topologies have G as their homeomorphism group. Thus, what we have shown amounts to proving the weaker versions of Zeeman's conjectures.

1. INTRODUCTION

The principle that the topological structure of Minkowski space should be such that its homeomorphism group is isomorphic to the group generated by the inhomogeneous Lorentz group and dilatations (call this group G) has led many[1-3] to suggest new topologies for Minkowski space. It is evident that while selecting a topology for Minkowski space, there is no reason why one topology should be given preference over another as long as both satisfy the prescription above, i.e., the homeomorphism group of each is G. Normally, one would try to avoid any discrimination and expect that the infimum and the supremum of all such topologies having G as their homeomorphism group might be of significance. It would then seem desirable to solve the problem of enumerating all topologies on Minkowski space having G as their homeomorphism group.

Indeed this problem is a particular case of a much more general problem posed by Everett and Ulam[4-5]: Given a group H and a set M, which topologies on M, have H as their homeomorphism group? Although a general answer to this question is not yet given, it seems from the example of Minkowski space that it would be very difficult to construct all topologies with the given property. In the case of Minkowski space, this target of constructing all such topologies (at least in principle) has not been achieved although it seems to be in sight. Since the cones associated with the indefinite fundamental form are invariant under G, it would be desirable, as a first step, to construct topologies which arise in a most natural manner from the cone structure. This paper is intended to be a contribution in this direction.

In his paper Zeeman[1] conjectured that (a) the finest topology on Minkowski space that induces the one-dimensional Euclidean topology on every timelike line has G as its homeomorphism group (we call it the time topology) and (b) the finest topology on Minkowski space that induces the three-dimensional Euclidean topology on every spacelike hyperplane also has G as its group of homeomorphisms (we call it the space topology). The two topologies we have considered in this paper, which we have named as the t-topology and the s-topology, respectively, are weaker than those in (a) and (b) and induce the Euclidean topology on timelike lines and spacelike hyperplanes, respectively. Moreover, the homeomorphism group in each case is G. Therefore, what we have done in this paper amounts to proving the weaker versions of Zeeman's conjectures. For convenience, we adopt the same notation and terminology as in our previous paper[3] (hereafter referred to as I). Technically, the proofs are not very complicated as in the case of fine topology[1] or the space topology.[3]

Once again, the tools used in the course of the proof are only the rudiments of topology.

2. DEFINITION AND PROPERTIES OF THE t-TOPOLOGY

Definition: Let $N_\epsilon^t(x) = N_\epsilon^E(x) \cap \mathbf{C}^T(x)$. The t-topology is defined by specifying a local base $\mathfrak{N}(x)$ of neighborhoods at each point x of M as follows:

$$\mathfrak{N}(x) = \{N_\epsilon^t(x) : \epsilon > 0\}.$$

Let M^t denote the set M equipped with the t-topology. One can always define the same topology by defining a countable local base at each point by taking rational ϵ's.

It follows at once from the definition that the t-topology is finer than the Euclidean topology and hence Hausdorff. It induces the one-dimensional Euclidean topology on every timelike line and the discrete topology on every lightlike line, light cone, and spacelike hyperplane. It is also not difficult to show that the t-topology is neither normal nor locally compact. If we define the time topology on M as the finest topology that induces the Euclidean topology on every timelike line, then, it is easy to show that the time topology is strictly finer than the t-topology. Consider, for example, a sequence $\{t_n\}$ of distinct timelike lines passing through a point z. Choose a point $z_n \in t_n$, such that $d(z_n, z) \to 0$ as $n \to \infty$. Let $Z = \{z_n\}$. We shall show that Z is closed in the time topology. It is enough to prove that $T \cap Z$ is a finite set (and hence closed), where T is any timelike line. Suppose to the contrary that $T \cap Z$ is an infinite set; then T being complete (as a metric space), the sequence of points $T \cap Z$ must converge to a point of $T \cap Z$, but this point must be z since the space is Hausdorff. Therefore, T passes through z; but, then, $T \cap Z$ is at most a singleton by our choice. This is a contradiction and our point is proved. Z^c is then open in the time topology (Z^c denotes the complement of Z). On the other hand, Z^c is not open in the t-topology, because any basic open set $N_\epsilon^t(z)$ in the t-topology about the point z will intersect Z.

3. HOMEOMORPHISMS OF M^t

We shall show in this section that the group of homeomorphisms of M^t is G. In view of Theorem

1 of I, it is enough[6] to prove that every homeomorphism of M^t is either $<$-order preserving or $<$-order reversing. We start by proving a lemma.

Lemma 1: Let h be a homeomorphism of M^t and $x < y$; then there exists a point $z \in [x,y]$ such that either (i) $hx < hz$ or (ii) $hz < hx$. Moreover, in case (i), $hx < hz'$ for all $z' \in [x,z]$ and similarly, in case (ii), $hz' < hx$ for all $z' \in [x,z]$.

Proof: h continuous implies that $h^{-1}(N^t_\delta(hx))$ is an open set containing x. Since sets of the form $N^t_\epsilon(x)$ form a local base at x, it follows that, for some $\epsilon > 0$, $N^t_\epsilon(x) \subset h^{-1}(N^t_\delta(hx))$. If we choose $z \in [x,y] \cap N^t_\epsilon(x)$, then clearly (i) or (ii) is satisfied.

To prove the second part of our assertion, suppose that (ii) is satisfied, i.e., $hz < hx$. Assume to the contrary that $hx < hz'$ for some $z' \in [x,z]$. Let $(x,z]$ denote the open-closed interval on the timelike line passing through x and z. Since the induced topology on $(x,z]$ is Euclidean, $(x,z]$ is connected. Therefore, z and z' belong to the same component of $N^t_\epsilon(x) - \{x\}$ (which is clearly disconnected and has two components), whereas hz and hz' belong to different components of the image. Since h is a homeomorphism, this is a contradiction. A similar proof applies to case (i).

Lemma 2: Let h be a homeomorphism of M^t and $x < y_1, y_2$, where y_1 and y_2 are on different timelike lines passing through x. Let z_1 and z_2 be chosen in $[x,y_1]$ and $[x,y_2]$, respectively, as in Lemma 1; then, either (i) $hx < hz_1$ and $hx < hz_2$ or (ii) $hz_1 < hx$ and $hz_2 < hx$; i.e., both z_1 and z_2 are either mapped to the forward time cone or the backward time cone at hx.

Proof: Suppose to the contrary that $hz_1 < hx < hz_2$. Let $\delta < 0$ be such that hz_1 and hz_2 belong to $N^t_\delta(hx)$; then, there exists $\epsilon > 0$ such that $N^t_\epsilon(x) \subset h^{-1}(N^t_\delta(hx))$ or, equivalently, $h(N^t_\epsilon(x)) \subset N^t_\delta(hx)$. Choose $z_1' \in (x,z_1] \cap N^t_\epsilon(x)$ and $z_2' \in (x,z_2] \cap N^t_\epsilon(x)$. By Lemma 1, $hz_1' < hx < hz_2'$. Note that z_1' and z_2' belong to the same component of $N^t_\epsilon(x) - \{x\}$, whereas hz_1' and hz_2' belong to different components of the image. This is a contradiction and the lemma is proved.

Lemma 2 implies that h is locally $<$-order preserving or $<$-order reversing. That this is globally true follows from the following lemma.

Lemma 3: Let $x < y$ and z be chosen as in Lemma 1; then $hx < hz$ implies $hx < hy$. On the other hand, $hz < hx$ implies $hy < hx$.

Proof: Due to the symmetry, it is enough to prove the first case only. Assume therefore that $hx < hz$.

For any point $p \in [x,y]$, choose $z_{p_1} \in [p,y]$ and $z_{p_2} \in [x,p]$ as in Lemma 1. Since $z_{p_2} < p < z_{p_1}$, it follows from Lemma 2 that either $hz_{p_2} < hp < hz_{p_1}$ or $hz_{p_1} < hp < hz_{p_2}$ (*). Cover $[x,y]$ by the intervals $\{(z_{p_1}, z_{p_2})\}$, $p \in [x,y]$. By using compactness of $[x,y]$, it is possible to choose a finite covering. Thus we get a finite sequence of points

$$x = p^0 < p^1 < p^2 < \cdots < p^n = y.$$

From our assumption that $hx < hz$ and the statement (*) above, it follows that

$$hx = hp^0 < hp^1 < hp^2 < \cdots < hp^n = hy.$$

A similar argument will show that if $hz < hx$, then $hy < hx$. This completes the proof of the lemma.

With the help of Lemmas 2 and 3, one can now proceed exactly as in the proof of Lemma 11 of I to show that h is either $<$-order preserving or $<$-order reversing, and then a proof identical to that of Theorem 2 (Sec. 6 of I) together with Zeeman's theorem[7] (that the group of $<$-automorphisms of M is G) will imply the following.

Theorem 1: The group of homeomorphisms of M^t is G.

Before we proceed on to the derivation of the group of homeomorphisms of the s-topology, it is worthwhile to note that the derivation of the homeomorphism group of the t-topology is quite straightforward. This is expected for two reasons. First, the topology is first countable (it is also second countable) and the basic open sets at each point are explicitly known. This leads to considerable simplicity in the proof. Secondly, the basic open sets are directly related to the time cones and hence to the partial order structure of the set M. Both points are essential. Indeed, if one drops first countability—for example, if one deals with the time topology—then one runs into enormous difficulties and, in that case, the conjecture (a) (Sec. 1) still remains to be proved. The advantage obtained from the second point is also considerable. It will shortly appear from Sec. 5 that the derivation of the homeomorphism group in the case of the s-topology is rather involved (although it is first countable) due to the simple fact that its basic open sets are not directly related to the partial order $<$.

4. DEFINITION AND PROPERTIES OF THE s-TOPOLOGY

Definition: Let $N^s_\epsilon(x) = N^E_\epsilon(x) \cap C^S(x)$. The s-topology on M is defined by specifying a countable local base at each point x of M as follows:

$$\mathfrak{N}(x) = \{N^s_\epsilon x);\ \epsilon > 0 \text{ and } \epsilon \text{ is rational}\}.$$

Let M equipped with the s-topology be denoted by M^s. It is easy to see that such a topology is finer than the Euclidean topology and hence Hausdorff. It induces the Euclidean topology on every spacelike hyperplane and the discrete topology on every timelike or lightlike line. However, it is not the finest topology with this property. For example, it is strictly weaker than the space topology, which is defined as the finest topology on M with respect to which the induced topology on every spacelike hyperplane is Euclidean (Proof in Sec. 4 of I). Like the t-topology, the s-topology possesses none of the good properties such as compactness, normality, etc.

5. HOMEOMORPHISMS OF THE s-TOPOLOGY

In what follows, closure of a set A in the s-topology will be denoted simply by $\bar{A}$.

Lemma 4: Let $N_\epsilon^s(x)$ be a basic open set of the s-topology about the point x and L any lightlike line passing through x; then, there is a point $z \in L$ such that $[x,z] \subset \overline{N_\epsilon^s(x)}$.

Proof: Take $z \in L$ such that $d(x,z) = \frac{1}{2}\epsilon$; we claim that $[x,z] \subset \overline{N_\epsilon^s(x)}$. Suppose to the contrary that there is a point z' in $[x,z]$ such that $z' \notin \overline{N_\epsilon^s(x)}$. Take a spacelike hyperplane H through z'; $H \cap \overline{N_\epsilon^s(x)}$ is closed in the induced Euclidean topology and must contain z'. Contradiction.

Lemma 5: Let h be a homeomorphism of M^s and L a light ray (i.e., a lightlike line) emerging from x; then, there is a point z in L with $x <. z$, such that $h[x,z] \subset C^L(hx)$.

Proof: $N_\infty^s(x) \cap L = \{x\}$ implies that $h(N_\infty^s(x)) \cap hL = \{hx\}$. Since $h(N_\infty^s(x))$ is an open set containing hx, it contains a basic open set of the form $N_\delta^s(hx)$ and, clearly,

$$N_\delta^s(hx) \cap hL = \{hx\} \tag{1}$$

Consider now the inverse image of $N_\delta^s(hx)$. Applying the same argument as above, we have $\epsilon > 0$ such that $N_\epsilon^s(x) \subset h^{-1}(N_\delta^s(hx))$ or, equivalently, $h(N_\epsilon^s(x)) \subset N_\delta^s(hx)$. Since h is a homeomorphism, we also have

$$h(\overline{N_\epsilon^s(x)}) = \overline{h(N_\epsilon^s(x))} \subset \overline{N_\delta^s(hx)}.$$

Choose $z' \neq x$ in L as in Lemma 4 so that $[x,z'] \subset \overline{N_\epsilon^s(x)}$. Then we have

$$h[x,z'] \subset h(\overline{N_\epsilon^s(x)}) \subset \overline{N_\delta^s(hx)}. \tag{2}$$

From (1), we also have

$$N_\delta^s(hx) \cap h[x,z'] = \{hx\}. \tag{3}$$

(2) and (3) together imply that $h(x,z']$ belongs to the boundary of the set $\overline{N_\delta^s(hx)}$, where boundary means the topological boundary, in this case, $\overline{N_\delta^s(hx)} - N_\delta^s(hx)$.

It is evident that the boundary of the set $\overline{N_\delta^s(hx)}$ is a union of two sets A and B, where (i) A consists of points contained in the light cone at hx and (ii) B consists of points contained in the space cone at hx at a Euclidean distance δ from hx. Take a basic open set $N_\lambda^s(hx)$ about hx with $2\lambda = \delta$. Now choose a suitable $\eta < \epsilon$ such that $h(N_\eta^s(x)) \subset N_\lambda^s(hx)$. If we choose $z \in \overline{N_\eta^s(x)} \cap [x,z']$, then

$$h(x,z] \subset h(\overline{N_\eta^s(x)}) = \overline{h(N_\eta^s(x))} \subset \overline{N_\lambda^s(hx)} \tag{4}$$

Moreover,

$$h(x,z] \subset h(x,z'] \subset A \cup B. \tag{5}$$

Therefore, from (4) it follows that $h(x,z] \subset (A \cup B) \cap \overline{N_\lambda^s(hx)} \subset A$. Since $A \subset C^L(hx)$, it follows that $h[x,z] \subset C^L(hx)$ and this completes the proof.

Corollary: Similarly, there exists a z' in L with $z' <. x$ such that $h[z',x] \subset C^L(hx)$.

The following lemma will show that $h[x,z]$ is contained in a lightlike line through hx.

Lemma 6: Let L and z be as in Lemma 5; then, $h[x,z] \subset L(hx,hz)$, where $L(hx,hz)$ denotes the line passing through the points hx and hz.

Proof: For each $y \in [x,z]$, choose z_y and z'_y as in Lemma 5 and its corollary; then, $h(z_y,z'_y) \subset C^L(hy)$. By Lemma 5, we also have $h(z_y,z'_y) \subset C^L(hx)$. Thus, $h(z_y,z'_y) \subset C^L(hx) \cap C^L(hy) = L(hx,hy)$. Note that $[x,z]$ is compact in the Euclidean topology and $\{(z_y,z'_y)\}$, $y \in [x,z]$ is an open covering of $[x,z]$. It is therefore possible to choose a finite covering which would give rise to a finite sequence of points

$$x = y_0 <. y_1 <. y_2 <. \cdots <. y_n = z,$$

such that each (z_{y_i},z'_{y_i}) is mapped by h into $L_i = L(hx,hy_i)$. Moreover,

$$(z_{y_i},z'_{y_i}) \cap (z_{y_{i+1}},z'_{y_{i+1}}) \neq \emptyset$$

for each i. This implies that, for each i, L_i and L_{i+1} meet at points other than hx. Since each L_i is contained in $C^L(hx)$, it follows that

$$L_1 = L_2 = L_3 = \cdots = L_n = L(hx,hz).$$

The proof is therefore complete.

Corollary: A similar result also applies to $[x,z']$, where z' is determined as in the corollary to Lemma 5.

Lemma 6 only tells us that for any point $p \in [x,z]$ with $x <. z$, it is true that either $hx <. hp$ or $hp <. hx$, i.e., the point p is either mapped to the forward light cone or the backward light cone at hx. Could it happen that, for two points p and p' in $[x,z]$, $hp' <. hx <. hp$? To eliminate this possibility, we have the following lemma.

Lemma 7: If h and z are as in Lemma 2, then there is a point $p \in [x,z]$, such that $h[x,p]$ is either contained in the forward light cone or the backward light cone at hx.

Proof: Let L^+ and L^- denote the subsets of $L(hx,hz)$ contained in the forward and backward light cone, respectively. Suppose now to the contrary that the lemma is false; then, for every $p \in [x,z]$, $h[x,p]$ meets L^+ and L^-. Since $\overline{N_\epsilon^s(x)} \cap [x,z] \supset [x,p]$ for some p in $[x,z]$, it follows that, for every $\epsilon > 0$, $h(\overline{N_\epsilon^s(x)})$ meets L^+ and L^-; consequently, every Euclidean ϵ-neighborhood of hx will contain points of L^+ and L^-. Now choose $\delta > 0$ such that $h^{-1}(N_\delta^s(hx)) \subset N_\epsilon^s(x)$. Take a 3-simplex Δ^3 in $N_\delta^s(hx) \cap H$, where H is a spacelike hyperplane passing through the point hx. Choose a point $q \in L^+ \cap h[x,z]$ such that the faces of the 4-simplex Δ^4 obtained by joining q to Δ^3 are all spacelike. This can always be done by taking q sufficiently near hx in the Euclidean sense. The topology induced on the boundary $\partial\Delta^4$ of the 4-simplex

$\triangle^4$ is then Euclidean and, therefore, $\partial\triangle^4 \cong S^3$, where S^3 denotes the 3-sphere. Hence,

$$g : \mathrm{id} \circ h^{-1} : S^3 \cong \partial\triangle^4 \xrightarrow{h^{-1}} M^s \xrightarrow{\mathrm{id}} M^E$$

(id denotes the identity map from $M^s \longrightarrow M^E$), being a one-to-one continuous map from a compact space to a Hausdorff space is an imbedding.[8] Moreover, by our choice, $h^{-1}(\triangle^3) \subset C^S(x)$.

Let $\lambda = d(hx, q)$ and $2\mu = \lambda$. Let $\nu > 0$ be such that $h(N_\nu^s(x)) \subset N_\mu^s(hx)$ and let $n \in N_\nu^s(x) \cap [x, z]$. By assumption, $h[x, n]$ meets both L^+ and L^- and by construction, $\partial\triangle^4$ meets $h[x, n]$ only at the point hx. Therefore, $g(\partial\triangle^4)$ meets $[x, n]$ only at x. On the other hand, both components of $M^s - \partial\triangle^4$ contain points of $h[x, n]$ or, equivalently, both components of $g(M^s - \partial\triangle^4) = M^E - g(\partial\triangle^4) = M^E - S^3$ will contain points of $[x, n]$. This is a contradiction and the proof of the lemma is complete.

Our next object is to prove that the relation $<.$ is either preserved or reversed locally by a homeomorphism.

Lemma 8: Let L and L' be two lightlike lines passing through x, and let p and p' be determined in L and L', respectively, as in Lemma 7, such that $x <. p$ and $x <. p'$; then $h[x, p]$ and $h[x, p']$ are either both contained in the forward light cone at hx or both in the backward light cone at hx.

Proof: Suppose to the contrary that $h[x, p]$ is contained in the forward light cone at hx and $h[x, p']$ is contained in the backward light cone at hx. Choose $\delta > 0$ such that $h^{-1}(N_\delta^s(hx)) \subset N_\epsilon^s(x)$ for some $\epsilon > 0$. It is then possible to choose a three-dimensional ball B in $N_\delta^s(hx) \cap H$, where H is a spacelike hyperplane passing through hx, such that $h^{-1}B \subset N_\epsilon^s(x)$. Choose a 3-simplex $\triangle^3$ in B. From the arguments of Lemma 7, it is clear that any Euclidean ϵ-neighborhood of hx contains points of $h[x, p]$; therefore, choose a suitable q in $h[x, p]$ such that all faces of the 4-simplex $\triangle^4$ which is obtained by joining the point q to $\triangle^3$ are contained in spacelike hyperplanes. Clearly, the topology induced on the boundary $\partial\triangle^4$ of the 4-simplex $\triangle^4$ by the s-topology is Euclidean, i.e., $\partial\triangle^4$ is a homeomorph of S^3.

Now $M^E - \partial\triangle^4$ has two components. Let U be the interior component of $M^E - \partial\triangle^4$, then U meets $h[x, p]$ and does not meet $h[x, p']$. Without loss of generality, one can assume $h(x, p] \subset U$ or, equivalently, $(x, p] \subset h^{-1}U$. Moreover, as in the case of Lemma 7, the map $g = \mathrm{id} \circ h^{-1}$,

$$g : S^3 \cong \partial\triangle^4 \xrightarrow{h^{-1}} M^s \xrightarrow{\mathrm{id}} M^E,$$

is an imbedding. Therefore, $g(\partial\triangle^4)$ is a homeomorph of S^3. By our choice, $g(\triangle^3)$ is contained in $C^S(x)$; it is therefore clear that $g(\partial\triangle^4)$ will meet $(x, p']$ and, consequently, gU will meet $(x, p']$. This is a contradiction which proves the lemma.

Corollary: A similar proof will imply that if $p <. x <. p'$, then $hp <. hx <. hp'$ or $hp' <. hx <. hp$; in other words, if p and p' are oppositely oriented with respect to $<.$, then hp and hp' are oppositely oriented with respect to hx.

Remark: The procedure of constructing a 4-simplex $\triangle^4$ and making a disconnection by $g(\triangle^4)$ in Lemmas 7 and 8 is the same as in the case of fine topology.[1] It has only been adapted here to the case of M^s.

We are now in a position to prove that the order relation $\ll$ is either preserved or reversed by a homeomorphism of M^s.

Lemma 9: Let $x \ll y$ and h a homeomorphism of M^s; then either $hx \ll hy$ or $hy \ll hx$.

Proof: Take a point $p \in C^L(x)$ such that $x <. p <. y$; then $[x, p] \cup [p, y]$ is compact in the Euclidean topology. Moreover, at each point of $[x, p] \cup [p, y]$, the relation $<.$ is either preserved or reversed. Therefore, using Euclidean compactness and Lemma 8, one has $hx \ll hy$ or $hy \ll hx$. [The procedure is exactly the same as in the case of Lemma 3 (Sec. 3)]. This completes the proof of the lemma.

Now, repeating the same procedure as in Lemma 11 and Theorem 2 of Sec. 6 of I, one can easily prove the following:

Theorem 2: The group of homeomorphisms of M^s is G.

ACKNOWLEDGMENTS

The author is grateful to Professor R. D. Norman and Professor A. J. Coleman of Queen's University Mathematics Department for their invaluable academic and financial assistance during his stay at Queen's. A scholarship awarded by the Canadian Commonwealth Scholarship and Fellowship Administration is gratefully acknowledged.

---

1 E. C. Zeeman, Topology 6, 161 (1967),
2 S. Nanda, Ph. D. Thesis, Queen's University at Kingston.
3 S. Nanda, J Math. Phys. 12, 394 (1971).
4 C. J. Everett and S. M. Ulam, Bull. Am. Math. Soc. 54, 285t, 646 (1948).
5 S. M. Ulam, *A Collection of Mathematical Problems* (Interscience, New York, 1960).
6 The proof in that case will proceed in the same way as in Theorem 2 of I.
7 E. C. Zeeman, J. Math. Phys. 5, 490 (1964).
8 J. L. Kelly, *General Topology* (Van Nostrand, New York, 1955), p. 141.

Lattice Green's Function for the Orthorhombic Lattice in Terms of the Complete Elliptic Integral

Tsuyoshi Horiguchi,*§ Yoshitake Yamazaki,† and Tohru Morita‡§
Tohoku University, Sendai, Japan
(Received 12 November 1970)

The real and imaginary parts of the orthorhombic lattice Green's function at the origin are expressed as a sum of simple integrals of the complete elliptic integrals of the first kind. In order to give the expressions for all values of the variable from $-\infty$ to $+\infty$, use is made of the method of analytic continuation. The results of the numerical computations are shown by figures.

1. INTRODUCTION

We consider the lattice Green's function at the origin for the orthorhombic lattice ($a:b:c = 1:x:y$, $\alpha = \beta = \gamma = 90°$) with arbitrary γ_1, γ_2, and γ_3:

$$G(t) = \frac{1}{\pi^3}\int_0^\pi dx \int_0^\pi dy \int_0^\pi dz \frac{1}{t - \gamma_1 \cos x - \gamma_2 \cos y - \gamma_3 \cos z}, \tag{1.1}$$

where $t = s - i\epsilon$; s takes on real values and ϵ is an infinitesimal positive number. This integral is an even function of γ_1, γ_2 and γ_3 and hence we may assume that all of γ_1, γ_2 and γ_3 are positive without loss of generality. For the simple cubic lattice, all of γ_1, γ_2 and γ_3 are equal to unity, and for the tetragonal lattice, any two of them are equal to unity and the other is arbitrary. The function (1.1) for these lattices was expressed as a sum of simple integrals of the complete elliptic integral of the first kind, by Morita and Horiguchi.[1,2] For the rectangular and the square lattices, the lattice Green's function were studied by Katsura and Inawashiro,[3] using the Mellin–Barnes type integral at an arbitrary lattice point. At the origin, the Green's functions for these two-dimensional lattices were expressed by the complete elliptic integral.[3,4]

We shall consider (1.1) for $-\infty < s < \infty$. Function (1.1) is real at $s \geq \gamma_1 + \gamma_2 + \gamma_3$. For this region, the function $G(t)$ is readily expressed as an integral of the complete elliptic integral of the first kind, by using the expression of the rectangular lattice Green's function, given in a preceding paper of two of the present authors.[4] The function defined by (1.1) as a complex function of the complex variable t is analytic on the whole complex t plane, excluding the real axis from $-(\gamma_1 + \gamma_2 + \gamma_3)$ to $\gamma_1 + \gamma_2 + \gamma_3$. Hence an expression of the function $G(t)$ on the whole t plane can be attained by the procedure of the analytic continuation from the above-mentioned expression, which is useful at $s > \gamma_1 + \gamma_2 + \gamma_3$.

In Sec. 2, in order to make our discussion self-contained, we give the expressions for the complete elliptic integral of the first kind and its analytic continuation[1] at the values of modulus on the real and imaginary axis in terms of the complete elliptic integral of modulus between zero and unity. In Secs. 3 and 4, they are used to express the real and imaginary parts of $G(s - i\epsilon)$ at all the real values of s from $-\infty$ to $+\infty$. The resulting expressions have the form of a sum of simple integrals of the complete elliptic integral of the first kind. The results of the numerical calculation are shown by figures.

2. THE COMPLETE ELLIPTIC INTEGRAL OF THE FIRST KIND AS A COMPLEX FUNCTION OF THE MODULUS[1]

The complete elliptic integral of the first kind $\mathbf{K}(k)$ as a complex function of the complex modulus k is defined by[5]

$$\mathbf{K}(k) = \int_0^{\pi/2} d\theta \frac{1}{\sqrt{1 - k^2 \sin^2\theta}}. \tag{2.1}$$

This function is an even function of k and $\mathbf{K}(k^*) = \mathbf{K}(k)^*$. As a consequence, $\mathbf{K}(k)$ is real when k is pure imaginary. The expansions in powers of k and $k' = (1 - k^2)^{1/2}$ are given, respectively, as follows:

$$\mathbf{K}(k) = \frac{\pi}{2}\sum_{n=0}^{\infty}\left(\frac{(\frac{1}{2})_n}{n!}\right)^2 k^{2n}, \tag{2.2}$$

$$\mathbf{K}(k) = -\frac{2}{\pi}\mathbf{K}(k')\ln k' + \sum_{n=0}^{\infty}\left(\frac{(\frac{1}{2})_n}{n!}\right)^2 [\psi(n+1) - \psi(n+\tfrac{1}{2})]k'^{2n}, \tag{2.3}$$

where

$$(\tfrac{1}{2})_n = \Gamma(n+\tfrac{1}{2})/\Gamma(\tfrac{1}{2}), \qquad \psi(z) = \Gamma'(z)/\Gamma(z).$$

The function $\mathbf{K}(k)$ has branch points at $k = \pm 1$. Expressions (2.1)–(2.3) are analytic on the Riemann surface excluding the branch cuts connecting $+1$ and $+\infty$ and -1 and $-\infty$, respectively, on the

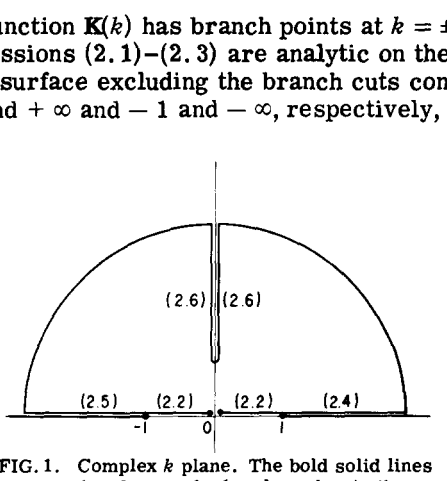

FIG. 1. Complex k plane. The bold solid lines between 1 and $+\infty$ and -1 and $-\infty$ denote the branch cuts for the complete elliptic integral of the first kind. The numbers of the equation useful at each portion near the axis are referred to.

real axis. We call this part of the Riemann surface as the sheet I. In the following calculation, we need only the function just above the real axis and along the imaginary axis on the sheet I, and do not need its analytic continuation to other sheets through the cut. The sheet I with the branch cuts connecting $+1$ and $+\infty$ and -1 and $-\infty$ and possible values of k occurring in Secs. 3 and 4, are shown in Fig. 1. Expressions (2.1)–(2.3) are valid on the sheet I. Equation (2.2) is used on the real axis $k = k_R + i\epsilon$, where $|k_R| < 1$ and $\epsilon \gtrsim 0$. The function just above the real axis $k = k_R + i\epsilon$, where $|k_R| > 1$ and $\epsilon \gtrsim 0$, is given from the one at $|k| < 1$ with the aid of the formula

$$\mathbf{K}(k) = \frac{1}{k}\left[\mathbf{K}\left(\frac{1}{k}\right) + i\mathbf{K}\left(\frac{\sqrt{k^2-1}}{k}\right)\right], \tag{2.4}$$

where $k_R > 1$ and

$$\mathbf{K}(k) = \frac{1}{|k|}\left[\mathbf{K}\left(\frac{1}{k}\right) - i\mathbf{K}\left(\frac{\sqrt{k^2-1}}{k}\right)\right], \tag{2.5}$$

where $k_R < -1$.

The function on the imaginary axis $k = \pm\epsilon + ik_I$, where $\epsilon \gtrsim 0$ is given as

$$\mathbf{K}(ik_I) = (1 + k_I^2)^{-1/2}\,\mathbf{K}[k_I(1 + k_I^2)^{-1/2}]. \tag{2.6}$$

In the following section, (2.4)–(2.6) are used instead of (2.2) according to the value of the modulus k.

3. ORTHORHOMBIC LATTICE GREEN'S FUNCTION; FORMULAS I

The Green's function defined by (1.1) is real at $\gamma_1 + \gamma_2 + \gamma_3 \leq s$ and is readily expressed from the previous paper[4] in terms of the complete elliptic integral of the first kind as follows:

$$G(t) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}}\int_0^\pi dx\; k\mathbf{K}(k), \tag{3.1}$$

where

$$k = \left(\frac{4\gamma_2\gamma_3}{(t - \gamma_1\cos x)^2 - (\gamma_3 - \gamma_2)^2}\right)^{1/2}. \tag{3.2}$$

We note that the expression (3.1) is valid on the whole complex t plane. We put $t = s - i\epsilon$, where s is a real variable and ϵ is an infinitesimal positive number.

When $\gamma_1 + \gamma_2 + \gamma_3 \leq s$, k is real and lies between 0 and 1, and Eq. (3.1) is appropriate for the numerical calculation.

The function $G(s - i\epsilon)$ is complex at $-(\gamma_1 + \gamma_2 + \gamma_3) < s < \gamma_1 + \gamma_2 + \gamma_3$. We denote the real and imaginary parts as $G_R(s)$ and $G_I(s)$, respectively,

$$G(s - i\epsilon) = G_R(s) + iG_I(s). \tag{3.3}$$

One can easily confirm that $G_R(s)$ is an odd function of s and $G_I(s)$ is an even function:

$$G_R(-s) = -G_R(s), \qquad G_I(-s) = G_I(s), \tag{3.4}$$

for instance, with the aid of the definition (1.1). Hence we have only to consider the range $0 < s < \gamma_1 + \gamma_2 + \gamma_3$ in the following. The formulas for this range are derived from (3.1) with (3.2) by the analytic continuation. In (3.1) with (3.2), γ_2 and γ_3 are symmetric and hence we assume that $\gamma_2 < \gamma_3$ without loss of generality in the following. The derived formulas are different according as γ_1 is larger than both of γ_2 and γ_3 or less than them or between them. We shall give formulas separately for these three cases:

(I) $0 < \gamma_1 < \gamma_2 < \gamma_3$,
(II) $0 < \gamma_2 < \gamma_3 < \gamma_1$,
(III) $0 < \gamma_2 < \gamma_1 < \gamma_3$.

The results obtained for each case specified by inequalities between γ's are valid also in the limit when the inequalities are replaced by equalities, though the equality signs are not written explicitly. The same applies for the other cases considered in the following.

In this section, we consider Case (I). It consists of two cases, $\gamma_3 < \gamma_1 + \gamma_2$ and $\gamma_1 + \gamma_2 < \gamma_3$. We first consider the former.

Case (IA): $0 < \gamma_1 < \gamma_2 < \gamma_3 < \gamma_1 + \gamma_2$: k defined by (3.2) takes values on the ranges

$$\begin{aligned} &0 < k < 1,\\ &1 < k < +\infty,\\ &\quad k = ik_I, k_I \text{ is real},\\ &-\infty < k < -1,\\ &-1 < k < 0, \end{aligned}$$

according as

$$\begin{aligned} &\gamma_3 + \gamma_2 < s - \gamma_1\cos x < +\infty,\\ &\gamma_3 - \gamma_2 < s - \gamma_1\cos x < \gamma_3 + \gamma_2,\\ &\gamma_2 - \gamma_3 < s - \gamma_1\cos x < \gamma_3 - \gamma_2,\\ &-\gamma_2 - \gamma_3 < s - \gamma_1\cos x < \gamma_2 - \gamma_3,\\ &\quad -\infty < s - \gamma_1\cos x < -\gamma_2 - \gamma_3, \end{aligned} \tag{3.5}$$

respectively. In Fig. 1, the expressions useful for $\mathbf{K}(k)$ in the respective regions are shown. The regions satisfying the above inequalities are shown in Figs. 2, 4, and 6, according to the relative magnitudes of γ_1, γ_2, and γ_3, where the above five regions correspond to the regions which are

above the line $(x = x_1)$,
above the line $(x = x_2)$ and below the line $(x = x_1)$,
above the line $(x = x_3)$ and below the line $(x = x_2)$,
above the line $(x = x_4)$ and below the line $(x = x_3)$,
below the line $(x = x_4)$,

respectively. The values of x between the regions are given by

$$x_1 = \cos^{-1}[(s - \gamma_3 - \gamma_2)/\gamma_1], \tag{3.6}$$

$$x_2 = \cos^{-1}[(s - \gamma_3 + \gamma_2)/\gamma_1], \tag{3.7}$$

$$x_3 = \cos^{-1}[(s + \gamma_3 - \gamma_2)/\gamma_1], \tag{3.8}$$

$$x_4 = \cos^{-1}[(s + \gamma_3 + \gamma_2)/\gamma_1]. \tag{3.9}$$

For the present case, we have Fig. 2(A). For a fixed value of s, an integration is taken over x from 0 to π, when $\cos x$ varies from +1 to −1. Considering the figure, we subdivide the present case as follows:

(ii) $\gamma_3 + \gamma_2 - \gamma_1 < s < \gamma_3 + \gamma_2 + \gamma_1$,

(iii) $\gamma_3 - \gamma_2 + \gamma_1 < s < \gamma_3 + \gamma_2 - \gamma_1$,

(iv) $-\gamma_3 + \gamma_2 + \gamma_1 < s < \gamma_3 - \gamma_2 + \gamma_1$,

(v) $0 < s < -\gamma_3 + \gamma_2 + \gamma_1$.

We notice that, for the values of s belonging to Case (ii), we have regions for $0 < k < 1$ and $1 < k < +\infty$ according as $x_1 < x < \pi$ and $0 < x < x_1$; for s belonging to Case (iii), we have only one region where $1 < k < +\infty$ for $0 < x < \pi$; and so on. Figure 3 shows the values of k taken for the respective case. The formulas appropriate for $\mathbf{K}(k)$ at the respective ranges of k are found from Fig. 1. Applying the formulas (2.2) and (2.4)–(2.6) of $\mathbf{K}(k)$ in (3.1) for the respective ranges, one obtains the following formulas:

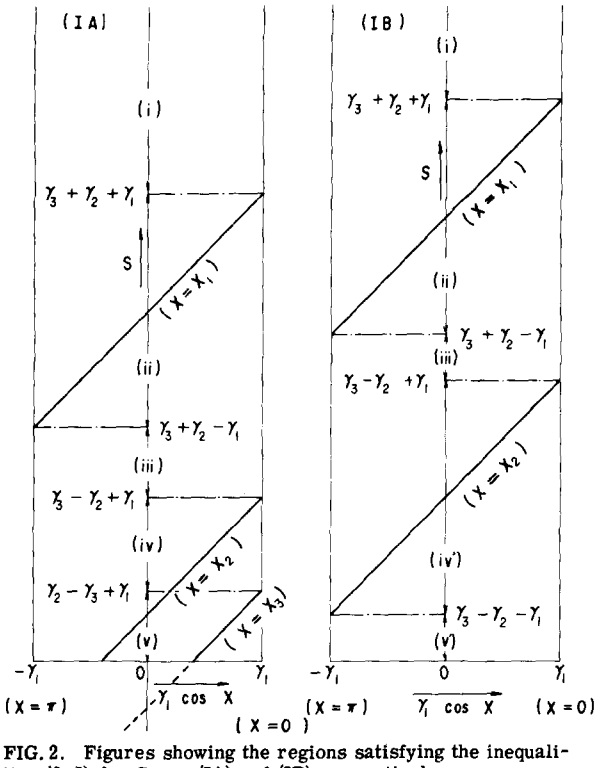

FIG. 2. Figures showing the regions satisfying the inequalities (3.5), for Cases (IA) and (IB), respectively.

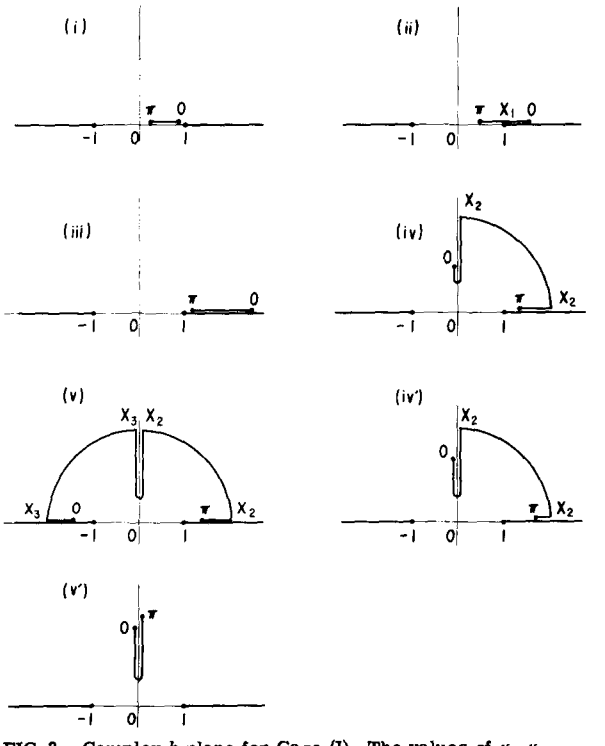

FIG. 3. Complex k plane for Case (I). The values of x_1, x_2, and x_3 are given by Eqs. (3.6)–(3.8).

Case (ii):

$$G_R(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}}\left[\int_0^{x_1} dx\, \mathbf{K}\left(\frac{1}{k}\right) + \int_{x_1}^{\pi} dx\, k\mathbf{K}(k)\right] \tag{3.10}$$

and

$$G_I(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}}\int_0^{x_1} dx\, \mathbf{K}\left(\frac{\sqrt{k^2-1}}{k}\right). \tag{3.11}$$

Case (iii):

$$G_R(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}}\int_0^{\pi} dx\, \mathbf{K}\left(\frac{1}{k}\right) \tag{3.12}$$

and

$$G_I(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}}\int_0^{\pi} dx\, \mathbf{K}\left(\frac{\sqrt{k^2-1}}{k}\right). \tag{3.13}$$

Case (iv):

$$G_R(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}}\int_{x_2}^{\pi} dx\, \mathbf{K}\left(\frac{1}{k}\right) \tag{3.14}$$

and

$$G_I(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}}\left[\int_0^{x_2} dx \frac{k_I}{\sqrt{1+k_I^2}}\mathbf{K}\left(\frac{k_I}{\sqrt{1+k_I^2}}\right) + \int_{x_2}^{\pi} dx\, \mathbf{K}\left(\frac{\sqrt{k^2-1}}{k}\right)\right], \tag{3.15}$$

where

$$k_I = -ik = \left(\frac{4\gamma_2\gamma_3}{(\gamma_3-\gamma_2)^2 - (s-\gamma_1\cos x)^2}\right)^{1/2}. \tag{3.16}$$

Case (v):

$$G_R(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}}\left[-\int_0^{x_3} dx\, \mathbf{K}\left(\frac{1}{|k|}\right) + \int_{x_2}^{\pi} dx\, \mathbf{K}\left(\frac{1}{k}\right)\right] \tag{3.17}$$

and

$$G_I(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}}\left[\int_0^{x_3} dx\, \mathbf{K}\left(\frac{\sqrt{k^2-1}}{|k|}\right) + \int_{x_3}^{x_2} dx \frac{k_I}{\sqrt{1+k_I^2}}\mathbf{K}\left(\frac{k_I}{\sqrt{1+k_I^2}}\right) + \int_{x_2}^{\pi} dx\, \mathbf{K}\left(\frac{\sqrt{k^2-1}}{k}\right)\right]. \tag{3.18}$$

Case (IB): $0 < \gamma_1 < \gamma_2 < \gamma_1 + \gamma_2 < \gamma_3$: For this case, inequalities (3.5) are shown in Fig. 2(B). From this figure, we see that the subdivisions for this case are

$$\begin{aligned}
&\text{(ii)} \quad \gamma_3 + \gamma_2 - \gamma_1 < s < \gamma_3 + \gamma_2 + \gamma_1,\\
&\text{(iii)} \quad \gamma_3 - \gamma_2 + \gamma_1 < s < \gamma_3 + \gamma_2 - \gamma_1,\\
&\text{(iv')} \quad \gamma_3 - \gamma_2 - \gamma_1 < s < \gamma_3 - \gamma_2 + \gamma_1,\\
&\text{(v')} \quad 0 < s < \gamma_3 - \gamma_2 - \gamma_1.
\end{aligned}$$

The values of k for the respective cases are shown in Fig. 3. For Cases (ii), (iii), and (iv'), we have Eqs. (3.10) and (3.11), Eqs. (3.12) and (3.13), and Eqs. (3.14) and (3.15), respectively. For Case (v'), we have

$$G_R(s) = 0 \tag{3.19}$$

and

$$G_I(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}} \int_0^{\pi} dx \, \frac{k_I}{\sqrt{1+k_I^2}} \mathbf{K}\left(\frac{k_I}{\sqrt{1+k_I^2}}\right). \tag{3.20}$$

4. ORTHORHOMBIC LATTICE GREEN'S FUNCTION: FORMULAS II AND III

We consider at first the case in which γ_1 is larger than γ_2 and γ_3, and then we have $0 < \gamma_2 < \gamma_3 < \gamma_1$. The formulas are given separately for the two cases, i.e., $\gamma_1 < \gamma_2 + \gamma_3$ and $\gamma_2 + \gamma_3 < \gamma_1$.

Case (IIA). $0 < \gamma_2 < \gamma_3 < \gamma_1 < \gamma_2 + \gamma_3$: The inequalities (3.5) are shown in Fig. 4(A) for this case. Now subdivisions are as follows:

$$\begin{aligned}
&\text{(ii)} \quad \gamma_3 - \gamma_2 + \gamma_1 < s < \gamma_3 + \gamma_2 + \gamma_1,\\
&\text{(iii)} \quad -\gamma_3 + \gamma_2 + \gamma_1 < s < \gamma_3 - \gamma_2 + \gamma_1,\\
&\text{(iv)} \quad \gamma_3 + \gamma_2 - \gamma_1 < s < -\gamma_3 + \gamma_2 + \gamma_1,\\
&\text{(v)} \quad 0 < s < \gamma_3 + \gamma_2 - \gamma_1.
\end{aligned}$$

The ranges of k for the respective cases are found from Fig. 4(A) and shown on Fig. 5. The values of x_1, x_2, x_3, and x_4 are given by (3.6)–(3.9). For Case (ii), we have Eqs. (3.10) and (3.11), for the real and imaginary parts, respectively. For Case (iii), one has

$$G_R(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}} \left[\int_{x_2}^{x_1} dx \, \mathbf{K}\left(\frac{1}{k}\right) + \int_{x_1}^{\pi} dx \, k\mathbf{K}(k)\right] \tag{4.1}$$

and

$$G_I(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}} \left[\int_0^{x_2} dx \, \frac{k_I}{\sqrt{1+k_I^2}} \mathbf{K}\left(\frac{k_I}{\sqrt{1+k_I^2}}\right) + \int_{x_2}^{x_1} dx \, \mathbf{K}\left(\frac{\sqrt{k^2-1}}{k}\right)\right]. \tag{4.2}$$

For Case (iv), one has

$$G_R(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}} \left[-\int_0^{x_3} dx \, \mathbf{K}\left(\frac{1}{|k|}\right) + \int_{x_2}^{x_1} dx \, \mathbf{K}\left(\frac{1}{k}\right) + \int_{x_1}^{\pi} dx \, k\mathbf{K}(k)\right] \tag{4.3}$$

and

$$G_I(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}} \left[\int_0^{x_3} dx \, \mathbf{K}\left(\frac{\sqrt{k^2-1}}{|k|}\right) + \int_{x_3}^{x_2} dx \, \frac{k_I}{\sqrt{1+k_I^2}} \mathbf{K}\left(\frac{k_I}{\sqrt{1+k_I^2}}\right) + \int_{x_2}^{x_1} dx \, \mathbf{K}\left(\frac{\sqrt{k^2-1}}{k}\right)\right]. \tag{4.4}$$

For Case (v), one has

$$G_R(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}} \left[-\int_0^{x_3} dx \, \mathbf{K}\left(\frac{1}{|k|}\right) + \int_{x_2}^{\pi} dx \, \mathbf{K}\left(\frac{1}{k}\right)\right] \tag{4.5}$$

and

$$G_I(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}} \left[\int_0^{x_3} dx \, \mathbf{K}\left(\frac{\sqrt{k^2-1}}{|k|}\right) + \int_{x_3}^{x_2} dx \, \frac{k_I}{\sqrt{1+k_I^2}} \mathbf{K}\left(\frac{k_I}{\sqrt{1+k_I^2}}\right) + \int_{x_2}^{\pi} dx \, \mathbf{K}\left(\frac{\sqrt{k^2-1}}{k}\right)\right]. \tag{4.6}$$

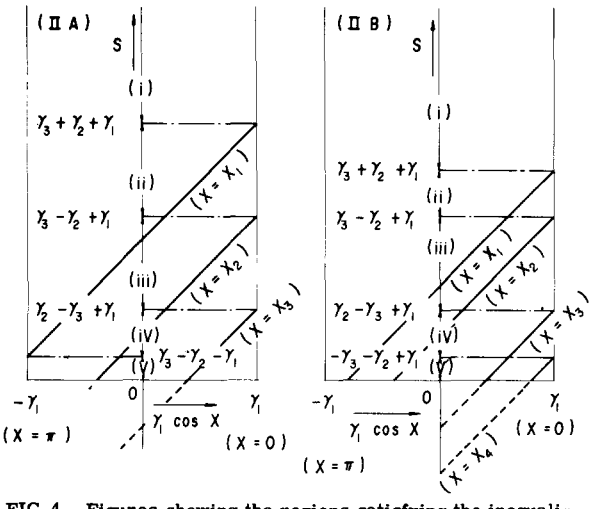

FIG. 4. Figures showing the regions satisfying the inequalities (3.5), for Cases (IIA) and (IIB), respectively.

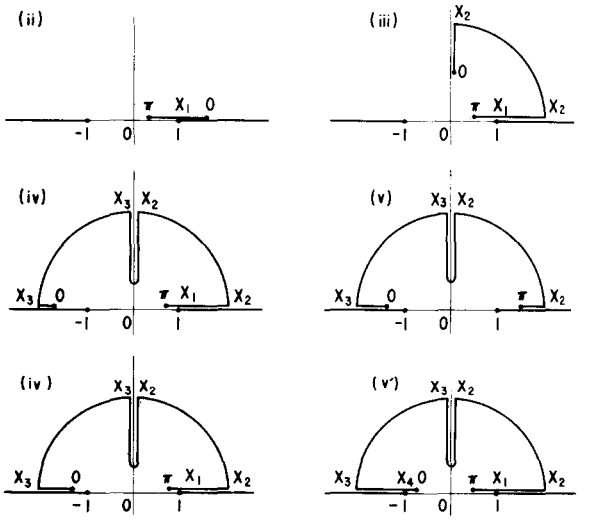

FIG. 5. Complex k plane for Case (II). The values of x_1, x_2, x_3, and x_4 are given by (3.6)–(3.9).

Case (IIB): $0 < \gamma_2 < \gamma_3 < \gamma_2 + \gamma_3 < \gamma_1$: For this case, the inequalities (3.5) are shown in Fig. 4(B). Now the range $0 < s < \gamma_1 + \gamma_2 + \gamma_3$ is subdivided as follows:

$$\begin{aligned}
&\text{(ii)} && \gamma_3 - \gamma_2 + \gamma_1 < s < \gamma_3 + \gamma_2 + \gamma_1,\\
&\text{(iii)} && -\gamma_3 + \gamma_2 + \gamma_1 < s < \gamma_3 - \gamma_2 + \gamma_1,\\
&\text{(iv')} && -\gamma_3 - \gamma_2 + \gamma_1 < s < -\gamma_3 + \gamma_2 + \gamma_1,\\
&\text{(v')} && 0 < s < -\gamma_3 - \gamma_2 + \gamma_1.
\end{aligned}$$

Figure 5 shows that, for Cases (ii), (iii), and (iv'), the values of k are restricted to the same regions as corresponding cases (ii), (iii), and (iv) of (IIA), and one obtains the same expressions (3.10) and (3.11), (4.1) and (4.2), and (4.3) and (4.4), respectively.

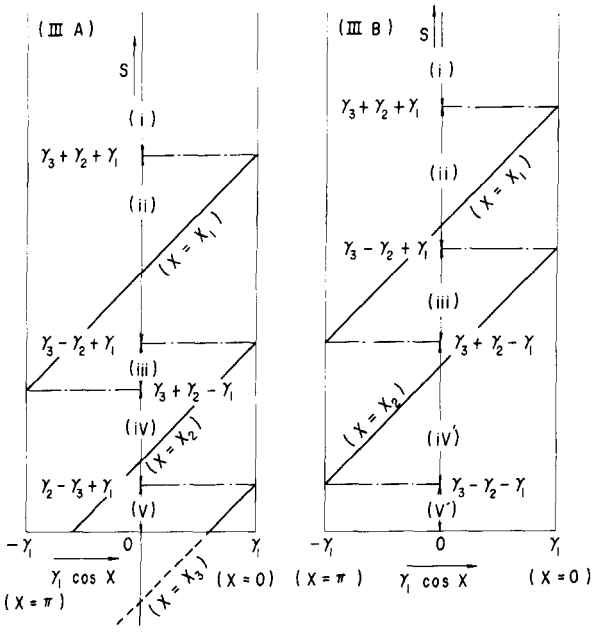

FIG. 6. Figures showing the regions satisfying the inequalities (3.5), for Cases (IIIA) and (IIIB), respectively.

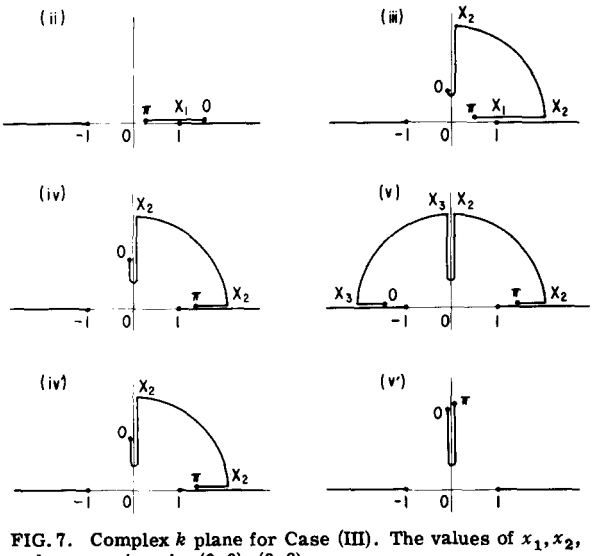

FIG. 7. Complex k plane for Case (III). The values of x_1, x_2, and x_3 are given by (3.6)–(3.8).

For Case (v'), one has

$$\begin{aligned}
G_R(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}}\Bigg[&-\int_0^{x_4} dx\; k\mathbf{K}(|k|)\\
&-\int_{x_4}^{x_3} dx\; \mathbf{K}\left(\frac{1}{|k|}\right)+\int_{x_2}^{x_1} dx\; \mathbf{K}\left(\frac{1}{k}\right)\\
&+\int_{x_1}^{\pi} dx\; k\mathbf{K}(k)\Bigg]
\end{aligned} \tag{4.7}$$

and

$$\begin{aligned}
G_I(s) = \frac{1}{\pi^2\sqrt{\gamma_2\gamma_3}}\Bigg[&\int_{x_4}^{x_3} dx\; \mathbf{K}\left(\frac{\sqrt{k^2-1}}{|k|}\right)\\
&+\int_{x_3}^{x_2} dx\; \frac{k_I}{\sqrt{1+k_I^2}}\,\mathbf{K}\left(\frac{k_I}{\sqrt{1+k_I^2}}\right)\\
&+\int_{x_2}^{x_1} dx\; \mathbf{K}\left(\frac{\sqrt{k^2-1}}{k}\right)\Bigg].
\end{aligned} \tag{4.8}$$

In the next place, we consider the case in which γ_1 is smaller than one of γ_2 and γ_3, and larger than the other. We assume that $\gamma_2 < \gamma_3$ without loss of generality, and hence we have $0 < \gamma_2 < \gamma_1 < \gamma_3$. The formulas are given for the two cases, $\gamma_3 < \gamma_1 + \gamma_2$ and $\gamma_1 + \gamma_2 < \gamma_3$, separately. Figure 6 shows the inequalities (3.5) for these cases.

Case (IIIA): $0 < \gamma_2 < \gamma_1 < \gamma_3 < \gamma_1 + \gamma_2$: By Fig. 6(A), we see that the subdivisions for this case are as follows:

$$\begin{aligned}
&\text{(ii)} && \gamma_3 - \gamma_2 + \gamma_1 < s < \gamma_3 + \gamma_2 + \gamma_1,\\
&\text{(iii)} && \gamma_3 + \gamma_2 - \gamma_1 < s < \gamma_3 - \gamma_2 + \gamma_1,\\
&\text{(iv)} && -\gamma_3 + \gamma_2 + \gamma_1 < s < \gamma_3 + \gamma_2 - \gamma_1,\\
&\text{(v)} && 0 < s < -\gamma_3 + \gamma_2 + \gamma_1.
\end{aligned}$$

The ranges of k for the respective cases are shown in Fig. 7. For Cases (ii), (iii), (iv), and (v), we have Eqs. (3.10) and (3.11), Eqs. (4.1) and (4.2), Eqs. (3.14) and (3.15), and Eqs. (3.17) and (3.18), respectively.

Case (IIIB): $0 < \gamma_2 < \gamma_1 < \gamma_1 + \gamma_2 < \gamma_3$: By Fig. 6(B), the range $0 < s < \gamma_1 + \gamma_2 + \gamma_3$ is subdivided as follows:

$$\begin{aligned}
&\text{(ii)} && \gamma_3 - \gamma_2 + \gamma_1 < s < \gamma_3 + \gamma_2 + \gamma_1,\\
&\text{(iii)} && \gamma_3 + \gamma_2 - \gamma_1 < s < \gamma_3 - \gamma_2 + \gamma_1,\\
&\text{(iv')} && \gamma_3 - \gamma_2 - \gamma_1 < s < \gamma_3 + \gamma_2 - \gamma_1,\\
&\text{(v')} && 0 < s < \gamma_3 - \gamma_2 - \gamma_1.
\end{aligned}$$

The ranges of k for these cases are shown on Fig. 7. They show that the formulas for Cases (ii), (iii), and (iv') are the same as the ones for Cases (ii), (iii), and (iv) of (IIIA), respectively. For Case (v'), we have Eqs. (3.19) and (3.20).

5. RESULTS OF NUMERICAL COMPUTATIONS AND REMARKS

We easily see that Eq. (1.1) is symmetric for the exchange of three parameters γ_1, γ_2, and γ_3. Therefore, we expect that formulas (IIA) and (IIIA) give the same results as formula (IA) when the parameters γ_1, γ_2, and γ_3 are interchanged in their roles. The situation is the same for formulas (IB),

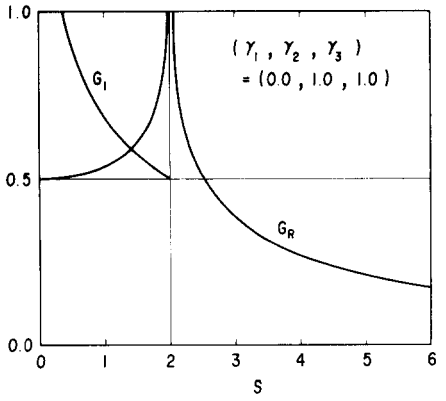

FIG. 8. The values of the lattice Green's function at the origin, $G(s - i\epsilon)$, for the orthorhombic lattice. Parameters γ_1, γ_2, and γ_3 are 0.0, 1.0, and 1.0. G_R and G_I indicate the real and imaginary parts, $G_R(s)$ and $G_I(s)$, respectively.

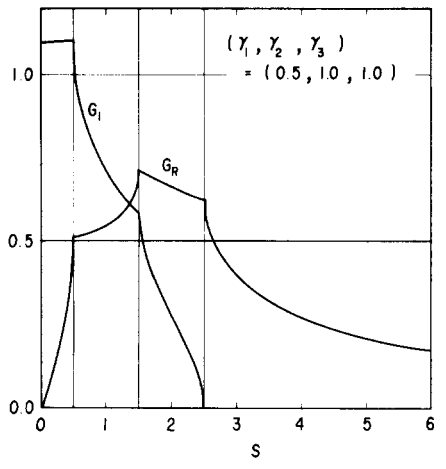

FIG. 9. $G_R(s)$ and $G_I(s)$ for the case where $(\gamma_1, \gamma_2, \gamma_3) = (0.5, 1.0, 1.0)$. Cf. the caption of Fig. 8.

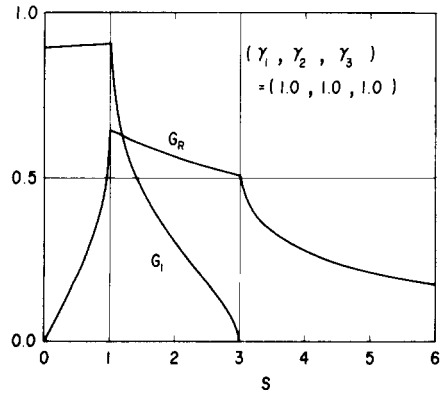

FIG. 10. $G_R(s)$ and $G_I(s)$ for the case where $(\gamma_1, \gamma_2, \gamma_3) = (1.0, 1.0, 1.0)$. Cf. the caption of Fig. 8.

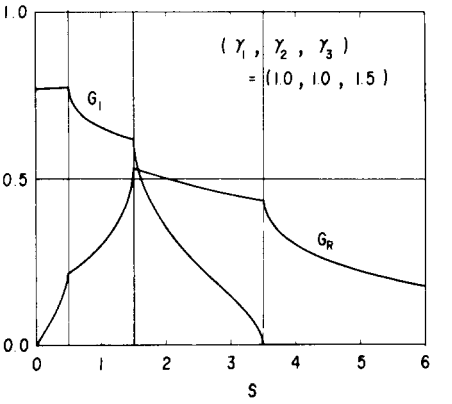

FIG. 11. $G_R(s)$ and $G_I(s)$ for the case where $(\gamma_1, \gamma_2, \gamma_3) = (1.0, 1.0, 1.5)$. Cf. the caption of Fig. 8.

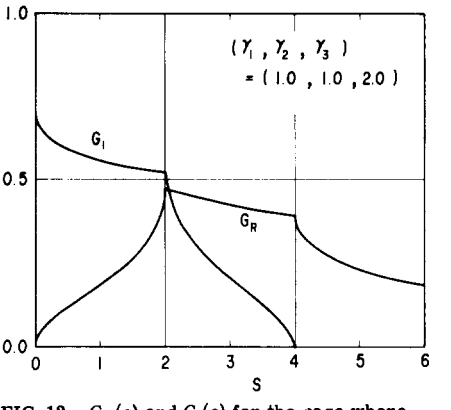

FIG. 12. $G_R(s)$ and $G_I(s)$ for the case where $(\gamma_1, \gamma_2, \gamma_3) = (1.0, 1.0, 2.0)$. Cf. the caption of Fig. 8.

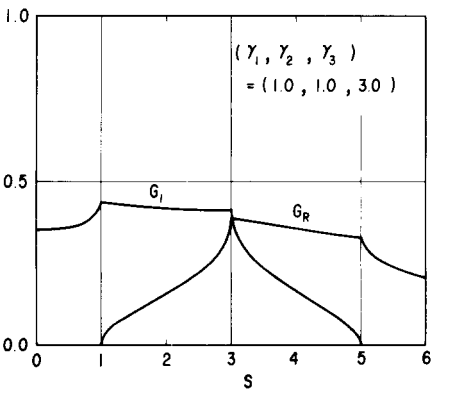

FIG. 13. $G_R(s)$ and $G_I(s)$ for the case where $(\gamma_1, \gamma_2, \gamma_3)$ (1.0, 1.0, 3.0). Cf. the caption of Fig. 8.

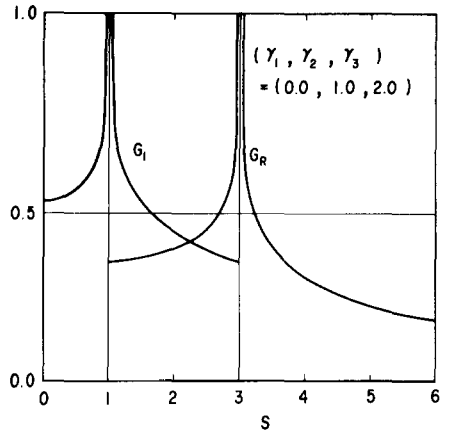

FIG. 14. $G_R(s)$ and $G_I(s)$ for the case where $(\gamma_1,\gamma_2,\gamma_3) = (0.0, 1.0, 2.0)$. Cf. the caption of Fig. 8.

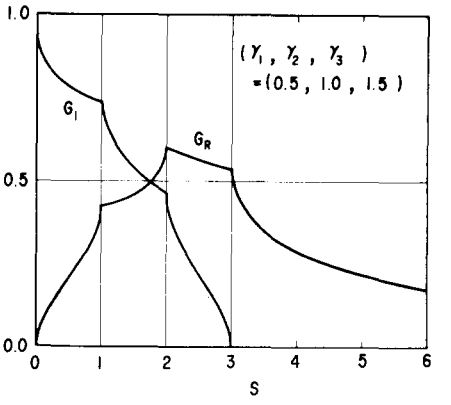

FIG. 16. $G_R(s)$ and $G_I(s)$ for the case where $(\gamma_1,\gamma_2,\gamma_3) = (0.5, 1.0, 1.5)$. Cf. the caption of Fig. 8.

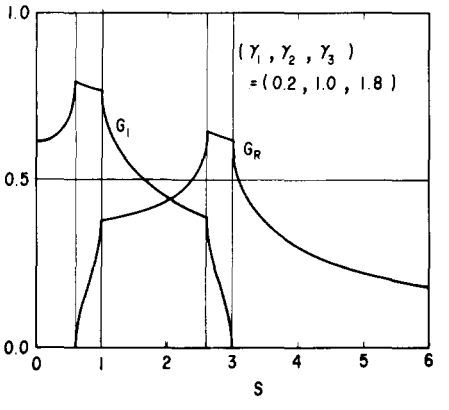

FIG. 15. $G_R(s)$ and $G_I(s)$ for the case where $(\gamma_1,\gamma_2,\gamma_3) = (0.2, 1.0, 1.8)$. Cf. the caption of Fig. 8.

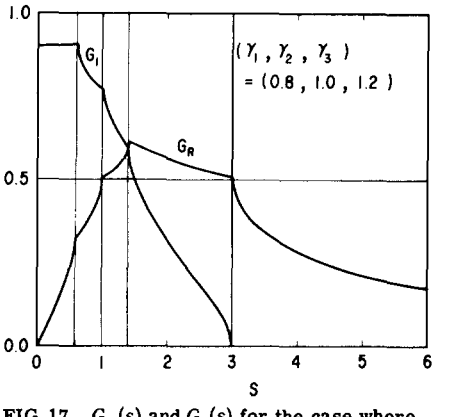

FIG. 17. $G_R(s)$ and $G_I(s)$ for the case where $(\gamma_1,\gamma_2,\gamma_3) = (0.8, 1.0, 1.2)$. Cf. the caption of Fig. 8.

(IIB), and (IIIB). In fact, we can confirm those facts by numerical calculations. However the numerical calculations of formula (I) are found to be less laborious than those for formulas (II) and (III). The curves of $G_R(s)$ and $G_I(s)$ are shown in Figs. 8–17 for a number of sets of parameters γ_1, γ_2, and γ_3.

As noticed for the cases of the simple cubic and tetragonal lattices, the expressions for the Green's function along an axis are obtained simply by multiplying the factor $\cos lx$ to the integrands of all the integrals occurring in our formulas. Here we must notice that all the formulas (I), (II), and (III) are needed for the calculation of the orthorhombic Green's function along all of the axes.

---

* Department of Physics, Faculty of Science.
† Department of Applied Physics, Faculty of Engineering.
‡ Department of Applied Science, Faculty of Engineering.
§ Present address: Department of Physics, Ohio University, Athens, Ohio 45701.

1 T. Morita and T. Horiguchi, J. Math. Phys. **12**, 981 (1971).
2 T. Morita and T. Horiguchi, J. Phys. Soc. Japan **30**, 957 (1971).
3 S. Katsura and S. Inawashiro, J. Math. Phys. **12**, (1971).
4 T. Morita and T. Horiguchi, J. Math. Phys. **12**, 986 (1971).
5 Bateman Manuscript Project, *Higher Transcendental Functions*, edited by A. Erdélyi (McGraw-Hill, New York, 1953), Vol. II, p. 317ff, and Vol. I, p. 105ff. W. Magnus and F. Oberhettinger, *Formeln und Satze für die Mathematische Physik* (Springer-Verlag, Berlin, 1948).

On Representations of the Conformal Group Which When Restricted to Its Poincaré or Weyl Subgroups Remain Irreducible

J. Mickelsson* and J. Niederle
Institute of Physics, Czechoslovak Academy of Sciences, Na Slovance 2, Prague 8
(Received 20 October 1970)

Unitary irreducible representations of the conformal group which, when restricted to its Poincaré or Weyl (= Poincaré group extended by dilatations) subgroups, remain irreducible are found. In particular it is proved that the continuous spin representations ($p_\mu \cdot p^\mu = 0$) of the Poincaré group cannot be extended to the conformal group and that, on the other hand, a known extension in the discrete spin case is unique (up to a unitary equivalence). Similar results hold for Weyl group for which, in addition, extensions also exist in the case $p_\mu \cdot p^\mu \neq 0$. Namely, each unitary irreducible representation of the Weyl group characterized by the sign of $p_\mu \cdot p^\mu$ ($\neq 0$) and by invariants of the corresponding little group can be extended to a one-parameter family of irreducible representations of the conformal group. Finally, it is shown that, besides the above mentioned extensions of unitary irreducible representations of the Weyl group, there are no others.

1. INTRODUCTION

In elementary particle physics, there has been a renewed interest in the conformal group.[1-8] The main motivation for it comes from the following features of the conformal group (which are, however, closely related to each other):

i) The conformal group is the lowest dimensional semi-simple group containing the Poincaré group. Hence, it may be considered as a possible candidate for a generalization of the Poincaré group.[2]

(ii) Since it is locally isomorphic to $SO(4, 2)$ or $SU(2, 2)$ which contains various 'physical' groups [e.g. $SO(4, 1)$, $SO(4)$, $SO(3, 1)$, $SO(2, 1)$, etc.], it serves as a spectrum generating group for hydrogenlike atoms and a successful model in hadronic spectroscopy as well.[3]

(iii) It is the largest group preserving locally the light cone and thus an exact space–time symmetry group for massless particles.[4]

(iv) The conformal group appears to have relevance also to massive particles. Based on our interpretation of five generators (in particular a dilatation) which in addition to the Poincaré generators form a Lie algebra of the conformal group, it may be treated either as an approximate symmetry group[5,9] or as an exact symmetry group.[6,10].

The last two aspects of the conformal group give rise to two problems in field theory—a rigorous definition of conformal invariance of field equations and a classification of the conformally invariant field equations. These questions were discussed in Refs. 7 and 8, where conditions under which the various equations are conformally invariant were derived. In our work we solve a related problem. We find unitary irreducible representations of the conformal group $[SU(2, 2)]$ which remain irreducible when restricted to the Poincaré group $\mathcal{P}$ (see Sec. 2) or the Weyl group $\mathcal{W}$ (Sec. 3)[11]—the largest group which maps (one-to-one) Minkowski space into itself preserving the causal order of vectors. This is necessary in order to classify all conformally invariant field equations in which the field transforms irreducibly under the Poincaré or Weyl group.

2. IRREDUCIBLE REPRESENTATIONS OF $SU(2, 2)\downarrow\mathcal{P}$

We have already mentioned that the conformal group is locally isomorphic to group $SU(2, 2)$. Various classes of irreducible representations (IR) of $SU(2, 2)$ have been studied;[12] however, there is no complete list of all. On the other hand, such a list exists for the (proper orthochronous) Poincaré group $\mathcal{P}$.[13] Thus in order to find all UIR's of $SU(2, 2)$ which when restricted to $\mathcal{P}$ are irreducible, we proceed as follows: We take a complete list of UIR's of $\mathcal{P}$ and show which of them can be extended to UIR of $SU(2, 2)$—that is we prove in which cases there exists extra operators acting on our carrier space of IR of $\mathcal{P}$ such that they can be identified with the generators of $SU(2, 2)$.

First let us remember that the Lie algebra of $SU(2, 2)$ is given by:

$$[M_{\kappa\lambda}, M_{\mu\nu}] = i(g_{\lambda\mu}M_{\kappa\nu} - g_{\kappa\mu}M_{\lambda\nu} - g_{\lambda\nu}M_{\kappa\mu} + g_{\kappa\nu}M_{\lambda\mu}), \tag{2.1}$$

$$[P_\mu, P_\nu] = 0, \tag{2.2}$$

$$[P_\lambda, M_{\mu\nu}] = i(g_{\lambda\mu}P_\nu - g_{\lambda\nu}P_\mu), \tag{2.3}$$

$$[K_\mu, K_\nu] = 0, \tag{2.4}$$

$$[K_\mu, M_{\lambda\nu}] = i(g_{\mu\lambda}K_\nu - g_{\mu\nu}K_\lambda), \tag{2.5}$$

$$[K_\mu, P_\nu] = 2i(g_{\mu\nu}D - M_{\mu\nu}), \tag{2.6}$$

$$[D, \mathbf{M}_{\mu\nu}] = 0, \tag{2.7}$$

$$[D, P_\mu] = iP_\mu, \tag{2.8}$$

$$[D, K_\mu] = -iK_\mu, \tag{2.9}$$

Here, $M_{\mu\nu}(M_{\mu\nu} = -M_{\nu\mu})$ and P_μ are generators of the Poincaré group $\mathcal{P}$, which together with a dilatation D and special conformal transformations K_μ form the Lie algebra of $SU(2, 2)$.[14] Because of the relation $e^{iaD}P_\mu \cdot P^\mu e^{-iaD} = e^{-2a}P_\mu \cdot P^\mu$ (a real), it is clear that representations which are irreducible under both $\mathcal{P}$ and $SU(2, 2)$ cannot be characterized by an eigenvalue of invariant $P_\mu \cdot P^\mu$ which is different from zero. Thus we may restrict ourselves on representations of $\mathcal{P}$ with $P_\mu \cdot P^\mu = m^2 = 0$.[15] There exist twenty classes of IR's of $\mathcal{P}$ with $p_\mu \cdot p^\mu = 0$ and $p_\mu \neq 0$.[13] They are characterized by four or five invariants as shown in Table I. Here, W_μ is the Pauli–Lubanski vector ($W_\mu = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}M^{\nu\rho}P^\sigma$), $w_\mu w^\mu$ is an invariant of the little group $E(2)$, S_E is a sign of energy ($S_E = (P_0/|P_0|)$), S_{W_0} is analogously a sign of the eigenvalue W_0 related to the second invariant of $E(2)$, and $|\lambda|$ is

TABLE I. Irreducible representations of $\mathcal{P}$ with $p_\mu \cdot p^\mu = 0, p_\mu \neq 0$.

Class	Invariants $p_\mu p^\mu, p_\mu \neq 0$	$w_\mu w^\mu$	S_E	S_{w_0}	$\lvert\lambda\rvert$	S_M	Representation of $\mathcal{P}$
1-8	0	0	±1	±1	$0, 1, 2, \cdots$, or $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \cdots$.	—	unitary
9-12	0	$j > 0, j$-real	±1	—	—	±1	unitary
13-16	0	$-j < 0, j$-real	±1	—	—	±1	non-unitary
17-20	0	α, α-complex	±1	—	—	±1	non-unitary

a fixed proportionality number, $W_\mu = \lvert\lambda\rvert P_\mu$. If $\lvert\lambda\rvert$ is an integer or a half-integer, the corresponding representation is single- or double-valued, respectively. In the case $w_\mu w^\mu \neq 0$, the IR of $\mathcal{P}$ is characterized by four invariants. It may again be single- or double-valued depending on the sign of $S_M = e^{2\pi i M_1} = e^{2\pi i M_2} = e^{2\pi i M_3}$, $M_i = \frac{1}{2}\epsilon_{ijk}M_{jk}$ —namely, if S_M is +1 or −1, respectively. IR of $\mathcal{P}$ characterized by $p_\mu p^\mu = w_\mu w^\mu = 0$ are called mass zero 'discrete spin' representations. The representations $p_\mu p^\mu = 0$ and $w_\mu w^\mu = j > 0$ are analogously called 'continuous spin' representations for the massless case.

If the mass $m^2 = p_\mu p^\mu = 0$ there is still another possibility. Namely p_μ is a null vector, i.e., $p_\mu = 0$. In this particular case IR's of $\mathcal{P}$ are characterized by two invariants of the Lorentz group $SO(3, 1)$[13]: $\mathbf{M}^2 - \mathbf{N}^2$ and $\mathbf{M}\cdot\mathbf{N}$,[16] that is by a pair of numbers (k_0, c) where $k_0 = 0, (1/2), 1, (3/2), \cdots$ and c is a complex number.[17] These representations are unitary (∞-dimensional) if either c is pure imaginary and $k_0 = 0, (1/2), 1, (3/2), \cdots$ (principal series) or $1 > c \geqq 0, c$ real and $k_0 = 0$ (supplementary series).

A. The Continuous Spin Case

A complete basis for IR of $\mathcal{P}$ characterized by $p_\mu p^\mu = 0, p_\mu \neq 0$, and $w_\mu w^\mu = j > 0$ is given by vectors[18]
$\lvert\mathbf{p}, \lambda\rangle, -\infty < p_i < \infty, \mathbf{p}^2 \neq 0, \lambda = 0, \pm 1, \pm 2, \cdots$ or

$\lambda = \pm(1/2), \pm(3/2), \cdots$, with the scalar product

$$\langle \mathbf{p}', \lambda' \lvert \mathbf{p}, \lambda\rangle = 2\lvert\mathbf{p}\rvert^3 \cdot \delta_{\lambda\lambda'} \cdot \delta^3(\mathbf{p} - \mathbf{p}').$$

Consider commutator (2.6):

$$[K_1, P_j] = -2i(\delta_{1j}D + M_{1j}). \tag{2.10}$$

From the explicit formulas for the generators given in Ref. 18, we see that M_{12} and M_{13} are diagonal in λ. Because of (2.7) there is a basis in which D and M_{1j} are diagonal in λ. Since P_μ are diagonal in λ, it follows from (2.10) that the operator $[K_1, f(\mathbf{P})]$, where f is a function of $\mathbf{P}$, is diagonal in λ. In particular, $f(\mathbf{P})$ can be taken as

$$f(\mathbf{P}) = (P_1^2 + P_2^2 + P_3^2)^{1/2} = P_0.$$

On the other hand from Eq. (2.6) follows

$$[K_1, P_0] = -2iM_{10}.$$

Since, according to Ref. 18, M_{10} connects states with λ differing by 0 and ±1 we get

$$[K_1, P_0] \neq [K_1, (P_1^2 + P_2^2 + P_3^2)^{1/2}].$$

The conclusion is that there exists no representation of $SU(2, 2)$ which, when restricted to $\mathcal{P}$, is UIR of $\mathcal{P}$ with continuous spin.

B. The Discrete Spin Case

In this case $p_\mu p^\mu = m^2 = j = 0$. Moreover, the helicity λ is fixed in UIR and equal to $0, \pm 1, \pm 2, \cdots$, or $\pm 1/2, \pm 3/2, \cdots$. Thus our basis of IR is only labeled by p_i. It is well known that these representations of $\mathcal{P}$ can be extended to the most degenerate discrete series of UIR of $SU(2, 2)$.[19] Let us show here that such an extension is unique.

Thus suppose there exist two extensions and that the corresponding five extra generators of $SU(2, 2)$ are K_μ, D and K'_μ, D', respectively. The most general action of the dilatation subgroup consistent with the commutation relation is

$$e^{iaD}\lvert\mathbf{p}\rangle = e^{iat}\lvert e^a\mathbf{p}\rangle, \tag{2.11}$$

where t is a constant which is real in UR. It then follows that the operator $D' - D$ is also a constant (times a unit operator). Let us denote

$$D' = D + t, \qquad A_\mu = K'_\mu - K_\mu$$

and let us show that necessarily $D' = D$ and $K'_\mu = K_\mu$. From expression (2.6) we have

$$[A_0, P_j] = [K'_0, P_j] - [K_0, P_j] = 0.$$

The operator A_0 is therefore diagonal in $\mathbf{p}$ and we may write

$$A_0\lvert\mathbf{p}\rangle = f(\mathbf{p})\lvert\mathbf{p}\rangle \text{ or } A_0 = f(\mathbf{P}),$$

where $f(\mathbf{P})$ is an operator valued distribution in $\mathbf{P}$. Since

$$[A_0, M_{ij}] = 0,$$

then $f(\mathbf{P}) = f(P)$, where $P = (\mathbf{P}\cdot\mathbf{P})^{1/2} = \pm P_0$. Further, from commutation relations we get

$$\langle \mathbf{p}' \lvert e^{-iaD} A_0 e^{iaD}\lvert\mathbf{p}\rangle = e^{-a}\langle \mathbf{p}' \lvert A_0\lvert\mathbf{p}\rangle = e^{-a} f(p) 2p^3\delta(\mathbf{p}' - \mathbf{p})$$

or according to (2.11)

$$\langle \mathbf{p}' \lvert e^{-iaD} A_0 e^{iaD}\lvert\mathbf{p}\rangle = \langle e^a\mathbf{p}' \lvert A_0\lvert e^a\mathbf{p}\rangle = f(e^a p)\, 2p^3\delta(\mathbf{p}' - \mathbf{p}).$$

Thus comparing rhs of the last equations we obtain

$$f(e^a p) = e^{-a} f(p)$$

so that $f(p) = b/p$, where b is some constant. From commutation relations (2.1), (2.5) we have

$$[M_{0j},[M_{0j},b/P]] = [M_{0j},[M_{0j},K'_0 - K_0]]$$
$$= -i[M_{0j},K'_j - K_j] = -K'_0 + K_0 = -b/P.$$

The lhs of this equation can be calculated by using the explicit form of M_{0j} given in Ref. 18. The result is

$$-(b/P) + (2bP_j^2/P^3).$$

Both expressions are not in contradiction only for $b = 0$, that is, if $K'_0 = K_0$. Commuting this with M_{0j}, one sees that $K'_j = K_j$, too.

Finally it follows from

$$2iD' = [K'_0, P_0] = [K_0, P_0] = 2iD$$

that $D' = D$. We conclude that a given mass-zero, discrete spin representations of $\mathcal{P}$ can be extended to IR of $SU(2,2)$ in one and only one way.

C. Null Vector Case

The fact that the UI representations of the Poincaré group with $p_\mu p^\mu = 0$ and $p_\mu = 0$ cannot be extended to representations of the conformal group is a trivial consequence of the simplicity of the latter.[20]

Finally let us remark that in the last two sections we have not used the integrability of special conformal transformations K_μ. Therefore, we have proved an even stronger result.

Theorem: The only representations of the conformal algebra which are

(i) integrable with respect to the Weyl subgroup,

(ii) unitary with respect to the Poincaré group,

(iii) and remain irreducible under the Poincaré group, are the most degenerate UIR of $SU(2,2)$ of discrete series (ladder representations) given in Ref. 12.

3. IRREDUCIBLE REPRESENTATIONS OF $SU(2,2) \downarrow \mathcal{W}$

The Lie algebra of the Weyl group $\{= T^4 \otimes [SO_0(3,1) \otimes D]\}$ is given by commutation relations (2.1)-(2.3), (2.7), and (2.8). The UIR's of $\mathcal{W}$ were studied in Ref. 21 and can be divided into five classes in accordance with orbits generated by a direct product of the universal covering of the Lorentz group with the dilatation, i.e., $SL(2,c) \otimes D$, in the group of characters of four translations T^4. We get the following five infinite-dimensional classes of UIR of $\mathcal{W}$[21]:

(i) Orbit $p_\mu = 0$. The UIR of $\mathcal{W}$ is characterized by the pair (k_0, c) related with two invariants of $SL(2,c)$ and by a real number s.

(ii) Orbit $p_\mu p^\mu > 0$ with either $p_0 > 0$ or $p_0 < 0$. The UIR of $\mathcal{W}$ is characterized by an eigenvalue of the invariant of $SU(2)$, that is by j which can be 0, (1/2), 1, (3/2), $\cdots$, and by the sign of the energy p_0.

(iii) Orbit $p_\mu p^\mu < 0$. The UIR of $\mathcal{W}$ is characterized by an invariant of $SU(1,1)$.

(iv) Orbit $p_\mu p^\mu = 0$ with either $p_0 > 0$ or $p_0 < 0$. The UIR of $\mathcal{W}$ is characterized by the sign of p_0, by the invariant j of $E(2)$, and, if $0 < j < \infty$, by the fact whether the helicity λ takes all integer or half-integer values.

(v) Orbit $p_\mu p^\mu = 0$ with either $p_0 > 0$ or $p_0 < 0$. The UIR of $\mathcal{W}$ is characterized by the sign of p_0, by invariant $j = 0$ of $E(2)$, by fixed helicity λ (an integer or half-integer number), and by the real parameter s.

The representations of classes (i) and (v) are irreducible under the universal covering of the Poincaré subgroup $\mathcal{P}$. All possible extensions from the *IR*'s of $\mathcal{P}$ to $SU(2,2)$ are already classified in Sec. 2, so that we are left with classes (ii), (iii) and (iv).

Exactly in the same way as in Sec. 2.1, it is possible to show that the representations of class (iv) cannot be extended to IR of $SU(2,2)$. Therefore, we only have to consider classes (ii) and (iii).

For each representation $\mathfrak{D}(G)$ of the little group G[$G = SU(2)$ or $SU(1,1)$; in the former case it is also necessary to specify the sign of p_0] we realize the representation $\mathfrak{U}(\mathfrak{D}(G))$ of $SU(2,2)$ induced by $\mathfrak{D}(G)$ as follows:

Consider the Hilbert space $\mathfrak{K}$ of square-integrable functions $\varphi, \psi, \cdots$ of a 4-momentum p_μ and a discrete variable λ with the scalar product

$$\langle \varphi | \psi \rangle = \sum_\lambda \int_\Delta \overline{\varphi(p_\mu, \lambda)}\, \psi(p_\mu, \lambda)\, d^4p. \tag{3.1}$$

Here: (a) If $G = SU(2)$, λ takes the values $-j, -j+1, \cdots, +j$ (j is the spin) and Δ consists of points for which $p_\mu p^\mu > 0$ and either $p_0 > 0$ or $p_0 < 0$, depending on the representation considered. (b) If $G = SU(1,1)$, Δ consists of points for which $p_\mu p^\mu < 0$ and λ takes the values $0, \pm 1, \pm 2, \cdots$, in the case of continuous series of representations of $SU(1,1)$ or $\lambda = m_0, m_0 + 1, \cdots$, or $-m_0, -m_0 - 1, \cdots$, in the case of discrete series. m_0 is a positive integer or half-integer.

In this space the Lie algebra of $\mathcal{W}$ is realized by the operators[22]

$$P_\mu = p_\mu, \tag{3.2}$$

$$D = i(p_\mu \partial^\mu + 2), \tag{3.3}$$

$$\begin{aligned} M_{12} &= -i(\mathbf{p} \times \nabla)_3 + R, \\ M_{23} &= -i(\mathbf{p} \times \nabla)_1 + (p_1/p + p_3)R, \\ M_{31} &= -i(\mathbf{p} \times \nabla)_2 + (p_2/p + p_3)R, \end{aligned} \tag{3.4}$$

$$\begin{aligned} M_{01} &= -i(p_0\partial_1 + p_1\partial_0) - \frac{p_0 p_2}{p(p+p_3)} R \\ &\quad - \frac{ms_E}{p^2}\left[\frac{p_1(p\cdot T)}{p+p_3} - pT_1\right], \\ M_{02} &= -i(p_0\partial_2 + p_2\partial_0) + \frac{p_0 p_1}{p(p+p_3)} R \\ &\quad - \frac{ms_E}{p^2}\left[\frac{p_2(p\cdot T)}{p+p_3} - pT_2\right], \\ M_{03} &= -i(p_0\partial_3 + p_3\partial_0) - \frac{ms_E}{p^2}(p\cdot T), \end{aligned} \tag{3.5}$$

where we have used the notations:

$$s_E = \text{sign}\,(p_0/|p_0|),$$
$$m = +\,(|p_\mu p^\mu|)^{1/2},$$
$$p = +\,(p_1^2 + p_2^2 + p_3^2)^{1/2},$$
$$(p\cdot T) = p_1 T_1 + p_2 T_2.$$

$\{T_1, T_2, R\}$ is a representation of the Lie algebra of G with the following commutation relations:

$$[T_1, R] = -iT_2,$$
$$[T_2, R] = iT_1, \tag{3.6}$$
$$[T_1, T_2] = i\delta R,\ \delta \equiv \text{sgn}(m^2/|m^2|).$$

We can now make an extension to the conformal algebra by setting

$$K_\mu = k_\mu + 2i(L_{\mu\nu} - M_{\mu\nu})\,(\partial^\nu + ib\,\frac{p^\nu}{m^2}), \tag{3.7}$$

where

$$L_{\mu\nu} = i(p_\mu g_{\nu\lambda}\partial^\lambda - p_\nu g_{\mu\lambda}\partial^\lambda), \tag{3.8}$$
$$k_\mu = p_\mu g_{\kappa\lambda}\partial^\kappa\partial^\lambda - 2g_{\mu\kappa}(p_\lambda\partial^\lambda + 2)\partial^\kappa - \frac{b^2 p_\mu}{m^2}. \tag{3.8'}$$

and b is an arbitrary constant. If we require the generators K_μ to be self adjoint, then b has to be real. In the spinless case ($L_{\mu\nu} = M_{\mu\nu}$), it is enough that b^2 is real.[23]

Let us remark that the commutation relations for the generators given above can be checked by a direct calculation. This is done easier in the 'spinor-basis' defined in Ref. 24 where the generators take a particularly simple form:

$$M_{\mu\nu} = L_{\mu\nu} + S_{\mu\nu},$$
$$K_\mu = k_\mu - 2iS_{\mu\nu}(\partial^\nu + ib(p^\nu/m^2),$$
$$P_\mu = p_\mu,$$
$$D = i(p_\mu\partial^\mu + 2).$$

Here,

$$(S_{01}, S_{02}, S_{03}) = i(S_{23}, S_{31}, S_{12}) = i\mathbf{S},$$
$$\mathbf{S} = \begin{cases}(T_1, T_2, R) \text{ if } p_\mu p^\mu > 0,\\ (iT_1, iT_2, R) \text{ if } p_\mu p^\mu < 0.\end{cases}$$

Next we shall prove the following theorem.

Theorem: Every representation of the conformal algebra acting in Hilbert space $\mathcal{H}$ defined in (3.1) which is integrable with respect to the subgroup $\mathcal{W}$, the restriction of which with respect to $\mathcal{W}$ being a unitary irreducible representation of the class (ii) or (iii), is equivalent to one representation of the family given in (3.1)–(3.7).

Proof: Let $\{M_{\mu\nu}, P_\mu, D\}$ span an integrable representation of the Lie algebra of $\mathcal{W}$ of class (ii) or (iii) acting on Hilbert space $\mathcal{H}$. Suppose that we have two extensions to the conformal algebra with additional generators K_μ and K'_μ, respectively.

Let us put

$$0_\mu = K'_\mu - K_\mu. \tag{3.9}$$

It is comfortable to work with the momentum–helicity basis

$$P_\mu |p_\mu, j\lambda> = p_\mu |p_\mu, j\lambda>,$$
$$\mathbf{P}\cdot\mathbf{M}|p_\mu, j\lambda> = p\lambda|p_\mu, j\lambda>,$$

where j labels the representations of the little group G.

We have the commutation relations

$$[P_\mu, 0_\nu] = [P_\mu, K'_\nu - K_\nu]$$
$$= 2i(M_{\nu\mu} - g_{\mu\nu}D) - 2i(M_{\nu\mu} - g_{\mu\nu}D) = 0,$$
$$[\mathbf{P}\cdot\mathbf{M}, 0_0] = \mathbf{P}\cdot[\mathbf{M}, 0_0] = 0.$$

Therefore, we can write

$$0_0|p_\mu, j\lambda> = f(p_\mu, \lambda)|p_\mu, j\lambda>. \tag{3.10}$$

Let us remember that Pauli–Lubanski 4-vector W_μ is defined by

$$W_\mu = \tfrac{1}{2}\,\epsilon_{\mu\nu\rho\sigma}M^{\nu\rho}P^\sigma \text{ or}$$
$$W_0 = \mathbf{P}\cdot\mathbf{M}, \quad \mathbf{W} = P_0\cdot\mathbf{M} - (\mathbf{P}\times\mathbf{N}), \tag{3.11}$$

where

$$\mathbf{M} = (M_{23}, M_{31}, M_{12}), \qquad \mathbf{N} = (M_{01}, M_{02}, M_{03}).$$

In the helicity basis we have

$$W_0|p_\mu, j\lambda> = \mathbf{P}\cdot\mathbf{M}|p_\mu, j\lambda> = p\lambda|p_\mu, j\lambda>$$

We can now rewrite Eq. (3.10) in operator notation:

$$0_0 = f(P_\mu, W_0).$$

Since 0_0 commutes with rotations, we can write

$$0_0 = f(M^2, P_0, W_0).$$

Now we use the fact that 0_μ transforms as a 4-vector under homogeneous Lorentz transformations. The most general 4-vector whose 0 component depends only on P_0 and W_0 is given by

$$0_\mu = f_1(M^2)P_\mu + f_2(M^2)W_\mu.^{25} \tag{3.12}$$

From the commutation relation (2.8) follows that

$$e^{iaD}P_\mu e^{-iaD} = e^{-a}P_\mu.$$

Therefore

$$e^{iaD}(P_\mu)^n e^{-iaD} = e^{-na}(P_\mu)^n \tag{3.13}$$

and taking into account (3.11)

$$e^{iaD}W_\mu e^{-iaD} = e^{-a}W_\mu. \tag{3.14}$$

From (2.9) it follows that

$$e^{iaD}0_\mu e^{-iaD} = e^a 0_\mu. \tag{3.15}$$

From this formula together with expression (3.13), it follows that 0_μ is a homogeneous function of degree -1 in the momenta P_μ, so that

$$0_\mu = \alpha\frac{P_\mu}{M^2} + \beta\frac{W_\mu}{M^2}. \tag{3.16}$$

Let us go to the spinor basis. Substituting expression (3.8) into $M_{\mu\nu} = L_{\mu\nu} + S_{\mu\nu}$ and putting the obtained expression for $M_{\mu\nu}$ into (3.11), we get

$$W_0 = \mathbf{p}\cdot\mathbf{S}, \quad \mathbf{W} = p_0\mathbf{S} - i(\mathbf{p}\times\mathbf{S}). \tag{3.17}$$

Keeping in mind that $S_{0j} = is_j$, we obtain

$$W_\mu = iS_{\mu\nu}p^\nu. \tag{3.18}$$

Suppose now that K_μ is given by expression (3.7) with $b = 0$. After a tedious calculation we receive

$$\begin{aligned}[K'_0, K'_1] &= [K_0 + 0_0, K_1 + 0_1] \\ &= \left[K_0 + \alpha\frac{p_0}{m^2} + i\beta S_{0\nu}\frac{p^\nu}{m^2},\right. \\ &\qquad \left. K_1 + \alpha\frac{p_1}{m^2} + i\beta S_{1\nu}\frac{p^\nu}{m^2}\right] \\ &= \frac{4i}{m^4}\left(\frac{\beta^2}{4} - \alpha\right)\{p_0p_2S_{21} + p_0p_3S_{31} + S_{02}p_1p_2 \\ &\qquad + S_{03}p_1p_3 + S_{01}(-p_2^2 - p_3^2)\}.\end{aligned} \tag{3.19}$$

This commutator is zero only if $S_{\mu\nu} = 0$ or $\beta^2/4 - \alpha = 0$. In both cases K'_μ belongs to the family given in (3.7). QED

ACKNOWLEDGMENTS

One of the authors (J.M.) would like to thank Dr. Pekárek for his hospitality at the Institute of Physics of the Czechoslovak Academy of Sciences in Prague and to express his gratitude to the Třinec Steel Scientific Foundation as well for a financial support. It is also a pleasure to thank Professor M. Flato and Dr. D. Sternheimer for stimulating discussions from which the problem has arisen.

---

* On leave of absence from the Research Institute for Theoretical Physics, Helsinki, Finland.

1 F. Gürsey, Nuovo Cimento **3**, 988 (1956); T. Fulton, F. Rohrlich, and L. Witten, Rev. Mod. Phys. **34**, 442 (1962); H. A. Kastrup, Ann. Phys. **7**, 388 (1962); L. Castell, Nucl. Phys. **4B**, 343 (1967); Abdus Salam and J. Strathdee, Phys. Rev. **184**, 1760 (1969); C. J. Isham, Abdus Salam and J. Strathdee, Phys. Letters **31B**, 300 (1970).

2 E. L. Hill, Phys. Rev. **84**, 1165 (1951); L. Castell, Nuovo Cimento **46A**, 1 (1966); M. Flato and D. Sternheimer, Compt. Rend. Acad. Sci. (Paris) **263A**, 935 (1966); H. A. Kastrup, Phys. Rev. **150**, 1183 (1966).

3 D. Bohm, M. Flato, D. Sternheimer, and J.-P. Vigier, Nuovo Cimento **38**, 1941 (1965); M. Flato, D. Sternheimer, and J.-P. Vigier, Compt. Rend. Acad. Sci. (Paris) **260**, 3869 (1965); I. A. Malkin and V. I. Manko, J. Nucl. Phys. (USSR) **3**, 372 (1966); Y. Nambu, Progr. Theoret. Phys. (Kyoto) Suppl. **37, 38**, 368 (1966); Phys. Rev. **160**, 1171 (1967); A. O. Barut and H. Kleinert, *ibid.* **156**, 1541 (1967); **161**, 1464 (1967); C. Fronsdal, *ibid.* **171**, 1811 (1968); H. Kleinert, *ibid.* **168**, 1827 (1968); R. W. Huff, Preprint of Univ. of California, Los Angeles, 1969.

4 E. Cunnigham, Proc. London Math. Soc. **8**, 77 (1910); H. Bateman, *ibid.* 223 (1910); P. A. M. Dirac, Ann. Math. **37**, 429 (1936); J. A. McLennan, Nuovo Cimento **10**, 1360 (1956); M. A. Buchdahl, *ibid.* **11**, 496 (1959); J. Wess, *ibid.* **18**, 1086 (1960); L. Gross, J. Math. Phys. **5**, 687 (1964); R. Penrose, Proc. Roy. Soc. (London) **284A**, 159 (1965).

5 H. A. Kastrup, Phys. Letters **3**, 78 (1962); Phys. Rev. **150**, 1183 (1966); N. F. Bali, D. D. Coon, and A. Katz, J. Math. Phys. **10**, 1939 (1969); D. D. Coon, *ibid.* **10**, 1314 (1969); M. Flato, P. Hillion, and J. Simon, Compt. Rend. Acad. Sci. (Paris) **268A**, 1065 (1969); L. Castell, Preprint of the Max-Planck-Institut für Physik und Astrophysik, 19/3/1970 München, 1970.

6 A. O. Barut, Phys. Rev. Letters **20**, 893 (1968); M. Flato and J. Simon, Compt. Rend. Acad. Sci. (Paris) **269A**, 1051 (1969).

7 M. Flato, J. Simon, and D. Sternheimer, Ann. Phys. (N.Y.) **60**, (1970) (to be published).

8 D. J. Gross and J. Wess, Preprint of the CERN, TH 1076, Genève, 1969; G. Mack and Abdus Salam, Ann. Phys. (N.Y.) **53**, 174 (1969).

9 An active interpretation—the approximate symmetry improves as the energies of involved massive particles increase (so that the rest masses are negligible in comparison).

10 A passive interpretation—dilatations are interpreted as changing the unit scale in a space–time dependent manner.

11 The Weyl group is a semidirect product of the positive dilatations with the (proper orthochronous) Poincaré group: $\mathfrak{W} = T^4 \otimes (SO_0(3,1) \otimes D)$

12 Y. Murai, Progr. Theoret. Phys. (Kyoto) **8**, 147 (1953); A. Esteve and P. G. Sona, Nuovo Cimento **32**, 473 (1964); I. M. Gel'fand and M. I. Graev, Izv. Akad. Nauk **29**, 1329 (1965); N. Limić, J. Niederle, and R. Raczka, J. Math. Phys. **7**, 1861; 2026 (1966); **8**, 1079 (1967); A. Kihlberg, W. F. Müller, and F. Halbwachs, Commun. Math. Phys. **3**, 194 (1966); J. Niederle, J. Math. Phys. **8**, 1921 (1967); N. Limič and J. Niederle, Ann. Inst. Henri Poincaré **9**, 327 (1968); Tsu Yao, J. Math. Phys. **8**, 1931 (1967); **9**, 1615 (1968).

13 Ju. M. Shirokov, Zh. Eksp. Teor. Fiz. **33**, 1196; 1208 (1957) [Sov. Phys.—JETP **6**, 919, 929 (1958)]; E. P. Wigner, Ann. Math. **40**, 149 (1939); V. Bergmann and E. P. Wigner, Proc. Math. Acad. Sci. (U.S.) **34**, 211 (1948).

14 We use metric $g_{00} = -g_{ii} = 1, i = 1, 2, 3, g_{\mu\nu} = 0$ for $\mu \neq \nu$. Latin indices run from 1 to 3 and Greek ones from 1 to 4.

15 We shall denote an operator and its representing operator in some representation by the same capital letter. The corresponding eigenvalues will be denoted by small letters.

16 $N_i = M_{0i}, i = 1, 2, 3$.

17 $\mathbf{M}^2 - \mathbf{N}^2 = k_0^2 + c^2 - 1$ and $\mathbf{M}\cdot\mathbf{N} = ik_0c$.

18 J. S. Lomont and H. E. Moses, J. Math. Phys. **3**, 405 (1962).

19 L. Gross, J. Math. Phys. **5**, 687 (1964); D. Sternheimer, J. Math. Pures Appl. **47**, 289 (1968); G. Mack and I. Todorov, J. Math. Phys. **10**, 2078 (1969).

20 No (nontrivial) representation of a simple group can be trivial on its continuous subgroup.

21 U. Ottoson, Ark. Fys. **33**, 523 (1967).

22 The $M_{\mu\nu}$'s are taken from the work of H. E. Moses, Ann. Phys. (N.Y.) **52**, 444 (1969) with some minor changes.

23 The spinless case was discussed earlier by L. Castell, Nucl. Phys. **4B**, 343 (1967).

24 H. Joos, Fortschr. Physik **10**, 65 (1962).

25 $W_\mu W^\mu$ and $P_\mu W^\mu$ are also scalars but $P_\mu W^\mu \equiv 0$ and $W_\mu W^\mu = j(j+1)M^2$, j being fixed in a UIR of the Weyl group.

On an Application of the Theory of Invariants

V. Syamala Devi
Department of Applied Mathematics, Andhra University, Waltair, India
(Received 14 September 1970)

The problem of finding highest-weight polynomials in certain chains of subgroups of the unitary group is shown to be related to finding semi-invariants of certain ground forms according to the theory of invariants developed by mathematicians long ago.

INTRODUCTION

The Lie algebras of the full linear group in n-dimensional space, as well as those of many of its Lie subgroups, can be realized as differential operators in certain indeterminates, and bases for integral representations of the linear group can be realized as polynomials in these indeterminates. Given any irreducible integral representation (IR) of the unitary group $U(n)$ in n dimensions, the problem of dividing the carrier space of the IR into subspaces, each of which is irreducibly invariant under a given subgroup of $U(n)$, is of much significance in problems of nuclear physics. It happens that this problem amounts to obtaining the semi-invariants of certain ground forms in multiple fields. The purpose of this article is to point out this connection and to work out one such problem of physical interest.

In Sec. 1, it is shown how the double binary (2, 1) form is related to one such problem of physical interest discussed recently by different methods.

In Sec. 2, the highest-weight vectors of a subgroup $R(3)$ belonging to the carrier space of an IR $(n, 1)$ of $U(5)$ are obtained and their relation to the complete system of covariants of the double binary (4, 1) form (which was discussed by Todd[1] in 1946) is shown.

In Sec. 3, we consider the binary n-ary $(r, 1)$ form subjecting each set of variables to the group of step transformations in the two-dimensional and n-dimensional space, respectively. Making use of both the Littlewood techniques[2] and the classical symbolic techniques, we obtain a complete irreducible system for the highest-weight vectors of the subgroup $R(3)$ belonging to the carrier space of the IR $(n, 1, 1)$ of $U(5)$, by considering the semi-invariants of a binary 5-ary (4, 1) form.

1. RELATION OF THE BASES IN THE CHAINS $U(3) \supset R(3)$ TO THE DOUBLE BINARY (2, 1) FORM

Consider the (2, 1) form

$$(a_{11}x^2 - 2a_{12}xy - 2a_{31}y^2)x' + (a_{12}x^2 - 2a_{22}xy - 2a_{32}y^2)y'. \quad (1.1)$$

A covariant of this form is a polynomial in the coefficients a_{ij} and the variables x, y, x', y', which is invariant with respect to any homogeneous linear transformation of the pair x, y and any homogeneous linear transformation of the pair x', y', independent of one another. A covariant is uniquely determined by its leading coefficient called the semi-invariant which is a polynomial in the a_{ij}.

A polynomial in a_{ij} is a semi-invariant if and only if it is annihilated by the operators[3]

$$C_{12} = a_{11}\frac{\partial}{\partial a_{12}} + a_{21}\frac{\partial}{\partial a_{22}} + a_{31}\frac{\partial}{\partial a_{32}},$$

$$L_+ = a_{21}\frac{\partial}{\partial a_{31}} + a_{22}\frac{\partial}{\partial a_{32}} - a_{11}\frac{\partial}{\partial a_{21}} - a_{12}\frac{\partial}{\partial a_{22}}.$$

A covariant may be supposed to be homogeneous both in the variables and the coefficients separately. The weight of a coefficient a_{ij} is defined to be $(i-1, j-1)$. The weight of a product of coefficients a_{ij} is the sum of the weights of its factors, the sum $(w_1, w_1') + (w_2, w_2')$ being defined as $(w_1 + w_2, w_1' + w_2')$. It would also follow that a semi-invariant is homogeneous and also isobaric. If a semi-invariant is of degree d and of weight (w, w'), then it is the leader of a covariant of orders n and n' respectively in the two pairs of variables, where $n = pd - 2w$ and $n' = p'd - 2w'$, p and p' being the orders of the ground form in the two pairs of variables. If S is a semi-invariant, then

$$[C_{12}, C_{21}]S = (C_{12}C_{21} - C_{21}C_{12})S = (C_{11} - C_{22})S = n'S \quad (1.2)$$

and

$$[L_+, L_-]S = (L_+L_- - L_-L_+)S = LS = (n/2)S, \quad (1.3)$$

where

$$C_{ss'} = \sum_{r=1}^{3} a_{rs}\frac{\partial}{\partial a_{rs'}}, \qquad C_r^{r'} = \sum_{s=1}^{2} a_{rs}\frac{\partial}{\partial a_{r's}} \quad (1.4)$$

and

$$L_+ = C_2^3 - C_1^2, \qquad L_- = C_3^2 - C_2^1, \qquad L = C_1^1 - C_3^3. \quad (1.5)$$

The covariants of (1.1) have a finite subsystem so that every covariant can be written as a polynomial in the elements of the subsystem. The complete irreducible system for the form (1.1) was first obtained in the last century. But the same complete system has been obtained recently[4] unconscious of the relation of the system to the problem of covariants, in a discussion of the fractional parentage coefficients in Nuclear Physics.

Consider the linear forms

$$a_{11}x' + a_{12}y', \qquad a_{21}x' + a_{22}y', \qquad a_{31}x' + a_{32}y'. \quad (1.6)$$

A polynomial in the coefficients a_{ij} is a simultaneous semi-invariant of these forms if and only if it is annihilated by the operator C_{12}. By Peano's theorem[5] the simultaneous covariants of the three forms can be obtained from those of the first two forms by Aronhold's process and the covariants of the first two forms other than the determinant $a_{11}a_{22} - a_{12}a_{21}$ can in their turn be obtained from those of the first form alone by Aronhold's process.

The operators $C_r^{r'}$, $r, r' = 1, 2, 3$, are then found to be nothing but the Aronhold operators of this

problem. Peano's theorem then implies in our present simple problem, that all the simultaneous semi-invariants of the forms (1.6) can be obtained by applying polynomials in $C_r^{r'}$ on the semi-invariants $a_{11}^{p}(a_{11}a_{22} - a_{12}a_{21})^{q}$. Since an Aronhold operator changes a semi-invariant into a semi-invariant and a covariant into a covariant and does not change the degree or order of a covariant, it follows that all semi-invariants which are of degree h_1 in the symbol 1 as second suffix and of degree h_2 in 2 as second suffix and hence are leading coefficients of covariants of degree $h_1 + h_2$ and weight h_2 are obtained by operating on $a_{11}^{h_1-h_2}(a_{11}a_{22} - a_{12}a_{21})^{h_2}$ by polynomials in $C_r^{r'}$. This shows that all the semi-invariants of degree $h_1 + h_2$ and weight h_2 span a linear space which is irreducibly invariant under the Lie algebra $\mathfrak{L}$ generated by the operators $C_r^{r'}$. Therefore these semi-invariants form a basis for the IR corresponding to the partition (h_1, h_2) of the group $U(3)$ whose Lie algebra $\mathfrak{L}$ is.

Now the operators L_+, L_-, and L span Lie algebra R which is isomorphic to that of $R(3)$, the rotation group in three dimensions and is a subalgebra of $\mathfrak{L}$. Since a polynomial in a_{ij} is a semi-invariant of the form (1.1) if and only if it is annihilated by C_{12} and L_+, it follows that each semi-invariant of (1.1) which is of degree $h_1 + h_2$ and weight h_2 is the highest-weight polynomial with respect to R of a certain subspace invariant under R, of the carrier space of the IR (h_1, h_2) of $U(3)$. The term "weight" in the expression "highest-weight polynomial" (hwp) is to be understood in a way quite different from the weight of a semi-invariant; weight here means the eigenvalue of the operator L corresponding to which the polynomial is the eigenvector of L.

These highest-weight vectors were obtained in Ref. 4 making use of a theorem of Littlewood[6] to determine which IR's of $R(3)$ occur when an IR of $U(3)$ restricted to the $R(3)$ subgroup, is reduced with respect to $R(3)$. The problem of obtaining the highest-weight vectors of these component IR's and the problem of finding which IR's of $R(3)$ occur as components when the restriction to $R(3)$ of a given IR of $U(3)$ is decomposed (branching problem) are dependent on one another. If the branching problem is solved, then the hwp can be obtained at once and if the hwp are known, then the branching problem is solved.

2. HWP IN THE CHAINS $U(5) \supset R(3)$ FROM THE DOUBLE BINARY (4, 1) FORM

The hwp in the $U(3) \supset R(3)$ chain (i.e., the hwp of the various IR's of $R(3)$ which occur when the $R(3)$ part of a given IR of $U(3)$ is decomposed into irreducible parts) are required for the evaluation of the fractional parentage coefficients in the configuration space of a single orbital in the p shell.[4] When the next shells are considered, one requires similar hwp in the $U(2j+1) \supset R(3)$ chains for $j = 1, 2, \cdots$. It will be found that the hwp in the $U(2j+1) \supset R(3)$ chain for IR's (h_1, h_2) of $U(2j+1)$ are identical with a set of semi-invariants of the $(2j, 1)$ double binary form

$$\sum_{i=1}^{2}\left(\sum_{v=-j}^{+j}(\underline{|j+v}\ \ \underline{|j-v+1})^{1/2}a_{vi}\ x^{j+v}y^{j-v}\right)x_i^1$$

The hwp in the chain $U(5) \supset R(3)$ will be found explicity only for the IR's $(h_1, 1)$ to save extensive numerical work, the procedure being clearly the same for any IR (h_1, h_2). For $j = 2$, the following operators (having the same commutation relations as the operators (1.5))

$$L_0 = 2(C_1^1 - C_5^5) + C_2^2 - C_4^4,$$

$$L_1 = 2(C_1^2 + C_4^5) + 3(C_3^4 + C_2^3),$$

$$L_{-1} = 2(C_2^1 + C_5^4) + 3(C_4^3 + C_3^2)$$

span a Lie algebra $\mathfrak{B}$ which is a subalgebra of the Lie algebra $\mathfrak{A}$ spanned by $C_j^i, i, j = 1, \ldots, 5$ of the unitary group $U(5)$. The hwp in the chain $U(5) \supset R(3)$ are precisely the semi-invariants of the (4, 1) form

$$\begin{aligned} f = {} & (\sqrt{2}a_{11}x^4 + 4a_{21}x^3y + 4\sqrt{3}a_{31}x^2y^2 \\ & + 8a_{41}xy^3 + 4\sqrt{2}a_{51}y^4)x' \\ & + (\sqrt{2}a_{12}x^4 + 4a_{22}x^3y \\ & + 4\sqrt{3}a_{32}x^2y^2 + 8a_{42}xy^3 + 4\sqrt{2}a_{52}y^4)y'. \quad (2.1) \end{aligned}$$

A polynomial in a_{ij} is a semi-invariant of the form (2.1) if and only if it is annihilated by L_1 and the operator

$$C_{12} = \sum_{r=1}^{5} a_{r1}\frac{\partial}{\partial a_{r2}}. \quad (2.2)$$

Similarly, a polynomial in a_{ij} is a semi-invariant of the linear forms

$$a_{i1}x' + a_{i2}y', \quad i = 1, \ldots, 5,$$

if and only if it is annihilated by the operator (2.2) Further, all the semi-invariants of degree $h_1 + h_2$ and weight h_2 of these forms are obtained by operating on

$$(a_{11})^{h_1-h_2}(a_{11}a_{22} - a_{12}a_{21})^{h_2}$$

by polynomials in the Aronhold operators C_j^i. Hence these semi-invariants span a space V which is irreducibly invariant under $\mathfrak{A}$ and every semi-invariant of (2.2) of degree $h_1 + h_2$ and having second component of weight equal to h_2 is the polynomial of highest-weight with respect to $\mathfrak{B}$ of a certain subspace V' of V, where V' is irreducibly invariant under $\mathfrak{B}$.

A complete system of covariants for the form (2.1) was obtained long back[1] but they will be obtained here using Littlewood's S function techniques because the chain $U(5) \supset R(3)$ is of much physical interest and would serve as an example of problems of similar kind which are in fact numerous. The Littlewood technique also provides one with the branching rules even before the semi-invariants are obtained and the determination of the branching rules for such chains has been a very difficult (almost impossible) task so far.

When the pairs of variables x, y and x', y' are subjected to independent linear transformations

$$\begin{pmatrix} x \\ y \end{pmatrix} = A\begin{pmatrix} X \\ Y \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = B\begin{pmatrix} X' \\ Y' \end{pmatrix}.$$

The products $x^4, x^3y, x^2y^2, xy^3, y^4$ are transformed by the induced matrix[2] $(A)^{\{4\}}$, and the row vector of the coefficients a_{ij} of the form (2.1) is transformed by the direct product of matrices $A^{\{4\}} \times B$. The homogeneous products of degree q of the coefficients are transformed by the qth induced matrix $(A^{\{4\}} \times B)^{\{q\}}$. Denoting the direct sum by $\sum'$, if

$$(A^{\{4\}} \times B)^{\{q\}} = \sum{}' A^{\{\lambda\}} \times B^{\{\mu\}}, \tag{2.3}$$

then to each term on the right-hand side corresponds a set of linear combinations of products of degree q in the coefficients which transform like the direct product of two tensors of types $\{\lambda\}$ and $\{\mu\}$. Writing the spur of $A^{\{4\}} \times B$ as $(\{4\}, \{1\})$, we have from (Ref. 2) $(\{4\}, \{1\}) \otimes \{n\} = \sum(\{4\} \otimes \{n_1, n_2\}, \{n_1, n_2\})$, where $\otimes$ is the Littlewood's "new multiplication" and the summation is over all partitions (n_1, n_2) of n into two parts because the matrices A and B transforming the variables of (2.1) are only 2×2 matrices.

It is not difficult to obtain the following expansions (some of these were obtained by Murnaghan[7]):

$$\begin{aligned}
&(\{4\}, \{1\}) \otimes \{1\} = (\{4\}, \{1\});\\
&(\{4\}, \{1\}) \otimes \{2\} = (\{8\}, \{6, 2\} + \{4, 4\}, \{2\}) + (\{7, 1\} + \{5, 3\}, \{1, 1\});\\
&(\{4\}, \{1\}) \otimes \{3\} = (\{12\} + \{10, 2\} + \{9, 3\} + \{8, 4\} + \{6, 6\}, \{3\}) + (\{11, 1\} + \{10, 2\} + \{9, 3\} + 2\{8, 4\} + \{7, 5\}, \{2, 1\});\\
&(\{4\}, \{1\}) \otimes \{4\} = (\{16\} + \{14, 2\} + \{13, 3\} + 2\{12, 4\} + 2\{10, 6\} + \{8, 8\}, \{4\})\\
&\qquad + (\{15, 1\} + \{14, 2\} + 2\{13, 3\} + 2\{12, 4\} + 3\{11, 5\} + 2\{10, 6\} + 2\{9, 7\}, \{3, 1\})\\
&\qquad + (\{4\} \otimes \{2, 2\}, \{2, 2\});\\
&(\{4\}, \{1\}) \otimes \{5\} = (\{4\} \otimes \{5\}, \{5\}) + (\{19, 1\} + \{18, 2\} + 2\{17, 3\} + 3\{16, 4\} + 3\{15, 5\} + 4\{14, 6\}\\
&\qquad + 4\{13, 7\} + 3\{12, 8\} + 2\{11, 9\} + \{10, 10\}, \{4, 1\}), + (\{4\} \otimes \{3, 2\}, \{3, 2\});\\
&(\{4\}, \{1\}) \otimes \{6\} = (\{4\} \otimes \{6\}, \{6\}) + (\{23, 1\} + \{22, 2\} + 2\{21, 3\} + 3\{20, 4\}\\
&\qquad + 4\{19, 5\} + 4\{18, 6\} + 6\{17, 7\} + 5\{16, 8\} + 5\{15, 9\} + 4\{14, 10\} + 3\{13, 11\}, \{5, 1\}\\
&\qquad + (\{4\} \otimes \{4, 2\}, \{4, 2\}) + (\{4\} \otimes \{3, 3\}, \{3, 3\}).
\end{aligned} \tag{2.4}$$

A covariant which is of type $(\{\lambda_1, \lambda_2\}, \{n_1, n_2\})$ has the leading term which is just the hwp of an IR of $\mathfrak{B}$ occurring in the reduction of the restriction to $\mathfrak{B}$ of the IR (n_1, n_2) of $\mathfrak{A}$. Hence in (2.4) each symbol $(\{4\} \otimes \{n_1, n_2\}, \{n_1, n_2\})$ may be interpreted as standing for the expression "the IR's of $\mathfrak{B}$ contained in the decomposition of the restriction to $\mathfrak{B}$ of the IR $\{n_1, n_2\}$ of $\mathfrak{A}$" and the right hand side gives the IRs.

The covariants of types $(\{\lambda_1, \lambda_2\}, \{n, 0\})$ of the form (2.1) are simply the covariants of the binary quartic for which the complete irreducible system is well known. So we will determine the complete system for covariants of types $(\{\lambda_1, \lambda_2\}, \{n, 1\})$, it being necessary to calculate expansions $\{4\} \otimes \{n_1, n_2\}$ as in (2.4) for some more values of n_1 and n_2 to find a complete system for covariants of all types.

A covariant of type $(\{\lambda_1, \lambda_2\}, \{n_1, n_2\})$ is of orders $(\lambda_1 - \lambda_2)$ and $(n_1 - n_2)$ in the two pairs of variables respectively, and is of degree $(n_1 + n_2)$. It will be convenient to denote a covariant of type $(\{\lambda_1, \lambda_2\}, \{n_1, n_2\})$ as $\binom{n_1, n_2}{\lambda_1 - \lambda_2}$. So from the expansions (2.4), $\binom{1,0}{4}$, $\binom{2,0}{4}$, $\binom{2,0}{0}$, $\binom{3,0}{6}$, $\binom{3,0}{0}$ are found to be irreducible covariants and the expansion of $(\{4\}, \{1\}) \otimes \{4\}$ gives no new irreducible covariants of types $(\{\lambda_1, \lambda_2\}, \{n, 0\})$. It may be easily verified that this set of covariants is the same as the well-known complete irreducible system for the binary quartic. So the completeness of this set among covariants of type $(\{\lambda_1, \lambda_2\}, \{n, 0\})$ is clear.

Coming to the covariants of type $(\{\lambda_1, \lambda_2\}, \{n, 1\})$, we see that the expansions (2.4) show that the following are irreducible covariants.

$$\binom{1,1}{6}, \binom{1,1}{2}, \binom{2,1}{8}, \binom{2,1}{4}, \binom{2,1}{4}', \binom{2,1}{2}, \binom{3,1}{4}, \binom{3,1}{4}', \binom{3,1}{2}, \binom{4,1}{4}, \binom{4,1}{0}. \tag{2.5}$$

The expansion $(\{4\}, \{1\}) \otimes \{5\}$ gives no new irreducible covariants. It can be shown now that if all covariants of types $(\{\lambda_1, \lambda_2\}, \{n-1, 1\})$ for $n - 1 \geq 4$ are expressible in terms of the covariants (2.5) and

$$\binom{1,0}{4}, \binom{2,0}{4}, \binom{2,0}{0}, \binom{3,0}{6}, \binom{3,0}{0}, \tag{2.6}$$

then all covariants of types $(\{\lambda_1, \lambda_2\}, \{n, 1\})$ are expressible in terms of the covariants (2.5) and (2.6) with the help of the following theorems.

The following statements regarding a (p, q) double binary form F can be proved[8] exactly in the same way as the corresponding statements for a single binary form.

(1) If the covariants of degree $(m-1)$ are all expressible as rational integral functions of $F, \phi_1, \ldots, \phi_k$ then every covariant of degree m is a linear combination of transvectants of the form

$$(K_{m-1}, F)^{r,s}, \quad r = 1, \ldots, p, \quad s = 1, \cdots, q,$$

where K_{m-1} is a product of $(F)^{\alpha}(\phi_1)^{\alpha_1} \cdots (\phi_k)^{\alpha_k}$ and is of degree $(m-1)$.

(2) If $K_{m-1} = VW$, where V and W are products of $F, \phi_1, \ldots, \phi_k$ but of smaller degree than K_{m-1} and if the two orders of W are, respectively, not less than r and s, then the transvectant $(K_{m-1}, F)^{r,s}$ may be excluded from investigation.

The following are the semi-invariants corresponding to the covariants (2.5) and (2.6)

$$\binom{10}{4} = a_{11}, \binom{20}{2} = 2\sqrt{2}a_{11}a_{31} - \sqrt{3}\,(a_{21})^2,$$

$$\binom{20}{0} = (a_{31})^2 + 2a_{11}a_{51} - 2a_{21}a_{41}$$

$$\binom{30}{3} = \sqrt{3}a_{11}a_{21}a_{31} - \sqrt{2}(a_{11})^2a_{41} - a_{21}^3/\sqrt{2}.$$

$$\binom{30}{0} = 2\sqrt{6}a_{11}a_{31}a_{51} - 3a_{11}a_{41}^3 - 3a_{21}^2a_{51} + \sqrt{6}a_{21}a_{31}a_{41} - \sqrt{2}/\sqrt{3}a_{31}^3,$$

$$\binom{11}{6} = \Delta_{12}^{12}, \binom{11}{2} = \sqrt{3}\Delta_{23}^{12} - \sqrt{2}\Delta_{14}^{12},$$

$$\binom{21}{8} = \sqrt{2}a_{11}\Delta_{13}^{12} - \sqrt{3}a_{21}\Delta_{12}^{12},$$

$$\binom{21}{4} = a_{11}\Delta_{15}^{12} - a_{21}\Delta_{14}^{12} + a_{31}\Delta_{13}^{12} - a_{41}\Delta_{12}^{12},$$

$$\binom{21}{4}^1 = \sqrt{3}a_{21}\Delta_{23}^{12} - 2\sqrt{2}a_{31}\Delta_{13}^{12} + 3\sqrt{2}a_{41}\Delta_{12}^{12},$$

$$\begin{aligned}\binom{31}{4} = {} & a_{21}a_{31}\Delta_{14}^{12} + a_{21}a_{41}\Delta_{13}^{12} + a_{31}a_{41}\Delta_{12}^{12} \\ & - \sqrt{6}(a_{11}a_{41}\Delta_{14}^{12} + a_{21}a_{51}\Delta_{12}^{12}) \\ & - \sqrt{3}/2a_{21}^2\Delta_{15}^{12} \\ & + 2(a_{11}a_{31}\Delta_{15}^{12} + a_{11}a_{51}\Delta_{13}^{12}) - a_{31}^2\Delta_{13}^{12},\end{aligned}$$

$$\begin{aligned}\binom{31}{4}^1 = {} & \sqrt{3}(a_{21}a_{31}\Delta_{14}^{12} + a_{31}a_{14}\Delta_{12}^{12}) \\ & - 1/\sqrt{3}(a_{21}a_{41}\Delta_{13}^{12} + 2a_{31}^2\Delta_{13}^{12}) \\ & - \sqrt{2}(a_{11}a_{41}\Delta_{14}^{12} \\ & + 2a_{21}a_{51}\Delta_{12}^{12}) + 4/\sqrt{3}a_{11}a_{51}\Delta_{13}^{12} \\ & + 1/\sqrt{2}(a_{21}a_{31}\Delta_{23}^{12} - a_{21}^2\Delta_{24}^{12}),\end{aligned}$$

$$\begin{aligned}\binom{31}{2} = {} & \sqrt{2}a_{11}^2\Delta_{45}^{12} + 1/\sqrt{2}a_{41}^2 - \sqrt{2}a_{31}a_{41}\Delta_{13}^{12} \\ & + 5\sqrt{2}/6a_{31}^2\Delta_{14}^{12} + 1/2\sqrt{3}a_{31}^2\Delta_{23}^{12} \\ & - 1/\sqrt{3}a_{21}a_{31}\Delta_{24}^{12} - \sqrt{3}a_{11}a_{21}\Delta_{35}^{12} \\ & - 1/\sqrt{3}a_{31}a_{51}\Delta_{12}^{12} + 1/\sqrt{2}a_{21}^2\Delta_{251}^{12}\end{aligned}$$

$$\begin{aligned}\binom{41}{4} = {} & 4\Delta_{14}^{12}a_{21}a_{31}^2 - 2\Delta_{13}^{12}a_{21}a_{31}a_{41} + 4\Delta_{12}^{12}a_{31}^2a_{41} \\ & - 4\sqrt{6}\Delta_{14}^{12}a_{11}a_{31}a_{41} + 2\sqrt{6}\Delta_{12}^{12}a_{21}a_{31}a_{51} \\ & - 4\sqrt{6}\Delta_{15}^{12}a_{21}^2a_{31} + 8\Delta_{15}^{12}a_{31}^2 - 4\Delta_{13}^{12}a_{11}a_{31}a_{51} \\ & - 2\Delta_{13}^{12}(a_{31})^3 - \sqrt{6}\Delta_{23}^{12}a_{21}^2a_{41} - \sqrt{6}\Delta_{24}^{12}a_{21}^2a_{31} \\ & + 6\Delta_{14}^{12}a_{21}^2a_{41} + 3\Delta_{25}^{12}a_{21}^3 + \sqrt{6}\Delta_{13}^{12}a_{21}^2a_{51} \\ & + \sqrt{6}\Delta_{23}^{12}a_{21}a_{31}^2 - 9\Delta_{12}^{12}a_{21}a_{41}^2 + 3\sqrt{6}\Delta_{13}^{12}a_{11}a_{41}^2,\end{aligned}$$

$$\begin{aligned}\binom{41}{0} = {} & 2\sqrt{2}a_{11}a_{31}a_{51}\Delta_{15}^{12} - 8\sqrt{2}\Delta_{14}^{12}a_{21}a_{31}a_{51} \\ & + 4\sqrt{2}\Delta_{13}^{12}a_{13}^2a_{51} - 8\sqrt{2}\Delta_{12}^{12}a_{31}a_{41}a_{51} \\ & - 3\sqrt{3}\Delta_{15}^{12}a_{11}a_{41}^2 - \sqrt{3}\Delta_{14}^{12}a_{21}a_{41}^2 \\ & - 3\sqrt{3}\Delta_{13}^{12}a_{31}a_{41}^2 \\ & + 3\sqrt{3}\Delta_{12}^{12}a_{41}^3 - \sqrt{3}\Delta_{15}^{12}a_{21}^2a_{51} \\ & + 7\sqrt{2}\Delta_{15}^{12}a_{21}a_{31}a_{41} \\ & - 3\sqrt{2}\Delta_{15}^{12}a_{31}^3 - 3\sqrt{3}\Delta_{13}^{12}a_{31}a_{41}^2 \\ & + 4\sqrt{3}\Delta_{14}^{12}a_{11}a_{41}a_{51} \\ & - 4\sqrt{2}\Delta_{13}^{12}a_{11}a_{51}^2 + 4\sqrt{3}\Delta_{12}^{12}a_{21}a_{51}^2 \\ & + 2\sqrt{2}\Delta_{13}^{12}a_{21}a_{41}a_{51} + 3\sqrt{3}\Delta_{24}^{12}a_{21}^2\Delta_5^1 \\ & - 2\sqrt{3}\Delta_{25}^{12}a_{21}^2a_{41} + \sqrt{2}\Delta_{25}^{12}a_{21}a_{31}^2 \\ & - 3\sqrt{3}\Delta_{23}^{12}a_{21}a_{31}a_{51} + 2\sqrt{2}\Delta_{23}^{12}a_{21}a_{41}^2 \\ & - \sqrt{2}\Delta_{24}^{12}a_{21}a_{31}a_{41},\end{aligned}$$

where $\Delta_{rs}^{12} = (a_{r1}a_{s2} - a_{r2}a_{s1})$. The above semi-invariants may be obtained by the method described in Ref. 2. They may also be obtained by writing down the Weyl basis vectors of the IR$(n, 1)$ of $U(5)$ and taking linear combinations of them so as to satisfy the equations $L_1P = 0$ and $L_0P = \frac{1}{2}(\lambda_1 - \lambda_2)$. The polynomial corresponding to the Weyl table $\begin{matrix} i & j & k & \cdots \\ & p & & \end{matrix}$ is $\Delta_{ip}^{12}a_{j1}a_{k1}\cdots$.

3. HWP IN THE CHAIN $U(5) \supset R(3)$ FROM THE BINARY-5-ARY (4, 1) FORM

To find the hwp in the chain $U(5) \supset R(3)$ for IR's of $U(5)$ corresponding to any partition $(h_1, \ldots, h_5)$, one has to solve the differential equations

$$C_{ii}P = h_iP, \quad C_{ij}P = 0, \quad i < j, \quad i, j = 1, \ldots, 5, \tag{3.1}$$

$$L_1P = 0, \quad LP = \alpha P \tag{3.2}$$

for polynomial solutions. For, it has been shown[9] that the polynomial solutions of (3.1) span a linear space W which is irreducibly invariant under the algebra $\mathfrak{A}$ and a solution of (3.1) is the hwp with weight α, of a subspace W which is irreducibly invariant under $\mathfrak{B}$ if and only if it satisfies (3.2) also. The required hwp are the invariants of the ground form

$$\sum_{i=1}^{5} x_i'(\sqrt{2}a_{1i}x_1^4 + 4a_{2i}x_1^3x_2 + 4\sqrt{3}a_{3i}x_1^2x_2^2 + 8a_{4i}x_1^3x_2 + 4\sqrt{2}a_{5i}x_2^4), \tag{3.3}$$

where the dashed and undashed variables are subjected to linear transformations separately and where both the dashed and undashed variables are subjected to transformations which are represented by upper triangular matrices and called step transformations by Weyl.[10]

The complete irreducible system of invariants under the group of step transformations for a finite number of linear forms in n variables $\sum_{j=1}^{n} a_{ij}x_j$, $i = 1, 2, \ldots$ was shown to be a_{i1}, Δ_{ij}^{12}, Δ_{ijk}^{123}, $\ldots$ $\Delta_{ij\cdots k}^{12\cdots n}$ by Weitzenböck[11] who called them semi-invariants; $\Delta_{ij\cdots k}^{12\cdots r} = \det|a_{i1}a_{j2}\cdots a_{kr}|$.

We will now proceed to develop some symbolic notation, making use of the already developed notation in invariant theory.

We state the following theorem whose proof is easy. Consider the binary n-ary form written symbolically as

$$f = (ax)^r(Ax'),$$

which is therefore of order r in the pair of variables $x = (x_1, x_2)$ and linear in the variables $x' = (x'_1, \ldots, x'_n)$. The invariants of this form under independent step transformations represented by upper triangular matrices, of the two sets of variables (x) and (x') separately, are all expressible as polynomials in the symbolic factors

$$a, [ab], A, [AB], [ABC], \ldots \underbrace{[AB \cdots K]}_{n \text{ symbols}},$$

where (ax) and (Ax') are symbolic factors of the first kind and $[ab]$, $[AB \cdots]$ are symbolic factors of the second kind and $aA, bB, \cdots$ are pairs of equivalent symbols.

We also state that such an invariant of degree m in the coefficients contains m equivalent pairs of symbols $aA, bB, \ldots, kK$ and the symbolic expression of the invariant is homogeneous of degree r in each of the small letters and homogeneous and linear in each of the capital letters.

We will refer to the above invariants as semi-invariants of f.

We define the width of a bracket factor $[AB \cdots]$ as the number of symbols it contains. A bracket of width zero is the number 1 and a bracket of width unity is a free symbol. Without loss of generality an invariant of the above form under the transformations mentioned, may be supposed to be homogeneous in each set of the coefficients $a_{i1}, a_{i2}, \ldots, a_{in}$ and hence each term of the invariant, when expressed symbolically, consists of the same number of bracket factors of each width.

Since the semi-invariants are polynomials in the coefficients of f and hence are polynomials in the symbols, we will define transvectants of symbolic factors as follows. The transvectant of index h, $0 \leq h \leq r$, is

$$(a^r, b^r)^h = [ab]^h a^{r-h} b^{r-h}.$$

The following transvectant of index 1, in capital letters, is defined as $(P[A'B' \cdots][A''B'' \cdots] \cdots, A)^1 = P[A'B' \cdots A][A''B'' \cdots] \cdots + P[A'B' \cdots][A''B'' \cdots A] \cdots + P[A'B' \cdots][A''B'' \cdots] \cdots [\cdots A]$, where the factors $[A'B' \cdots], [A''B'' \cdots]$ all have the same width w, and P is a product of factors each of which has width $> w$. We define $(P, A)^0 = PA$ and $(P + Q, A)^1 = (P, A)^1 + (Q, A)^1$, where P and Q are symbolic products in capital letters. We also define $([A'B' \cdots], A)^1 = 0$ when the left-hand side bracket has width equal to n. The transvectant of index (h, H) of a semi-invariant $P(a', b' \cdots)Q(A', B' \cdots)$ with a^rA, is defined as $(P(a', b' \cdots)Q(A', B' \cdots), a^rA)^{h,H} = (P(a', b' \cdots), a^r)^h (Q(A', B' \cdots), A)^H$, where $0 \leq h \leq r$ and $0 \leq H \leq 1$.

Let C_m be a semi-invariant of degree m and be a product of symbolic factors. Then the product contains m small letters $a_1, \ldots, a_{m-1}, a$ and m capital letters $A_1, \ldots, A_{m-1}, A$ and may be written as

$$C_m = P[a_1 a]^{\alpha_1}[a_2 a]^{\alpha_2} \cdots [a_{m-1} a]^{\alpha_{m-1}} a^\nu [A_i A_j \cdots A_k A],$$

where $\alpha_1 + \cdots \alpha_{m-1} + \nu = r$ and P does not contain the symbols a, A and all bracket factors in the capital letters contained in P have width not less than the width of $[A_i \cdots A_k A]$ (which may be of width 1 also). Then C_m is a term in the transvectant $(C_{m-1}, a^rA)^{h,H}$, where $h = r - \nu$ and $H = 0$ or 1 and $C_{m-1} = P a_1^{\alpha_1} \cdots a_{m-1}^{\alpha_{m-1}} [A_i \cdots A_k]$. Now

$$(C_{m-1}, a^rA)^{h,H} = (d_{m-1}, a^r)^h (D_{m-1}, A)^H,$$

where d_{m-1} is a product of symbols of small letters and D_{m-1} is a product of symbols of capital letters. Let $C_m = tT$, t and T being, respectively, symbolic products of small and capital letters. Then $(D_{m-1}, A)^H = T$ and

$$(C_{m-1}, a^rA)^{h,H} - C_m = [(d_{m-1}, a^r)^h - t]T = \Sigma(\bar{d}_{m-1}, a^r)^{h'} T,$$

where $h' < h$ and where each $\bar{d}_{m-1}$ is obtained from d_{m-1} by convolution, i.e., by replacing pairs of free symbols by bracket factors. Therefore,

$$C_m = (C_{m-1}, a^rA)^{h,H} + \Sigma(\bar{C}_{m-1}, a^rA)^{h',H},$$

where $\bar{C}_{m-1}$ is obtained from C_{m-1} by convolution with respect to the small letters and $h' < h$.

Now to find semi-invariants of f of degree q we must take in (2.3) A as a 2×2 upper triangular matrix and B as an $n \times n$ upper triangular matrix. On account of the one to one correspondence of the semi-invariants and covariants of a binary form, the decomposition of $(A^{\{r\}})^{\{\mu\}}$ and hence the calculation of $\{r\} \times \{\mu\}$ where (μ) is a partition of q, may be done as usual. Coming to the part $B^{\{\mu\}}$, we know all the semi-invariants of type $(\mu) = (\mu_1, \ldots, \mu_n)$ of n linear ground forms.

As an example, we will find a complete system for the semi-invariants of type $(n, 1, 1) = (\mu)$ in the previous paragraph, for the ground form (3.3), n being a positive integer. In their symbolic expression, these have in the capital letter part only one bracket factor of width 3; in other words, their capital letter part would be like $[ABC]\,DE \cdots$ where the symbols $D, E, \cdots$ are $(n-1)$ in number. We wil denote such a semi-invariant by $(^{n,1,1}_{\lambda_1-\lambda_2})$ where $\lambda_1 - \lambda_2 = 4d - 2\rho$, ρ being its weight with respect to the binary variables (definition of weight is given on p. 2) and $d = n + 2$ and $\{\lambda_1, \lambda_2\}$ is an S function occurring in the expansion of $\{4\} \otimes \{n, 1, 1\}$.

We have the following expansions:

$$\{4\} \otimes \{2, 1, 1\} = \{13, 3\} + \{12, 4\} + 2\{11, 5\} + \{10, 6\} + 2\{9, 7\},$$

$$\{4\} \otimes \{3, 1, 1\} = \{17, 3\} + \{16, 4\} + 3\{15, 5\} + 2\{14, 6\} + 4\{13, 7\} + 2\{12, 8\} + 3\{11, 9\},$$

$$\{4\} \otimes \{4, 1, 1\} = \{21, 3\} + \{20, 4\} + 3\{19, 5\} + 3\{18, 6\} + 5\{17, 7\} + 4\{16, 8\} + 6\{15, 9\} + 3\{14, 10\} + 4\{13, 11\},$$

$$\{4\} \otimes \{5, 1, 1\} = \{25, 3\} + \{24, 4\} + 3\{23, 5\} + 3\{22, 6\} + 6\{21, 7\} + 5\{20, 8\} + 8\{19, 9\}$$

$$+ 6\{18, 10\} + 8\{17, 11\} + 4\{16, 12\} + 5\{15, 13\},$$

$$\{4\} \otimes \{6, 1, 1\} = \{29, 3\} + \{28, 4\} + 3\{27, 5\} + 3\{26, 6\} + 6\{25, 7\} + 6\{24, 8\} + 9\{23, 9\} + \{22, 10\} + 11\{21, 11\} + 8\{20, 12\} + 10\{19, 13\} + 5\{18, 14\} + 6\{17, 15\}.$$

These suggest the semi-invariants of the following types to be irreducible and complete together with the semi-invariants (2.6):

$$\binom{1,1,1}{6}, \binom{1,1,1}{2}, \binom{2,1,1}{8}, \binom{2,1,1}{6}, \binom{2,1,1}{4}, \binom{2,1,1}{2}, \binom{2,1,1}{2}', \binom{3,1,1}{8}, \binom{3,1,1}{4}, \binom{3,1,1}{4}', \binom{3,1,1}{2}, \binom{3,1,1}{2}', \binom{4,1,1}{4}, \binom{4,1,1}{4}', \binom{4,1,1}{2}, \text{and } \binom{5,1,1}{4}. \tag{3.4}$$

We will show that if all semi-invariants of the type $\binom{m,1,1}{2\alpha}$ where $m \geq 6$ can be expressed as polynomials in the semi-invariants (3.4) and (2.6) then all semi-invariants of the type $\binom{m+1,1,1}{2\alpha}$ can be done so.

Now any semi-invariant of degree $(m + 3)$ is a linear combination of transvectants $(C_{m+2}, a^4A)^{h,H}$ where C_{m+2} is a semi-invariant of degree $(m + 2)$. When $H = 1$ in the above transvectant, we need to consider C_{m+2} which have just one bracket factor $[A'B']$ in which case the transvectant is of degree 3 only. Hence we have to consider transvectants $(C_{m+2}, a^4A)^{h,0}$, where C_{m+2} has one and only one of the semi-invariants (3.4) as a factor. In C_{m+2} we need take only $\binom{1,1,1}{2}$, $\binom{2,1,1}{2}$, $\binom{2,1,1}{2}'$, $\binom{3,1,1}{2}$, $\binom{3,1,1}{2}'$, $\binom{4,1,1}{2}$, $\binom{2,0}{0}$ and $\binom{3,0}{0}$ as factors because all others have their $2\alpha \geq 4$. (Refer to statement 2 on p. 8). But $\binom{2,0}{0}$ and $\binom{3,0}{0}$ come out as factors from these transvectants and the transvectants $(F, a^4A)^{h,0}$, where F is any one of the above semi-invariants except the last two, are all of the type $\binom{m,1,1}{2\alpha}$ where $m \leq 5$. Hence they are expressible in terms of (3.4) and (2.5). Therefore all the semi-invariants of type $\binom{m,1,1}{2\alpha}$ are expressible as polynomials in the semi-invariants (3.4) and (2.5), which are irreducible.

ACKNOWLEDGMENT

The author is grateful to Professor M. Moshinsky for suggesting the problems of construction of polynomial bases in various chains of groups.

1 J. A. Todd, Proc. Cambridge Phil. Soc. **42**, 206 (1946).
2 D. E. Littlewood, Phil. Trans. Roy. Soc. (London) **A239**, 305 (1944).
3 H. W. Turnbull, *The Theory of Determinants Matrices and Invariants* (Dover, New York, 1960), p. 227.
4 M. Moshinsky and V. Syamala Devi, J. Math. Phys. **10**, 455 (1969).
5 Reference 4, p. 261.
6 D. E. Littlewood, *Theory of Group Characters* (Oxford U.P., London, 1950).
7 F. E. Murnaghan, Proc. Natl. Acad. Sci. (U.S.) **40**, 821 (1954).
8 H. W. Turnbull, Proc. Roy. Soc. Edinburgh **42**, 147 (1921).
9 M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).
10 H. Weyl, *Classical Groups* (Princeton U.P., Princeton, N.J., 1946), p. 49.
11 R. Weitzenböck, *Invariantentheorie* (Groningen, 1923), p. 219.

Radial Matrix Elements of the Radial-Angular Factorized Hydrogen Atom

M. J. Cunningham
Physics Department, University of Canterbury, Christchurch, New Zealand
(Received 28 September 1970)

The hydrogen atom is factorized according to the scheme $O(4, 2) \supset O(2, 1) \times O(3)$ and the radial group $O(2, 1)$ studied. It is shown that $r^k D_{n/(n+q)}$, where D_a is a dilatation operator, is proportional to a tensor operator in this scheme, allowing a group theoretical study of the radial matrix elements of r^k, including an explanation of the Pasternack and Sternheimer selection rule.

I. INTRODUCTION

The selection rule, on hydrogenic radial integral discovered by Pasternack and Sternheimer,[1] namely

$$\int_0^\infty \frac{R_{nl} R_{nl'}}{r^s} r^2 dr = 0, \quad s = 2, 3, \cdots, |l - l'| + 1$$

has stimulated recently at least two attempts to explain it group theoretically. That of Swamy, Kulkarni and Biedenharn[2] is based on the $O(4)$ symmetry of the hydrogen atom and uses a complex recursive technique. Armstrong[3] approaches the problem more directly by showing that radiallike functions of two variables r and τ transform according to representations of the noncompact group $O(2, 1)$, and, in this scheme, he shows that positive and negative powers of r have tensorial transformation properties. Armstrong's scheme however is unsatisfactory for two important reasons. First, his two variable functions of r and τ never completely coincide with true one variable radial functions of r/n, where n is the principal quantum number. As one consequence of this, no treatment of off-diagonal matrix elements can be given. Secondly, Armstrong radiallike functions and their associated $O(2, 1)$ group stand isolated from any encompassing group scheme such as the $O(4, 2)$ model of Barut and Kleinert.[4] Kleinert[5] has written a paper containing an excellent review of this model.

$$+ 6\{18, 10\} + 8\{17, 11\} + 4\{16, 12\} + 5\{15, 13\},$$

$$\{4\} \otimes \{6, 1, 1\} = \{29, 3\} + \{28, 4\} + 3\{27, 5\} + 3\{26, 6\} + 6\{25, 7\} + 6\{24, 8\} + 9\{23, 9\} + \{22, 10\} + 11\{21, 11\} + 8\{20, 12\} + 10\{19, 13\} + 5\{18, 14\} + 6\{17, 15\}.$$

These suggest the semi-invariants of the following types to be irreducible and complete together with the semi-invariants (2.6):

$$\binom{1,1,1}{6}, \binom{1,1,1}{2}, \binom{2,1,1}{8}, \binom{2,1,1}{6}, \binom{2,1,1}{4}, \binom{2,1,1}{2}, \binom{2,1,1}{2}', \binom{3,1,1}{8}, \binom{3,1,1}{4}, \binom{3,1,1}{4}', \binom{3,1,1}{2}, \binom{3,1,1}{2}', \binom{4,1,1}{4}, \binom{4,1,1}{4}', \binom{4,1,1}{2}, \text{and } \binom{5,1,1}{4}. \tag{3.4}$$

We will show that if all semi-invariants of the type $\binom{m,1,1}{2\alpha}$ where $m \geq 6$ can be expressed as polynomials in the semi-invariants (3.4) and (2.6) then all semi-invariants of the type $\binom{m+1,1,1}{2\alpha}$ can be done so.

Now any semi-invariant of degree $(m + 3)$ is a linear combination of transvectants $(C_{m+2}, a^4A)^{h,H}$ where C_{m+2} is a semi-invariant of degree $(m + 2)$. When $H = 1$ in the above transvectant, we need to consider C_{m+2} which have just one bracket factor $[A'B']$ in which case the transvectant is of degree 3 only. Hence we have to consider transvectants $(C_{m+2}, a^4A)^{h,0}$, where C_{m+2} has one and only one of the semi-invariants (3.4) as a factor. In C_{m+2} we need take only $\binom{1,1,1}{2}$, $\binom{2,1,1}{2}$, $\binom{2,1,1}{2}'$, $\binom{3,1,1}{2}$, $\binom{3,1,1}{2}'$, $\binom{4,1,1}{2}$, $\binom{2,0}{0}$ and $\binom{3,0}{0}$ as factors because all others have their $2\alpha \geq 4$. (Refer to statement 2 on p. 8). But $\binom{2,0}{0}$ and $\binom{3,0}{0}$ come out as factors from these transvectants and the transvectants $(F, a^4A)^{h,0}$, where F is any one of the above semi-invariants except the last two, are all of the type $\binom{m,1,1}{2\alpha}$ where $m \leq 5$. Hence they are expressible in terms of (3.4) and (2.5). Therefore all the semi-invariants of type $\binom{m,1,1}{2\alpha}$ are expressible as polynomials in the semi-invariants (3.4) and (2.5), which are irreducible.

ACKNOWLEDGMENT

The author is grateful to Professor M. Moshinsky for suggesting the problems of construction of polynomial bases in various chains of groups.

1 J. A. Todd, Proc. Cambridge Phil. Soc. **42**, 206 (1946).
2 D. E. Littlewood, Phil. Trans. Roy. Soc. (London) **A239**, 305 (1944).
3 H. W. Turnbull, *The Theory of Determinants Matrices and Invariants* (Dover, New York, 1960), p. 227.
4 M. Moshinsky and V. Syamala Devi, J. Math. Phys. **10**, 455 (1969).
5 Reference 4, p. 261.
6 D. E. Littlewood, *Theory of Group Characters* (Oxford U.P., London, 1950).
7 F. E. Murnaghan, Proc. Natl. Acad. Sci. (U.S.) **40**, 821 (1954).
8 H. W. Turnbull, Proc. Roy. Soc. Edinburgh **42**, 147 (1921).
9 M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).
10 H. Weyl, *Classical Groups* (Princeton U.P., Princeton, N.J., 1946), p. 49.
11 R. Weitzenböck, *Invariantentheorie* (Groningen, 1923), p. 219.

Radial Matrix Elements of the Radial-Angular Factorized Hydrogen Atom

M. J. Cunningham
Physics Department, University of Canterbury, Christchurch, New Zealand
(Received 28 September 1970)

The hydrogen atom is factorized according to the scheme $O(4, 2) \supset O(2, 1) \times O(3)$ and the radial group $O(2, 1)$ studied. It is shown that $r^k D_{n/(n+q)}$, where D_a is a dilatation operator, is proportional to a tensor operator in this scheme, allowing a group theoretical study of the radial matrix elements of r^k, including an explanation of the Pasternack and Sternheimer selection rule.

I. INTRODUCTION

The selection rule, on hydrogenic radial integral discovered by Pasternack and Sternheimer,[1] namely

$$\int_0^\infty \frac{R_{nl} R_{nl'}}{r^s} r^2 dr = 0, \quad s = 2, 3, \cdots, |l - l'| + 1$$

has stimulated recently at least two attempts to explain it group theoretically. That of Swamy, Kulkarni and Biedenharn[2] is based on the $O(4)$ symmetry of the hydrogen atom and uses a complex recursive technique. Armstrong[3] approaches the problem more directly by showing that radiallike functions of two variables r and τ transform according to representations of the noncompact group $O(2, 1)$, and, in this scheme, he shows that positive and negative powers of r have tensorial transformation properties. Armstrong's scheme however is unsatisfactory for two important reasons. First, his two variable functions of r and τ never completely coincide with true one variable radial functions of r/n, where n is the principal quantum number. As one consequence of this, no treatment of off-diagonal matrix elements can be given. Secondly, Armstrong radiallike functions and their associated $O(2, 1)$ group stand isolated from any encompassing group scheme such as the $O(4, 2)$ model of Barut and Kleinert.[4] Kleinert[5] has written a paper containing an excellent review of this model.

These difficulties are met in this paper by considering the hydrogen atom factorized into its radial $[O(2,1)]$ and angular $[O(3)]$ parts, according to the scheme $O(4,2) \supset O(3) \times O(2,1)$ of Barut and Kleinert.[4,5] Here the true radial wavefunctions of the variable r/n transform according to a representation of the $O(2,1)$ group (Sec. II) and it is shown that the quantity $r^k D_{n/n+q}$ for all positive and negative integer k, where D_a is a dilatation operator defined by $D_a f(x) = f(ax)$, is proportional to the qth component of a tensor (Sec. III). The Pasternack and Sternheimer selection rule follows naturally and, since the Wigner–Eckart theorem is shown to hold (Sec. IV), calculation of the appropriate Clebsch–Gordan coefficients (Sec. V) allows matrix elements to be calculated. Some progress is made on the difficult problem of off-diagonal elements of r^k (Sec. V).

II. THE GROUP SCHEME

The 15 generators $L_{\mu\nu}$, $1 \le \mu < \nu \le 6$, of $O(4,2)$ are defined in terms of their action on the hydrogen atom wavefunction written in parabolic coordinates.[4,5] Their commutation relations are given by

$$[L_{\mu\nu} L_{\mu\lambda}] = i g_{\mu\mu} L_{\nu\lambda}, \quad g_{\alpha\alpha} = \begin{cases} 1, & 1 \le \alpha \le 4, \\ -1, & 5 \le \alpha \le 6. \end{cases}$$

A subgroup $O(3) \times O(2,1)$ can be formed with the $O(3)$ generators given by L_{12}, L_{23}, and L_{13} and the $O(2,1)$ generators given by L_{45}, L_{46}, and L_{56}. These last generators are formed into the linear combinations $K_\pm = L_{45} \mp i L_{46}$ and $K_0 = L_{56}$ to give the standard $O(2,1)$ commutation relations[6-8] of

$$[K_0 K_\pm] = \pm K_\pm \quad \text{and} \quad [K_+ K_-] = -2K_0.$$

Consider now the entire hydrogen atom wavefunction

$$|nlm\rangle = R_{nl}(x_n) Y_{lm}(\theta\varphi),$$

where

$$R_{nl}(x_n) = N_{nl} e^{-x_n} (2x_n)^l L^{2l+1}_{n-l-1}(2x_n),$$

$$N_{nl} = -(-1)^{n+l} \frac{2Z^{3/2}}{n^2} \left[\frac{(n-l-1)!}{(n+l)!}\right]^{1/2}, \quad (1)$$

$$x_n = \frac{Zr}{n},$$

and $Y^l_m(\theta\varphi)$ are the usual spherical harmonics. A Hilbert space is defined by the inner product

$$\int_{r^0}^{\infty} \int_\Omega R_{n'l'}(x_{n'}) Y_{l'm} x_{n'}^2 R_{nl}(x_n) Y_{lm} \frac{n'^2 n}{Z^3} dx_n d\Omega = \delta(nn')\delta(ll'), \quad (2)$$

where $d\Omega = \sin\theta d\theta d\varphi$.

In the $O(4.2) \supset O(2,1) \times O(3)$ scheme of Barut and Kleinert,[4] these wavefunctions for fixed m form a basis of representation of $O(2,1)$ since

$$K_\pm |nlm\rangle = [(n \mp l)(n \pm l \pm 1)]^{1/2} |n \pm 1 lm\rangle,$$
$$K_0 |nlm\rangle = n |nlm\rangle. \quad (3)$$

The representation has only a lower bound on the $O(2)$ quantum number (equal to $l+1$) and is shown below to be unitary and irreducible. Equation (3) implies that the realization in this Hilbert space of the generators of $O(2,1)$ is

$$K = \mp \left(\frac{n \pm 1}{n}\right) D_{n/(n\pm1)} \left(x_n \frac{\partial}{\partial x_n} \mp x_n \pm n + 1\right) \quad (4)$$

and $K_0 = n$, where D_a is a dilatation operator defined such that $D_a f(r) = f(ar)$, which implies that

$$D_{n/n'} f(x_n) = f(x_{n'}).$$

If we form the Casimir invariant

$$G = K_0^2 - \tfrac{1}{2}(K_+ K_- + K_- K_+),$$

we find that

$$\int_0^\infty \int_\Omega R_{nl} Y_{lm} x_n^2 G R_{nl} Y_{lm} n^3/Z^3 \, dx_n = \langle nlm | G | nlm \rangle = l(l+1). \quad (5)$$

However, Barut and Fronsdal[7] show that the eigenvalues of G are of the form $\Phi(\Phi+1)$. If a representation is bounded below, its lower bound is $-\Phi$. If Φ is negative, the representation is unitary and labeled $D^+_{-\Phi}$. All this implies that the wavefunctions $|nlm\rangle$ for fixed m form a basis for the unitary irreducible representation D^+_{l+1}.

III. THE TENSORIAL PROPERTIES OF r^k

In this section, it is shown that the quantity $r^k/[(n+q)^{k-1}] D_{n/(n+q)} A^l_{l'}(\theta\varphi)$, where q and k are integers, transforms as the qth component of an $O(2,1)$ tensor operator, the associated representation depending upon the size of k relative to -2. Here $A^l_{l'}$ is an operator defined such that

$$A^l_{l'}(\theta\varphi) Y_{lm} = Y_{l'm}. \quad (6)$$

To be specific, we can build $A^l_{l'}$ up as a power of products A^l_{l+1} or A^l_{l-1}, where

$$A^l_{l+1} = \left(\frac{2l+3}{(l+1-m)(l+1+m)(2l+1)}\right)^{1/2} \times \left((l+\tfrac{1}{2})\cos\theta + \sin\theta \frac{d}{d\theta}\right)$$

(see, for instance, Infeld and Hull).[8]

Consider

$$\langle n+q \pm 1 lm | \left[K_\pm, \frac{r^k}{(n+q)^{k-1}} D_{n/(n+q)} A^{l'}_l\right] | nl'm \rangle$$
$$= \mp \int_0^\infty \int_\Omega R_{n+q\pm1\,l} Y_{lm} x^2_{n+q\pm1} \left[\left(\frac{n+q\pm1}{n+q}\right) D_{n+q/(n+q\pm1)} \left(x_{n+q} \frac{\partial}{\partial x_{n+q}} \mp x_{n+q} \pm n \pm q + 1\right)(n+q) x^{k+2}_{n+q} D_{n/(n+q)}\right.$$
$$\left. - (n+q\pm1) x^{k+2}_{n+q\pm1} D_{n\pm1/(n+q\pm1)} \times \left(\frac{n\pm1}{n}\right) D_{n/(n\pm1)} \left(x_n \frac{\partial}{\partial x_n} \mp x_n \pm n + 1\right)\right] A^{l'}_l R_{nl'} Y_{l'm} \frac{(n+q\pm1)^2 n}{Z^{3+k}} dx_n d\Omega$$

$$= \int_0^\infty \int_\Omega R_{n+q\pm1}(n+q\pm1)x_{n+q\pm1}^{k+2}\left[\left(x_{n+q\pm1}\frac{\partial}{\partial x_{n+q\pm1}} \mp x_{n+q\pm1} + 2 \pm n \pm q + 1\right) - \left(x_{n+q\pm1}\frac{\partial}{\partial x_{n+q\pm1}} \mp x_{n+q\pm1} \pm n + 1\right)\right] D_{n/(n+q\pm1)}A_l^{l'}R_{nl'}Y_{l'm}\frac{(n+q\pm1)^2n}{Z^{3+k}}\,dx_n d\Omega$$

$$= \mp(k+2\pm q)\int_0^\infty \int_\Omega R_{n+q\pm1}Y_{lm}(n+q\pm1)\,x_{n+q\pm1}^{k+2}\,D_{n/(n+q\pm1)}A_l^{l'}\,R_{nl'}Y_{l'm}\,\frac{(n+q\pm1)^2n}{Z^3}\,dx_n d\Omega$$

$$= \mp(k+2\pm q)\left\langle n+q\pm1\,lm\left|\frac{r^k}{(n+q\pm1)^{k-1}}\,D_{n/(n+q\pm1)}A_{l'}^{l}\right|nl'm\right\rangle. \tag{7}$$

We see then that if $k+2 \leq 0$,

$$\left[K_\pm, \frac{r^k}{(n\mp k\mp 2)^{k-1}}\,D_{n/(n+k+2)}A_l^{l'}\right] = 0,$$

i.e., $k+2 \leq q \leq -k-2$, while if $k+2 > 0$,

$$\left[K_\mp, \frac{r^k}{(n\pm k\pm 2)^{k-1}}\,D_{n/(n\pm k\pm2)}A_l^{l'}\right] = 0.$$

Hence, if $k \leq -2$, we have a representation of finite dimension equal to $-2k-3$ labeled D_{-k-2}, while if $k > -2$, the representation is infinite dimensional and reducible, but not fully reducible; for while the spaces $q \geq k+2$ and $q \leq -k-2$ are invariant under the group operations, the space $-k-1 \leq q \leq k+1$ is not. This representation will be labeled as D'_{k+1}. Both of these representations are nonunitary since the eigenvalues of $K_\pm K_\pm$ are not positive definite.[7] This means we are free to normalize our tensor operators in an arbitrary manner, and a particular choice is made in Sec. V.

IV. THE WIGNER–ECKART THEOREM

A well-known method for proving the Wigner–Eckart theorem for $O(3)$ (see, for instance, Messiah[9]) sets up recursion relations between different matrix elements of the qth and $(q\pm1)$th components of an $O(3)$ tensor operator and shows that these are identical to the recursion relations for the corresponding $O(3)$ Clebsch-Gordan coefficients. This implies that Clebsch–Gordan coefficient and matrix element are proportional, the proportionality constant being the reduced matrix element. This proof in fact requires only that if $J_i|lm\rangle = a_i|lm_i\rangle$, then $[J_i T_m^l] = a_i T_{m_i}^l$, where J_i is an $O(3)$ generator and T_m^l an $O(3)$ tensor operator, and that the states being coupled up be orthonormal.

In the $O(2,1)$ case, we first write the states corresponding to the representations D_{-k-2} and D'_{k+1} as $|kq\rangle$, where the use of k and q distinguishes them from the states $|nl\rangle$ of D_{l+1}^+, and the corresponding tensor operator as $N_{kq}[r^k/(n+q)^{k-1}]D_{n/(n+q)}A_l^{l'}$, where N_{kq} is some arbitrary normalization. We have then that

$$K_\pm|kq\rangle = \pm(k+2\pm q)(N_{kq}/N_{kq\pm1})|kq\pm1\rangle$$

and

$$[K_\pm, T_q^k] = \pm(k+2\pm q)(N_{kq}/N_{kq\pm1})T_{q\pm1}^k.$$

K_0 in both cases gives an eigenvalue of q.

It therefore follows that since the states $|nl\rangle$ are orthonormal, the Wigner–Eckart theorem holds, and we can write

$$\langle lnm|T_q^k|l'n'm\rangle = \langle l||T^k||l'\rangle C_{n'qn}^{l'kl}, \tag{8}$$

where $C_{n'qn}^{l'kl}$ is an $O(2,1)$ Clebsch-Gordan coefficient coupling the states of $D_{l'+1}^+$ with the states of either D_{-k-2} or D'_{k+1}, normalized as above, to the states of D_{l+1}^+, while $\langle l|T^k|l'\rangle$ is the reduced matrix element and is independent of n, n', and q.

V. THE RACAH ALGEBRA

In this section, the Clebsch–Gordan coefficients appropriate to the representations we are considering are derived. The Pasternack and Sternheimer selection rule drops out naturally during the analysis. The reduced matrix elements of T^k are derived, allowing the diagonal matrix elements of r^k to be calculated. A brief discussion of the off-diagonal elements of r^k concludes the paper.

The most direct means of deriving the Clebsch–Gordan coefficients, namely that of coupling two representations and the contragradient of a third to an invariant is perhaps most well known in Bargmann's work,[10] but is due to van der Waerden.[11] Barut and Fronsdal[7] have used this technique to derive some $O(2,1)$ Clebsch–Gordan coefficients for unitary representations. This method has the advantage of not requiring an explicit form for the arbitrary normalization of the states of nonunitary representations.

To apply this technique, we must first realize the representations in terms of multispinors $N_{ab}\xi^a\eta^b$, where N_{ab} is a normalization constant and the quantity (ξ,η), called a spinor, forms a basis for the fundamental irreducible representation of $O(2,1)$. In this realization

$$K_0 = \tfrac{1}{2}\left(\xi\frac{\partial}{\partial\xi} - \eta\frac{\partial}{\partial\eta}\right), \qquad K_+ = \xi\frac{\partial}{\partial\eta},$$

$$K_- = -\eta\frac{\partial}{\partial\xi}. \tag{9}$$

So $K_0 N_{ab}\xi^a\eta^b = \tfrac{1}{2}(a-b)N_{ab}\xi^a\eta^b$ and $GN_{ab}\xi^a\eta^b = \tfrac{1}{4}(a-b)(a-b+2)N_{ab}\xi^a\eta^b$.

Hence, if the eigenvalues of G and K_0 are, respectively, $\Phi(\Phi+1)$ and m, then the states are represented by $N_{ab}\xi^{\Phi+m}\eta^{\Phi-m}$, where N_{ab} is shown by Barut and Fronsdal to be equal to $[(m-1-\Phi)!/(m+\Phi)!]^{1/2}$ for the unitary representation D_Φ^+. As remarked above, N_{ab} is arbitrary for nonunitary representations.

We now form an invariant coupling of two representations and a contragradient representation

defined as

$$\langle \Phi_m | = N_m^{-1} \xi^{\Phi - m} \eta^{\Phi + m} \tag{10}$$

giving

$$I = \sum_{m_1 m_2 m_3} C^{\Phi_1 \Phi_2 \Phi_3}_{m_1 m_2 m_3} N_1 \xi_1^{\Phi_1 + m_1} \eta_1^{\Phi_1 - m_1} N_2 \xi_2^{\Phi_2 + m_2} \eta_2^{\Phi_2 - m_2} \times N_3^{-1} \xi_3^{\Phi_3 - m_3} \eta_3^{\Phi_3 + m_3}, \tag{11}$$

where I is an invariant in the space of the polynomials of $\Pi_i \xi_i^{a_i} \eta_i^{b_i}$. However, since the representation matrices are unimodular, the only invariants in this space are the three determinants

$$\delta_1 = \xi_2 \eta_3 - \xi_3 \eta_2, \quad \delta_2 = \xi_3 \eta_1 - \xi_1 \eta_3, \quad \delta_3 = \xi_1 \eta_2 - \xi_2 \eta_1$$

and every monomial in these; see Bargmann.[10] Hence

$$\delta_1^{k_1} \delta_2^{k_2} \delta_3^{k_3} = \sum_{m_1 m_2 m_3} C^{\Phi_1 \Phi_2 \Phi_3}_{m_1 m_2 m_3} N_1 \xi_1^{\Phi_1 + m_1} \eta_1^{\Phi_1 - m_1} \times N_2 \xi_2^{\Phi_2 + m_2} \eta_2^{\Phi_2 - m_2} N_3 \xi_3^{\Phi_3 - m_3} \eta_3^{\Phi_3 + m_3}. \tag{12}$$

This implies

$$k_2 + k_3 = 2\Phi_1, \quad k_3 + k_1 = 2\Phi_2, \quad k_2 + k_1 = 2\Phi_3 \tag{13}$$

or equivalently

$$k_1 = \Phi_2 + \Phi_3 - \Phi_1, \quad k_2 = \Phi_3 + \Phi_1 - \Phi_2, \quad k_3 = \Phi_1 + \Phi_2 - \Phi_3. \tag{14}$$

We now note that

$$(\xi_a \eta_b - \xi_b \eta_a)^n = \sum_{s \geq 0} \binom{n}{s} (\xi_a \eta_b)^s (-\xi_b \eta_a)^{n-s}$$

or

$$\sum_{s \geq 0} \binom{n}{s} (\xi_a \eta_b)^{n-s} (-\xi_b \eta_a)^s$$

for all positive and negative integer n, so that by comparing coefficients, we have

$$N_1 N_2 N_3^{-1} C^{\Phi_1 \Phi_2 \Phi_3}_{m_1 m_2 m_3} = \sum_{pqr} (-1)^{p+q+r} \binom{k_1}{p} \binom{k_2}{q} \binom{k_3}{r}. \tag{15}$$

Consider first $k \leq -2$. In this case the representation D_{Φ_2} is finite dimensional, while $D^+_{\Phi_1}$ and $D^+_{\Phi_3}$ are infinite dimensional. Inspection of (12) shows that this requires k_1 and k_3 to be positive and k_2 negative. Taking these constraints and substituting

$$\Phi_1 = -l - 1, \quad \Phi_3 = -l' - 1, \quad \Phi_2 = -k - 2 = s - 2$$

into (14) gives the Pasternack and Sternheimer selection rule, namely

$$2 \leq s \leq |l - l'| + 1 \tag{16}$$

gives a vanishing matrix element.

To cast (15) into a useful form, we put $p = z$. From (12), we see that $p + q = \Phi_3 - m_3$ and $p + r = \Phi_2 + m_2$ so that (15) becomes

$$N_1 N_2 N_3^{-1} C^{l \; k \; l'}_{n \; n'-n \; n'} = \sum_z (-)^{k+l'+1+n'+z} \binom{l - l' - k - 2}{z} \times \binom{k - l - l'}{1 - l - k + n' + z} \times \binom{l' - l - k - 2}{l' - l - n' + n + z}, \quad k \leq -2. \tag{17}$$

This formula is quite adequate as it stands (after substituting for N), but it is of interest to note, with Armstrong, that it can be cast into a form that shows that this $O(2,1)$ Clebsch–Gordan coefficient is in fact equal, within a phase, to the $O(3)$ Clebsch–Gordan coefficients.

To do this, Eq. (17) is expanded in factorial notation (remembering that $k - l - l'$ is negative) and identities I and II derived in the Appendix are applied giving

$$A \left(\frac{(l+n)!(l'+n')!}{(k-q)!(k+q)!(n-l-1)!(n'-l'-1)!} \right)^{1/2} \sum_t (-1)^t \frac{(k+n-l'-1-t)!(k+n'+l-t)!}{t!(k-l'+l-t)!(n+l-t)!(n'-l'-1-t)!},$$

where A contains unimportant constant terms. N_2 has been chosen as $[(-k-2-q)!(-k-2+q)!]^{-1/2}$ which implies that $K_+ |kq\rangle = \mp [(k \mp q)(k \pm q \pm 1)]^{1/2} |kq \pm 1\rangle$. If the Clebsch–Gordan coefficients are now renormalized to satisfy the orthonormality condition

$$\sum_{qn} C^{l \, k \, l'}_{nq \, n'} C^{l \, k \, l''}_{n \, q \, n''} (-1)^{l+l'+q} = \delta(l'l'') \delta(n'n''), \tag{18}$$

derived by Armstrong, we finish with

$$C^{l \, k \, l'}_{n \, q \, n'} = (-1)^{k+q} \left(\frac{(l'-l-k-2)!(l+l'+k+2)!(l-l'-k-2)!}{(l+l'-k-1)!} \right)^{1/2} \times \left(\frac{(n'-l'-1)!(l'+n')!(-k-2-q)!(2l'+1)}{(-k-2+q)!(l+n)!(n-l-1)!} \right)^{1/2} \sum_t (-1)^{t+k+q} \binom{l'-n'+t}{-k-2-q} \binom{k+n'+l-t}{l'+n'} \binom{l+n}{t}$$

$$= a \langle kq, ln | l'n' \rangle, \tag{19}$$

where $a = 1$ if $l - l'$ is even and $-i$ if $l - l'$ is odd.

We consider now the case of $k > -2$. In this case, it is convenient to put $r = z$, giving

$$N_1N_2N_3^{-1}C^{l\ \ k\ \ l'}_{n\ n'-n\ n'} = \sum_z (-1)^{l'+1+n'+z}\binom{l-l'+k+1}{l-l'-n'+n+z} \times \binom{-2-l-l'-k-1}{-1-l'-k-1+n+z} \times \binom{l'-l+k+1}{z}. \tag{20}$$

The power of this technique is illustrated here, for Armstrong, using the recursive technique of Racah,[12] derives these particular Clebsch–Gordan coefficients for $q = n' - n$ equal to zero only.

Since in this case Φ_2 is positive, we cannot have both k_1 and k_3 negative. Thus, if we assume $l \geq l'$, we have two possibilities, namely both $l - l' + k + 1$ and $l' - l + k + 1$ positive or $l - l' + k + 1$ positive and $l' - l + k + 1$ negative. Since $q = 0$ is the case of most interest, we will consider the case of $k + q$ positive; thus the first possibility allows Eq. (19) to be transformed exactly as was Eq. (16) to give

$$\frac{(l-l'+k+2)!(l'+n')!}{(k+1+q)!}\sum_t \frac{(k+n'-l-t)!}{(k+2+l-n'+t)!(n'+l'-t)!(n'-l'-1-t)!(n-l-1-t)!t!}, \tag{21}$$

where each N_i has been put equal to 1. The second case transforms via identity III to the same expression.

Equation (20) is also proportional to the corresponding $O(3)$ Clebsch–Gordan coefficient since it is of the same form as that taken by Eq. (5) after the application of identity I.

The reduced matrix elements can be derived from first principles by considering the matrix element

$$\langle l+1\ l'm|T^k|l+1\ lm\rangle = C^{l\ \ k\ \ l'}_{l+1\ 0\ l+1}\langle l'||T^k||l\rangle$$

$$= \int_0^\infty\int_\Omega R_{l+1\ l'}Y_{l'm}\frac{r^{k+2}}{(l+1)^{k-1}}A^l_{l'}R_{l+1\ l}Y_{lm}dr d\Omega,\ l \geq l'. \tag{22}$$

The integral through Ω is one, and, since L_0^{2l+1} is a constant, Eq. (22) becomes a sum of integrals of the form

$$\int_0^\infty e^{-\alpha x}x^b dx = \frac{b!}{a^{b+1}}$$

provided that $b > -1$, i.e., provided that $l + l' + k + 2 \geq 0$.

The integral evaluates to

$$\tfrac{1}{2}\frac{1}{(2Z)^k}\left(\frac{1}{(2l+1)!(l+l'+1)!(l-l')!}\right)^{1/2}\frac{(l+l'+k+2)!(l-l'+k+1)!}{(k+1)!}, \quad k > -2,$$

$$\text{or } \tfrac{1}{2}\frac{1}{(2Z)^k}\left(\frac{1}{(2l+1)!(l+l'+1)!(l-l')!}\right)^{1/2}\frac{(-k-2)!(l+l'+k+2)!}{(l'-l-k-2)!}, \quad k \leq -2.$$

However, since

$$C^{l\ \ k\ \ l'}_{l+1\ 0\ l+1} = (-1)^{k+q}\left(\frac{(2l'+1)}{2l+1}\right)^{1/2}C^{l'\ \ k\ \ l}_{l+1\ 0\ l+1} = \frac{(l+l'+1)}{(l-l')!(k+1)!}\left(\frac{(l+l'+1)!}{(l-l')!(2l+1)!}\right)^{1/2}, \quad k > -2,$$

$$\text{or } \left(\frac{(2l'+1)(l-l'-k-2)!(l+l'+k+2)!(l+l'-k-1)!}{(2l+1)!(l'-l)!(l'-l-k-2)!(l+l'+1)!}\right)^{1/2}, \quad k \leq -2,$$

we have

$$\langle l'||T^k||l\rangle = \tfrac{1}{2}\frac{1}{(2Z)^k}\frac{(l+l'+k+2)!(l-l')!}{(l+l'+1)!}, \quad k > -2,$$

or

$$\tfrac{1}{2}\frac{1}{(2Z)^k}\left(\frac{(l+l'+k+2)!}{(l'-l-k-2)!(l-l'-k-2)!(l+l'-k-1)!(2l'+1)}\right)^{1/2}.$$

Since $D_{n/n} = 1$ and the integral through Ω is 1, T^k is proportional to a matrix element of r^k, diagonal in n, namely

$$\begin{aligned}\int_0^\infty R_n r^k R_{nl'}r^2dr &= \langle nlm|r^k|nl'm\rangle \\ &= n^{k-1}\langle nlm|T^k|nl'm\rangle, \quad k > -2, \\ &= n^{k-1}(-k-2)!\,\langle nlm|T^k|nl'm\rangle, \\ &\quad k \leq -2. \end{aligned} \tag{23}$$

Off-diagonal elements of r^k, however, are much harder to deal with. Placing a complete set of states in the expression

$$\langle n_p l_p m|N^k_{n_p n_p}r^k D_{n_q/n_p}\cdot D_{n_p/n_q}A^{l_q}_{l_p}|n_q l_q m\rangle$$

(where $N^k_{n_p n_q}$ is the normalization required to make $r^k D_{n_q/n_p}A^{l_q}_{l_p}$ into a tensor operator) will not

do because $r^k D_{n_q/n_p} A^{l_q}_{l_p}$ is a tensor operator between the states $|n_p l_p m\rangle$ and $|n_q l_q m\rangle$ only. A formal solution to this problem can be given by inverting the infinite matrix formed by putting a complete set of states in the expression $\langle n_p l_p m | T^k_{n_p - n_r} | n_r l_r m\rangle$. Thus

$$\begin{aligned}\langle n_p l_p m | T^k_{n_p-n_r} | n_r l_r m\rangle &\equiv \langle n_p l_p m | N^k_{n_p n_r} r^k D_{n_r/n_p} A^{l_r}_{l_p} | n_r l_r m\rangle \\ &= \sum_{n_q} \langle n_p l_p m | r^k A^{l_r}_{l_p} | n_q l_r m\rangle \\ &\quad \times \langle n_q l_r m | N^k_{n_p n_r} D_{n_r/n_p} | n_r l_r m\rangle,\end{aligned}$$

since D_{n_r/n_p} is diagonal in l and m.[4,5] So

$$\begin{aligned}\langle n_p l_p m | T^k_{n_p-n_r} | n_r l_r m\rangle &= \sum_{n_q} A_{qr} \int_0^\infty R_{n_p l_p} r^k R_{n_q l_r} r^2 dr,\end{aligned}$$

where A is the infinite matrix whose elements are

$$A_{qr} = \langle n_q l_r m | N^k_{n_p n_r} D_{n_r/n_p} | n_r l_r m\rangle.$$

This in principle is invertible to give

$$\int_0^\infty R_{n_p l_p} r^k R_{n_q l_q} r^2 dr = \sum_{n_q} A^{-1}_{qr} \langle n_p l_p m | T^k_{n_p - n_q} | n_q l_r m\rangle.$$

VI. CONCLUSIONS

In this paper, we have shown how to circumvent some of the difficulties encountered when one uses the "natural" group scheme for the hydrogen atom, namely $O(4,2) \supset O(3) \times O(2,1)$, to explain the selection rule of Pasternack and Sternheimer and to derive corresponding matrix elements. A complete treatment of off-diagonal matrix elements is not yet available, since in this group scheme it is $r^k D_{n/n'}$ and not r^k, as might be hoped, that is proportional to a tensor operator.

To apply these methods to more complicated atoms requires at least an approximate dynamical group that can be factorized into angular and radial parts. This would seem to be one of the directions in which group theoretical atomic spectroscopy could be heading.

ACKNOWLEDGMENT

The author would like to thank Dr. L. Armstrong for communicating his manuscript before publication and for pointing out some errors in the first manuscript.

APPENDIX: SOME COMBINATORIAL IDENTITIES

The addition theorem for binomial coefficients, namely

$$\sum_s \binom{x}{s}\binom{y}{z-s} = \binom{x+y}{s}$$

gives

(a) $$\frac{a!}{b!c!} = \sum_s \frac{(a-b)!(a-c)!}{(a-b-s)!(a-c-s)!(b+c-a+s)!s!}$$

if x and y are positive,

(b) $$\sum_s (-1)^s \frac{(a-s)!}{s!(b-s)!(c-s)!} = (-1)^c \frac{(a-c)!(b+c-a-1)!}{b!c!(b-a-1)!}$$

if y is negative and $b > a \geq c \geq 0$,

or

$$\frac{(a-b)!(a-c)!}{b!c!(a-b-c)!}$$

if y is negative and $a \geq b \geq 0, a \geq c \geq 0$,

(c) $$\frac{a!}{b!c!} = \sum_s \frac{(-1)^s (a-c)!(s-a+b-1)!}{(b-a-1)!s!(a-c-s)!(b-a+c+s)!}$$

if x is negative.

The following are the identities mentioned in this paper.

I. $$\sum_t \frac{(a-t)!}{(b-t)!(c-t)!(d-t)!(e+t)!t!} = \sum_{t,s} \frac{(a-b)!(a-c)!}{(a-b-s)!(a-c-s)!(b+c-a+s-t)!s!(d-t)!(e+t)!t!},$$

by (a),

$$= \frac{(a-b)!(a-c)!}{d!(d+e)!} \sum_s \frac{(b+c+d+e-a+s)!}{(b+c+e-a+s)!(b+c-a+s)!(a-b-s)!(a-c-s)!s!}$$

also by (a).

Equivalently

$$\sum_s \frac{(a+s)!}{(b+s)!(c+s)!(d-s)!(e-s)!s!} = \frac{(a-b)!(a-c)!}{d!e!} \sum_t \frac{(d+e+c-t)!}{(a-b-t)!(e+c-t)!(d+c-t)!(b-c+t)!t!}.$$

II. $$\sum_s (-1)^s \frac{(a-s)!(b-s)!}{s!(c-s)!(d-s)!(e-s)!} = \sum_{s,t} \frac{(-1)^s (a-c)!(a-d)!(b-s)!}{t!(a-c-t)!(a-d-t)!(c+d-a-s+t)!(e-s)!s!},$$

by (a),

$$= \sum_t \frac{(a-c)!(a-d)!(b-e)!(b+a-c-d-t)!}{e!(a-c-t)!t!(a-d-t)!(c+d-a+t)!(b+a-c-d-e-t)!}$$

by (b).

$$\text{III.} \sum_t \frac{(a-t)!}{(b-t)!(c-t)!(d-t)!(e+t)!t!} = \frac{(a-c)!}{d!(d+e)!(b-a-1)!} \times \sum_s \frac{(-1)^s(b-a+c+d+e+s)!(b-a-1+s)!}{(b-a+c+e+s)!s!(a-c-s)!(b-a+c+s)!}, \quad a-b \text{ negative},$$

by (c) and (a).

Equivalently

$$\sum_s \frac{(-1)^s(a+s)!(b+s)!}{(c+s)!(d+s)!(e-s)!s!} = \frac{(a-c)!(a-d)!b!}{e!} \sum_t \frac{(e+d-b-1-t)!}{(e+d-t)!(c-d+t)!(d-b-1-t)!(a-c-t)!}.$$

[1] S. Pasternack and R. M. Sternheimer, J. Math. Phys. 3, 1280 (1962).
[2] N. V. V. J. Swamy, R. G. Kulkarni and L. C. Biedenharn, J. Math. Phys. 11, 1165 (1970).
[3] L. Armstrong, Jr., J. Math. Phys. 12, 953 (1971).
[4] A. O. Barut and H. M. Kleinert, Phys. Rev. 156, 1541 (1967); 157, 1180 (1967); 160, 1149 (1967).
[5] H. M. Kleinert, Fortschr. Physik 16, (1968).
[6] V. Bargmann, Ann. Math. 48, 568 (1947).
[7] A. O. Barut and C. Fronsdal, Proc. Roy. Soc. (London) A287, 532 (1965).
[8] L. Infeld and T. E. Hull, Rev. Mod. Phys. 23, 21 (1951).
[9] A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1962), Vol. II.
[10] V. Bargmann, Rev. Mod. Phys. 34, 829 (1962).
[11] B. L. van der Waerden, *Die gruppentheoretische Methode in der Quantenmechanik* (Springer-Verlag, Berlin, 1932).
[12] G. Racah, Phys. Rev. 62, 438 (1942).

Example Related to the Foundations of Quantum Theory

E. B. Davies
St. John's College, Oxford, England

(Received 30 October 1970)

We construct a very simple example of a statistical system which satisfies almost all of the axioms used by various authors in the quantum logic approach to the foundations of quantum mechanics. Since the example is not quantum mechanical, it is seen that present arguments are a long way from characterising quantum mechanics in terms of a set of physically meaningful axioms.

1. INTRODUCTION

In the last ten years, there have been a large number of papers attempting to derive the accepted structure of quantum mechanics from certain axioms involving the existence of a quantum logic, a concept which first appeared in a paper by Birkhoff and von Neumann.[1] Typically these works start with a general discussion of the statistical nature of measurements and obtain axioms which are either claimed to be necessary to any physical theory of a statistical nature, or, more modestly, to be general physically meaningful assumptions which seem to be satisfied in quantum theory, but each of which could in principle be refuted by experiment. From the axioms, the authors attempt either to derive the Hilbert space structure of quantum mechanics, or at least a Jordan algebra whose states are the states of the physical system.

We mention a few papers among the very large number in the field. Pool[2] discusses a collection of axioms which are representative of those generally taken in studies of quantum logic. He does not claim to characterize quantum mechanics and, indeed, pathological systems of this degree of generality are known.[3] The papers by Zierler,[4] Piron,[5] and Gunson[6] have come close to a characterization of the von Neuman model of quantum mechanics in terms of physically plausible assumptions. The first two authors base their work on detailed assumptions about the partially ordered orthocomplemented set, while Gunson uses a quite different algebraic approach.

In this paper, we consider in detail one very simple example of a statistical system, whose properties indicate that quantum mechanics is only one of a large class of statistical theories with very similar general properties. Our example satisfies all the axioms of Pool but cannot be described in terms of any Hilbert space or Jordan algebra. Moreover, the example satisfies all the axioms of Gunson except for one technical condition which seems not to have any physical interpretation even though it is the crucial assumption in Gunson's arguments.

by (a),

$$= \sum_t \frac{(a-c)!(a-d)!(b-e)!(b+a-c-d-t)!}{e!(a-c-t)!t!(a-d-t)!(c+d-a+t)!(b+a-c-d-e-t)!}$$

by (b).

$$\text{III.} \sum_t \frac{(a-t)!}{(b-t)!(c-t)!(d-t)!(e+t)!t!} = \frac{(a-c)!}{d!(d+e)!(b-a-1)!} \times \sum_s \frac{(-1)^s(b-a+c+d+e+s)!(b-a-1+s)!}{(b-a+c+e+s)!s!(a-c-s)!(b-a+c+s)!}, \quad a-b \text{ negative},$$

by (c) and (a).

Equivalently

$$\sum_s \frac{(-1)^s(a+s)!(b+s)!}{(c+s)!(d+s)!(e-s)!s!} = \frac{(a-c)!(a-d)!b!}{e!} \sum_t \frac{(e+d-b-1-t)!}{(e+d-t)!(c-d+t)!(d-b-1-t)!(a-c-t)!}.$$

[1] S. Pasternack and R. M. Sternheimer, J. Math. Phys. 3, 1280 (1962).
[2] N. V. V. J. Swamy, R. G. Kulkarni and L. C. Biedenharn, J. Math. Phys. 11, 1165 (1970).
[3] L. Armstrong, Jr., J. Math. Phys. 12, 953 (1971).
[4] A. O. Barut and H. M. Kleinert, Phys. Rev. 156, 1541 (1967); 157, 1180 (1967); 160, 1149 (1967).
[5] H. M. Kleinert, Fortschr. Physik 16, (1968).
[6] V. Bargmann, Ann. Math. 48, 568 (1947).
[7] A. O. Barut and C. Fronsdal, Proc. Roy. Soc. (London) A287, 532 (1965).
[8] L. Infeld and T. E. Hull, Rev. Mod. Phys. 23, 21 (1951).
[9] A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1962), Vol. II.
[10] V. Bargmann, Rev. Mod. Phys. 34, 829 (1962).
[11] B. L. van der Waerden, *Die gruppentheoretische Methode in der Quantenmechanik* (Springer-Verlag, Berlin, 1932).
[12] G. Racah, Phys. Rev. 62, 438 (1942).

Example Related to the Foundations of Quantum Theory

E. B. Davies
St. John's College, Oxford, England

(Received 30 October 1970)

We construct a very simple example of a statistical system which satisfies almost all of the axioms used by various authors in the quantum logic approach to the foundations of quantum mechanics. Since the example is not quantum mechanical, it is seen that present arguments are a long way from characterising quantum mechanics in terms of a set of physically meaningful axioms.

1. INTRODUCTION

In the last ten years, there have been a large number of papers attempting to derive the accepted structure of quantum mechanics from certain axioms involving the existence of a quantum logic, a concept which first appeared in a paper by Birkhoff and von Neumann.[1] Typically these works start with a general discussion of the statistical nature of measurements and obtain axioms which are either claimed to be necessary to any physical theory of a statistical nature, or, more modestly, to be general physically meaningful assumptions which seem to be satisfied in quantum theory, but each of which could in principle be refuted by experiment. From the axioms, the authors attempt either to derive the Hilbert space structure of quantum mechanics, or at least a Jordan algebra whose states are the states of the physical system.

We mention a few papers among the very large number in the field. Pool[2] discusses a collection of axioms which are representative of those generally taken in studies of quantum logic. He does not claim to characterize quantum mechanics and, indeed, pathological systems of this degree of generality are known.[3] The papers by Zierler,[4] Piron,[5] and Gunson[6] have come close to a characterization of the von Neuman model of quantum mechanics in terms of physically plausible assumptions. The first two authors base their work on detailed assumptions about the partially ordered orthocomplemented set, while Gunson uses a quite different algebraic approach.

In this paper, we consider in detail one very simple example of a statistical system, whose properties indicate that quantum mechanics is only one of a large class of statistical theories with very similar general properties. Our example satisfies all the axioms of Pool but cannot be described in terms of any Hilbert space or Jordan algebra. Moreover, the example satisfies all the axioms of Gunson except for one technical condition which seems not to have any physical interpretation even though it is the crucial assumption in Gunson's arguments.

We do not suppose the example has any direct physical significance, but just use it to prove that the structure of quantum mechanics cannot be deduced from purely philosophical discussions of its statistical nature, and also that present arguments are a long way from reducing quantum theory to any collection of physically meaningful principles.

2. THE EXAMPLE

The particular example is set up to describe the internal degrees of freedom of a conceptual elementary particle, and is finite dimensional. The example can be conveniently described within the framework of Ref. 7, where the basic notion is the set of states.

We let V be the real three-dimensional vector space of triples $(\alpha_1, \alpha_2, \alpha_3)$, where $\alpha_i \in R$ for $i = 1, 2, 3$. V is given the positive cone

$$V^+ = \{(\alpha_1, \alpha_2, \alpha_3): 0 \leq |\alpha_1| + |\alpha_2| \leq \alpha_3\}$$

and the positive linear functional

$$\tau(\alpha_1, \alpha_2, \alpha_3) = \alpha_3.$$

The triple (V, V^+, τ) is called the *state space* and the set

$$K = \{\rho \in V^+: \tau(\rho) = 1\}$$

is called the set of *normalized states*. K is a compact convex set (a square) and it has four extreme points

$$a = (1, 0, 1), \qquad b = (0, 1, 1),$$
$$a' = (-1, 0, 1), \qquad b' = (0, -1, 1),$$

which we call *pure states*. It is convenient to work with unnormalized states (elements of V^+).

An *operation* is defined as a positive linear map T: $V \to V$ such that $\tau[T\rho] \leq \tau[\rho]$ for all $\rho \in V^+$, and is interpreted as a description of a test on an ensemble of copies comprising a state $\rho \in K$, where $\tau[T\rho]$, which lies between 0 and 1, gives the proportion of the ensemble being passed through the test and $(T\rho)/\tau[T\rho]$ gives the conditional output state provided that $\tau[T\rho] \neq 0$. The two notions of state and operation are all that are really necessary to describe a statistical system. However, to relate our example to the quantum logic approach, we need to discuss the idea of repeatability of measurements.

It was recognised by Birkhoff and von Neumann, and most subsequent authors, that the importance of projections in quantum mechanics is related to the repeatability hypothesis (see also Ref. 7). Now while position and spin measurements on single particles can be considered as repeatable, the great majority of measurements in physics are not of this type. For example, measurement of the energy of a bound electron in an excited state is done by observing photons emitted during transitions, so that this kind of observation is the opposite of repeatable. In second quantized theories there seem to be no repeatable observables in common use. For example, in quantum optics,[8] one can find the distribution of the number of photons arriving at a counter within a fixed time. The expected number of such photons can be calculated from a self-adjoint operator, but the distribution is not obtained from the spectral resolution of that operator. In high energy physics only the coefficients of the S matrix are ever actually measured, and for these the repeatability hypothesis is meaningless. Finally the fact that position observations for photons cannot be described in terms of projection-valued measures[9] seems to be related to the fact that observation of a photon involves its absorption and annihilation, so that repeatability is violated.

In spite of these doubts about the fundamental nature of the repeatability hypothesis, we now formulate it mathematically and determine its significance in our example.

We define a *filter* as a pair $\mathcal{E}_0, \mathcal{E}_1$ of operations satisfying

(i) $\tau[\mathcal{E}_0\rho + \mathcal{E}_1\rho] = \tau[\rho]$ for all $\rho \in V$,

(ii) $\mathcal{E}_i\mathcal{E}_j\rho = \delta_{ij}\mathcal{E}_j\rho$ for any i, j,

(iii) if $\tau[\mathcal{E}_i\rho] = \tau[\rho]$, then $\mathcal{E}_i\rho = \rho$ for $i = 0$ or 1.

A filter is interpreted as a procedure for dividing an ensemble into two subensembles depending on the result of some test. The first condition states that every member of the ensemble is put into one of the two subensembles. The second states that if the test is repeated, then the result is still the same and the states are not modified at all by the second measurement. The third states that if all copies of an ensemble give the same result in the test, then the state after the test is the same as that before.

There are two simple filters, which we call N and Y, given by

$$\left.\begin{aligned} N_0\rho = \rho, \quad N_1\rho = 0 \\ Y_0\rho = 0, \quad Y_1\rho = \rho \end{aligned}\right\} \quad \text{for all } \rho \in V.$$

To classify all filters, one has only to observe that, given a filter different from the above two, the sets

$$S_0 = \{\rho \in K: \mathcal{E}_1\rho = 0\}, \qquad S_1 = \{\rho \in K: \mathcal{E}_1\rho = \rho\}$$

are disjoint, nonempty, proper faces of K. Moreover, the linear functional

$$A(\rho) = \tau[\mathcal{E}_1\rho]$$

satisfies $0 \leq A(\rho) \leq 1$ for all $\rho \in K$ and

$$S_0 = \{\rho \in K: A(\rho) = 0\}, \qquad S_1 = \{\rho \in K: A(\rho) = 1\}.$$

Simple case considerations now show that S_0 and S_1 must be opposite vertices of K. The most general filter can now be written down as follows. Choose two opposite vertices of K, say a and a', and a linear functional A on V such that

$$\{a'\} = \{\rho \in K: A(\rho) = 0\}, \qquad \{a\} = \{\rho \in K: A(\rho) = 1\}.$$

Then define

$$\mathcal{E}_0(\rho) = \{\tau[\rho] - A(\rho)\}\, a', \qquad \mathcal{E}_1(\rho) = A(\rho) = a(\rho)a.$$

It is seen that the filter is uniquely determined by the functional A. However, there exist an infinite number of filters for each pair of opposite vertices of K.

We can attempt to make the set of filters into a partially ordered orthocomplemented set. The complementation should clearly be defined by

$$\mathcal{E}_0' = \mathcal{E}_1, \quad \mathcal{E}_1' = \mathcal{E}_0.$$

We consider two possible definitions of the implication relation, namely

(i) $\mathcal{E} \to {}_1\mathcal{F}$ means $\tau[\mathcal{E}_1\rho] \leq \tau[\mathcal{F}_1\rho]$ for all $\rho \in V^+$

(ii) $\mathcal{E} \to {}_2\mathcal{F}$ means if $\rho \in V^+$ and $\tau[\mathcal{E}_1\rho] = \tau[\rho]$, then $\tau[\mathcal{F}_1\rho] = \tau[\rho]$.

The first of these definitions was used by Piron[5] and the second by Pool.[2] It may be seen that, in this example, there exist two different filters $\mathcal{E}, \mathcal{F}$ such that $\mathcal{E} \to {}_2\mathcal{F}$ and $\mathcal{F} \to {}_2\mathcal{E}$, so that $\to_2$ does not define a partial ordering. The first definition is statistical rather than logical in nature, and turns the set of filters into a partially ordered orthocomplemented lattice.

We have shown in Ref. 7 that in quantum mechanics the filters correspond one-to-one to the projections and that in this case the definitions $\to_1$ and $\to_2$ coincide; indeed, this was remarked by Birkhoff and von Neumann.[1] Within this context our system therefore has significant differences from quantum mechanics. This can be overcome by defining the set of questions to correspond to a subset of the set of filters in some arbitrary but convenient manner. We now discuss one particular choice of the set of questions.

We define V, V^+, τ, and K as before. We define Q to be the following set of functionals on V:

$$0(\rho) = 0, \quad 1(\rho) = \tau[\rho],$$
$$A(\alpha, \beta, \gamma) = \tfrac{1}{2}\alpha + \tfrac{1}{2}\gamma, \quad B(\alpha, \beta, \gamma) = \tfrac{1}{2}\beta + \tfrac{1}{2}\gamma,$$
$$A' = 1 - A, \quad B' = 1 - B.$$

For $X, Y \in Q$ we define complementation by $X' = 1 - X$ and implication $X \to Y$ by $X(\rho) \leq Y(\rho)$ for all $\rho \in V^+$. This turns Q into an orthocomplemented lattice. We define $P: Q \times K \to [0, 1]$ by

$$P(X, \rho) = X(\rho).$$

We define an operation $\mathcal{E}_X: V \to V$ for every $X \in Q$ as follows: $\mathcal{E}_0(\rho) = 0$ and $\mathcal{E}_1(\rho) = \rho$ for all $\rho \in V$. For any other $X \in Q$, let x_X be the unique normalized state such that $P(X, x_X) = 1$ and define $\mathcal{E}_X$ by

$$\mathcal{E}_X(\rho) = P(X, \rho)x_X.$$

It may be easily verified that K, Q, P is an event-state structure in the sense of Pool[2] and that $K, Q, P, \mathcal{E}$ is an event-state-operation structure, except for a slight technical difference in the definition of an operation.

Our example also satisfies all the axioms of Gunson[6] except for his Axiom A. 10 which states essentially:

"There exists a one-to-one linear map $\sigma: V \to V'$ which maps the cone V^+ onto a dense subset of the positive cone of V', taking pure states into atomic propositions and such that $\langle g, \sigma(f)\rangle \leq 1$ where f, g are pure states, equality being attained only for $f = g$."

However, our example does satisfy this axiom provided the single word "dense" is deleted. Gunson's density assumption seems not to have any physical interpretation, but to be a mathematical device introduced in order to use the theory of self-adjoint cones. We remark also that even if there is a one-to-one correspondence between pure states and atomic propositions, we see no physical reason why this should have a one-to-one linear extension as Gunson supposes, even though our example does have this property.

Finally we remark that our example is by no means exceptional. The set K may be replaced by any convex body in any number of dimensions or even by any compact convex set in a locally convex topological linear space. By suitable choices of K, one may obtain classical probability theory, quantum mechanics, or a large variety of other statistical theories. We do not, however, maintain that any conceivable statistical theory can be represented in this manner.

ACKNOWLEDGMENTS

The author would like to thank J. T. Lewis for his criticism of an earlier version of this paper, and the referee for some very valuable comments.

[1] G. Birkhoff and J. von Neumann, Ann. Math. **37**, 823 (1936).
[2] J. C. T. Pool, Commun. Math. Phys. **9**, 118 (1968).
[3] R. J. Greechie, Caribbean J. Science Math. **1**, 15.
[4] N. Zierler, Pacific J. Math. **11**, 1151 (1961).
[5] C. Piron, Helv. Phys. Acta **37**, 439 (1964).
[6] J. Gunson, Commun. Math. Phys. **6**, 262 (1967).
[7] E. B. Davies and J. T. Lewis, Commun. Math. Phys. **17**, 239 (1970).
[8] J. R. Klauder and E. C. G. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1968).
[9] A. S. Wightman, Rev. Mod. Phys. **34**, 845 (1962).

Uniform Asymptotic Solution of the Radial Schrödinger Equation for an Electron Bound in a Central Field Potential*

Erich Zauderer
Polytechnic Institute of Brooklyn, Brooklyn, New York
(Received 15 September 1970)

Assuming Z (the nuclear charge) is large, it is shown, using a theory of McKelvey, how a uniform asymptotic solution of the radial Schrödinger equation for an electron bound in a central field potential and for a related problem may be obtained. Requiring the asymptotic solutions to be bounded leads to a determination of the energy expansion parameters.

1. INTRODUCTION

The radial Schrödinger equation for an electron bound in a central field potential has been studied by Iafrate and Mendelsohn,[1,2] using a method which combines a large-Z (the nuclear charge) asymptotic expansion theory with the conventional perturbation theory approach. Applying the stretching and matching method (occasionally referred to as the method of matched asymptotic expansions) familiar in fluid dynamics, they match the "outer" large-Z asymptotic expansion to the "inner" perturbation theory expansion, thereby specifying unknown terms in both expansions. This method has also been applied by Mendelsohn[3] to a repulsive central field potential problem for a one-electron ion.

In this note we observe that the relevant Schrödinger equations for the above problems belong to a class of equations of the form

$$\frac{d^2u}{dx^2} - [Z^2\phi^2(x) + Zx^{-1}\Psi(x, Z) + x^{-2}\tau]u(x) = 0, \quad (1)$$

for which McKelvey[4] obtained uniform asymptotic expansions when Z is a large parameter. McKelvey's approach is based on the related equation technique developed by Langer for turning point problems and results in an asymptotic expansion of $u(x)$ in terms of Whittaker functions and their derivatives. On applying McKelvey's method to the aforementioned Schrödinger equations, it will become necessary to expand the energy in inverse powers of Z in order that the solutions remain bounded. The resulting asymptotic expansions are uniformly valid in the "inner" region, i.e., near the origin, as well as in the "outer" region, and do not require any "matching" to be carried out. The expansions will be compared with those obtained by the matched asymptotic expansion and will be seen to be closely related because of the special form of the problems. We do not examine any particular case in detail or obtain explicit uniform asymptotic expansions, but rather rely on the method of McKelvey and the general form of his asymptotic solution to draw certain conclusions. Detailed results have been obtained by Iafrate and Mendelsohn[1-3] for these problems using their method, and they have compared their results to other work.

2. THE UNIFORM ASYMPTOTIC EXPANSION

The radial Schrödinger equation for an electron bound in a generalized central field potential of the form

$$V(r) = -\frac{Z}{r}B(\lambda r) = -\frac{Z}{r}\sum_{j=0}^{\infty} B_j(\lambda r)^j, \quad B_0 \equiv 1 \quad (2)$$

can be written as

$$-\frac{1}{2}\frac{d^2\psi}{dr^2} - \frac{1}{r}\frac{d\psi}{dr} - \frac{Z}{r}B(\lambda r)\psi + \frac{1}{2}\frac{l(l+1)}{r^2}\psi = E\psi. \quad (3)$$

We also consider the Schrödinger equation for a one-electron ion in a repulsive central field potential of the type considered by Mendelsohn[3]

$$-\frac{1}{2}\frac{d^2\psi}{dr^2} - \frac{1}{r}\frac{d\psi}{dr} - \frac{Z}{r}\psi + \lambda r^N\psi = E\psi, \quad (4)$$

where N is a positive integer. To bring (3) and (4) into the general form (1), we set

$$\psi(r) = (1/r)v(r) \quad (5)$$

to obtain for (3)

$$v''(r) + [2E + (2Z/r)B(\lambda r) - l(l+1)/r^2]v(r) = 0, \quad (6)$$

and for (4) an equation like (4) except that the first derivative term is absent. As $\lambda \to 0$ in (4) and (6), we expect that the energy E reduces to that for the hydrogen atom. That is, $E \to -\frac{1}{2}Z^2/n^2$ as $\lambda \to 0$ in (6), where n is the principal quantum number and $E \to -\frac{1}{2}Z^2$ in (4), where only the ground state is considered. We assume that E has an expansion of the form

$$E = -\tfrac{1}{2}\epsilon^2 Z + \sum_{m=-1}^{\infty} \epsilon_m Z^{-m}, \quad (7)$$

with ϵ and ϵ_m as yet unspecified. Letting $X = \epsilon r$ in (6), dividing out ϵ^2, and using (7), we have

$$v''(x) - \left[Z^2 - \frac{2Z}{\epsilon x}B\left(\frac{\lambda}{\epsilon}x\right) + \frac{x\sum_{m=-1}^{\infty}\epsilon_m Z^{-m}}{\epsilon^2 x} + \frac{l(l+1)}{x^2}\right]v(x) = 0. \quad (8)$$

A similar equation with the x^{-2} term absent results from the modified form of (4).

Equation (8) is now in the form (1), and we may apply McKelvey's theory to it. We have divided out ϵ^2 in (8) since $\phi^2(x)$ is assumed to be normalized so that $\phi(0) = 1$. Further, $\Psi(x, Z)$ in (1) has the form

$$\Psi(x, Z) = \psi(x) + \sum_{j=0}^{\infty}\frac{\psi_j(x)}{Z^{j+1}}, \quad (9)$$

where $\psi(x)$ and $\psi_j(x)$ are nonsingular at the origin and analytic, so that $\Psi(x, Z)$ corresponds to the two middle terms in the bracket in (8), implying that $B(\lambda r)$ must be nonsingular and analytic. Also, the term τ in (1), which is required to be constant, corresponds to $l(l+1)$ in (8). Similar statements apply to the modified form of (4) except that $\tau = 0$ in that case.

Applying McKelvey's method to (8), where Z (the nuclear charge) is assumed to be large, we set

$$\xi(x, Z) = 2Zx \tag{10}$$

and obtain as the related equation

$$\frac{d^2M(\xi)}{d\xi^2} - \left(\frac{1}{4} - \frac{k(Z)}{\xi} + \frac{l(l+1)}{\xi^2}\right) M(\xi) = 0, \tag{11}$$

where

$$k(Z) = K + \sum_{m=0}^{\infty} \frac{K_m}{Z^{m+1}} \tag{12}$$

with the K_m to be specified and

$$K = 1/\epsilon. \tag{13}$$

For the modified form of (4) we also obtain (11), but with the ξ^{-2} term absent. In terms of a suitable solution $M(\xi)$ of (11), a uniform asymptotic solution of (8) or the modified form of (4) is obtained from McKelvey's theory in the form

$$v(x) = F(x, Z)M(\xi) + Z^{-1}G(x, Z)\left(\xi \frac{dM(\xi)}{d\xi}\right), \tag{14}$$

where F and G are expanded in the form (9) and depend on the parameters $\epsilon, \epsilon_m, K_m$, and B_j. McKelvey specifies F and G so that they are bounded at the origin and analytic. This is done by exhibiting a set of equations for quantities which comprise F and G and showing how appropriate, unique choices for the K_m lead to boundedness and analyticity for F and G.

Introducing two linearly independent solutions of Whittaker's equation (11) into (14) yields a fundamental set of uniform solutions of (1) or more specifically (8) in our case. The type of problems that may be solved for (1) depends on the behavior of solutions of (11). In our problems for the Schrödinger equations (3) and (4), we require that ψ be bounded. In view of the analyticity and boundedness of F and G in (14), this implies that we require bounded solutions for the Whittaker equation (11). This can only be achieved for values of $k(Z)$ such that

$$k(Z) = r + (l + 1) \equiv n, \tag{15}$$

where r is an integer and it will be seen in the following that n is the principal quantum number. For a continuous domain of values of Z, (15) can be possible only if

$$k(Z) \equiv K \tag{16}$$

so that all the $K_m = 0$ and $\epsilon = 1/n$ in view of (13). The solutions of (10) then take the form

$$M(\xi) = \xi^{l+1} \exp[-\xi/2] L^{2l+1}_{K-(l+1)}(\xi) \tag{17}$$

apart from a multiplicative constant. The $L^{\beta}_{\alpha}(\xi)$ are Laguerre polynomials.

Since all the K_m must vanish, we can no longer assure the analyticity and boundedness of F and G at the origin unless the extraordinary case prevails that the K_m may be taken to be zero at each step of their specification. However, the energy expansion terms ϵ_m are as yet unspecified and are available to take the place of the K_m in McKelvey's equations for specifying the analytic and bounded behavior of F and G at the origin. We have carried out the necessary calculations for the first few ϵ_m for the modified form of (4). It was seen that they can be uniquely determined and that they permit the first few terms in F and G to be specified as in McKelvey's theory. That is, they take over the role of the K_m. The results obtained agree with those of Mendelsohn.[3] Similar calculations for (8) appear to be extremely laborious because of the particular method used by McKelvey in obtaining the uniform asymptotic expansions. Recently, simplifications have been achieved in the problem of obtaining formal uniform asymptotic expansions for a large class of turning point problems,[5,6] and the ideas applied there, can no doubt be used in McKelvey's problem. For this reason, we have not tried to determine explicitly the ϵ_m, F, and G, which result for Eq. (8). However, we return to this question in the next section when we compare the uniform expansion method with that of matched asymptotic expansions.

3. DISCUSSION AND CONCLUSION

In the large-Z expansion theory of Mendelsohn[1-3] a formal "outer" expansion of the form

$$\psi(r) = \exp[-ZS(r)] \sum_{m=0}^{\infty} a_m(r) Z^{-m} \tag{18}$$

is assumed to satisfy (3) or (4). Introducing (18) into (3) or (4) and setting coefficients of different powers of Z equal to zero yields a set of equations for $S(r)$ and the $a_m(r)$. The appropriate solution of the equation for $S(r)$ is

$$S(r) = r/n, \tag{19}$$

where n is the principal quantum number. The $a_m(r)$ can be solved for recursively and determined up to a set of constants. They may become singular at the origin where (18) is not necessarily expected to be valid. The conventional perturbation theory approach is identified by Mendelsohn with a stretching of the coordinate r near the origin, i.e., introducing a new variable $\hat{\xi}$ given by

$$\hat{\xi} = Zr \tag{20}$$

and expanding the solution $\Psi(\hat{\xi})$ of the stretched differential equation in the form

$$\Psi(\xi) = \sum_{j=0}^{\infty} \Psi_j(\hat{\xi}) \omega^j, \tag{21}$$

where $\omega = \lambda/Z$ for (3) and $\omega = \lambda/Z^{N+2}$ for (4). This gives the "inner" expansion. In addition, via a perturbation theory argument, E/Z^2, where E is the energy term in (3) and (4), is also expanded in powers of ω as in (21). This expansion of E is used not only in obtaining (21) but also in obtaining (18). Requiring the Ψ_j to be nonsingular near the origin leads to a unique determination of the energy expansion coefficients to arbitrary order. The equations satisfied by the Ψ_j are closely related to (11) except that for $j > 0$ they are inhomogeneous equations. Once the perturbation expansion (21) is

obtained, undetermined constants in the $a_m(r)$ and $\Psi_j(\hat{\xi})$ are specified by applying the matching procedure. It was observed by Iafrate and Mendelsohn in the cases they considered that the perturbation theory result to a given order is obtained by using the large-Z expansion (18) to some order and expressing it in terms of the stretched variable $\hat{\xi}$. We note that both $\psi(r)$ and $\Psi(Zr)$ contain the multiplicative factor $\exp[-(Z/n)r]$.

The uniform asymptotic expansion (14) is seen to involve the two variables $x = r/n$ and $\xi = 2Zx = 2Zr/n = 2\hat{\xi}/n$, so that it simultaneously involves the unstretched variable r and the stretched variable $\hat{\xi}$. This dependence on two variables is a common feature of uniform asymptotic expansions and distinguishes them from the "inner" and "outer" expansions of the theory of matched asymptotic expansions [Eqs. (21) and (18) in our case], each of which involves only one variable for one-dimensional problems. By expressing the uniform expansion in the stretched and unstretched variables separately, one expects to retrieve the inner and outer expansions, respectively. In general, the solutions of the related equations of the uniform expansion theory involve higher transcendental functions such as Whittaker or Airy functions which have markedly different behavior for small and large values of their arguments. In our special case, the functions $M(\xi)$ and $M'(\xi)$ have the simple form of the product of an exponential term with linear exponent and a polynomial. Expressed in terms of r, the exponential occurring in $M(\xi)$ and $M'(\xi)$ is $\exp[-(Z/n)r]$, which is identical to that arising in (18) and (21). Letting $\xi = 2Zr/n$, we may write the uniform expansion for ψ in the form

$$\psi(r) = \exp\left(-\frac{Zr}{n}\right)\left[\frac{1}{r}F(x,Z)M(\xi)e^{\xi/2} + \frac{1}{Zr}G(x,Z)\left(\xi\frac{dM(\xi)}{d\xi}\right)e^{\xi/2}\right]_{\xi=2Zr/n} \quad (22)$$

and rearrange the terms in the brackets in decreasing powers of Z. This gives rise to an "outer" expansion form (18). It suggests that we may begin by directly using the "outer" expansion and requiring the solution to be nonsingular at the origin. On the other hand, if we set $r = \hat{\xi}/Z$ and $\xi = 2\hat{\xi}/n$ in (22) and expand the terms in the second bracket in the form (21), we should get the perturbation theory result of Iafrate and Mendelsohn. Now, in the uniform expansion for (4) the Laguerre Polynomial that enters in (17) is $L_0^1 \equiv 1$ so that the distinction between the stretched and the unstretched variable form of (22) involves only a linear term in ξ which enters in $M(\xi)$ and $M'(\xi)$. In that case, the uniform expansion is practically identical to the "outer" expansion form (18). For the problem involving (3), there are two Laguerre polynomials, $L_{n-(l+1)}^{2l+1}(\xi)$ and $L_{n-1-(l+1)}^{2l+1}(\xi)$, which enter into (14). In the perturbation expansions obtained by Iafrate and Mendelsohn,[1,2] a number of Laguerre polynomials enter, all having the same upper index $2l + 1$. Via recurrence properties of the Laguerre polynomials, they can all be expressed in terms of the above Laguerre polynomials. Since, in expanding (22) in terms of $\hat{\xi}$, all the functions must be expanded in power series and every polynomial can be expressed in terms of Laguerre polynomials, it is difficult to relate the perturbation theory expansion to the uniform expansion in a direct way. Again, however, the uniform expansion is very similar to the "outer" expansion (18) since only polynomials in the stretched variables enter in (22).

In conclusion, we observe that the uniform asymptotic solutions of (3) and (4) are not identical in form to either the "outer" or "inner" expansions of Iafrate and Mendelsohn. However, in view of the simple form of the solutions $M(\xi)$ of the related equations, given in (17), the uniform expansion is very similar in form to the "outer" expansion (18). Thus, while we have not proven directly for all cases that bounded uniform solutions can be obtained by a unique specification of the energy expansion coefficients, the results of Iafrate and Mendelsohn imply that it is so. The ease in matching the "inner" and "outer" expansions to any order that was experienced by Iafrate and Mendelsohn can be explained by the similarity in the uniform and the "outer" expansion forms. That is, the "outer" expansion is, in fact, valid up to the origin, with appropriate choices of the energy expansion coefficients ϵ_m. Thus a need for "matching" is obviated as we essentially have a uniform "outer" expansion. Finally, we note that to apply McKelvey's method to (6) we need only assume *a priori* that $E = O(Z^2)$ and not necessarily the expansion given in (7). However, if we then require bounded solutions, the K_m must be zero, for all m. This implies the given equation must have a form which, in fact, leads to the choices $K_m = 0$ in the process of their specification. This can be achieved, in general, only if E is expanded in the form (7) with appropriate values of ϵ_m.

---

* Permanent address: Polytechnic Institute of Brooklyn, Brooklyn, N.Y. This work was supported in part by the National Science Foundation under Grant No. GP-12835 and was completed at the Weizmann Institute of Science.

1 G. J. Iafrate and L. B. Mendelsohn, Phys. Rev. 182, 244 (1969).
2 G. J. Iafrate and L. B. Mendelsohn, Phys. Rev. A 2, 561 (1970).
3 L. B. Mendelsohn, Phys. Rev. 160, 16 (1967).
4 R. W. McKelvey, Trans. Am. Math. Soc. 91, 410, (1959).
5 R. E. Langer, Bol. Soc. Mat. Mex. 5, 1 (1960).
6 R. Y. S. Lynn and J. B. Keller, Commun. Pure Appl. Math. 23, 379 (1970).

On Canonical Commutation Relations and Infinite-Dimensional Measures

Gerhard C. Hegerfeldt
Institut für Theoretische Physik, Universität Göttingen, Göttingen, Germany
(Received 20 November 1970)

Recently, Araki has proved ray continuity for the measure-space realization of the CCR's. In this paper, stronger continuity properties of this realization are derived, from which Araki's result follows as a corollary. It is shown that $V(g)$ is continuous on complete metric subspaces of $\mathfrak{V}_\pi$ for any metric stronger than the weak topology. Surprisingly no further assumptions on the measure are needed. If $\mathfrak{V}_\pi$ is an F- or LF-space, $V(g)$ is continuous on the whole of $\mathfrak{V}_\pi$. These results are equivalent to certain continuity properties of the Radon–Nikodym derivative of the measure. Via these, it is then shown that every quasi-invariant measure on a nuclear F- or LF-space, such as S or $\mathfrak{D}$, is a superposition of ergodic measures. In the derivation, their close connection with irreducible representations of the CCR's is exploited.

1. INTRODUCTION

It has been known for some time that the canonical commutation relations (CCR's) of quantum field theory or of statistical mechanics for infinitely many degrees of freedom are closely related to quasi-invariant measures on infinite-dimensional vector spaces.[1–4] Even though the explicit form of the measure is not always easy to obtain, its very existence does lead to new structural insights into the CCR's.[5,6]

Given two real linear spaces $\mathfrak{V}_\Phi$ and $\mathfrak{V}_\pi$ and a bilinear form (f,g) on $\mathfrak{V}_\Phi \times \mathfrak{V}_\pi$, a family of unitary operators $\{U(f), V(g); f \in \mathfrak{V}_\Phi, g \in \mathfrak{V}_\pi\}$ in a Hilbert space $\mathfrak{H}$ are said to satisfy the Weyl commutation relations if

$$\begin{aligned} U(f_1)U(f_2) &= U(f_1 + f_2) \\ V(g_1)V(g_2) &= V(g_1 + g_2) \\ V(g)U(f) &= \exp\{i(f,g)\}U(f)V(g). \end{aligned} \tag{1.1}$$

This family is called a *representation of the* CCR's *over* $\mathfrak{V}_\Phi \times \mathfrak{V}_\pi$ if the bilinear form is nondegenerate and if $U(\lambda f)$ and $V(\lambda g)$ are strongly continuous in the real variable λ ("ray continuity"). In what follows the nondegeneracy of the bilinear form is not important and will not be assumed.

Now consider for the moment a normed measure on $\mathfrak{V}_\Phi^*$, the algebraic dual of some real linear space $\mathfrak{V}_\Phi$ (a subspace of $\mathfrak{V}_\Phi^*$ will also do; detailed measure-theoretic definitions will follow in Sec. 2). Let $\mathfrak{V}_\pi$ be a subspace of $\mathfrak{V}_\Phi^*$ and let μ be $\mathfrak{V}_\pi$-quasi-invariant. Then, writing $F(f) = (f, F)$ for the action of a linear functional $F \in \mathfrak{V}_\Phi^*$, one can define unitary operators $U(f)$ and $V(g)$ in the function space $L^2(\mathfrak{V}_\Phi^*, \mu)$ by

$$(U(f)\varphi)(F) = e^{i(f,F)}\varphi(F), \tag{1.2}$$

$$(V(g)\varphi)(F) = \left(\frac{d\mu(F+g)}{d\mu(F)}\right)^{1/2} \varphi(F+g). \tag{1.3}$$

These operators satisfy the Weyl commutation relations, and unitarity follows from $\mu(\mathfrak{V}_\Phi^*) = 1$. Furthermore, the vector $\Omega \in L^2(\mathfrak{V}_\Phi^*, \mu)$,

$$\Omega(F) \equiv 1, \tag{1.4}$$

is cyclic for $\{U(f)\}$. Conversely, Araki[1] has shown that every representation of the CCR's with cyclic $\{U(f)\}$ can be brought into this form, except for an additional phase factor $a_g(F)$ in the expression for $V(g)$, satisfying

$$a_{g_1}(F)a_{g_2}(F + g_1) = a_{g_1+g_2}(F). \tag{1.5}$$

Hence Eqs. (1.2) and (1.3) are, so to say, the standard measure-space realization of the CCR's or Weyl relations for given μ. Only quite recently, however, Araki[7] has shown that these operators satisfy ray continuity. For $U(f)$ this follows trivially by Lebesgue's bounded convergence theorem

$$\|U(\lambda_n f)\varphi - \varphi\|^2 = \int |e^{i\lambda_n(f,F)} - 1|^2 |\varphi(F)|^2 d\mu(F) \to 0$$

for $\lambda_n \to 0$, but for $V(g)$ this is not at all obvious.

In this paper we investigate this standard realization of the CCR's for stronger continuity properties of $V(g)$. For $U(f)$ one can again apply Lebesgue's bounded convergence theorem to obtain stronger continuity properties if, instead of $\mathfrak{V}_\Phi^*$, one uses a topological dual.[8]

It turns out that, if a subspace of $\mathfrak{V}_\pi$ is a complete metric space for some metric topology stronger than the weak topology, then $V(g)$ is already continuous on the subspace in this metric. Surprisingly, no further assumptions on $\mathfrak{V}_\Phi$ and μ are needed, nor do separability conditions enter. From this, continuity of $V(g)$ on all of $\mathfrak{V}_\pi$ follows if $\mathfrak{V}_\pi$ is an F- or LF-space, e.g., a Hilbert space, the space S or $\mathfrak{D}$ of Schwartz. Araki's[7] result on ray continuity is also a simple corollary (a short and direct proof is also given).

At the end of Sec. 3 the equivalence of these results to continuity properties of the Radon–Nikodym is pointed out.

Finally, the results are applied to the decomposition of quasi-invariant measures into ergodic ones. The latter correspond to irreducible representations of the CCR's. Recently, Hegerfeldt[8] has shown that every representation of the CCR's of the Gårding–Wightman type and every continuous representation over a nuclear test function space such as S or $\mathfrak{D}$ can be decomposed into a direct integral of irreducible representations. Applications to ergodic measures were also given. Now we use these results and the continuity properties of $V(g)$ to show that every quasi-invariant measure on a nuclear F- or LF-space is a superposition of ergodic measures.

2. BASIC NOTIONS

For the convenience of the reader, we review some basic notions and lemmas which will be of use in later sections. For additional information see Umemura[9] and Hegerfeldt.[6]

Let $\mathfrak{V}$ be a real vector space, with elements $f, \ldots$.

Denote by $\mathcal{U}^*$ its algebraic dual, consisting of all linear functionals on $\mathcal{U}$. Let $\mathcal{U}^\dagger$ be a linear subspace of $\mathcal{U}^*$, with elements $F, \ldots$, and construct a family of subsets of $\mathcal{U}^\dagger$ in the following way. First consider subsets of $\mathcal{U}^\dagger$ defined by an $F_0 \in \mathcal{U}^\dagger$, $f \in \mathcal{U}$, and a positive number λ as

$$\Theta(F_0, f, \lambda) = \{F \in \mathcal{U}^\dagger;\ |(f, F - F_0)| < \lambda\}, \quad (2.1)$$

where (f, F) denotes the action of the linear functional F on f. These sets form a base for the weak topology on $\mathcal{U}^\dagger$. Denote the σ-ring generated by these sets by $\mathfrak{B}(\mathcal{U}^\dagger, \mathcal{U})$. Its elements are often called the *Borel sets of* $\mathcal{U}^\dagger$.

Now let $\mathcal{U}_1$ be a subspace of $\mathcal{U}$. We denote by $\mathfrak{B}(\mathcal{U}^\dagger, \mathcal{U}_1)$ the σ-ring in $\mathcal{U}^\dagger$ generated by all sets $\Theta(F_0, f, \lambda)$ with $f \in \mathcal{U}_1$. Evidently $\mathfrak{B}(\mathcal{U}^\dagger, \mathcal{U}_1)$ is a sub-σ-ring of $\mathfrak{B}(\mathcal{U}^\dagger, \mathcal{U})$. If $\mathcal{W}$ is a finite-dimensional subspace of $\mathcal{U}$, then the sets in $\mathfrak{B}(\mathcal{U}^\dagger, \mathcal{W})$ are called *cylinder sets with base in* $\mathcal{W}$. Geometric visualization is easy. Denote by $\mathcal{W}^\perp$ the subspace of $\mathcal{U}^\dagger$ orthogonal to $\mathcal{W}$. Then such a cylinder set A consists of cosets $A = A + \mathcal{W}^\perp$. Since $\mathcal{U}^\dagger/\mathcal{W}^\perp$ is finite dimensional (it is isomorphic to $\mathcal{W}$ if $\mathcal{U}^\dagger$ separates $\mathcal{U}$, i.e., if $\mathcal{U}^\dagger$ is large enough), the representative points of A in $\mathcal{U}^\dagger/\mathcal{W}^\perp$ form a finite-dimensional Borel set, the "base". Denote by $\eta_{\mathcal{W}}$ the canonical map of $\mathcal{U}^\dagger$ onto $\mathcal{U}^\dagger/\mathcal{W}^\perp$,

$$\eta_{\mathcal{W}}: \mathcal{U}^\dagger \longrightarrow \mathcal{U}^\dagger/\mathcal{W}^\perp. \quad (2.2)$$

Then $\mathfrak{B}(\mathcal{U}^\dagger, \mathcal{W})$ consists of all sets of the form $\eta_{\mathcal{W}}^{-1}(\tilde{A})$, with $\tilde{A} \subset \mathcal{U}^\dagger/\mathcal{W}^\perp$ being a Borel set in the finite-dimensional space $\mathcal{U}^\dagger/\mathcal{W}^\perp$. It is clear that $\mathfrak{B}(\mathcal{U}^\dagger, \mathcal{U})$ is the σ-algebra generated by all cylinder sets.

From Eq. (2.1), one sees that $\mathfrak{B}(\mathcal{U}^\dagger, \mathcal{U})$ and $\mathfrak{B}(\mathcal{U}^\dagger, \mathcal{W})$ are invariant under translations by arbitrary elements of $\mathcal{U}^\dagger$.

By a measure μ on $\mathcal{U}^\dagger$, we mean a countably additive positive set function on the Borel sets of $\mathcal{U}^\dagger$. The restriction of μ to $\mathfrak{B}(\mathcal{U}^\dagger, \mathcal{W})$ is denoted by $\mu_{\mathcal{W}}$. From the preceding discussion it is clear that $\mu_{\mathcal{W}}$ defines a measure $\tilde{\mu}_{\mathcal{W}}$ on the finite-dimensional space $\mathcal{U}^\dagger/\mathcal{W}^\perp$ by

$$\tilde{\mu}_{\mathcal{W}}(\tilde{A}) = \mu_{\mathcal{W}}(\eta_{\mathcal{W}}^{-1}(\tilde{A})). \quad (2.3)$$

A measure ν on $\mathcal{U}^\dagger$ is called *absolutely continuous* with respect to μ, $\nu \ll \mu$, if $\mu(A) = 0$ implies $\nu(A) = 0$, for $A \in \mathfrak{B}(\mathcal{U}^\dagger, \mathcal{U})$. Note also then that $\nu_{\mathcal{W}} \ll \mu_{\mathcal{W}}$ and $\tilde{\nu}_{\mathcal{W}} \ll \tilde{\mu}_{\mathcal{W}}$.

Let $\mathcal{U}_\pi$ be a subset of $\mathcal{U}^\dagger$ (in general, $\mathcal{U}_\pi$ will be a linear subspace). For given μ and $g \in \mathcal{U}_\pi$, one obtains a translated measure μ^g by

$$\mu^g(A) \equiv \mu(A + g). \quad (2.4)$$

The measure μ is called $\mathcal{U}_\pi$-*quasi-invariant* if $\mu(A) = 0$ implies $\mu^g(A) = 0$ and vice versa. Note that in this case $\mu_{\mathcal{W}}$ is also $\mathcal{U}_\pi$-quasi-invariant and that $\tilde{\mu}_{\mathcal{W}}$ is quasi-invariant under $\eta_{\mathcal{W}}(\mathcal{U}_\pi)$, the image of $\mathcal{U}_\pi$ in $\mathcal{U}^\dagger/\mathcal{W}^\perp$.

Now let μ be a bounded measure on $\mathcal{U}^\dagger$, $\mu(\mathcal{U}^\dagger) = 1$ say, and consider the Hilbert space of all complex-valued μ-square-integrable functions on $\mathcal{U}^\dagger$, $\mathcal{K} = L^2(\mathcal{U}^\dagger, \mathfrak{B}(\mathcal{U}^\dagger, \mathcal{U}), \mu)$. The functions of the form $\varphi(F) = e^{i(f,F)}$, $f \in \mathcal{U}$, are total in $\mathcal{K}$. Hence the vector $\Omega \in \mathcal{K}$,

$$\Omega(F) \equiv 1, \quad (2.5)$$

is cyclic for the operators $\{U(f); f \in \mathcal{U}\}$ defined as multiplication by $e^{i(f,F)}$. If $\mathcal{W}$ is a finite-dimensional subspace of $\mathcal{U}$, we define $\mathcal{K}_{\mathcal{W}}$ to be the closed subspace of $\mathcal{K}$ obtained by restricting f to $\mathcal{W}$,

$$\mathcal{K}_{\mathcal{W}} = \overline{\{e^{i(f,F)}; f \in \mathcal{W}\}}. \quad (2.6)$$

Note that, if $\psi \in \mathcal{K}_{\mathcal{W}}$, then $\psi(F)$ is constant within $\mathcal{W}^\perp$-cosets, and is thus essentially a function on the finite-dimensional space $\mathcal{U}^\dagger/\mathcal{W}^\perp$. Let $P_{\mathcal{W}}$ be the projector onto $\mathcal{K}_{\mathcal{W}}$,

$$P_{\mathcal{W}}\mathcal{K} = \mathcal{K}_{\mathcal{W}}. \quad (2.7)$$

The family of all finite-dimensional subspaces of $\mathcal{U}$ forms a directed set by inclusion since, for given $\mathcal{W}_1$ and $\mathcal{W}_2$, there is always a $\mathcal{W}$ such that $\mathcal{W} \supset \mathcal{W}_1$ and $\mathcal{W} \supset \mathcal{W}_2$. Hence this family can be used as an index set for generalized sequences (nets[10]). One has the following simple result.

Lemma 2.1: $\mathrm{st} - \lim_{\mathcal{W}\uparrow} P_{\mathcal{W}} = 1$, i.e., given $\varphi \in \mathcal{K}$ and $\epsilon > 0$, there is a $\mathcal{W}_0$ such that for all $\mathcal{W} \supset \mathcal{W}_0$ one has $\|P_{\mathcal{W}}\varphi - \varphi\| < \epsilon$.

Proof: By the cyclicity of $\{U(f)\}$ there exists a $\mathcal{W}_0$ and a $\psi_{\mathcal{W}_0} \in \mathcal{K}_{\mathcal{W}_0}$ such that

$$\|\varphi - \psi_{\mathcal{W}_0}\| < \epsilon/2.$$

If $\mathcal{W} \supset \mathcal{W}_0$, one has $P_{\mathcal{W}}\psi_{\mathcal{W}_0} = \psi_{\mathcal{W}_0}$ and thus

$$\|P_{\mathcal{W}}\varphi - \varphi\| \leq \|P_{\mathcal{W}}(\varphi - \psi_{\mathcal{W}_0})\| + \|P_{\mathcal{W}}\psi_{\mathcal{W}_0} - \varphi\| < \epsilon.$$

QED

The next lemma is contained in Umemura[9] [Sec. 3, Eq. (5)] and Hegerfeldt[5] [Eq. (6.2)].

Lemma 2.2: Let $\nu \ll \mu$ with $\nu(\mathcal{U}^\dagger) < \infty$, and put $a(F) = [d\nu(F)/d\mu(F)]^{1/2}$. Then $a \in \mathcal{K} = L^2_\mu$. Putting $P_{\mathcal{W}}\, a = a_{\mathcal{W}}$, one has, for each finite-dimensional $\mathcal{W}$ and μ, almost all $F \in \mathcal{U}^\dagger$,

$$0 \leq a_{\mathcal{W}}(F) \leq (d\nu_{\mathcal{W}}(F)/d\mu_{\mathcal{W}}(F))^{1/2}. \quad (2.8)$$

Proof: One has

$$\int \frac{d\nu}{d\mu} d\mu = \nu(\mathcal{U}^\dagger) < \infty.$$

Hence $a \in \mathcal{K}$. Now let A be a cylinder set, $A \in \mathfrak{B}(\mathcal{U}^\dagger, \mathcal{W})$, and let χ_A be the characteristic function. Then one has $\chi_A \in \mathcal{K}_{\mathcal{W}}$ and

$$0 \leq \langle \chi_A, a\rangle = \langle \chi_A, a_{\mathcal{W}}\rangle.$$

Hence $0 \leq a_{\mathcal{W}}(F)$ for μ, almost all F since $a_{\mathcal{W}} \in \mathcal{K}_{\mathcal{W}}$. From $P_{\mathcal{W}}\chi_A a = \chi_A a_{\mathcal{W}}$ and $\|P_{\mathcal{W}}(\chi_A a)\| \leq \|\chi_A a\|$, one obtains

$$\int \chi_A(F)\, |a_{\mathcal{W}}(F)|^2 d\mu(F) \leq \int \chi_A \frac{d\nu}{d\mu} d\mu = \int \chi_A d\nu_{\mathcal{W}} = \int \chi_A \frac{d\nu_{\mathcal{W}}}{d\mu_{\mathcal{W}}} d\mu.$$

Since $a_{\mathfrak{W}}(F)$ and $d\nu_{\mathfrak{W}}/d\mu_{\mathfrak{W}}$ are constant within $\mathfrak{W}^{\perp}$-cosets, Eq. (2.8) follows. QED

For sequences the next lemma has been proved by Umemura.[9] For nets the procedure is similar.

Lemma 2.3: Let ν and a be as in Lemma 2.2. Then a is the normlimit in $\mathfrak{K} = L^2_{\mu}$ of $(d\nu_{\mathfrak{W}}/d\mu_{\mathfrak{W}})^{1/2}$,

$$\lim_{\mathfrak{W}\uparrow} \int \left| \left\{ \frac{d\nu_{\mathfrak{W}}(F)}{d\mu_{\mathfrak{W}}(F)} \right\}^{1/2} - a(F) \right|^2 d\mu(F) = 0. \quad (2.9)$$

Proof: Define $b_{\mathfrak{W}}(F)$ by orthogonal decomposition

$$a = a_{\mathfrak{W}} + b_{\mathfrak{W}}.$$

By Lemma 2.1, $b_{\mathfrak{W}} \to 0$ in norm, and by orthogonality

$$\int \left(\frac{d\nu_{\mathfrak{W}}}{d\mu_{\mathfrak{W}}} - (a_{\mathfrak{W}})^2 \right) d\mu = \int \left(\frac{d\nu}{d\mu} - (a_{\mathfrak{W}})^2 \right) d\mu = \|b_{\mathfrak{W}}\|^2. \quad (2.10)$$

By Eq. (2.8) one has

$$\frac{d\nu_{\mathfrak{W}}}{d\mu_{\mathfrak{W}}} - (a_{\mathfrak{W}})^2 \geq \left[\left(\frac{d\nu_{\mathfrak{W}}}{d\mu_{\mathfrak{W}}} \right)^{1/2} - a_{\mathfrak{W}} \right]^2,$$

and with Eq. (2.10) this yields

$$\left\| \left(\frac{d\nu_{\mathfrak{W}}}{d\mu_{\mathfrak{W}}} \right)^{1/2} - a \right\| \leq \left\| \left(\frac{d\nu_{\mathfrak{W}}}{d\mu_{\mathfrak{W}}} \right)^{1/2} - a_{\mathfrak{W}} \right\| + \|b_{\mathfrak{W}}\| \leq 2\|b_{\mathfrak{W}}\| \longrightarrow 0. \quad \text{QED}$$

The last two lemmas play a central role in the next section, where μ will be a quasi-invariant measure and where ν will be identified with μ^g.

3. CONTINUITY OF $V(g)$ AND OF THE RADON-NIKODYM DERIVATIVE

Throughout this section $\mathfrak{V}_{\Phi}$ will denote a real linear space and $\mathfrak{V}^{\dagger}_{\Phi}$ a subspace of the algebraic dual $\mathfrak{V}^{*}_{\Phi}$; furthermore, $\mathfrak{V}_{\pi}$ will denote a linear subspace of $\mathfrak{V}^{\dagger}_{\Phi}$ and μ a $\mathfrak{V}_{\pi}$-quasi-invariant normed measure on $\mathfrak{V}^{\dagger}_{\Phi}$. It is not assumed that $\mathfrak{V}_{\pi}$ separates $\mathfrak{V}_{\Phi}$, i.e., it may happen that, for some $f \in \mathfrak{V}_{\Phi}$, $(f, g) = 0$ for all $g \in \mathfrak{V}_{\pi}$. The weak topology $w(\mathfrak{V}_{\pi}, \mathfrak{V}_{\Phi})$ on $\mathfrak{V}_{\pi}$ is defined as usual: $g_{\alpha} \to g_0$ weakly if and only if $(f, g_{\alpha}) \to (f, g_0)$ for each $f \in \mathfrak{V}_{\Phi}$. As before, $\mathfrak{W}$ will always denote a finite-dimensional subspace of $\mathfrak{V}_{\Phi}$ and $\eta_{\mathfrak{W}}$ the canonical homorphism of $\mathfrak{V}^{\dagger}_{\Phi}$ onto $\mathfrak{V}^{\dagger}_{\Phi}/\mathfrak{W}^{\perp}$.

For greater transparency, we divide our procedure into several steps.

Lemma 3.1: Let $\psi \in \mathfrak{K} = L^2_{\mu}$, where μ is $\mathfrak{V}_{\pi}$-quasi-invariant and normed. Then

$$\left\langle \psi, \left(\frac{d\mu^g_{\mathfrak{W}}}{d\mu_{\mathfrak{W}}} \right)^{1/2} \right\rangle \equiv \int d\mu(F) \bar{\psi}(F) \left(\frac{d\mu_{\mathfrak{W}}(F+g)}{d\mu_{\mathfrak{W}}(F)} \right)^{1/2}$$

is weakly continuous in $g \in \mathfrak{V}_{\pi}$, for each finite-dimensional $\mathfrak{W} \subset \mathfrak{W}_{\Phi}$.

Proof: As before, denote by $\tilde{\mu}_{\mathfrak{W}}$ the measure on $\mathfrak{V}^{\dagger}_{\Phi}/\mathfrak{W}^{\perp}$ induced by $\mu_{\mathfrak{W}}$. We have, for some integers $m \leq n < \infty$, that $\eta_{\mathfrak{W}}(\mathfrak{V}^{\dagger}_{\Phi}) \equiv \mathfrak{V}^{\dagger}_{\Phi}/\mathfrak{W}^{\perp}$ can be identified with R^n and $\eta_{\mathfrak{W}}(\mathfrak{V}_{\pi})$ with $R^m \subset R^n$. Let A be a Borel set in R^n with characteristic function $\chi_A(x)$. Letting λ vary over R^m and noting that $\chi_A(x + \lambda)$ is measurable over $R^m \times R^n$, one finds by Fubini's theorem,

$$\begin{aligned} \sigma(A) &\equiv \int_{R^m} d^m\lambda \int_{R^n} d\tilde{\mu}_{\mathfrak{W}}(x) \chi_A(x+\lambda) \\ &= \int_{R^n} d\tilde{\mu}_{\mathfrak{W}}(x) \int_{R^m} d^m\lambda \chi_A(x+\lambda). \end{aligned} \quad (3.1)$$

From the second equality, one finds that $\sigma(A + \lambda_0) = \sigma(A)$ and, for bounded sets $I_m \subset R^m$,

$$\sigma(I_m \times R^{n-m}) < \infty. \quad (3.2)$$

Hence σ defines an R^m-invariant σ-finite measure on R^n. Furthermore, $\sigma(A) = 0$ implies $\tilde{\mu}_{\mathfrak{W}}(A + \lambda) = 0$ for almost all λ, and so $\tilde{\mu}_{\mathfrak{W}}(A) = 0$ by quasi-invariance. Conversely, $\tilde{\mu}_{\mathfrak{W}}(A) = 0$ implies $\sigma(A) = 0$, and hence σ and $\tilde{\mu}_{\mathfrak{W}}$ are equivalent. Thus, if $d\tilde{\mu}_{\mathfrak{W}}(x) = \rho(x) d\sigma(x)$, one has

$$\frac{d\tilde{\mu}_{\mathfrak{W}}(x+\lambda)}{d\tilde{\mu}_{\mathfrak{W}}(x)} = \frac{\rho(x+\lambda)}{\rho(x)}. \quad (3.3)$$

Now, putting $x = \eta_{\mathfrak{W}}(F)$ and $\eta_{\mathfrak{W}}(g) = \tilde{g}$, we find

$$\begin{aligned} &\left\| \left(\frac{d\mu^g_{\mathfrak{W}}}{d\mu_{\mathfrak{W}}} \right)^{1/2} - \left(\frac{d\mu^{g_0}_{\mathfrak{W}}}{d\mu_{\mathfrak{W}}} \right)^{1/2} \right\|^2 \\ &= \int_{R^n} \left| \left(\frac{d\tilde{\mu}_{\mathfrak{W}}(x+\tilde{g})}{d\tilde{\mu}_{\mathfrak{W}}(x)} \right)^{1/2} - \left(\frac{d\tilde{\mu}_{\mathfrak{W}}(x+\tilde{g}_0)}{d\tilde{\mu}_{\mathfrak{W}}(x)} \right)^{1/2} \right|^2 d\tilde{\mu}_{\mathfrak{W}}(x) \\ &= \int_{R^n} \left| [\rho(x+\tilde{g})]^{1/2} - [\rho(x+\tilde{g}_0)]^{1/2} \right|^2 d\sigma(x). \end{aligned} \quad (3.4)$$

If $g \to g_0$ weakly, then $\tilde{g} \to \tilde{g}_0$ in R^m and the right-hand side converges to zero [for continuous functions this follows from Lebesgue's bounded convergence; these functions are dense in $L^2(R^n, \sigma)$]. From this the statement follows immediately. QED

The next lemma is the corner-stone of this section. The results announced in the Introduction will be derived directly from it.

Lemma 3.2: Let $\mu, \mathfrak{V}_{\pi}, \mathfrak{K}, \Omega$, etc., be defined as before and let $V(g)$ be given by Eq. (1.3). Let $\hat{\mathfrak{V}}_{\pi}$ be a subset of $\mathfrak{V}_{\pi}$, and let τ be a topology on $\hat{\mathfrak{V}}_{\pi}$ for which (f, g) is continuous on $\hat{\mathfrak{V}}_{\pi}$, for each $f \in \mathfrak{V}_{\Phi}$, i.e., $\tau \geq w(\hat{\mathfrak{V}}_{\pi}, \mathfrak{V}_{\Phi})$. Let $0 \leq \psi_0(F) \in \mathfrak{K}_{\mathfrak{W}_0}$ for some $\mathfrak{W}_0$. Then $\langle \psi_0, V(g)\Omega \rangle$ is an upper semi-continuous function on $\hat{\mathfrak{V}}_{\pi}$. If, furthermore, $\langle \Omega, V(g)\Omega \rangle$ is continuous at some $g_0 \in \hat{\mathfrak{V}}_{\pi}$, then $\langle \psi, V(g)\varphi \rangle$ is continuous at g_0 for arbitrary $\varphi, \psi \in \mathfrak{K}$.

Proof: Defining a^g and $a^g_{\mathfrak{W}}$ analogous to Lemma 2.2, with $\nu = \mu^g$, one has, for $\mathfrak{W} \supset \mathfrak{W}_0$ and for each $g \in \mathfrak{V}_{\pi}$,

$$\langle \psi_0, V(g)\Omega \rangle = \langle \psi_0, a^g_{\mathfrak{W}} \rangle \leq \langle \psi_0, (d\mu^g{}_{\mathfrak{W}}/d\mu_{\mathfrak{W}})^{1/2} \rangle. \quad (3.5)$$

By Lemma 2.3, $\langle \psi_0, V(g)\Omega \rangle$ is the limit of the right-hand side; hence it is also the infimum over

all $\mathcal{W} \supset \mathcal{W}_0$. The right-hand side is τ-continuous on $\hat{\mathcal{U}}_\pi$ by Lemma 3.1. But an infimum of continuous functions is upper semicontinuous.[11]

Now assume that, on $\hat{\mathcal{U}}_\pi$, $\langle \Omega, V(g)\Omega\rangle$ is τ-continuous at g_0, and let $\psi_0 \in \mathcal{K}_{\mathcal{W}_0}$ for some $\mathcal{W}_0$, $0 \leq \psi(F) \leq 1$. Then, if $\psi_1 = \Omega - \psi_0$, one also has $\psi_1 \in \mathcal{K}_{\mathcal{W}_0}$ and $0 \leq \psi_1(F) \leq 1$. Thus

$$\langle \Omega, V(g)\Omega\rangle = \langle \psi_0, V(g)\Omega\rangle + \langle \psi_1, V(g)\Omega\rangle \qquad (3.6)$$

is the sum of two upper semicontinuous functions. Now, if one has two upper semicontinuous functions $h_1(x)$ and $h_2(x)$ on some space and if $h_3 = h_1 + h_2$ is continuous at $x = x_0$, one obtains, from the relation

$$\overline{\lim_{x \to x_0}} h_i(x) = h_i(x_0), \qquad i = 1, 2,$$

valid for upper semicontinuous functions,[11] and, from the continuity of $h_3(x)$ at x_0,

$$\begin{aligned} 0 &= \overline{\lim_{x \to x_0}} h_3(x) - \underline{\lim_{x \to x_0}} h_3(x) \\ &\geq h_1(x_0) + h_2(x_0) - \underline{\lim_{x \to x_0}} h_1(x) - \overline{\lim_{x \to x_0}} h_2(x) \\ &= \overline{\lim_{x \to x_0}} h_1(x) - \underline{\lim_{x \to x_0}} h_1(x) \geq 0. \end{aligned}$$

Hence the oscillation of $h_1(x)$ at $x = x_0$ vanishes, and thus $h_1(x)$ is continuous at x_0. Therefore, $\langle \psi_0, V(g)\Omega\rangle$ is continuous at g_0.

Since the set of allowed vectors ψ_0 is total in $\mathcal{K}$ and since $V(g)$ is unitary, $\langle \psi, V(g)\Omega\rangle$ is continuous at g_0 for all $\psi \in \mathcal{K}$. If $\varphi = U(f)\Omega$, one has

$$\langle \psi, V(g)\varphi\rangle = e^{i(f,g)}\langle U(-f)\psi, V(g)\Omega\rangle,$$

which is also continuous at g_0. Since Ω is cyclic for $U(f)$, one has continuity for all $\varphi \in \mathcal{K}$. QED

In the preceding Lemma, $\hat{\mathcal{U}}_\pi = \mathcal{U}_\pi$ is, of course, allowed. It may happen that $\langle \Omega, V(g)\Omega\rangle$ has no point of continuity in the weak topology of $\mathcal{U}_\pi$. The simplest example is furnished by the Fock representation, for which $\mathcal{U}_\Phi$ and $\mathcal{U}_\pi$ can be identified with a separable Hilbert space $\mathcal{U}$ and where $\mathcal{U}_\Phi^\dagger = \mathcal{U}^*$. The measure μ is a Gaussian measure. One has

$$\langle \Omega, V_F(g)\Omega\rangle = e^{-(g,g)/4},$$

which is clearly not weakly continuous on $\mathcal{U}$.

However, one knows that an upper semicontinuous function does have points of continuity in the case of certain Baire spaces and in particular for all complete metric spaces. From this the next theorem follows immediately.

Theorem 3.1: Let $\mathcal{U}_\Phi$ be a real linear space, and let $\mathcal{U}_\Phi^\dagger$ be a subspace of the algebraic dual $\mathcal{U}_\Phi^*$. Let $\mathcal{U}_\pi$ be a subspace of $\mathcal{U}_\Phi^\dagger$, and let μ be a $\mathcal{U}_\pi$-quasi-invariant measure on the Borel sets of $\mathcal{U}_\Phi^\dagger$. Consider a subspace $\hat{\mathcal{U}}_\pi \subset \mathcal{U}_\pi$ and assume that $\hat{\mathcal{U}}_\pi$ can be made a complete metric topological vector space[12] by some metric topology τ for which (f,g) is continuous on $\hat{\mathcal{U}}_\pi$, for each $f \in \mathcal{U}_\Phi$. Then the map $g \mapsto V(g)$ defined by Eq. (1.3) is strongly continuous on $\hat{\mathcal{U}}_\pi$.

Proof: An upper semicontinuous function of a complete metric space has a dense set of points of continuity.[13] Hence $\langle \Omega, V(g)\Omega\rangle$ and $\langle \psi, V(g)\varphi\rangle$ are continuous at some $g_0 \in \hat{\mathcal{U}}_\pi$, for all ψ, $\varphi \in \mathcal{K}$, by Lemma 3.2. But then one has continuity everywhere on $\hat{\mathcal{U}}_\pi$, since, if $g \to g_1$, one has $g - g_1 + g_0 \to g_0$ in $\mathcal{U}_\pi$, and thus

$$\langle \psi, V(g)\varphi\rangle \to \langle \psi, V(g_0)V(g_1 - g_0)\varphi\rangle = \langle \psi, V(g_1)\varphi\rangle.$$

Since $V(g)$ is unitary, weak continuity implies strong continuity. QED

This result allows some immediate applications. Recall that a Fréchet (F-) space is a complete metric locally convex vector space and that an LF-space $\mathcal{E}$ is a strict inductive limit of F-spaces, i.e., a countable union of subspaces $\mathcal{E}_n$ which, by the topology of $\mathcal{E}$, are F-spaces. A sequence in $\mathcal{E}$ converges if and only if it is contained in some $\mathcal{E}_n$ and converges there. Examples of F-spaces are Hilbert spaces and the space S of Schwartz of rapidly decreasing infinitely differentiable functions. As an LF-space we mention the space $\mathcal{D}$ of Schwartz of infinitely differentiable functions with compact support.

Corollary 3.1: Let $\mathcal{U}_\pi$ be an F-space or an LF-space for some topology $\tau \geq w(\mathcal{U}_\pi, \mathcal{U}_\Phi)$. Then the map $g \to V(g)$ is strongly τ-continuous on $\mathcal{U}_\pi$.

Proof: If $\mathcal{U}_\pi$ is an F-space, put $\hat{\mathcal{U}}_\pi = \mathcal{U}_\pi$ in Theorem 3.1. If $\mathcal{U}_\pi = \cup_n \mathcal{E}_n$, then put $\hat{\mathcal{U}}_\pi = \mathcal{E}_n$ in Theorem 3.1. Then $g \mapsto V(g)$ is continuous on each $\mathcal{E}_n$ and thus, by the properties of inductive limits, also on $\mathcal{U}_\pi$. QED

We note that for $U(f)$ a similar result holds. If $\mathcal{U}_\Phi$ is an F-space or LF-space and $\mathcal{U}_\Phi^\dagger = \mathcal{U}_\Phi'$, then $f \mapsto U(f)$ is strongly continuous. This follows by Lebesgue's bounded convergence.[8]

A special case of Theorem 3.1 is obtained if $\hat{\mathcal{U}}_\pi$ is any finite-dimensional subspace of $\mathcal{U}_\pi$, in particular if $\hat{\mathcal{U}}_\pi$ is one dimensional, i.e., if $\hat{\mathcal{U}}_\pi = \{\lambda g\}$ for some $g \in \mathcal{U}_\pi$. The Euclidean metric satisfies all conditions of Theorem 3.1, and one obtains ray continuity, a result derived by Araki[7] in a completely different way.

Corollary 3.2 (Araki): $V(\lambda g)$ is strongly continuous in λ for fixed $g \in \mathcal{U}_\pi$.

A short and direct proof of Corollary 3.2, which is not based on Theorem 3.1 nor on intricate Hilbert space techniques, can be given as follows. If $g \neq 0$, there is an $f_1 \in \mathcal{U}_\Phi$ satisfying $(f_1, g) = 1$.

If one defines $Y \equiv \{F \in \mathcal{U}_\Phi^\dagger; (f_1, F) = 0\}$ and if $\mathfrak{B}_Y \equiv \mathfrak{B}(Y, \mathcal{U}_\Phi)$ denotes the Borel sets in Y, $\mathfrak{B}_0$ those of $\mathcal{U}_\Phi^\dagger$, and $\mathfrak{B}_R$ those of $R = \{\lambda g\}$, then $\mathcal{U}_\Phi^\dagger$ is the direct vector sum of R and Y. Moreover, it follows from the definitions in Sec. 2 that $(\mathcal{U}_\Phi^\dagger, \mathfrak{B}_0) = (R, \mathfrak{B}_R) \times (Y, \mathfrak{B}_Y)$, as already noted in Ref. 7. Hence μ is an R-quasi-invariant measure on $R \times Y$. For any Borel set A in $R \times Y$, with characteristic function $\chi_A(x,y)$, $\chi_A(x + \lambda, y)$ is measurable on $R \times R \times Y$. As in Eq. (3.1), one can define a σ-finite R-invariant measure σ on $R \times Y$ which is equivalent to μ. Now the map

$$\psi \mapsto \hat{\psi} \equiv (d\mu/d\sigma)^{1/2}\psi$$

is an isometry of $L^2(\mathcal{U}_\Phi^\dagger, \mu)$ onto $L^2(R \times Y, \sigma)$ under which $V(\lambda g)$ becomes translation of x by λ. If $\hat{\psi}(x,y) = \varphi_1(x)\varphi_2(y)$ with $\varphi_1(x)$ continuous, one has $\|\hat{\psi}(x + \lambda_n, y) - \hat{\psi}(x,y)\|_\sigma \to 0$ for $\lambda_n \to 0$, by Lebesgue's bounded convergence. Since these functions are total, one obtains strong continuity of the translations in L^2_σ and thus of $V(\lambda g)$ in L^2_μ. QED

A slightly stronger result than Theorem 3.1 and Corollary 3.1 can be derived under the additional assumption that $\mathcal{H} = L^2_\mu$ is separable. In this case one can use sequences instead of nets. One knows that if $\{h_n(x)\}$ is a sequence of continuous functions converging pointwise to a function $h(x)$, then the points of discontinuity of $h(x)$ form a meager (first category) set.[14] Hence, if $\hat{\mathcal{U}}_\pi$ is of second category in itself for some topology $\tau \geq w(\hat{\mathcal{U}}_\pi, \mathcal{U}_\Phi)$, then $V(g)$ is strongly continuous on $\hat{\mathcal{U}}_\pi$.

To conclude this section, we note that the above results can be reformulated as statements on continuity properties of the Radon–Nikodym derivative $d\mu^g/d\mu$. Strong continuity of $V(g)$ implies that, for $g \to g_0$,

$$\|V(g)\Omega - V(g_0)\Omega\|^2 = \int \left| \left(\frac{d\mu^g}{d\mu}\right)^{1/2} - \left(\frac{d\mu^{g_0}}{d\mu}\right)^{1/2} \right|^2 d\mu \to 0. \quad (3.7)$$

This means that $(d\mu^g/d)^{1/2}$, regarded as a function of g with values in L^2_μ, is norm continuous. From the inequality

$$\left(\int \left| \frac{d\mu^g}{d\mu} - \frac{d\mu^{g_0}}{d\mu} \right| d\mu\right)^2 \leq 4 \int \left| \left(\frac{d\mu^g}{d\mu}\right)^{1/2} - \left(\frac{d\mu^{g_0}}{d\mu}\right)^{1/2} \right|^2 d\mu, \quad (3.8)$$

which is obtained by means of Schwartz's inequality, it follows that $d\mu^g/d\mu$, regarded as a function of g with values in L^1_μ, is norm continuous. From this follows, conversely, strong continuity of $V(g)$ if (f,g) is continuous in g for each $f \in \mathcal{U}_\Phi$.

4. ERGODIC DECOMPOSITION OF QUASI-INVARIANT MEASURES

Parts of this section have already been obtained by Hegerfeldt[8] through the decomposition of representations of the CCR's into irreducible ones. The preceding section can be used to extend those results considerably.

A $\mathcal{U}_\pi$-quasi-invariant measure μ on $\mathcal{U}_\Phi^\dagger$ is called $\mathcal{U}_\pi$-*ergodic* if there is no nontrivial $\mathcal{U}_\pi$-quasi-invariant measure on $\mathcal{U}_\Phi^\dagger$ which is absolutely continuous with respect to μ. This means, if $\nu \ll \mu$ and if ν is $\mathcal{U}_\pi$-quasi-invariant, then either $\nu \equiv 0$ or ν is equivalent to μ. Geometrically, ergodicity of ν means that $\mathcal{U}_\Phi^\dagger$ cannot be decomposed into two subsets of positive measure which are invariant under translations by elements of $\mathcal{U}_\pi$, modulo μ-null sets. Thus ergodic measures are in a certain sense minimal.

Now we consider a real linear topological space $\mathcal{U}$ and take for $\mathcal{U}^\dagger$ the topological dual $\mathcal{U}'$. If (f,g) is a nondegenerate continuous bilinear form on $\mathcal{U} \times \mathcal{U}$, one can embed $\mathcal{U}$ in $\mathcal{U}'$ by regarding each $g \in \mathcal{U}$ as a continuous linear functional on $\mathcal{U}$. Nuclear spaces play an important role in the following. The only property of nuclearity exploited here is the Bochner–Minlos theorem, which states that every continuous positive-definite function on a nuclear space is the Fourier transform of a bounded measure on the dual.[9]

The next theorem solves the problem of ergodic decomposability of a $\mathcal{U}$-quasi-invariant measure on the dual $\mathcal{U}'$ of a nuclear F- or LF-space completely. It extends Theorem 6.5 of Ref. 8.

Theorem 4.1: Let $\mathcal{U}$ be a nuclear F- or LF-space, and let (f,g) be a nondegenerate continuous bilinear form on $\mathcal{U} \times \mathcal{U}$ so that $\mathcal{U}$ can be regarded as a subspace of $\mathcal{U}'$. If μ is a $\mathcal{U}$-quasi-invariant bounded measure on $\mathcal{U}'$, then there is a standard Borel space Z, a bounded positive measure ρ on Z, for each $\zeta \in Z$ a $\mathcal{U}$-ergodic measure μ_ζ on $\mathcal{U}'$ with $\mu_\zeta(\mathcal{U}') = \mu(\mathcal{U}')$ such that, for each Borel set A in $\mathcal{U}'$,

$$\mu(A) = \int d\rho(\zeta)\mu_\zeta(A).$$

Proof: By the Corollary 3.1, the continuity conditions of Lemma 6.2 of Ref. 8 are satisfied. Then the statement follows immediately from Theorem 6.5 of that reference. QED

This result can be sharpened somewhat by extending Theorem 6.4 of Ref. 8 with the help of Corollary 3.2. It suffices to demand nuclearity for $\mathcal{U}_\pi$ only.

Corollary 4.1: Let $\mathcal{U}$ be a separable F- or LF-space, let $\mathcal{U}_\pi$ be a subspace of $\mathcal{U}'$, and let μ be a bounded $\mathcal{U}_\pi$-quasi-invariant measure on $\mathcal{U}'$.

If, for some topology τ finer than $w(\mathcal{U}_\pi, \mathcal{U})$, $\mathcal{U}_\pi$ is a nuclear F- or LF-space, then there is a standard Borel space Z, a bounded positive measure ρ on Z, for each $\zeta \in Z$ a $\mathcal{U}_\pi$-ergodic measure μ_ζ on $\mathcal{U}'$ with $\mu_\zeta(\mathcal{U}') = \mu(\mathcal{U}')$ such that, for each Borel set A in $\mathcal{U}'$,

$$\mu(A) = \int d\rho(\zeta)\mu_\zeta(A).$$

Proof: Let $V(g)$ be defined as in Eq. (1.3). Then $\langle \Omega, V(g)\Omega \rangle$ is a positive-definite function which, by Corollary 3.1, is continuous for the nuclear topology of $\mathcal{U}_\pi$. Hence, by the Bochner–Minlos theorem, there exists a normed measure ν on $\mathcal{U}'_\pi$, such that

$$\langle \Omega, V(g)\Omega \rangle = \int_{\mathcal{U}'_\pi} d\nu(\xi) e^{i(g,\xi)}, \quad (4.1)$$

where ξ runs through $\mathcal{U}'_\pi$. Now one can insert literally the proof of Theorem 6.4 of Ref. 8. Using Eq. (4.1) it is then shown in exactly the same way as in Lemma 4.2 and Eq. (4.3) of Ref. 8 that the measures appearing in this proof are $\mathcal{U}_\pi$-quasi-invariant. QED

It is clear that Theorem 4.1 and Corollary 4.1 also hold for σ-finite measures.

[1] H. Araki, J. Math. Phys. 1, 492 (1960).
[2] H. Fukotome, Progr. Theoret. Phys. (Kyoto) 23, 983 (1960).
[3] I. M. Gel'fand and N. Y. Vilenkin, *Generalized Functions* (Academic, New York, 1964), Vol. 4 (last chapter).
[4] G. C. Hegerfeldt and O. Melsheimer, Commun. Math. Phys. 12, 304 (1969).
[5] G. C. Hegerfeldt, Habilitationsschrift, Universität Marburg, 1968. Cf. also J. Math. Phys. 10, 1681 (1969).
[6] G. C. Hegerfeldt, J. Math. Phys. 12, 167 (1971).
[7] H. Araki, Commun. Math. Phys. 20, 9 (1971).
[8] G. C. Hegerfeldt, Nuovo Cimento B4, 225 (1971).
[9] Y. Umemura, Publ. Res. Inst. Math. Sci. A 1, 1 (1966).
[10] N. Dunford and J. T. Schwartz, *Linear Operators, Part I* (Interscience, New York, 1966), p. 26; J. L. Kelley, *General Topology* (Van Nostrand, New York, 1955), Chap. 2.
[11] N. Bourbaki, *General Topology I, II* (Hermann, Paris) and (Addison-Wesley, Reading, Mass., 1966), Chap. IV, Sec. 6.
[12] Addition in $\hat{\mathcal{U}}_\pi$ must be continuous.
[13] G. Choquet, *Lectures on Analysis*, edited by J. Marsden, T. Lance, and S. Gelbart (Benjamin, New York, 1969), Vols. I–III, Secs. 7.6 and 8.3.
[14] Reference 11, Chap. IX, Sec. 5, Exercise 22.

Application of the Theory of Orlicz Spaces to Statistical Mechanics. I. Integral Equations

Ing-Yih S. Cheng*

and

John J. Kozak

Department of Chemistry, The University of Notre Dame, Notre Dame, Indiana 46556

(Received 9 November 1970)

In this paper it is suggested that there may exist a fundamental relationship between the variables of thermodynamics, the operators associated with certain nonlinear integral equations of statistical mechanics, and the properties of a class of convex functions, called N functions, investigated by Krasnosel'skii and Rutickii. In particular, it is pointed out that the most general theoretical framework within which all these problems can be studied is that provided by the theory of Orlicz spaces. In the first part of our study, presented here, it is shown that the existence of solutions to certain nonlinear integral equations, derived either from the BBKYG hierarchy or from the grand partition function using a variational approach, can be established with some generality. The relationship between our results and those obtained by Ruelle is discussed.

I. INTRODUCTION

The class of problems with which we shall be concerned in this paper is best introduced by recalling the dependence of pressure on density diagramed in Fig. 1. We shall also refer to the dependence of density on fugacity shown in Fig. 2. One of the main problems in statistical mechanics is the attempt to determine the equations describing these curves for specific physical systems. For example, the equation of state in the grand canonical ensemble is given by

$$pV = kT \ln \Xi \tag{1}$$

and the evaluation of Ξ for a system gives a trajectory similar to that shown in Fig. 1.

In our study of this problem, we shall adopt a slightly different point of view. Rather than considering the curve itself, we instead focus attention on the area under the curve and then examine whether this area can be provided with a well-defined topological basis. We observe that the function $p(\rho)$ plotted in Fig. 1 is right continuous for $\rho \geq 0$, positive for $\rho > 0$, and nondecreasing. Furthermore, $p(\rho)$ satisfies the conditions

$$p(0) = 0 \tag{2a}$$

and,

$$p(\infty) = \lim_{\rho \to \infty} p(\rho) = \infty. \tag{2b}$$

If one abstracts the specific properties characterizing the thermodynamic variable $p(\rho)$, one realizes that these properties are precisely the ones used by M. A. Krasnosel'skii and Y. B. Rutickii in defining a general class of convex functions called N functions.[1,2] These functions, designated $M(u)$, admit the representation

$$M(u) = \int_0^{|u|} p(t)dt, \tag{3}$$

and, in the most general case, can have a graph of the form shown in Fig. 3. This figure serves to illustrate that this class of convex functions is general enough to include functions whose associated trajectories are either concave upward or concave downward, or have discontinuities. Thus, returning to thermodynamics, we note that an N function can be associated not only with the pressure as a function of density (as diagramed in Fig. 1), namely,

$$M(u) = \int_0^{|u|} p(\rho)d\rho,$$

but also with the density as a function of fugacity (as diagramed in Fig. 2), say

$$M'(u) = \int_0^{|u|} \rho(z)dz.$$

From the above discussion, it is clear that the value of the N function itself is given by the area of the corresponding curvilinear trapezoid.

In developing the theory of N functions, it is convenient to introduce another function called a complementary N function. To do this, a new function $q(s)$ with $s \geq 0$ is defined as follows:

$$q(s) = \sup_{p(t) \leq s} t. \tag{4}$$

In more usual language, $q(s)$ is the inverse function of $p(t)$ when $p(t)$ is continuous and monotonically in-

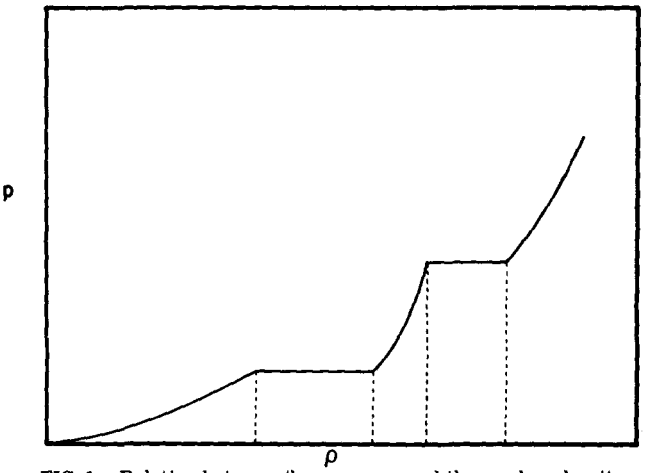

FIG. 1. Relation between the pressure and the number density at a given temperature.

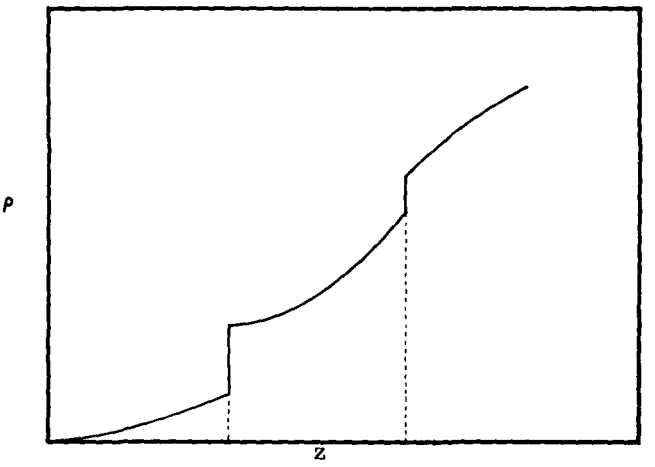

FIG. 2. Relation between the number density and the fugacity at a given temperature.

creasing. In general, $q(s)$ is called the right inverse of $p(t)$. Notice that the function $q(s)$ possesses the same properties as the function $p(t)$, since it is positive for $s > 0$, nondecreasing, and satisfies the conditions

$$q(0) = 0, \tag{5a}$$

$$q(\infty) = \lim_{s\to\infty} q(s) = \infty. \tag{5b}$$

The N function complementary to $M(u)$ is then defined as

$$N(v) = \int_0^{|v|} q(s)ds. \tag{6}$$

The relationship between these two functions can be displayed graphically as shown in Fig. 4. The simplest example of a complementary pair of N functions is the following:

$$M_1(u) = \frac{|u|^\alpha}{\alpha}, \quad \alpha > 1, \tag{7a}$$

$$N_1(v) = \frac{|v|^\beta}{\beta}, \quad \frac{1}{\alpha} + \frac{1}{\beta} = 1. \tag{7b}$$

For the problems treated in this paper, especially those studied in Secs. II and III, the following, strongly nonlinear, N function will be used:

$$M_2(u) = e^{|u|} - |u| - 1. \tag{8a}$$

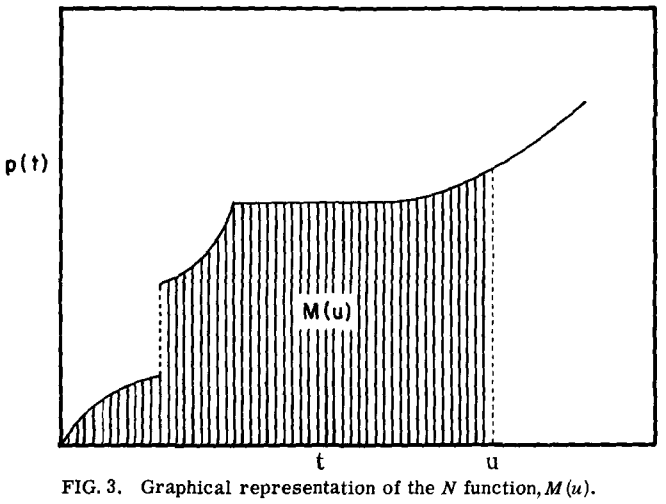

FIG. 3. Graphical representation of the N function, $M(u)$.

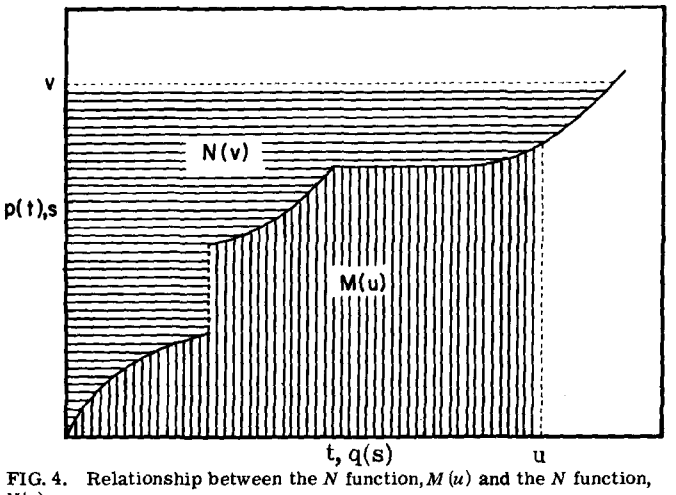

FIG. 4. Relationship between the N function, $M(u)$ and the N function, $N(v)$.

The N function complementary to $M_2(u)$ is

$$N_2(v) = (1 + |v|)\, ln(1 + |v|) - |v|. \tag{8b}$$

The main idea developed by Krasnosel'skii and Rutickii in the theory of Orlicz spaces is to consider general properties associated with pairs of N functions and in particular to consider not merely power functions [for example, Eqs. (7)], but exponential functions as well [for example, Eqs. (8)]. This program is carried out by defining a class of all real-valued functions defined on a bounded, closed set G, and denoted by $L_M(G)$, such that

$$\int_G M[u(x)]dx < \infty. \tag{9}$$

Equation (9) defines the "pseudo-norm" of u. The class $L_N(G)$ and its pseudo-norm are defined analogously. It should be emphasized that the spaces $L_M(G)$ and $L_N(G)$ are not, in general, linear spaces and the "pseudo-norms" have few of the properties of norms. Accordingly, one defines $L_M^*(G)$ to be the set of all function $u(x)$ satisfying

$$\left|\int_G u(x)v(x)dx\right| < \infty \tag{10}$$

for all $v(x) \in L_N(G)$. The set $L_N^*(G)$ is defined analogously. The major step taken by Krasnosel'skii and Rutickii is their demonstration that these sets are, in fact, linear spaces and can be provided with norms in such a way that $L_M^*(G)$ and $L_N^*(G)$ become a conjugate pair of reflexive Banach spaces. Then, $L_M^*(G)$ is called an Orlicz space and the norm is defined by the relation

$$\|u\|_M = \text{Sup}\left|\int_G u(x)v(x)dx\right|, \tag{11}$$

where the supremum is taken over all $v \in L_N$ with pseudo-norm less than or equal to unity.

Having sketched the general ideas underlying N functions and Orlicz spaces, it is apparent from a study of Figs. 1 and 2 in relation to Fig. 4, that several lines of research could be pursued at this point. For example, one could make extensive use of the algebraic properties of N functions, as typified by the Young inequality

$$uv \leq M(u) + N(v) \tag{12}$$

which, written in terms of thermodynamic variables becomes

$$\rho p \leq M(\rho) + N(p). \tag{13}$$

On the other hand, given the norm defined by Eq. (11), it might be recalled that Krasnosel'skii and Rutickii have proved several theorems on the existence of solutions to certain nonlinear integral equations with exponential nonlinearities. Since this latter approach provides the closest connection with work already in the literature, we consider here the application of these existence theorems to certain approximate integral equations derivable from Eq. (1), and defer the algebraic analysis to a separate paper.

To see the close connection between N functions, nonlinear integral equations with exponential nonlinearities, and thermodynamics, we recall the expression

for Ξ given by Morita and Hiroike[3]:

$$\ln\Xi = \int d\mathbf{R}\rho(\mathbf{R})\ln z^*(\mathbf{R}) - \int d\mathbf{R}\rho(\mathbf{R})[\ln\rho(\mathbf{R}) - 1] + \tfrac{1}{2}\int\int d\mathbf{R}_1 d\mathbf{R}_2\rho(\mathbf{R}_1)\rho(\mathbf{R}_2)\,[1 + v(\mathbf{R}_1,\mathbf{R}_2)] \times \ln[1 + b(\mathbf{R}_1,\mathbf{R}_2)] - \tfrac{1}{2}\int\int d\mathbf{R}_1 d\mathbf{R}_2\rho(\mathbf{R}_1)\rho(\mathbf{R}_2) \times \{[1 + v(\mathbf{R}_1,\mathbf{R}_2)]\ln[1 + v(\mathbf{R}_1,\mathbf{R}_2)] - v(\mathbf{R}_1,\mathbf{R}_2)\} + - \cdots, \quad (14)$$

where

$$z^*(\mathbf{R}) = z\,\exp^{-\Psi(\mathbf{R})/kT},$$
$$b(\mathbf{R}_1,\mathbf{R}_2) = \exp^{-u(\mathbf{R}_1,\mathbf{R}_2)/kT} - 1,$$
$$v(\mathbf{R}_1,\mathbf{R}_2) = g(\mathbf{R}_1,\mathbf{R}_2) - 1.$$

Here, $\Psi(\mathbf{R})$ is the external field (set to zero in what follows), $u(\mathbf{R}_1,\mathbf{R}_2)$ is the pair potential between particles 1 and 2, $\rho(\mathbf{R})$ is the singlet distribution function, and $g(\mathbf{R}_1,\mathbf{R}_2)$ is the pair-correlation function. Provided all terms on the right-hand side of Eq. (14) are taken into account, this equation is an exact representation of Ξ. If one uses the variational condition

$$\left(\frac{\delta\ln\Xi}{\delta\rho(\mathbf{R})}\right) z^*, b, v = 0$$

in conjunction with Eq. (14), one obtains the following expression:

$$\ln z^*(\mathbf{R}_1) = \ln\rho(\mathbf{R}_1) - \int d\mathbf{R}_2\rho(\mathbf{R}_2)\{[1 + v(\mathbf{R}_1,\mathbf{R}_2)] \times \ln[1 + b(\mathbf{R}_1,\mathbf{R}_2)] - [1 + v(\mathbf{R}_1,\mathbf{R}_2)] \times \ln[1 + v(\mathbf{R}_1,\mathbf{R}_2)] + v(\mathbf{R}_1,\mathbf{R}_2)\} - + \cdots \quad (15)$$

As it stands, Eq. (15) represents an exact integral equation for $\rho(\mathbf{R})$ if $v(\mathbf{R}_1,\mathbf{R}_2)$ is known. In practice, one neglects all terms beyond the second on the right-hand side of Eq. (15), thus forming the hypernetted chain approximation. Using this approximation, the integral equation (15) can be written in the form

$$\Phi(\mathbf{R}_1) + \int_V k(\mathbf{R}_1,\mathbf{R}_2)\ \exp[\Phi(\mathbf{R}_2)]d\mathbf{R}_2 = 0, \quad (16)$$

where

$$\Phi(\mathbf{R}) = \ln\frac{\rho(\mathbf{R})}{z}, \quad (17)$$

$$K(\mathbf{R}_1,\mathbf{R}_2) = K_{HNC}(\mathbf{R}_1,\mathbf{R}_2) = \beta z u(\mathbf{R}_{12})g^{(2)}(\mathbf{R}_1,\mathbf{R}_2) + z\{[1 + v(\mathbf{R}_1,\mathbf{R}_2)]\ln[1 + v(\mathbf{R}_1,\mathbf{R}_2)] - v(\mathbf{R}_1,\mathbf{R}_2)\}. \quad (18)$$

Various other integral equations for the singlet distribution function can be derived, which can be cast into the Hammerstein form[4] of Eq. (16), but with different kernels. Several of these kernels have been discussed in Ref. 5, together with the approximations involved in their formulation. For example, starting with the first equation of the Kirkwood coupling-parameter hierarchy, one can derive an integral equation having the same overall structure as Eq. (16), but with the kernel

$$K_{KJ}(\mathbf{R}_1,\mathbf{R}_2) = \beta z u(R_{12})\int_0^1 d\xi g^{(2)}(\mathbf{R}_1,\mathbf{R}_2,\xi), \quad (19)$$

where ξ is the Kirkwood coupling parameter. Alternatively, Green[6] has shown that an equation similar to Eq. (16) can be derived starting from the *YBG* equation; here, the kernel is

$$K_G(\mathbf{R}_1,\mathbf{R}_2) = \beta c'\int_{R_{12}}^{\infty} dr u'(r)g^{(2)}(r). \quad (20)$$

Finally, one can start from an approximate mean-field representation of the singlet distribution function suggested by Brout[7] and derive an equation similar to Eq. (16), but with the kernel

$$K_B(\mathbf{R}_1,\mathbf{R}_2) = \beta z u^{(1)}(R_{12})g_0^{(2)}(R_{12}). \quad (21)$$

In this expression, $g_0^{(2)}(R_{12})$ is the pair-correlation function of a reference system and $u^{(1)}(R_{12})$ represents a perturbation potential.

The feature which is common to all of these approximations is that the structure of the nonlinear function appearing in the associated integral equation remains unchanged. That is, although the kernel changes from one representation to another, the nonlinearity remains exponential in each case. This suggests that if general results on the behavior of thermodynamic systems are to be inferred from a study of well-defined approximations, one should focus one's attention on existence theorems for nonlinear integral equations with exponential nonlinearities, since apparently it is this feature which is preserved in all of the above approximation schemes. The natural framework for discussing such equations is that of Orlicz spaces, developed by Krasnosel'skii and Rutickii and introduced above. The main point is that very different requirements must be satisfied in order to prove existence of solutions for integral equations with exponential, as opposed to power, nonlinearities. It is this feature which distinguishes the present work from that presented in Ref. 5, since in that paper the development was formulated explicitly within the framework of L^p space. This is not to say that exponential nonlinearities were not considered in Ref. 5; however, to insure that the associated nonlinear integral equation was defined within L^p space, it was necessary to bound the nonlinear function, thus limiting the generality of the results. A further distinction between the material presented in Ref. 5 and that given here is that Ref. 5 was concerned with the problem of bifurcation, whereas our primary interest here is the application of theorems on existence.

We present in Sec. II a detailed analysis of the conditions which must be satisfied by the various kernels introduced above in order that complete continuity of the associated operator be guaranteed. Since the details of the theorems proved by Krasnosel'skii and Rutickii, which we use extensively, have been presented in English translation only in a monograph and one review article (although their principal results have been summarized in various other sources[8]), we have taken the caution to state their theorems fully. For convenience, the theorems cited are labeled as in Ref. 2. In Sec. III we focus on the complete continuity of the nonlinear operator associated with the nonlinear integral equation under study. Then, in Sec. IV we present the central result of this investigation in the form of an existence theorem of Eq. (16). Various aspects of this result are discussed in Sec. V, and several comments are made regarding the relationship between our results and those obtained by

Ruelle.[9] Finally, a summary of the Δ_2, Δ^1, Δ_3, and Δ^2 conditions, useful in discussing the "growth" of nonlinear functions, is presented in the Appendix.

II. THE COMPLETE CONTINUITY OF THE KERNEL

We begin our discussion by examining those conditions which must be satisfied in order that the kernel be completely continuous. Let us associate with the kernel the following operator:

$$\mathfrak{K}\Phi(\mathbf{R}_1) = \int_V K(\mathbf{R}_1, \mathbf{R}_2)\Phi(\mathbf{R}_2)d\mathbf{R}_2.$$

The theorem which will serve as the basis of our discussion can now be stated.

Theorem 4.1: Let $M(u)$ be an N function, and let its complementary N function, $N(v)$, satisfy the Δ' condition. If $K(\mathbf{R}_1, \mathbf{R}_2) \in \hat{E}_M$ that is, if

$$\int_V\int_V M[\alpha K(\mathbf{R}_1, \mathbf{R}_2)]d\mathbf{R}_1 d\mathbf{R}_2 < \infty \tag{22}$$

for all α, then the operator $\mathfrak{K}$ acts from L_N^* to L_M^* and is completely continuous.

In the inequality (22), V refers to the (finite) volume in which the system is contained; more precisely, it is a bounded, closed set in a finite-dimensional space. As noted in the preceding section, the Δ' conditions, as well as the other Δ conditions used below, are defined in the Appendix.

As one example of an N function whose complementary N function satisfies the Δ'condition, we cite $M_2(u)$ and $N_2(v)$, Eqs. (8a) and (8b), respectively. Also, we note that even though the complementary N function corresponding to the N function $[M_3(u) = e^{u^2} - 1]$ cannot be written down explicitly, it can be demonstrated that $N_3(v)$ satisfies the Δ' condition by using the properties of $M_3(u)$. For these two cases, if $|K(\mathbf{R}_1, \mathbf{R}_2)|$ is finite on $V \times V$, then condition (22) is satisfied.

Let us now consider the kernels introduced in Sec. I. For convenience, we shall discuss the Kirkwood-Jancovici kernel first. Equation (19) can be written in the form

$$K_{KJ}(\mathbf{R}_1, \mathbf{R}_2) = \beta z \int_0^1 d\xi\, u(R_{12})g^{(2)}(R_{12}, \xi). \tag{23}$$

We will assume that $u(R_{12}) \to \infty$ as $R_{12} \to 0$. From the theory of distribution functions,[10] it is known that

$$g^{(2)}(R_{12}, \xi) \sim \exp[-\xi u(R_{12})/kT],$$

and so in the limit $R_{12} \to 0$ it is reasonable to conclude that the product

$$u(R_{12})g^{(2)}(R_{12}, \xi)$$

is dominated by the exponential term. That is, we conclude

$$\lim_{R_{12}\to 0} u(R_{12})g^{(2)}(R_{12}\xi) = 0.$$

We next assume that $u(R_{12}) \to 0$ as $R_{12} \to \infty$. In the Kirkwood-Jancovici approximation, the liquid pair-correlation function is chosen to fix the kernel, and for this choice $g^{(2)}(R_{12})$ approaches unity in the limit $R_{12} \to \infty$. Obviously, for this case

$$\lim_{R_{12}\to\infty} u(R_{12})g_l^{(2)}(R_{12}, \xi) = 0.$$

If instead a pair-correlation function characteristic of a periodic phase is chosen to fix the kernel, then $g^{(2)}(R_{12}, \xi)$ will oscillate with increasing R_{12}. As long as the amplitude of $g^{(2)}(R_{12}, \xi)$ is bounded, however, condition (22) will be satisfied. In the case of intermediate R_{12}, that is, $0 < R_{12} < \infty$, it is necessary to impose the Ruelle-Fisher bound[11] to insure compliance with (22). Briefly, we require that

$$U(\mathbf{R}_1, \ldots, \mathbf{R}_n) = \sum_{1\le i<j\le n} u(R_{ij}) \ge -nB, \tag{24}$$

for some $B \ge 0$ and $n = 1, 2, 3, \cdots$. This bound is necessary if we wish to discuss real, physical systems since otherwise an infinitely deep potential well might exist, and the whole system would collapse. Condition (24) is a lower bound on $u(R_{12})$, and since for finite $u(R_{12})$, $g^{(2)}(R_{12}, \xi)$ is finite, we conclude that the product

$$u(R_{12})g^{(2)}(R_{12}, \xi)$$

is bounded for intermediate R_{12}. Given these remarks, we conclude that the kernel K_{KJ} is finite on $V \times V$. Since the inequality (22) is then satisfied, we can state that the Kirkwood-Jancovici kernel is completely continuous on some Orlicz space.

In studying the Green kernel (20), we consider explicitly a potential that can be represented in the form

$$u(R_{12}) = a\left(\frac{1}{R_{12}}\right)^m - b\left(\frac{1}{R_{12}}\right)^n, \tag{25}$$

where $m > n \ge 1$ and $a > 0$, $b > 0$. By constructing the derivative corresponding to (25), it is seen that in the limit $R_{12} \to 0$, $u'(R_{12}) \to -\infty$. However, by the same argument as above, $g^{(2)}(R_{12})$ should dominate $u'(R_{12})$ in this limit, and so we conclude that

$$\lim_{R_{12}\to 0} u'(R_{12})g^{(2)}(R_{12}) = 0.$$

Turning next to intermediate values of $\mathbf{R}_{12}$, we note that the potential (25) attains its minimum value when

$$R_{12} = \sqrt[m-n]{\frac{am}{bn}}$$

and

$$\min\{u(R_{12})\} = a\left(\frac{bn}{am}\right)^{\frac{m}{m-n}} - b\left(\frac{bn}{am}\right)^{\frac{n}{m-n}}. \tag{26}$$

The derivative $u'(R_{12})$ constructed from (30) has as its maximum value,

$$\max\{u'(R_{12})\} = -am\left(\frac{bn(n+1)}{am(m+1)}\right)^{\frac{m+1}{m-n}} + bn\left(\frac{bn(n+1)}{am(m+1)}\right)^{\frac{n+1}{m+n}}. \tag{27}$$

Hence, by considering a potential for which the Ruelle-Fisher bound is satisfied, we have the product $u'(R_{12})g^{(2)}(R_{12})$ bounded for intermediate values of R_{12}. Finally, we consider the case $R_{12} \to \infty$. If in the relation (20) we choose a liquid pair-correlation function, $g_l^{(2)}(R_{12})$ then in the limit $R_{12} \to \infty$, $g^{(2)}(R_{12}) \to 1$, and so $u'(R_{12})g^{(2)}(R_{12}) \to u'(R_{12})$. But

$$u'(R_{12}) = -am\left(\frac{1}{R_{12}}\right)^{m+1} + bn\left(\frac{1}{R_{12}}\right)^{n+1}$$

goes to zero at least as fast as R_{12}^{-2}, so $\int_a^\infty u'(\mathbf{R}_{12})d\mathbf{R}_{12}$ converges. The above argument still holds if we choose a periodic pair-correlation function, provided the amplitude of $g^{(2)}(R_{12})$ is bounded. Thus, in any case, we conclude that

$$\int_{R_{12}}^{\infty} dru'(r)g^{(2)}(r)$$

is finite on $V \times V$ and the condition (22) is satisfied.

The conditions which must be satisfied in order that the Brout kernel be finite on $V \times V$ can be established easily, given the arguments presented in the preceding paragraphs. Furthermore, the first term in (17), $\beta z u(R_{12})g^{(2)}(R_{12})$, is similar in structure to the Kirkwood–Jancovici kernel which is finite on $V \times V$. The second term in (17) may be written as

$$z[g(R_{12}) \ln g(R_{12}) - g(R_{12}) + 1].$$

Provided we impose the Ruelle–Fisher bound, this term is also finite on $V \times V$ for all values of R_{12}, so that $\rho^{(2)}(R_{12}) < \infty$.

III. COMPLETE CONTINUITY OF THE NONLINEAR OPERATOR

We shall associate with the nonlinear function appearing in Eq. (16) an operator $\mathfrak{f}$ defined by

$$\mathfrak{f}\,\Phi(R) = \exp[\Phi(R)].$$

We note further that the integral appearing in this equation can also be associated with an operator

$$\mathfrak{A}\,\Phi(\mathbf{R}_1) = \int_V K(\mathbf{R}_1, \mathbf{R}_2)\ \exp[\Phi(\mathbf{R}_2)]d\mathbf{R}_2 \tag{28}$$

so that,

$$\mathfrak{A} = \mathfrak{K}\,\mathfrak{f}. \tag{29}$$

We will assume that the nonlinear function appearing in Eq. (16) can be characterized by the following inequality:

$$|f[\mathbf{R}, \Phi(\mathbf{R})]| \leq b(\mathbf{R}) + G(|\Phi|), \qquad \mathbf{R} \in V_1 \quad -\infty < \Phi < \infty, \tag{30}$$

where we specify that $G(\Phi)$ is continuous and monotonically increasing for $\Phi \geq 0$ and $b(\mathbf{R}) \geq 0$ is a measurable function. In order to insure the complete continuity of the operator $\mathfrak{A}$, it will be necessary to require that $\mathfrak{f}$ is continuous and bounded and $\mathfrak{f} : L_M^* \to L_N^*$ and the operator $\mathfrak{K}$ is completely continuous and $\mathfrak{K} : L_N^* \to L_M^*$. Then, for the full nonlinear operator, we have that $\mathfrak{A} : L_M^* \to L_M^*$ and has a bounded norm; that is, if θ is the null vector, then

$$\|\mathfrak{A}\Phi\|_M < c, \qquad \Phi \in T_\rho(\theta)_M, \tag{31}$$

where $T_\rho(\theta)_M$ refers to a sphere around θ with radius ρ of the space L_M^* and c is a constant which depends only on the functions $b(\mathbf{R})$ and $G(\Phi)$, and the number a,

$$\int\int M[K(\mathbf{R}_1, \mathbf{R}_2)]d\mathbf{R}_1 d\mathbf{R}_2 = a < \infty. \tag{32}$$

As is indicated in Ref. 2, these conditions will all be satisfied if

$$N[\mu G(\nu\Phi)] \leq M(\Phi) \tag{33}$$

for large Φ, where $\mu, \nu > 0$. In the case that f is strongly nonlinear, G should increase faster than a power function. Thus, by the inequality (33), the N function $M(u)$ should grow faster than a power function. Such N functions, generally speaking, satisfy the Δ_3 or Δ^2 condition (see Appendix).

We now state explicitly, in the form of two theorems, the (sufficient) conditions for which the nonlinear operator $\mathfrak{A}$ can be shown to be completely continuous.

Theorem 4.4: Let the following conditions be satisfied:

(a) For all α, let

$$\int_V \int_V M[\alpha K(\mathbf{R}_1, \mathbf{R}_2)]d\mathbf{R}_1 d\mathbf{R}_2 < \infty,$$

where $M(u)$ is some N function which satisfies the Δ_3 condition and where the complementary $N(v)$ satisfies the Δ' condition.

(b) Let the function $f(\mathbf{R}, \Phi)$ satisfy the inequality

$$|f(\mathbf{R}, \Phi)| \leq b(\mathbf{R}) + G(|\Phi|), \quad \mathbf{R} \in V_1 \quad -\infty < \Phi < \infty,$$

where

$$\int_V |b(\mathbf{R})|M^{(-1)}(|b(\mathbf{R})|)d\mathbf{R} < \infty$$

and $G(\Phi)$ is continuous and monotonically increasing for $\Phi \geq 0$.

(c) Let a constant γ exist such that for large Φ

$$G(\gamma\Phi) \leq M(\Phi).$$

Then, the operator $\mathfrak{A} = \mathfrak{K}\mathfrak{f}$ is completely continuous in some sphere of the space L_M^*.

Theorem 4.5: Let the following conditions be satisfied:

(a) For all α, let

$$\int_V \int_V M[\alpha K(\mathbf{R}_1, \mathbf{R}_2)]d\mathbf{R}_1 d\mathbf{R}_2 < \infty,$$

where the N function $M(u)$ satisfies the Δ^2 condition.

(b) Let Condition (b) of Theorem 4.4 hold.

(c) Let an N function $Q(u)$ exist such that for large u

$$G[Q(\Phi)] \leq M(\Phi).$$

Then, the operator $\mathfrak{A}$ is completely continuous in the whole space L_M^*.

The difference between the conclusions of the two theorems is that Theorem 4.4 establishes the complete continuity of $\mathfrak{A}$ in some sphere, say $T_\rho(\theta)_M$, of the space L_M^* while Theorem 4.5 gives complete continuity of $\mathfrak{A}$ in the whole space L_M^*. Although both theorems can be applied to establish the existence of at least one solution to the integral equation under study, it is sometimes more convenient to have complete continuity in the whole space L_{M^*}.

Let us now consider these theorems, one at a time, starting with Theorem 4.4. To apply this theorem, let $M(u) = e^{|u|} - |u| - 1$.

It is easy to show that $M(u)$ satisfies the Δ_3 condition, and that the complementary N function $N(v)$

satisfies the Δ' condition. From the arguments presented in Sec. II, Condition (a) is satisfied for all the kernels. Now, let $b(\mathbf{R}) \equiv 0$ and $G(\Phi) = \exp[\Phi]$, then clearly

$$|f(\mathbf{R}, \Phi(\mathbf{R}))| \leq b(\mathbf{R}) + G(|\Phi|)$$

and

$$\int_V |b(\mathbf{R})| M^{(-1)}(|b(\mathbf{R})|) < \infty,$$

since $M^{(-1)}(0) = 0$. Thus, Condition (b) of Theorem 4.4 is also satisfied. Now let $\gamma = \frac{1}{2}$; then,

$$G(\gamma\Phi) = \exp\left[\frac{\Phi}{2}\right] \leq \exp[|\Phi|] - |\Phi| - 1 \quad \text{for large } \Phi.$$

Therefore, Condition (c) is also satisfied, and hence the conclusion that $\mathfrak{A}$ is completely continuous in some sphere of the space L_M^*, for $M(u) = e^{|u|} - |u| - 1$.

Turning to Theorem 4.5, let

$$M(u) = e^{u^2} - 1.$$

Then $M(u)$ satisfies the Δ^2 condition, and from Sec. II, we conclude that Condition (a) is satisfied. If we identify $b(\mathbf{R}), G(\Phi)$ as before, then Condition (b) of this theorem is also satisfied. Finally, let $Q(u) = \frac{1}{2}|u|^2$. Then,

$$G[Q(\Phi)] = \exp\left[\frac{|\Phi|^2}{2}\right] \leq e^{\Phi^2} - 1 \quad \text{for large } \Phi,$$

and Condition (c) of Theorem 4.5 is satisfied. Therefore, we conclude that $\mathfrak{A}$ is completely continuous in the whole space L_M^*, where

$$M(u) = e^{u^2} - 1.$$

IV. EXISTENCE OF SOLUTIONS OF THE NONLINEAR INTEGRAL EQUATION

We begin this section by recalling that the complete continuity of the operator $\mathfrak{A}$ can be used to establish the existence of solutions of the integral equation under study. From the preceding sections we have the result that the operator $\mathfrak{A}$ is completely continuous in some sphere, say $T_\rho(\theta)_M$, of the space L_M^*. In order to cast our integral equation into standard form, let us identify

$$\mathfrak{K} = \lambda \mathfrak{K}', \tag{34}$$

and

$$\mathfrak{B} = \mathfrak{K}' \mathfrak{f}, \tag{35}$$

where $\beta z = -\lambda$. With these definitions, the integral equation (16) can be written as

$$\lambda \mathfrak{B}\Phi = \Phi. \tag{36}$$

The complete continuity of $\mathfrak{A}$ implies the complete continuity of $\mathfrak{B}$. Therefore,

$$\sup_{\Phi \in T_\rho(\theta)} \|\mathfrak{A}\Phi\|_M = d, \tag{37}$$

where d is some constant. Now, let $0 < \lambda \leq \frac{\rho}{d}$; then,

$$\sup_{\Phi \in T_\rho(\theta)} \|\mathfrak{B}\Phi\|_M = \sup_{\Phi \in T_\rho(\theta)} \|\lambda \mathfrak{B}\Phi\|_M \leq |\lambda| \sup_{\Phi \in T_\rho(\theta)} \|\mathfrak{B}\Phi\| \leq \rho. \tag{38}$$

Thus, for $|\lambda|$ small enough, the operator $\mathfrak{A}$ is completely continuous, and $\mathfrak{A} : T_\rho(\theta)_M \to T_\rho(\theta)_M$. We recall from Schauder's fixed point theorem that if the operator $\mathfrak{A}$ is completely continuous and maps the sphere $T_\rho(\theta)_M$ into itself, then there exists at least one solution of the associated equation, Eq. (36), in $T_\rho(\theta)_M$. This, in turn, implies that there exists at least one solution of the nonlinear integral equation for the singlet distribution function, Eq. (16).

It should be noted that the use of different ρ in the above condition on existence allows different upper bounds on the parameter $|\lambda|$. If we apply Theorem 4.4, the variation of ρ is bounded by the theorem itself. On the other hand, by applying Theorem 4.5, ρ can take on any finite value such that $\mathfrak{A}$ is completely continuous in $T_\rho(\theta)_M$. Therefore, in searching for an upper bound in λ that still guarantees the existence of a solution, we may be able to formulate a more favorable upper bound using Theorem 4.5, since there exists a larger range of λ for which the existence of at least one solution is insured.

The physical interpretation of this result may now be stated. Since, $\beta = 1/kT$ and $z \sim$density, the condition that βz be small enough means that the system is at high temperature and/or low densities. Under these conditions, the integral equation for the singlet distribution function has at least one solution. This result may be compared with the one obtained by Ruelle[9] in his analysis of the Kirkwood–Salsburg integral equations. Further comment on this point will be deferred until the following section.

V. DISCUSSION

In this paper we have discussed the conditions which must be satisfied in order to insure the existence of at least one solution to a particular integral equation, Eq. (16). Our interest in this equation was stimulated by the observation that regardless of whether one started from the BBKYG hierarchy or from the grand partition function and used a variational approach, the integral equation obtained for the singlet distribution function could always be cast into the Hammerstein form with an exponential nonlinearity. The kernel in the Hammerstein representation was found to depend on the particular approximation scheme used to effect closure. Hence, the point of view was taken that if one hoped to infer general results on thermodynamic systems from a study of approximate equations for the singlet distribution function, it was inadmissible to weaken the exponential nonlinearity. The various kernels which arise have a certain overall similarity in structure in that an intermolecular potential energy function is always multiplied by a pair-correlation function, and it is this feature which allows one to prove the condition on complete continuity of the associated operator $\mathfrak{K}$, provided one assumes the Ruelle–Fisher bound. The overall framework within which this problem was studied was that provided by Krasnosel'skii and Rutickii, namely the theory of Orlicz spaces, and in this paper we have focused on that part of the theory which deals with strongly nonlinear functions.

It is to be noticed that in choosing the $M(\Phi)$ used in Theorems 4.4 and 4.5, there is a tradeoff between the "good" properties of $\mathfrak{K}$ and those of $\mathfrak{f}$. That is, in establishing the complete continuity of $\mathfrak{K}$, we want

$M(\Phi)$ to be small for large Φ, whereas in establishing the continuity and boundedness of $\mathfrak{f}$, we want $M(\Phi)$ to be large for large Φ. In our problem, the strong nonlinearity of $\mathfrak{f}$ requires that $M(\Phi)$ grow faster than a power function, and, as a consequence, we cannot relax the strong restrictions on the singularities of the kernel. Of course, if one is willing to consider the case that $M(\Phi)$ grows slower than a power function, then far more general results than those obtained in Sec. IV can be proved. For example, one can prove that the existence of a solution of $\lambda\mathfrak{A}\Phi = \Phi$ is guaranteed for all λ. However, the weakening of the nonlinearity leads to equations which, though more tractable, may have nothing to do with statistical mechanics.

There is, of course, one approach that can be pursued in order to obtain more general conditions on existence, and possibly on uniqueness as well. That approach is to restrict even further the singularities of the kernel. For example, if we restrict the kernel to be symmetric and positive definite, then the following theorem[2] can be proved:

Theorem: Let the symmetric, positive definite kernel satisfy the condition

$$\int_V\int_V \exp|K(\mathbf{R}_1,\mathbf{R}_2)|^{1+\epsilon}d\mathbf{R}_1 d\mathbf{R}_2 < \infty, \qquad \epsilon > 0.$$

Let

$$|f(\mathbf{R},\Phi)| \leq b + e^{\alpha\Phi}, \qquad \alpha > 0, \mathbf{R} \in V, -\infty < \Phi < \infty.$$

If λ_0 is the largest eigenvalue of the self-adjoint operator

$$\mathfrak{A}\Phi(\mathbf{R}_1) = \int_V K(\mathbf{R}_1,\mathbf{R}_2) f[\mathbf{R}_2, \Phi(\mathbf{R}_2)]d\mathbf{R}_2$$

and

$$\frac{1}{\lambda}\int_0^u f(\mathbf{R},\Phi)d\Phi \leq \tfrac{1}{2}a\Phi^2 + b, \qquad \mathbf{R} \in V, -\infty < \Phi < \infty,$$

where $a < 1/\lambda_0$ and $b > 0$, then the equation

$$\frac{1}{\lambda}\int_V K(\mathbf{R}_1,\mathbf{R}_2)[f(\mathbf{R}_2, \Phi(\mathbf{R}_2))]d\mathbf{R}_2 = \Phi(\mathbf{R}_1)$$

has at least one solution $\Phi_0(\mathbf{R})$ which satisfies the condition

$$\int_V \exp|\Phi_0(\mathbf{R})|d\mathbf{R} < \infty.$$

In terms of Eq. (16), this result states that if one of the kernels, Eqs. (17), (19), (20), or (21) can be proved to be symmetric and positive definite, and in addition, satisfies the conditions on complete continuity stated in Sec. II, then there exists at least one solution $\Phi(\mathbf{R})$ such that

$$\frac{1}{z}\int_V \rho(\mathbf{R})d\mathbf{R} < \infty.$$

Since, for realistic potentials, the kernel will have both positive and negative parts depending on the range of R_{12} under study, it is clear that this theorem is probably too restrictive. In fact, using the Brout kernel and choosing $u^{(1)}(R_{12})$ to be a "soft" repulsive potential and $g^{(2)}(R_{12})$ to be the pair-correlation function for a system of hard spheres, this result amounts to saying that there exists at least one solution to the nonlinear integral equation for the singlet distribution function for a system interacting with purely repulsive forces.

In assessing the theoretical results presented in this paper, it is well to keep in mind that the paucity of more general theorems on the existence and uniqueness of solutions of the nonlinear integral equation (16) for systems interacting with realistic potentials is a reflection of the fundamental work that has yet to be done in developing the basic theory of Orlicz spaces. For example, the integral equation for the singlet distribution function could be studied only if V represented a bounded, closed set in a finite-dimensional space, that is, for a system of molecules enclosed in a finite volume. To the best of our knowledge, nothing whatever is known, either in the theory of Orlicz spaces or in the simpler theory of L^p spaces, on the existence properties of nonlinear equations for measure $\mu(V) = \infty$.

Finally, we comment on the relationship between the work presented here and the results obtained by Ruelle[9] in his study of the Kirkwood–Salsburg integral equations. In the paper of Ruelle, the integral equations studied are essentially linear, whereas those investigated here are nonlinear. In both cases, it has been found possible to introduce a well-defined norm on a Banach space (here, an Orlicz space). Because his study deals with linear integral equations, Ruelle can prove uniqueness as well as existence, whereas here only existence of solutions can be established. Furthermore, because of the way the problem is formulated, it is not necessary for Ruelle to introduce a closure, that is, an approximation for decoupling the hierarchy, whereas for the various representations studied in this paper, this was necessary. It is in this sense that the integral equations studied in this paper are approximate, while those studied in Ruelle's paper are not.

In view of the above remarks, it might well be asked why it was necessary to go to a more difficult theory, i.e., nonlinear integral equations defined on Orlicz space, when an essentially simpler theory, linear integral equations defined on a Banach space, yields exact results on existence *and* uniqueness. This question might be answered in several ways. First of all, as Ruelle implies, it is difficult to see how his bound on fugacity which gives existence and uniqueness can be generalized to higher densities and lower temperatures, whereas using the theory of Orlicz spaces at least there exists a mechanism for generalizing the bound obtained on Eq. (16); one searches for a function $M(u)$ for which a "tradeoff" between "good" and "bad" properties of kernel and nonlinear function can be effected. Secondly, several workers have taken the point of view that a phase transition can in some sense be analyzed by considering changes in the solutions of certain nonlinear integral equations derived from the BBKYG hierarchy.[12,13,5] In particular, one considers the possibility of bifurcation of solutions of Eq. (16) corresponding to certain eigenvalues of the kernel. The phenomenon of bifurcation for completely continuous kernels is a property that is restricted to nonlinear equations, and hence it would appear that the linearized system of equations studied by Ruelle might not be the best framework for investigating the general theory of phase transitions. Finally, in view of our observation that the dependence of pressure on density for physical systems, Fig. 1, essentially defines an N function, it is possible that relationships between thermodynamics and the under-

lying statistical-mechanical theory might be clearly displayed if one insisted on working in a Banach space appropriate to both classes of problems, that is, Orlicz space. It is this program of research which will be pursued in a forthcoming publication, wherein the algebraic aspects of thermodynamic theory will be developed using the theory of N functions.

APPENDIX

Let $M_1(u)$ and $M_2(u)$ be N functions. We write $M_1(u) \prec M_2(u)$ if there exist positive constants u_0 and k such that

$$M_1(u) \leq M_2(ku), \quad u \geq u_0.$$

Two N functions $M_1(u)$ and $M_2(u)$ are said to be equivalent,

$$M_1(u) \sim M_2(u),$$

if $M_1(u) \prec M_2(u)$ and $M_2(u) \prec M_1(u)$.

An N function $M(u)$ is said to satisfy the Δ_2 condition if there exists constants $k > 0, u_0 \geq 0$, such that

$$M(2u) \leq kM(u), \quad u \geq u_0.$$

An N function $M(u)$ is said to satisfy the Δ' condition it there exist positve constants c and u_0, such that

$$M(uv) \leq cM(u)M(v), \quad u, v \geq u.$$

An N function $M(u)$ is said to satisfy the Δ_3 condition if

$$M(u) \sim |u| M(u).$$

Finally, an N function $M(u)$ satisfies the Δ^2 condition if

$$M(u) \sim M^2(u).$$

---

* Reilly Fellow

1 M. A. Krasnosel'skii and Y. B. Rutickii, *Convex Functions and Orlicz Spaces* (P. Noordhoff, Groningen, 1961).
2 M. A. Krasnosel'skii and Y. B. Rutickii, Am. Math. Soc. Trans. **60**, 51 (1967).
3 T. Morita and K. Hiroike, Progr. Theoret. Phys. (Kyoto) **25**, 537 (1961).
4 A. Hammerstein, Acta Math. **54**, 117 (1930).
5 J. D. Weeks, S. A. Rice, and J. J. Kozak, J. Chem. Phys. **52**, 2416 (1970).
6 Professor K. Hiroike in a private communication to Professor S. A. Rice has drawn attention to H. S. Green's derivation, given in *Molecular Theory of Fluids* (North Holland, Amsterdam, 1952).
7 R. Brout, *Phase Transitions* (Benjamin, New York, 1965).
8 C. L. Dolph and G. J. Minty, in *Nonlinear Integral Equations*, edited by P. M. Anselone (University of Wisconsin Press, Madison, 1964).
9 D. Ruelle, Ann. Phys. (N.Y.) **25**, 109 (1963).
10 T. L. Hill, *Statistical Mechanics* (McGraw-Hill, New York 1957).
11 See, for example, D. Ruelle in *The Equilibrium Theory of Classical Fluids*, edited by H. L. Frisch and J. L. Lebowitz (Benjamin, New York, 1964).
12 A. A. Vlasov, *Many-Particle Theory and Its Application to Plasma* (Gordon and Breach, New York, 1961).
13 A. G. Khachaturyan, Fiz. Tverd. Tela 5, 26 (1963) [Sov. Phys. Solid State 5, 16 (1963)].

Wave Propagation in Certain One-Dimensional Random Media

John E. Molyneux*

Department of Mechanical and Aerospace Sciences, University of Rochester, Rochester, New York 14627

Receipt date (Received 15 July 1971)

In this paper we investigate the stochastic ordinary differential equation $u'' + k_0^2[1 + \epsilon y(t)]u = 0$ with $y(t)$ a random process. Two specific types of process $y(t)$ are considered. Both of these arise from a bounded mapping $y(t) = f(x(t))$ of a countable state space Markov process $x(t)$. Exact equations are derived for the statistical moments of $u(t)$, and the behavior of the first two moments is discussed in the limit of small ϵ. A description of the layered media to which our results apply is given and a comparison of our exact results with certain perturbation methods is made.

I. INTRODUCTION

In this paper, we will determine the statistical moments of the solution matrix $U(t) = [u_{ij}(t)]$, $i,j = 1,2$, of the initial value problem

$$\frac{dU}{dt} = A(t,\epsilon)U, \quad t \in [0,\infty), \quad \epsilon > 0, \quad U(0) = I_2, \quad (1)$$

where $I_2 = [\delta_{ij}]$, $A(t,\epsilon) = \{\delta_{i1}\delta_{j2} - k_0^2[1 + \epsilon y(t)]\delta_{i2}\delta_{j1}\}$, δ_{ij} is the Kronecker delta, and $y(t)$ is a stochastic process. The particular type of stochastic process which we will consider has the form $y(t) = f(x(t))$, where $x(t)$ is a countable state space Markov process and f is a real valued function which is periodic in the following sense. If $\{a_0, a_1, a_2, \ldots\}$ is the set of possible values for $x(t)$, then there exists an integer $M > 0$ such that $f(a_{k+M}) = f(a_k)$ for all $k = 0, 1, 2, \ldots$. In this case, the range of y is a finite set of real numbers $\{b_0, b_1, b_2, \ldots, b_{M-1}\}$, where the $b_k = f(a_k)$ are not necessarily distinct.

For $x(t)$ a finite state space Markov process, Eq. (1) has been studied by McKenna and Morrison.[1,2] In Sec. II we use their results to derive partial differential equations for functions $G_k(U,t), k = 0, 1, \ldots, M-1$, where $G_k(U,t)du_{11}\cdots du_{22} = \text{Prob}\{u_{ij} < u_{ij}(t) < u_{ij} + du_{ij}, i,j = 1,2, y(t) = b_k\}$. For a general countable state space Markov process these equations contain additional unknown functions, but we have found two types of process (which we call types I and II) for which the G_k can be determined.

The principal results of the paper are obtained in Sec. III. In this section we derive the equations satisfied by the statistical moments of $U(t)$ when $x(t)$ is a

lying statistical-mechanical theory might be clearly displayed if one insisted on working in a Banach space appropriate to both classes of problems, that is, Orlicz space. It is this program of research which will be pursued in a forthcoming publication, wherein the algebraic aspects of thermodynamic theory will be developed using the theory of N functions.

APPENDIX

Let $M_1(u)$ and $M_2(u)$ be N functions. We write $M_1(u) \prec M_2(u)$ if there exist positive constants u_0 and k such that

$$M_1(u) \leq M_2(ku), \quad u \geq u_0.$$

Two N functions $M_1(u)$ and $M_2(u)$ are said to be equivalent,

$$M_1(u) \sim M_2(u),$$

if $M_1(u) \prec M_2(u)$ and $M_2(u) \prec M_1(u)$.

An N function $M(u)$ is said to satisfy the Δ_2 condition if there exists constants $k > 0, u_0 \geq 0$, such that

$$M(2u) \leq kM(u), \quad u \geq u_0.$$

An N function $M(u)$ is said to satisfy the Δ' condition it there exist positve constants c and u_0, such that

$$M(uv) \leq cM(u)M(v), \quad u, v \geq u.$$

An N function $M(u)$ is said to satisfy the Δ_3 condition if

$$M(u) \sim |u| M(u).$$

Finally, an N function $M(u)$ satisfies the Δ^2 condition if

$$M(u) \sim M^2(u).$$

---

* Reilly Fellow

1 M. A. Krasnosel'skii and Y. B. Rutickii, *Convex Functions and Orlicz Spaces* (P. Noordhoff, Groningen, 1961).
2 M. A. Krasnosel'skii and Y. B. Rutickii, Am. Math. Soc. Trans. **60**, 51 (1967).
3 T. Morita and K. Hiroike, Progr. Theoret. Phys. (Kyoto) **25**, 537 (1961).
4 A. Hammerstein, Acta Math. **54**, 117 (1930).
5 J. D. Weeks, S. A. Rice, and J. J. Kozak, J. Chem. Phys. **52**, 2416 (1970).
6 Professor K. Hiroike in a private communication to Professor S. A. Rice has drawn attention to H. S. Green's derivation, given in *Molecular Theory of Fluids* (North Holland, Amsterdam, 1952).
7 R. Brout, *Phase Transitions* (Benjamin, New York, 1965).
8 C. L. Dolph and G. J. Minty, in *Nonlinear Integral Equations*, edited by P. M. Anselone (University of Wisconsin Press, Madison, 1964).
9 D. Ruelle, Ann. Phys. (N.Y.) **25**, 109 (1963).
10 T. L. Hill, *Statistical Mechanics* (McGraw-Hill, New York 1957).
11 See, for example, D. Ruelle in *The Equilibrium Theory of Classical Fluids*, edited by H. L. Frisch and J. L. Lebowitz (Benjamin, New York, 1964).
12 A. A. Vlasov, *Many-Particle Theory and Its Application to Plasma* (Gordon and Breach, New York, 1961).
13 A. G. Khachaturyan, Fiz. Tverd. Tela 5, 26 (1963) [Sov. Phys. Solid State 5, 16 (1963)].

Wave Propagation in Certain One-Dimensional Random Media

John E. Molyneux*

Department of Mechanical and Aerospace Sciences, University of Rochester, Rochester, New York 14627

Receipt date (Received 15 July 1971)

In this paper we investigate the stochastic ordinary differential equation $u'' + k_0^2[1 + \epsilon y(t)]u = 0$ with $y(t)$ a random process. Two specific types of process $y(t)$ are considered. Both of these arise from a bounded mapping $y(t) = f(x(t))$ of a countable state space Markov process $x(t)$. Exact equations are derived for the statistical moments of $u(t)$, and the behavior of the first two moments is discussed in the limit of small ϵ. A description of the layered media to which our results apply is given and a comparison of our exact results with certain perturbation methods is made.

I. INTRODUCTION

In this paper, we will determine the statistical moments of the solution matrix $U(t) = [u_{ij}(t)]$, $i,j = 1,2$, of the initial value problem

$$\frac{dU}{dt} = A(t,\epsilon)U, \quad t \in [0,\infty), \quad \epsilon > 0, \quad U(0) = I_2, \quad (1)$$

where $I_2 = [\delta_{ij}]$, $A(t,\epsilon) = \{\delta_{i1}\delta_{j2} - k_0^2[1 + \epsilon y(t)]\delta_{i2}\delta_{j1}\}$, δ_{ij} is the Kronecker delta, and $y(t)$ is a stochastic process. The particular type of stochastic process which we will consider has the form $y(t) = f(x(t))$, where $x(t)$ is a countable state space Markov process and f is a real valued function which is periodic in the following sense. If $\{a_0, a_1, a_2, \ldots\}$ is the set of possible values for $x(t)$, then there exists an integer $M > 0$ such that $f(a_{k+M}) = f(a_k)$ for all $k = 0, 1, 2, \ldots$. In this case, the range of y is a finite set of real numbers $\{b_0, b_1, b_2, \ldots, b_{M-1}\}$, where the $b_k = f(a_k)$ are not necessarily distinct.

For $x(t)$ a finite state space Markov process, Eq. (1) has been studied by McKenna and Morrison.[1,2] In Sec. II we use their results to derive partial differential equations for functions $G_k(U,t), k = 0, 1, \ldots, M-1$, where $G_k(U,t)du_{11}\cdots du_{22} = \text{Prob}\{u_{ij} < u_{ij}(t) < u_{ij} + du_{ij}, i,j = 1,2, y(t) = b_k\}$. For a general countable state space Markov process these equations contain additional unknown functions, but we have found two types of process (which we call types I and II) for which the G_k can be determined.

The principal results of the paper are obtained in Sec. III. In this section we derive the equations satisfied by the statistical moments of $U(t)$ when $x(t)$ is a

type I or II process, and we study in detail specific examples of each type of process. For each of the processes studied, we determine explicitly the behavior of the first two moments of $U(t)$ in the limit of small ϵ, and we compare these results with those obtained by various perturbation techniques.

Equation (1) is, of course, equivalent to the second order ordinary differential equation $u'' + k_0^2[1 + \epsilon y(t)]u = 0$ with the initial conditions $u(0) = 1$, $u'(0) = 0$ or $u(0) = 0$, $u'(0) = 1$. Assuming t has dimensions of length and letting $k_0^2 = \omega^2/c^2$, where c is the velocity of light, we can interpret u as the spacial part of the time harmonic (circular frequency ω) electric field in a one-dimensional random medium with refractive index $n^2(t) = 1 + \epsilon y(t)$. Since $y(t)$ can take only a finite set of values, the medium is composed of layers of random thickness. The particular example of the type I process which we treat in detail in Sec. III corresponds to a medium in which the refractive index of the layers alternately increases and decreases by increments ϵ between the extreme values $1 - \frac{1}{2}\epsilon K, 1 + \frac{1}{2}\epsilon K$, where K is a positive integer and $M = 2K$. The graph of $n^2(t)$, therefore, looks something like a jagged random sine wave. The type II process which we study corresponds to a medium whose refractive index is bounded by the values $1 \pm \epsilon K'$ (where K' is a positive integer and $M = 4K'$), alternately increases and decreases by increments ϵ between extreme values that on the average decrease with increasing t. The graph of $n^2(t)$ in this case resembles a jagged, randomly damped, random sine wave.

II. GOVERNING EQUATIONS

We first consider the definition of the stochastic process $y(t)$ appearing in Eq. (1). Let $x(t)$ be a countable state space Markov process defined for $t \in T = [t_0, \infty)$, $t_0 \leq 0$, and assuming values in an arbitrary set $S = \{a_j : j \in J\}$, where J is $Z^+ = \{0, 1, 2, \ldots\}$ or $Z = \{\ldots -2, -1, 0, 1, 2, \ldots\}$. All distribution functions of $x(t)$ are completely determined by the initial distribution[3]

$$p_j^0 = P\{x(t_0) = a_j\}, \quad j \in J \tag{2}$$

and the conditional probabilities that $x(t_2) = a_j$ given that $x(t_1) = a_i$, $t_0 \leq t_1 \leq t_2$, which we denote by

$$p_{ij}(t_1, t_2) = P\{x(t_2) = a_j | x(t_1) = a_i\}, \quad i, j \in J. \tag{3}$$

It is convenient for our purposes to assume that the probabilities p_{ij} defining the process $x(t)$ arise in the following way. We suppose that there exist quantities $\lambda_{ij}(t)$ satisfying

$$\lambda_{ij}(t) \geq 0, \quad i \neq j, \quad \lambda_{ii}(t) \leq 0, \tag{4}$$

$$\sum_{k \in J} \lambda_{ik}(t) = 0 \quad \text{for every } t \in T, \tag{5}$$

and for $\Delta t \to 0$ we define

$$p_{ii}(t, t + \Delta t) = 1 + \lambda_{ii}(t)\Delta t + O(\Delta t), \tag{6}$$

$$p_{jk}(t, t + \Delta t) = \lambda_{jk}(t)\Delta t + O(\Delta t). \tag{7}$$

Then, assuming that the $\lambda_{ij}(t)$ are bounded, it is shown in Feller[4] that Eqs. (6) and (7) together with the Chapman–Kolmogorov identity uniquely determine the probabilities $p_{ij}(t_1, t_2)$. In addition to Eqs. (4) and (5) and the boundedness assumption, we shall require that the $\lambda_{ij}(t)$ satisfy

$$\lambda_{ij}(t) = \lambda(i - j, t) \tag{8}$$

and that there exist nonnegative integers L and N such that

$$\lambda_{ij}(t) \neq 0 \quad \text{only for } -L \leq i - j \leq N. \tag{9}$$

Given a Markov process $x(t)$ whose state space is $S = \{a_j, j \in J\}$, we define the stochastic process $y(t)$ as follows. Let f be a real valued function defined on S and assume that there exists a positive integer M such that $f(a_{k+M}) = f(a_k)$, $k \in J$. Now, let $y(t) = f(x(t))$, $t \in T$. Then the range of $y(t)$, $R \equiv y(T) = f(S) = \{b_0, b_1, \ldots, b_{M-1}; b_k = f(a_k)\}$ is a finite set and all distribution functions of y are uniquely determined. Indeed, let J_i denote either $Z_i^+ = \{k \in Z^+; f(a_k) = b_i\}$ or $Z_i = \{k \in Z; f(a_k) = b_i\}$; $p_j(t) \equiv \sum_{\nu \in J} p_\nu^0 p_{\nu j}(t_0, t) = P\{x(t) = a_j\}$. Then it is easy to see that

$$P\{y(t) = b_i\} = \sum_{k \in J_i} P\{x(t) = a_k\} = \sum_{k \in J_i} p_k(t), \tag{10}$$

$$P\{y(t_1) = b_{i_1}, y(t_2) = b_{i_2}, \ldots, y(t_n) = b_{i_n}\} = \sum_{k_1 \in J_{i_1}} \cdots \sum_{k_n \in J_{i_n}} p_{k_1}(t_1) p_{k_1 k_2}(t_1, t_2) \cdots p_{k_{n-1} k_n}(t_{n-1}, t_n),$$

$$t_1 \leq t_2 \leq \cdots \leq t_n, \quad n = 2, 3, \ldots. \tag{11}$$

The process $y(t)$ is, therefore, completely defined by Eqs. (10) and (11) for any given $\lambda_{ij}(t)$ satisfying Eqs. (4), (5), (8), and (9). We now turn to the derivation of the equations for the moments of $U(t)$.

Let $E_j(U, t)$ be the event (ω-set) $E_j(U, t) = \{u_{\alpha\beta} \leq u_{\alpha\beta}(t) < u_{\alpha\beta} + du_{\alpha\beta}, \ \alpha, \beta = 1, 2, \ x(t) = a_j\}$ and define

$$F_j(U, t) du_{11} du_{21} du_{12} du_{22} = P(E_j(U, t)), \quad j \in J. \tag{12}$$

Using the phase space method of McKenna and Morrison[1,2] or the methods described in Frisch[5] it is possible to show rigorously that the F_j are the weak solutions of the partial differential equations

$$\frac{\partial F_j}{\partial t} + \sum_{\alpha, \beta, \nu = 1}^{2} a_{\alpha\nu}^{(j)} u_{\nu\beta} \frac{\partial F_j}{\partial u_{\alpha\beta}} - \sum_{k \in J} \lambda_{kj}(t) F_k = 0, \quad j \in J, \tag{13}$$

where $a_{\alpha\beta}^{(j)} = \delta_{\alpha 1}\delta_{\beta 2} - k_0^2[1 + \epsilon f(a_j)]\delta_{\alpha 2}\delta_{\beta 1}$, and that the F_j satisfy the initial conditions

$$F_j(U, 0) = p_j(0)\delta(u_{11} - 1)\delta(u_{21})\delta(u_{12})\delta(u_{22} - 1). \tag{14}$$

For completeness, we shall give a brief hueristic derivation of Eq. (13). Let Δt be positive and consider the event $E_j(U, t + \Delta t)$. This event can occur in the following mutually exclusive ways. First, with probability $1 + \lambda_{jj}(t)\Delta t + o(\Delta t)$, $x(t + \Delta t) = a_j$ and x does not change between t and $t + \Delta t$. In this case,

$$u'_{\alpha\beta} \equiv u_{\alpha\beta}(t + \Delta t) = u_{\alpha\beta}(t) + \sum_{\nu = 1}^{2} a_{\alpha\nu}^{(j)} u_{\nu\beta}(t)\Delta t + o(\Delta t), \tag{15}$$

and we note that the Jacobian $|\partial u'_{\alpha\beta}/\partial u_{\sigma\nu}| = 1 + o(\Delta t)$. Second, with probability $\lambda_{kj}(t)\Delta t + o(\Delta t)$, we have $x(t) = a_k$, $k \neq j$, $x(t + \Delta t) = a_j$, and x changes value just once in $[t, t + \Delta t]$. In this case,

$$u'_{\alpha\beta} \equiv u_{\alpha\beta}(t + \Delta t) = u_{\alpha\beta}(t) + o(\Delta t) \tag{16}$$

and the Jacobian is $1 + o(\Delta t)$. Third, with probability $o(\Delta t)$, x changes value two or more times in $[t, t + \Delta t]$. Expressing $E_j(U, t + \Delta t)$ as the union of these events and taking probabilities, we find that

$$F_j(U,t+\Delta t)du_{11}\cdots du_{22} = \Big(F_j(U - A^{(j)}U\Delta t, t)[1 + \lambda_{jj}(t)\Delta t] + \Delta t\sum_{\substack{k\in J\\ k\neq j}} \lambda_{kj}(t)F_k(U,t) + o(\Delta t)\Big)du_{11}\cdots du_{22}, \quad (17)$$

where $A^{(j)} = [a^{(j)}_{\alpha\beta}]$. We obtain the first term on the right-hand side of Eq. (17) by using the facts (i) that the inverse of the transformation $u'_{\alpha\beta} = \psi_{\alpha\beta}(u_{\sigma\upsilon})$, where $\psi_{\alpha\beta}$ is given by Eq. (15), is obtained by reversing the sign of Δt in this equation, (ii) that $u_{\alpha\beta}(t + \Delta t)$ belongs to an interval $I_{\alpha\beta} = (u_{\alpha\beta}, u_{\alpha\beta} + du_{\alpha\beta})$ if and only if $u_{\alpha\beta}(t)$ belongs to $\psi^{-1}_{\alpha\beta}(I_{\alpha\beta}) \equiv J_{\alpha\beta}$, and (iii) that $P\{u_{\alpha\beta}(t) \in J_{\alpha\beta}, x(t) = a_j, x(t + \Delta t) = a_j, x(t)$ does not change in $(t, t + \Delta t)\} = P\{u_{\alpha\beta}(t) \in J_{\alpha\beta}, x(t) = a_j\}$ $P\{x(t + \Delta t) = a_j$, and x does not change in $(t, t + \Delta t) | x(t) = a_j\} = F_j(U - A^{(j)}U\Delta t, t)[1 + \lambda_{jj}(t)\Delta t + O(\Delta t)]$ $du_{11}\cdots du_{22}$. The second term is obtained by similar reasoning. If we expand $F_j(U - A^{(j)}U\Delta t, t)$ in a Taylors series, divide Eq. (17) by $du_{11}\cdots du_{22}$, and allow Δt to approach zero, we obtain Eq. (13). The initial conditions (14) are an obvious consequence of the definition of $F_j(U, t)$ and the fact that $U(0) = I_2$.

Since there are an infinite number of F_j, Eq. (13) is of little practical value. Let us consider, therefore, the event

$$E^{(k)}(U,t) = \bigcup_{j\in J_k} E_j(U,t) = \{u_{\alpha\beta} \le u_{\alpha\beta}(t) < u_{\alpha\beta} + du_{\alpha\beta}, \alpha, \beta = 1,2, y(t) = b_k\}$$

and define

$$G_k(U,t)du_{11}\cdots du_{22} = P(E^{(k)}(U,t)) = \Big(\sum_{j\in J_k} F_j(U,t)\Big)du_{11}\cdots du_{22}. \quad (18)$$

We note that for all $j \in J_k$, $a^{(j)}_{\alpha\beta} = b^{(k)}_{\alpha\beta} = \delta_{\alpha 1}\delta_{\beta 2} - k_0^2(1 + \epsilon b_k)\delta_{\alpha 2}\delta_{\beta 1}$. Hence, summing Eq. (13) over all $j \in J_k$, we obtain

$$\frac{\partial G_k}{\partial t} + \sum_{\alpha,\beta,\nu=1} b^{(k)}_{\alpha\nu}u_{\nu\beta}\frac{\partial G_k}{\partial u_{\alpha\beta}} - S_k = 0, \quad k = 0,1,2,\ldots,M-1, \quad (19)$$

where

$$S_k = \sum_{j\in J_k}\sum_{m\in J} F_m\lambda_{mj}(t). \quad (20)$$

A similar summation applied to Eq. (14) gives

$$G_k(U,0) = q_k(0)\delta(u_{11} - 1)\delta(u_{21})\delta(u_{12})\delta(u_{22} - 1), \quad (21)$$

where $q_k(0) \equiv P\{y(0) = b_k\}$ is given by Eq. (10). Equations (19) cannot, in general, be solved for G_k since the S_k are also unknown. However, we have found two cases in which S_k can be expressed entirely in terms of λ_{ij} and G_l. First, if $J = Z^+$ and $N = 0$ in Eq. (9), it is shown in Appendix A that, for $M > L$,

$$S_n = \sum_{k=0}\lambda(k-n)G_k + \sum_{k=M-L+n}^{M-1}\lambda(k-n-M)G_k, \quad n = 0,1,\ldots L-1, \quad (22)$$

$$S_n = \sum_{k=n-L}^{n}\lambda(k-n)G_k, \quad n = L, L+1,\ldots,M-1.$$

where $\lambda(k - n) \equiv \lambda_{kn}(t) = \lambda(k - n, t)$. Second, if $J = Z$, it is shown in Appendix A that for $M > L + N$,

$$S_n = \sum_{m=0}^{n+N}\lambda(m-n)G_m + \sum_{m=1}^{L-n}\lambda(-m-n)G_{M-m}, \quad n = 0,1,\ldots,L-1,$$

$$S_n = \sum_{m=n-L}^{n+N}\lambda(m-n)G_m, \quad n = L, L+1,\ldots,M-N-1, \quad (23)$$

$$S_n = \sum_{m=n-L}^{M-1}\lambda(m-n)G_m + \sum_{m=0}^{n-(M-N)}\lambda(m-n+M)G_m, \quad n = M-N,\ldots,M-1.$$

In these two cases, therefore, we may determine G_k, $k = 0, 1, \ldots, M - 1$ as the (weak) solutions of Eq. (19) satisfying the initial conditions of Eq. (21). The moments of $U(t)$ may be calculated from $G_k(U, t)$ using the formulas

$$\langle u_{\alpha_1\beta_1}(t)\cdots u_{\alpha_n\beta_n}(t)\rangle = \sum_{k=0}^{M-1}\langle u_{\alpha_1\beta_1}(t)\cdots u_{\alpha_n\beta_n}(t)\rangle_k, \quad \alpha_i,\beta_i = 1,2, \quad (24)$$

where

$$\langle u_{\alpha_1\beta_1}(t)\cdots u_{\alpha_n\beta_n}(t)\rangle_k \equiv \int_{-\infty}^{\infty} G_k(U,t)u_{\alpha_1\beta_1}\cdots u_{\alpha_n\beta_n}du_{11}\cdots du_{22} \quad (25)$$

and where we use brackets to denote statistical average. Because Eq. (19) is homogeneous in $u_{\alpha\beta}$, we may calculate the nth order moments of Eq. (24) directly. In fact, it is obvious that if for each $k = 0, 1, \ldots, M - 1$ we multiply the equation for G_k by the $\binom{n+3}{3}$ distinct nth order products $u_{\alpha_1\beta_1}\cdots u_{\alpha_n\beta_n}$ and integrate over all $u_{11}, \ldots, u_{22}$, we will obtain $M\binom{n+3}{3}$ coupled first order ordinary differential equations for the $M\binom{n+3}{3}$ distinct nth order moments in Eq. (25). The initial conditions $\langle u_{\alpha_1\beta_1}(0)\cdots u_{\alpha_n\beta_n}(0)\rangle = q_k(0)\delta_{\alpha_1\beta_1}\delta_{\alpha_2\beta_2}\cdots\delta_{\alpha_n\beta_n}$ for these equations are obtained from Eq. (21).

III. APPLICATIONS

In this section we consider specific examples of stochastic processes $x(t)$ and $y(t) = f(x(t))$ for which Eqs. (22) and (23) are valid. A process for which $J = Z^+$, $N = 0$, and S_n is given by Eq. (22) will be called a type I process. A process for which $J = Z$ and S_n is given by Eq. (23) will be called a type II process.

The class of type I processes which we shall study may be described as follows. We take $a_j = j, j \in Z^+$, $L = 1$, $N = 0$, $\lambda_{ii} = \lambda(0) = -\lambda$, $\lambda_{i,i+1} = \lambda(-1) = \lambda$, where λ is a positive constant. Then $x(t)$ is the Poisson process and $p_{ij}(\tau) \equiv p_{ij}(t, t + \tau) = 0$ for $j < i$, $p_{ij}(\tau) = [(j - i)!]^{-1}(\lambda\tau)^{j-i}\exp(-\lambda\tau)$ for $j \ge i$. We assume that $p^0_j = \delta_{j0}$. For construction of the process $y(t)$, we take $M = 2K$, let $g(j) = j$, $j = 0, 1, \ldots, K$, $g(j) = 2K - j$, $j = K + 1, \ldots, 2K - 1$, $g(j + 2K) = g(j)$, $j \in Z^+$, and define $f(j) = g(j) - \frac{1}{2}K$. The elements of the set $\{b_0, b_1, \ldots, b_{2K-1}\}$ of possible values of $y(t)$ are $b_0 = -K/2, b_i = b_{2K-i} = i - K/2, i = 1, 2, \ldots, K$. Various properties of $y(t)$ are derived in Appendix B.

From Eq. (22), we find that $S_0 = -\lambda G_0 + \lambda G_{2K-1}$ and $S_n = \lambda G_{n-1} - \lambda G_n$, $n = 1,2,\ldots,2K-1$. Thus, the equations for G_k are

$$\frac{\partial G_0}{\partial t} + \mathcal{L}_0 G_0 + \lambda G_0 - \lambda G_{2K-1} = 0$$
$$\frac{\partial G_k}{\partial t} + \mathcal{L}_k G_k + \lambda G_k - \lambda G_{k-1} = 0, \quad k = 1,2,\ldots,2K-1, \tag{26}$$

where

$$\mathcal{L}_m \equiv \sum_{\alpha,\beta,\nu=1}^{2} b_{\alpha\nu}^{(m)} u_{\nu\beta} \frac{\partial}{\partial u_{\alpha\beta}} = \sum_{\beta=1}^{2} \left\{ u_{2\beta} \frac{\partial}{\partial u_{1\beta}} - \kappa_m u_{1\beta} \frac{\partial}{\partial u_{2\beta}} \right\},$$
$$\kappa_m \equiv k_0^2(1 + \epsilon b_m).$$

If we multiply Eq. (26) by $\prod_{i=1}^{n} u_{\alpha_i\beta_i}$ and integrate over $u_{11},\ldots,u_{22}$, we may derive ordinary differential equations for the moments

$$U^{(k)}_{\alpha_1\beta_1\cdots\alpha_n\beta_n}(t) = \langle u_{\alpha_1\beta_1}(t)\cdots u_{\alpha_n\beta_n}(t)\rangle_k,$$

$\alpha_i, \beta_i = 1,2, k = 0,1,\ldots,2K-1$. In fact, we may show by elementary manipulation that

$$\int \prod_{i=1}^{n} u_{\alpha_i\beta_i} \mathcal{L}_k G_k \, du_{11}\cdots du_{22}$$
$$= -\sum_{i=1}^{n} \delta_{\alpha_i 1} U^{(k)}_{\alpha_1\beta_1\cdots\alpha_{i-1}\beta_{i-1}2\beta_i\alpha_{i+1}\cdots\alpha_n\beta_n} + \kappa_k \sum_{i=1}^{n} \delta_{\alpha_i 2} U^{(k)}_{\alpha_1\beta_1\cdots\alpha_{i-1}1\beta_i\alpha_{i+1}\beta_{i+1}\cdots\alpha_n\beta_n}, \tag{27}$$

and from this equation and Eq. (26) it is clear that for each fixed set of values $\beta_1,\beta_2,\ldots,\beta_n$ we must solve a system of $2K\cdot 2^n$ first-order equations for $U^{(k)}_{\alpha_1\beta_1\cdots\alpha_n\beta_n}$, $\alpha_i = 1,2$, $k = 0,1,\ldots,2K-1$. For $n = 1$, we let $U_{\alpha\beta}$ be the $2K \times 1$ column vector $[U^{(0)}_{\alpha\beta}, U^{(1)}_{\alpha\beta},\ldots, U^{(2K-1)}_{\alpha\beta}]^{\sim}$ (where $\sim$ indicates transpose) and obtain the two vector equations

$$U'_{1\beta} = \lambda T U_{1\beta} + U_{2\beta}, \quad U'_{2\beta} = -DU_{1\beta} + \lambda T U_{2\beta}, \tag{28}$$

where $' \equiv d/dt$, $T = [-\delta_{ij} + \delta_{i,j+1} + \delta_{i0}\delta_{j,2K-1}]$, $D = [\kappa_i\delta_{ij}]$, $i,j = 0,1,2,\ldots,2K-1$. The initial conditions for Eq. (28) are

$$U_{\alpha\beta}(0) = \delta_{\alpha\beta}[q_0(0), q_1(0),\ldots,q_{2K-1}(0)]^{\sim}. \tag{29}$$

Obviously, Eq. (28) is equivalent to the single vector equation

$$U'_\beta = M_1 U_\beta, \quad M_1 = \begin{bmatrix} \lambda T & I \\ -D & \lambda T \end{bmatrix}, \quad U_\beta = [U_{1\beta}, U_{2\beta}]^{\sim}, \tag{30}$$

where I is the $2K \times 2K$ identity matrix. Since M_1 is a $4K \times 4K$ constant coefficient matrix, Eq. (30) has the fundamental solution $\exp[tM_1]$. In order to study the behavior of the moments $\langle u_{\alpha\beta}(t)\rangle$, we must therefore find the characteristic values of M_1. We will consider this problem later in this section and will turn now to the development of equations for the higher-order moments.

For $n = 2$, we define the $2K \times 1$ column vectors $U_{\alpha_1\beta_1\alpha_2\beta_2} = [U^{(0)}_{\alpha_1\beta_1\alpha_2\beta_2},\ldots,U^{(2K-1)}_{\alpha_1\beta_1\alpha_2\beta_2}]^{\sim}$ and use Eqs. (26) and (27) to obtain the four vector equations

$$U'_{1\beta_1 1\beta_2} = \lambda T U_{1\beta_1 1\beta_2} + U_{1\beta_1 2\beta_2} + U_{2\beta_1 1\beta_2}, \tag{31}$$

$$U'_{1\beta_1 2\beta_2} = -DU_{1\beta_1 1\beta_2} + \lambda T U_{1\beta_1 2\beta_2} + U_{2\beta_1 2\beta_2}, \tag{32}$$

$$U'_{2\beta_1 1\beta_2} = -DU_{1\beta_1 1\beta_2} + \lambda T U_{2\beta_1 1\beta_2} + U_{2\beta_1 2\beta_2}, \tag{33}$$

$$U'_{2\beta_1 2\beta_2} = -DU_{1\beta_1 2\beta_2} - DU_{2\beta_1 1\beta_2} + \lambda T U_{2\beta_1 2\beta_2}, \tag{34}$$

with initial conditions

$$U_{\alpha_1\beta_2\alpha_2\beta_2}(0) = \delta_{\alpha_1\beta_1}\delta_{\alpha_2\beta_2}[q_0(0), q_1(0),\ldots,q_{2K-1}(0)]^{\sim}. \tag{35}$$

As we have previously remarked, Eq. (1) is equivalent to the two initial value problems $u''_\beta + k_0^2(1 + \epsilon y(t))u_\beta = 0$, $u_\beta(0) = \delta_{1\beta}$, $u'_\beta(0) = \delta_{2\beta}$, where $u_\beta(t) = u_{1\beta}(t)$, $u'_\beta(t) = u_{2\beta}(t)$, $\beta = 1,2$. If we wish to calculate the moments $\langle u_{\beta_1}(t)\cdots u_{\beta_n}(t)\rangle$ and $\langle u'_{\beta_1}(t)\cdots u'_{\beta_n}(t)\rangle$, we have to find only $U^{(k)}_{1\beta_1 1\beta_2\cdots 1\beta_n}$ and $U^{(k)}_{2\beta_1 2\beta_2\cdots 2\beta_n}$. We shall restrict our considerations to this problem. Then, since only the sum $U_{1\beta_1 2\beta_2} + U_{2\beta_1 1\beta_2}$ occurs in Eqs. (31) and (34) and since the coefficients of $U_{1\beta_1 2\beta_2}$ and $U_{2\beta_1 1\beta_2}$ are the same on the right-hand sides of Eqs. (32) and (33), we may sum Eqs. (32) and (33) and obtain the system

$$V'_{1\gamma_2} = \lambda T V_{1\gamma_2} + V_{2\gamma_2},$$
$$V'_{2\gamma_2} = -2DV_{1\gamma_2} + \lambda T V_{2\gamma_2} + 2V_{3\gamma_2}, \tag{36}$$
$$V'_{3\gamma_2} = -DV_{2\gamma_2} + \lambda T V_{3\gamma_2},$$

where $\gamma_2 = (\beta_1,\beta_2)$, $V_{1\gamma_2} = U_{1\beta_1 1\beta_2}$, $V_{2\gamma_2} = U_{1\beta_1 2\beta_2} + U_{2\beta_1 1\beta_2}$, $V_{3\gamma_2} = U_{2\beta_1 2\beta_2}$. Equation (36) is equivalent to the single $6K \times 1$ vector equation

$$V'_{\gamma_2} = M_2 V_{\gamma_2}, \quad M_2 = \begin{bmatrix} \lambda T & I & 0 \\ -2D & \lambda T & 2I \\ 0 & -D & \lambda T \end{bmatrix},$$
$$V_{\gamma_2} = [V_{\gamma_2}, V_{2\gamma_2}, V_{3\gamma_2}]^{\sim}, \tag{37}$$

which has the fundamental solution $\exp[tM_2]$. Thus, the behavior of the second moments is determined by the characteristic values of M_2.

The procedures used in obtaining Eq. (37) may be generalized quite readily to give expressions for all higher-order moments. We define the $2K \times 1$ column vectors $U_{\alpha_1\beta_1\cdots\alpha_n\beta_n} = [U^{(0)}_{\alpha_1\beta_1\cdots\alpha_n\beta_n},\ldots,U^{(2K-1)}_{\alpha_1\beta_1\cdots\alpha_n\beta_n}]^{\sim}$, the n-tuple $\gamma_n = (\beta_1,\beta_2,\ldots,\beta_n)$ and the $2K \times 1$ column vectors $V_{m\gamma_n} = \sum'_{\alpha_1\cdots\alpha_n} U_{\alpha_1\beta_1\cdots\alpha_n\beta_n}$, $m = 1,2,\ldots,n+1$, where $\sum'$ denotes a sum over all $\alpha_i = 1,2$ such that $\sum_{i=1}^{n} \alpha_i = n + m - 1$. Then, if $V_{\gamma_n} \equiv [V_{1\gamma_n}, V_{2\gamma_n},\ldots,V_{n+1,\gamma_n}]^{\sim}$, we may show, using Eqs. (26) and (27), that V_{γ_n} satisfies

$$V'_{\gamma_n} = M_n V_{\gamma_n}, \tag{38}$$

where M_n is an $(n+1) \times (n+1)$ array of $2K \times 2K$ matrices

$$M_n = [M_{ij}], \quad i,j = 1,2,\ldots,n+1, \tag{39}$$

and

$$M_{ij} = \lambda T\delta_{ij} + iI\delta_{i+1,j} - (n+1-j)D\delta_{i,j+1}. \quad (40)$$

The initial conditions for Eq. (38) are

$$V_{\gamma_n}(0) = [V_{1\gamma_n}(0),\ldots,V_{n+1,\gamma_n}(0)]^{\sim},$$
$$V_{m\gamma_n}(0) = \left(\sum_{\alpha_1\cdots\alpha_n}{}' \delta_{\alpha_1\beta_1}\delta_{\alpha_2\beta_2}\cdots\delta_{\alpha_n\beta_n}\right) \times[q_0(0),q_1(0),\ldots,q_{2K-1}(0)]^{\sim}. \quad (41)$$

Again, the fundamental solution of Eq. (38) is $\exp[tM_n]$ and the behavior of the nth moments is obtained from the characteristic values of M_n.

If W_j, $j = 1,2,\ldots,n+1$, is a $2K \times 1$ column vector and $W = [W_1, W_2, \ldots, W_{n+1}]^{\sim}$ is an $(n+1)2K \times 1$ column vector, then, using Eqs. (39) and (40), we see that the characteristic value problem $M_nW = sW$ takes the form

$$W_2 = -(\lambda T - sI)W_1,$$
$$mW_{m+1} = -(\lambda T - sI)W_m + (n-m+2)DW_{m-1}, \quad m = 2,3,\ldots,n, \quad (42)$$
$$(\lambda T - sI)W_{n+1} = DW_n.$$

In the case $n = 1$, Eq. (42) is equivalent to the characteristic value problem

$$[(\lambda T - sI)^2 + D]W_1 = 0, \quad (43)$$

and for the second, third, and fourth moments we obtain

$$[\mathcal{A}_s^3 + 2(\mathcal{A}_sD + D\mathcal{A}_s)]W_1 = 0, \quad n = 2, \quad (44)$$

$$[\mathcal{A}_s^4 + 3(\mathcal{A}_s^2D + D\mathcal{A}_s^2) + 4\mathcal{A}_sD\mathcal{A}_s + 9D^2]W_1 = 0, \quad n = 3, \quad (45)$$

$$[\mathcal{A}_s^5 + 4(\mathcal{A}_s^3D + D\mathcal{A}_s^3) + 6(\mathcal{A}_s^2D\mathcal{A}_s + \mathcal{A}_sD\mathcal{A}_s^2) + 16D\mathcal{A}_sD + 24(\mathcal{A}_sD^2 + D^2\mathcal{A}_s)]W_1 = 0, \quad n = 4, \quad (46)$$

where $\mathcal{A}_s = \lambda T - sI$. For arbitrary n, Eq. (42) yields a characteristic value problem of the form $\mathcal{P}_n(\mathcal{A}_s, D)W_1 = 0$, where $\mathcal{P}_n$ is a polynomial of degree $n+1$, $\mathcal{P}_n = \mathcal{A}_s^{n+1} + n(\mathcal{A}_s^nD + D\mathcal{A}_s^n) + \cdots$, but we have not found the general expression for $\mathcal{P}_n$. We recall that $D = [\kappa_i\delta_{ij}] = [k_0^2\delta_{ij} + \epsilon k_0^2 b_i\delta_{ij}] \equiv k_0^2I + \epsilon B$. Hence, the solutions s, W_1 of Eqs. (43)–(46) will depend on the parameter ϵ. For example, substituting for D in Eqs. (43) and (44), we obtain

$$[(\mathcal{A}_s + ik_0I)(\mathcal{A}_s - ik_0I) + \epsilon B]W_1 = 0, \quad (47)$$

$$[\mathcal{A}_s(\mathcal{A}_s + 2ik_0I)(\mathcal{A}_s - 2ik_0I) + 2\epsilon(\mathcal{A}_sB + B\mathcal{A}_s)]W_1 = 0. \quad (48)$$

For $\epsilon = 0$ the solutions of these equations [and of Eqs. (45) and (46)] are easy to obtain. In fact, as we show in Appendix C, the matrix T has $2K$ distinct eigenvalues $\tau_0, \tau_1, \ldots, \tau_{2K-1}$, where $\tau_n = \theta_n - 1$ and θ_n are the $2K$th roots of unity, and a corresponding set of $2K$ orthogonal eigenvectors $X_0, X_1, \ldots X_{2K-1}$. Thus, there are $4K$ distinct values of s which satisfy Eq. (47); we denote them by $s_{0\nu}^{(\alpha)} = (-1)^{\alpha}ik_0 + \lambda\tau_\nu$, $\alpha = 0,1$, $\nu = 0,1,\ldots,2K-1$. Similarly, there are $6K$ distinct solutions of Eq. (48), $s_{0\nu}^{(\beta)} = 2ik_0\beta + \lambda\tau_\nu$, $\beta = -1,0,1$. A completely analogous result holds for Eqs. (45) and (46) when $\epsilon = 0$.

For $\epsilon > 0$ we have used (see Appendix C) a perturbation analysis to obtain the characteristic values of Eqs. (47) and (48). We assume

$$s = \sum_{n=0}^{\infty} s_n\epsilon^n, \qquad W_1 = \sum_{n=0}^{\infty} Y_n\epsilon^n. \quad (49)$$

Then, substituting Eq. (49) into Eq. (47), we find that s_0 takes the $4K$ values given in the previous paragraph, that $s_1 = 0$, and that

$$s_2 = s_{2\mu}^{(\alpha)} = (4K\lambda k_0^{-1})^{-2}\lambda \times \sum_{\nu=0}^{2K-1}{}' \left[(\theta_\nu - \theta_\mu)\left(1 + (-1)^{\alpha}\frac{i\lambda}{2k_0}(\theta_\nu - \theta_\mu)\right)\right]^{-1} \times \csc^4\left(\frac{(\nu-\mu)\pi}{2K}\right), \quad (50)$$

where the prime on the summation sign indicates that the terms for which $\nu - \mu = 0, \pm2, \pm4, \ldots$ are omitted. A similar calculation performed with Eq. (48) shows that in the case of the second moment $s_0 = s_{0\nu}^{(\beta)} = 2ik_0\beta + \lambda\tau_\nu$, $\beta = -1,0,1$, $s_1 = 0$ and that

$$s_2 = s_{2\mu}^{(\beta)} = -ik_0K^{-2}(12|\beta| - 4)^{-1}\sum_{\nu=0}^{2K-1}{}' \times\left[\left((1-|\beta|) - i\frac{\lambda}{2k_0}(\theta_\nu - \theta_\mu)\right)\left((1-|\beta|+\beta) + i\frac{\lambda}{2k_0}(\theta_\nu - \theta_\mu)\right)\right]^{-1}\left(\beta + i\frac{\lambda}{4k_0}(\theta_\nu - \theta_\mu)\right) \times \csc^4\left(\frac{(\nu-\mu)\pi}{2K}\right), \quad (51)$$

where the prime on the summation has the same meaning as in Eq. (50)

The expressions of Eq. (49) are readily justified. Consider, for example, the first moment. In order to find the characteristic values of Eq. (47), we must find the roots of the equation $F_K(\epsilon, s) = \det[(\mathcal{A}_s + ik_0I)(\mathcal{A}_s - ik_0I) + \epsilon B] \equiv \det[\mathfrak{F}]$. Letting Q be the matrix whose columns are the characteristic vectors of T, we find that (since $|\det Q| = 1 \neq 0$) $F_K = \det[Q^{*\sim}\mathfrak{F}Q] = \det(\operatorname{diag}\{[s - (\lambda\tau_\nu - ik_0)][s - (\lambda\tau_\nu + ik_0)]\} + \epsilon Q^{*\sim}BQ)$, where we show in Appendix C that the elements (X_μ, BX_ν) of the matrix $Q^{*\sim}BQ$ are zero if $\mu = \nu$ and $\frac{1}{2}[(-1)^{\mu-\nu} - 1]k_0^2(2K)^{-1}\csc^2[(\mu-\nu)\pi(2K)^{-1}]$ if $\mu \neq \nu$. Obviously $F_K(\epsilon, s)$ is a polynomial in ϵ and s is therefore an analytic function of these variables. From the last expression for F_K we see that

$$F_K(0,s) = \prod_{\substack{\alpha=0,1 \\ \mu=0,1,\ldots,2K-1}} (s - s_{0\mu}^{(\alpha)}).$$

Thus, $F_K(0, s)$ has a simple zero at each of the $4K$ points $s_{0\mu}^{(\alpha)}$. By the implicit function theorem for functions of two complex variables[6] there are circular neighborhoods $N_\mu^{(\alpha)}(s_{0\mu}^{(\alpha)})$, $N_\mu^{(\alpha)}(0)$ of $s = s_{0\mu}^{(\alpha)}$ and $\epsilon = 0$, respectively, such that $F_K(\epsilon, s)$ has a unique root $s_\mu^{(\alpha)}(\epsilon)$ in $N_\mu^{(\alpha)}(s_{0\mu}^{(\alpha)})$ for any given ϵ in $N_\mu^{(\alpha)}(0)$ and such that $s_\mu^{(\alpha)}(\epsilon)$ is single-valued and analytic on $N_\mu^{(\alpha)}(0)$ and satisfies the condition $s_\mu^{(\alpha)}(0) = s_{0\mu}^{(\alpha)}$. Hence, for ϵ in $N(0) = \bigcap_{\alpha,\mu} N_\mu^{(\alpha)}(0)$ all roots of $F_K(\epsilon, s)$ are analytic in ϵ and the expansion of s in Eq. (49) is justified. The components of W_1 corresponding to any particular root of F_K are determined by Eq. (47), and so it is ob-

vious that these components are also analytic for ϵ in a neighborhood of $\epsilon = 0$.

Let us consider the behavior of the first moments for $K = 1,2$ in the limit of small ϵ. In this limit we may take $s_\mu^{(\alpha)}(\epsilon) \cong s_{0\mu}^{(\alpha)} + \epsilon^2 s_{2\mu}^{(\alpha)} = (-1)^\alpha i k_0 + \lambda\tau_\mu + \epsilon^2 s_{2\mu}^{(\alpha)}$. Since $\mathrm{Re}(\tau_\mu) < 0$ for $\mu \neq 0$, all the $s_\mu^{(\alpha)}(\epsilon)$ give an exponential decay as $t \to \infty$ with the possible exception of $s_0^{(\alpha)}(\epsilon)$ (recall that $\tau_0 = 0$). For $K = 1$ we find that

$$s_0^{(\alpha)}(\epsilon) \cong (-1)^\alpha i k_0 - \epsilon^2 (32)^{-1} \lambda (k_0/\lambda)^2 [1 + (\lambda/k_0)^2]^{-1} \times [1 + (-1)^\alpha i (\lambda/k_0)]. \quad (52)$$

In this case $y(t)$ is a random telegraph wave with values $\pm\frac{1}{2}$, and Eq. (52) shows that the solution corresponding to $s_0^{(\alpha)}(\epsilon)$ also decays. This result has been obtained by several authors.[2,7] The case $K = 2$ is somewhat more interesting. In this case we find

$$s_0^{(\alpha)}(\epsilon) \cong (-1)^\alpha i k_0 - \epsilon^2 (16)^{-1} \lambda (k_0/\lambda)^2 \times [1 + 4^{-1}(\lambda/k_0)^4]^{-1} \{[1 - \tfrac{1}{2}(\lambda/k_0)^2] + (-1)^\alpha i (\lambda/k_0)\}. \quad (53)$$

Equation (53) shows that the solutions corresponding to $s_0^{(\alpha)}(\epsilon)$ decay as $t \to \infty$ if $(\lambda/k_0) < \sqrt{2}$ and increase exponentially as $t \to \infty$ if $(\lambda/k_0) > \sqrt{2}$.

It is of some interest to compare the solutions of Eqs. (52) and (53) with those obtained by applying perturbation methods to the governing equation. For $K = 1$ it is well known[2,7] that the first-order smoothing approximation[5] predicts the characteristic values of s exactly. For $K = 2$ this method gives a characteristic equation for s which is of the sixth order and thus obviously cannot be exact. However, in the limit $\epsilon \to 0$ two of the roots of this equation correspond to those given by Eq. (53), while the remaining roots give an exponential decay as $t \to \infty$. The two-time expansion of Papanicolau and Keller,[8] when applied to the cases $K = 1,2$, also predicts the values of s given by Eqs. (52) and (53).

Finally, let us consider briefly the behavior of the second moments for the case $K = 2$. In the limit of small ϵ we may take $s_\mu^{(\beta)}(\epsilon) \cong s_{0\mu}^{(\beta)} + \epsilon^2 s_{2\mu}^{(\beta)} = 2ik_0\beta + \lambda\tau_\mu + \epsilon^2 s_{2\mu}^{(\beta)}$, where $\beta = -1, 0, 1$, $\mu = 0, 1, \ldots, 2K-1$ and $s_{2\mu}^{(\beta)}$ is given by Eq. (51). Again, since $\mathrm{Re}(\tau_\mu) < 0$, $\mu \neq 0$, all the $s_\mu^{(\beta)}(\epsilon)$ give an exponential decay as $t \to \infty$ with the possible exception of $s_0^{(\beta)}(\epsilon)$. For $K = 2$, we find from Eq. (51) that

$$s_{20}^{(0)} = 8^{-1}\lambda(k_0/\lambda)[1 + 4^{-1}(\lambda/k_0)^4]^{-1}[1 + 2^{-1}(\lambda/k_0)^2], \quad (54)$$

$$s_{20}^{(1)} = s_{20}^{(-1)*} = -4^{-1}\lambda(k_0/\lambda)^2[1 + 4^{-1}(\lambda/k_0)^4]^{-1} \times \{[1 - 4^{-1}(\lambda/k_0)^2 + 8^{-1}(\lambda/k_0)^4] + 2^{-1}i(\lambda/k_0)\}. \quad (55)$$

Equation (54) indicates that the second moments will increase exponentially as $t \to \infty$. As in the case of the first moments, the two-time expansion of Papanicolau and Keller predicts the values for s given by Eqs. (54) and (55).

We now turn to a description of a class of stochastic processes of type II. We take $a_j = j$, $j \in Z$, $L = N = 1$, $\lambda_{i,i+1} = \lambda(-1) = \lambda_{i-1,i} = \lambda(1) = \lambda$, $\lambda_{ii} = -2\lambda$, where λ is a positive constant. We show in Appendix D that, for this process, $p_{ij}(\tau) \equiv p_{ij}(t, t+\tau) = \exp(-2\lambda\tau) I_{j-i}(2\lambda\tau)$, where I_n is the modified Bessel function[9] of order n. We assume that $p_j^0 = \delta_{j0}$. For construction of the process $y(t)$, we take $M = 2K$, where $K = 2K'$ and $K' = 1, 2, 3, \ldots$, $f(j) = j$, $0 \leq j \leq K'$, $f(j) = 2K' - j$, $K' + 1 \leq j \leq 2K'$, $f(-j) = -f(j)$, and $f(j + 4K') = f(j)$. The elements of the set $\{b_0, b_1, \ldots, b_{4K'-1}\}$ of possible values of $y(t)$ are $b_0 = b_{2K'} = 0$ and $b_j = b_{2K'-j} = j$, $b_{2K'+j} = b_{4K'-j} = j$, $j = 1, 2, \ldots, K'$.

If we use Eq. (23) to calculate S_ν for the type II process described above and substitute the result into Eq. (19), we find that

$$\frac{\partial G_0}{\partial t} + \mathcal{L}_0 G_0 - \lambda(G_1 - 2G_0 + G_{2K-1}) = 0,$$

$$\frac{\partial G_\nu}{\partial t} + \mathcal{L}_\nu G_\nu - \lambda(G_{\nu-1} - 2G_\nu + G_{\nu+1}) = 0, \quad \nu = 1, 2, \ldots, 2K-2, \quad (56)$$

$$\frac{\partial G_{2K-1}}{\partial t} + \mathcal{L}_{2K-1} G_{2K-1} - \lambda(G_{2K-2} - 2G_{2K-1} + G_0) = 0.$$

It is clear from the similarity of Eq. (56) to Eq. (26) that the various order moments of $u_{\alpha\beta}(t)$ will satisfy equations of exactly the same form as we derived in the case of the type I process. In fact, the only difference will be that of $2K \times 2K$ matrix T will be replaced everywhere by the $2K \times 2K = 4K' \times 4K'$ matrix $\bar{T}$ defined by

$$\bar{T} = (-2\delta_{ij} + \delta_{ij-1} + \delta_{i,j+1} + \delta_{i0}\delta_{j,4K'-1} + \delta_{i,4K'-1}\delta_{j0}), \quad i,j = 0, 1, \ldots, 4K'-1. \quad (57)$$

Thus, the nth moments will be determined by solving Eq. (38) with M_n an $(n+1) \times (n+1)$ array of $2K \times 2K = 4K' \times 4K$ matrices, $M_n = [M_{ij}]$, $i,j = 1, 2, \ldots, n+1$, and $M_{ij} = \lambda\bar{T}\delta_{ij} + iI\delta_{i+1,j} - (n+1-j)D\delta_{ij+1}$. The time behavior of the nth moments will again be determined by the characteristic values of M_n. The procedure used for solving this characteristic value problem in the case of the type I process may be carried through exactly as before and, in the case of the first and second moments, leads to Eqs. (47) and (48) with $\mathcal{A}_s = \lambda\bar{T} - sI$.

The solutions of Eqs. (47) and (48) for $\epsilon = 0$ may be obtained by solving the characteristic problem for $\bar{T}$. Since $\bar{T}$ is a real symmetric matrix, all its eigenvalues, $\bar{\tau}_0, \bar{\tau}_1, \ldots, \bar{\tau}_{4K'-1}$, are real. However, they are no longer distinct as in the case of T. We show in Appendix D that

$$\bar{\tau}_\nu = -2[1 - \cos(\nu\pi/2K')], \quad \nu = 0, 1, \ldots, 4K'-1. \quad (58)$$

Hence, there are two simple roots, $\bar{\tau}_0$ and $\bar{\tau}_{2K'}$, and $2K'-1$ double roots. The system of $4K'$ orthonormal characteristic vectors $\bar{X}_\nu$ corresponding to these characteristic values is given by Eq. (D15) of Appendix D.

For $0 < \epsilon \ll 1$ we determine the characteristic values of Eqs. (47) and (48) by assuming the expansions of Eq. (49) for s and W_1 and using the perturbation analysis given in Appendix C. The details of the analysis are slightly different in this case due to the fact that $\bar{T}$ has repeated characteristic roots. For example, when we calculate $s_{2\mu}^{(\alpha)}$ for the first moment,

we find that because $\bar{\tau}_\mu = \bar{\tau}_{4K'-\mu}$, both $c^{(\alpha)}_{\mu,4K'-\mu}$ and $c^{(\alpha)}_{\mu,4K'-\mu}$ of Eq. (C14) are undetermined. However, $s_{2\mu}$ may still be found from Eq. (C17) since both $(\bar{X}_\mu, B\bar{X}_\mu)$ and $(\bar{X}_\mu, B\bar{X}_{4K'-\mu})$ vanish. The justification of the perturbation expansions can be carried through as before with slight modifications due to the existence of repeated roots.

For the first moment we find that

$$s_0 = s^{(\alpha)}_{0\mu} = (-1)^\alpha ik_0 + \lambda\bar{\tau}_\mu, \quad \alpha = 0,1, \quad \mu = 0,1,\ldots,4K'-1 \tag{59}$$

$$s_1 = 0 \tag{60}$$

From Eq. (C17) and the fact that $\bar{\tau}_\nu - \bar{\tau}_\mu = 2[\cos(\nu\pi/2K') - \cos(\mu\pi/2K')] \equiv 2(c_\nu - c_\mu)$, we find that

$$s_2 = s^{(\alpha)}_{2\mu} = (8\lambda k_0^2)^{-1} \sum_{\nu}{}'_{0}^{4K'-1} \{(c_\nu - c_\mu)[1 + (\lambda/k_0)^2(c_\nu - c_\mu)^2\}^{-1} \times [1 - (-1)^\alpha i(\lambda/k_0)(c_\nu - c_\mu)](\bar{X}_\mu, B\bar{X}_\nu)^2, \tag{61}$$

where the prime on the summation indicates that terms for which $\mu + \nu = 0, \pm 2, \pm 4, \ldots$ are omitted and where the scalar products $(\bar{X}_\mu, B\bar{X}_\nu)$ are given in Appendix D, Eq. (D16). In the limit of small ϵ, Eqs. (59)–(61) show that the characteristic values $s^{(\alpha)}_\mu(\epsilon)$ of Eq. (47) are given by $s^{(\alpha)}_\mu(\epsilon) \cong (-1)^\alpha ik_0 + \lambda\bar{\tau}_\mu + \epsilon^2 s^{(\alpha)}_{2\mu}$. As in the case of the type I process, $s^{(\alpha)}_\mu(\epsilon)$ for $\mu \neq 0$ lead to exponential decay as $t \to \infty$. For the type II process, however, Eq. (61) shows that $s^{(\alpha)}_{20} < 0$. Thus, $s^{(\alpha)}_0(\epsilon)$ also leads to exponential decay as $t \to \infty$, and there is no possibility in this case that the first moments will increase with time.

Turning to the second moments, we find that

$$s_0 = s^{(\beta)}_{0\nu} = 2ik_0\beta + \lambda\bar{\tau}_\nu, \quad \beta = -1,0,1, \quad \nu = 0,1,\ldots,4K'-1, \tag{62}$$

$$s_1 = 0 \tag{63}$$

and from Eq. (3.25), we obtain

$$s_2 = s^{(\beta)}_{2\nu} = [k_0^3(6|\beta| - 2)]^{-1} \sum_{\mu=0}{}'^{4K'-1} \{[1 + (\lambda/k_0)^2(c_\mu - c_\nu)^2] \times [(1 - |\beta|) + (\lambda/k_0)^2(c_\mu - c_\nu)^2]\}^{-1} \times \{(\lambda/k_0)(c_\mu - c_\nu)[(\lambda/k_0)^2(c_\mu - c_\nu)^2 + (1 + |\beta|)] - i\beta[2(1 - |\beta|) + (\lambda/k_0)^2(c_\mu - c_\nu)^2]\}(\bar{X}_\mu, B\bar{X}_\nu)^2,$$

where the prime on the summation has the same meaning as in Eq. (61).

If we approximate the characteristic values of Eq. (48) by $s^{(\beta)}_\nu(\epsilon) \cong 2ik_0\beta + \lambda\bar{\tau}_\nu + \epsilon^2 s^{(\beta)}_{2\nu}$, we see that $\mathrm{Re}[s^{(\beta)}_\nu(\epsilon)] < 0, \nu = 0,1,\ldots,4K'-1$, and the second moments decay exponentially as $t \to \infty$. This behavior should be compared with the results for the type I process (with $K = 2$) given in Eq. (54) where an exponential increase was found.

IV. SUMMARY

We have determined the statistical moments of the solutions to the second order ordinary differential equations $u'' + k_0^2[1 + \epsilon y(t)]u = 0$, where $y(t)$ is one of two particular types (I or II) of stochastic process. The two types of process which we have considered in detail have the common properties: (1) $y(t)$ is derived from a countable state space Markov process; (2) $y(t)$ is bounded and can take only a finite number of values. For the type I process, $y(t)$ oscillates between its maximum and minimum values, and in the interval between two adjacent extreme values $y(t)$ is monotonic. For the type II process, $y(t)$ again oscillates but its successive maxima and minima decrease on the average.

Corresponding to the different general appearance of the type I and II processes as discussed above, we have found rather striking differences in the behavior of the first two statistical moments of the solution process of our differential equation. For the type I process we have shown that if the ratio λ/k_0 is large enough, the first moments may increase exponentially as $t \to \infty$, while for the type II process the first moments always decrease for any value of λ/k_0. If we assume t is a length coordinate, then λ/k_0 is essentially the ratio of the wavelength of the propagating wave $u(t)$ to the correlation length of the random process $y(t)$, and we can interpret these phenomena as follows: For the type I process, phase cancellations always occur if λ/k_0 is small and the average wave always decays, but phase additions may occur for λ/k_0 large and the average wave may increase. For the type II process, fluctuations of $y(t)$ away from zero (long wavelength phenomena) always cause phase cancellations and result in decay of the average wave. The three-level ($K = 2$) type I process gave second moments which increased exponentially as $t \to \infty$ while the second moments for the type II process always decreased. For the type I process we see, therefore, that the destructive interference which caused decay of the average wave (for λ/k_0 small) is always erased when the wave is multiplied by itself before averaging. In the case of the type II process, fluctuations in y continue to cause the second moments to decay.

It is clear that the two problems we have treated in detail are directly applicable to the study of the propagation of time harmonic electromagnetic waves in one-dimensional layered media. We hope also that the exact solutions we have obtained may be useful in testing various perturbation methods for stochastic equations.

APPENDIX A

In order to derive Eq. (22), we note that for the case under consideration $J_\nu = Z^+_\nu = \{\nu + kM : k = 0,1,2,\ldots\}$ and therefore

$$G_\nu = \sum_{\mu \in Z^+_\nu} F_\mu = \sum_{l=0} F_{\nu + lM}.$$

If $\sigma_\mu \equiv \sum_{k=0}^{\infty} F_k \lambda_{k\mu}$, then since $\lambda_{k\mu} = \lambda(k - \mu)$ is nonzero only in the range $-L \leq k - \mu \leq 0$, we find that

$$\sigma_\mu = \sum_{m=-\mu}^{0} \lambda(m) F_{\mu+m}, \quad 0 \leq \mu \leq L-1,$$

$$\sigma_\mu = \sum_{m=-L}^{0} \lambda(m) F_{\mu+m}, \quad \mu \geq L.$$

S_ν may be expressed in the form

$$\sum_{\mu \in Z^+_\nu} \sigma_\mu = \sum_{l=0}^{\infty} \sigma_{\nu + lM}, \quad \nu = 0,1,\ldots,M-1.$$

If we assume $M > L$, we obtain, for $\nu = 0, 1, \ldots, L-1$,

$$\begin{aligned} S_\nu &= \sum_{m=-\nu}^{0} \lambda(m) F_{\nu+m} + \sum_{l=1}^{\infty} \sum_{m=-L}^{0} \lambda(m) F_{\nu+lM+m} \\ &= \sum_{m=-\nu}^{0} \lambda(m) \sum_{l=0}^{\infty} F_{\nu+lM+m} + \sum_{m=-L}^{-\nu-1} \sum_{l=1}^{\infty} \lambda(m) F_{\nu+lM+m} \\ &= \sum_{m=-\nu}^{0} \lambda(m) G_{m+\nu} + \sum_{m=-L}^{-\nu-1} \lambda(m) G_{\nu+m+M} \end{aligned} \tag{A1}$$

and, for $\nu = L, L+1, \ldots$,

$$S_\nu = \sum_{l=0}^{\infty} \sum_{m=-L}^{0} \lambda(m) F_{\nu+lM+m} = \sum_{m=-L}^{0} \lambda(m) G_{\nu+m}. \tag{A2}$$

Equations (A1) and (A2) give Eq. (22) after suitable changes of summation indices.

Equation (23) is derived by similar elementary manipulations. In this case $J_\nu = Z_\nu = \{\nu + kM : k = 0, \pm 1, \pm 2, \ldots\}$ and $G_\nu = \sum_{\mu \in Z_\nu} F_\mu = \sum_{l=-\infty}^{\infty} F_{\nu+lM}$. The quantity σ_μ introduced in the preceding paragraph takes the form

$$\sigma_\mu = \sum_{k=-\infty}^{\infty} F_k \lambda_{k\mu} = \sum_{m=-L}^{N} \lambda(m) F_{\mu+m}, \quad \mu = 0, \pm 1, \pm 2, \ldots,$$

since in this case $\lambda_{k\mu} = \lambda(k-\mu)$ is zero outside the range $-L \leq k - \mu \leq N$. Then, for $\nu = 0, 1, \ldots, M-1$,

$$S_\nu = \sum_{l=-\infty}^{\infty} \sigma_{\nu+lM} = \sum_{l=-\infty}^{\infty} \sum_{m=-L}^{N} \lambda(m) F_{\nu+lM+m}. \tag{A3}$$

We assume that $M > L + N$ and we find, for example, that

$$\begin{aligned} S_0 &= \sum_{l=-\infty}^{\infty} \left(\sum_{m=1}^{L} \lambda(-m) F_{-m+lM} + \sum_{m=0}^{N} \lambda(m) F_{m+lM} \right), \\ S_0 &= \sum_{m=1}^{L} \lambda(-m) G_{M-m} + \sum_{m=0}^{N} \lambda(m) G_m. \end{aligned} \tag{A4}$$

Similar calculations give the remaining expressions in Eq. (23).

APPENDIX B

For a Poisson process $x(t)$ starting at $t_0 \leq 0$, the probability that $x(t) = j$, $j \in Z^+$, is given by[10]

$$p_j(t) = p_{0j}(t_0, t_0 + \tau) = (j!)^{-1} (\lambda\tau)^j e^{-\lambda\tau}, \quad \tau = t - t_0. \tag{B1}$$

Thus, for the process $y(t)$ considered in Sec. III, the probability $q_i(t)$ that $y(t) = b_i$, $i = 0, 1, \ldots, 2K-1$, is given by

$$q_i(t) = \sum_{\nu \in Z_i^+} p_\nu(t) = \sum_{\nu=0}^{\infty} p_{i+2\nu K}(t) = \sum_{\nu=0}^{\infty} [(i+2\nu K)!]^{-1} \times (\lambda\tau)^{i+2\nu K} e^{-\lambda\tau}. \tag{B2}$$

In order to sum the series appearing in (B2), we introduce the notation $\theta_n = \exp[in\pi/K]$, $n = 0, 1, \ldots, 2K-1$. Then, since θ_n are the $2K$th roots of unity, it is easy to show that

$$(2K)^{-1} \sum_{n=0}^{2K-1} \theta_n^m = \begin{cases} 1, & m = 2\nu K \\ 0, & m \neq 2\nu K \end{cases}. \tag{B3}$$

Now we define

$$\begin{aligned} H_0(x) &= (2K)^{-1} \sum_{n=0}^{2K-1} \exp(\theta_n x) = (2K)^{-1} \sum_{n=0}^{2K-1} \sum_{m=0}^{\infty} \frac{1}{m!} \theta_n^m x^m \\ &= (2K)^{-1} \sum_{m=0}^{\infty} \frac{1}{m!} \left(\sum_{n=0}^{2K-1} \theta_n^m \right) x^m. \end{aligned} \tag{B4}$$

Because of the identities in Eq. (B3),

$$H_0(x) = \sum_{\nu=0}^{\infty} ((2\nu K)!)^{-1} x^{2\nu K}. \tag{B5}$$

Comparing Eqs. (B2) and (B5), we see that $q_0(t) = e^{-\lambda\tau} H_0(\lambda\tau)$.

If we define

$$H_i(x) = \int_0^x H_{i-1}(x')\,dx', \quad i = 1, 2, \ldots, 2K-1, \tag{B6}$$

we find from Eqs. (B2), (B4), and (B5) that

$$\begin{aligned} H_i(x) &= (2K)^{-1} \sum_{n=0}^{2K-1} \theta_n^{-i} \exp(\theta_n x) \\ &= \sum_{\nu=0}^{\infty} [(i+2\nu k)!]^{-1} x^{i+2\nu K}, \end{aligned} \tag{B7}$$

$$q_i(t) = e^{-\lambda\tau} H_i(\lambda\tau). \tag{B8}$$

The expectation value of $y(t)$ may be found from Eq. (B8). By definition,

$$\begin{aligned} \langle y(t) \rangle &= \sum_{i=0}^{2K-1} b_i q_i(t) \\ &= e^{-\lambda\tau} \left\{ \sum_{i=1}^{K-1} i \left[H_i(\lambda\tau) + H_{2K-i}(\lambda\tau) \right] + K H_K(\lambda\tau) \right\}, \end{aligned} \tag{B9}$$

where we use the fact that $b_i = b_{2K-1} = i - K/2$, $i = 0, 1, \ldots, K$. If we use the identity $\sum_{n=1}^{N-1} n x^n = [x(1-x^N) - Nx^N(1-x)](1-x)^2$, $x \neq 1$, to obtain the formula

$$\sum_{m=1}^{K-1} m(\theta_\nu^m + \theta_\nu^{-m}) = \begin{cases} K(K-1), & \nu = 0 \\ -K, & \nu = 2, 4, \ldots, 2K-2, \\ K - \csc^2(\nu\pi/2K), & \nu = 1, 3, \ldots, 2K-1 \end{cases} \tag{B10}$$

we may calculate the sums that appear in Eq. (B9). After some manipulation, we obtain

$$\langle y(t) \rangle = -(2K)^{-1} e^{-\lambda\tau} \sum_{\nu=1}^{K} \csc^2[(2\nu-1)\pi/2K] \times \exp(\theta_{2\nu-1}\lambda\tau). \tag{B11}$$

The joint probabilities $q_{ij}(t_1, t_2) = P\{y(t_1) = b_i, y(t_2) = b_j\}$, $t_2 > t_1$, may be found from

$$\begin{aligned} q_{ij}(t_1, t_2) &= \sum_{\nu \in Z_j^+} \sum_{\mu \in Z_j^+} p_\nu(t_1) p_{\nu\mu}(t_1, t_2) \\ &= \sum_{\nu \in Z_i^+} \sum_{\mu \in Z_j^+} p_\nu(t_1) p_{\mu-\nu}(t_2 - t_1 + t_0). \end{aligned} \tag{B12}$$

Changing summation indices from ν, μ to $\nu, \sigma = \mu - \nu$, we may write Eq. (B11) as

$$q_{ij}(t_1, t_2) = q_i(t_1) \sum_{\sigma \in Z_{j-i}} p_\sigma(t_2 - t_1 + t_0), \tag{B13}$$

where $Z_{j-i} = \{j - i + 2\sigma K : \sigma \in Z\}$. Since $p_\sigma = 0$ for $\sigma < 0$, Eq. (B13) gives

$$q_{ij}(t_1,t_2)=\begin{cases} q_i(t_1)\sum_{\sigma=0}^{\infty} p_{j-i+2\sigma K}(t_2-t_1+t_0)=q_i(t_1)q_{j-i}(t_2-t_1-t_0), & j\geq i \\ q_i(t_1)\sum_{\sigma=0}^{\infty} p_{j-i+2K+2\sigma K}(t_2-t_1+t_0)=q_i(t_1)q_{j-i+2K}(t_2-t_1+t_0), & j<i \end{cases} \tag{B14}$$

Since $y(t)$ can take only $K+1$ distinct values $i-K/2$, $i=0,1,\ldots,K$, it is convenient for most purposes to use the probabilities $Q_{ij}(t_1,t_2)\equiv P\{y(t_1)=i-K/2, y(t_2)=j-K/2\}$. These may be obtained in a straightforward manner from Eq. (B14). We have, for example,

$$\begin{aligned} Q_{00}(t_1,t_2)&=q_0(t_1)q_0(t_2-t_1+t_0),\\ Q_{KK}(t_1,t_2)&=q_K(t_1)q_0(t_2-t_1+t_0),\\ Q_{iK}(t_1,t_2)&=q_i(t_1)q_{K-i}(t_2-t_1+t_0)\\ &\quad+q_{2K-i}(t_1)q_{K+i}(t_2-t_1+t_0)\\ Q_{0j}(t_1,t_2)&=q_0(t_1)[q_j(t_2-t_1+t_0)\\ &\quad+q_{2K-j}(t_2-t_1+t_0),\\ Q_{0K}(t_1,t_2)&=q_0(t_1)q_K(t_2-t_1+t_0),\\ Q_{ij}(t_1,t_2)&=q_i(t_1)[q_{j-i}(t_2-t_1+t_0)\\ &\quad+q_{2K-i-j}(t_2-t_1+t_0)]\\ &\quad+q_{2K-i}(t_1)[q_{i+j}(t_2-t_1+t_0)\\ &\quad+q_{2K-j+i}(t_2-t_1+t_0)], \end{aligned} \tag{B15}$$

where $1\leq i\leq j\leq K-1$. The probabilities in Eq. (B14) may be used to find the second moment $\langle y(t_1)y(t_2)\rangle$. By definition this quantity is

$$\begin{aligned} \langle y(t_1)y(t_2)\rangle &=\sum_{i,j=0}^{2K-1} b_i b_j q_{ij}(t_1,t_2)\\ &=\sum_{i,j=0}^{2K-1} ijQ_{ij}(t_1,t_2)-(K/2)^2. \end{aligned} \tag{B16}$$

It is instructive to obtain an explicit expression for the second moment in the limit $t_0\to-\infty$. In this case, it is clear from Eq. (B8) that $q_i(t)\to(2K)^{-1}$, $i=0,1,\ldots,2K-1$ and that $Q_{ij}(t_1,t_2)$ depends only on $s\equiv t_2-t_1$. From Eq. (B15), we obtain

$$\begin{aligned} R(s)&\equiv\langle y(t_1)y(t_1+s)\rangle\\ &=K^2Q_{KK}+\sum_{n=1}^{K-1}n^2Q_{nn}+2K\sum_{n=1}^{K-1}nQ_{nK}\\ &\quad+2\sum_{n=1}^{K-2}\sum_{m=n+1}^{K-1}nmQ_{nm}-(K/2)^2. \end{aligned} \tag{B17}$$

Simplification of the sums appearing in Eq. (B16) is a lengthy but straightforward process. Using Eq. (B10) we find that

$$R(s)=(2K)^{-2}\sum_{\nu=1}^{K}\exp[-(1-\theta_{2\nu-1})\lambda s]\csc^4\left((2\nu-1)\frac{\pi}{2K}\right) \tag{B18}$$

For $K=1,2$, $R(s)$ is given by the formulas

$$\begin{aligned} R(s)&=4^{-1}e^{-2\lambda s}, \quad K=1\\ R(s)&=2^{-1}e^{-\lambda s}\cos(\lambda s), \quad K=2. \end{aligned} \tag{B19}$$

APPENDIX C

If $X=[x_0,x_1,\ldots,x_{2K-1}]^{\sim}$, the characteristic value problem $TX=\tau X$ for the matrix $T=(-\delta_{ij}+\delta_{i,j+1}+\delta_{i0}\delta_{j,2K-1})$ leads to the equations

$$\begin{aligned} -x_0+x_{2K-1}&=\tau x_0,\\ x_{m-1}-x_m&=\tau x_m, \quad m=1,2,\ldots,2K-1 \end{aligned} \tag{C1}$$

and the characteristic equation

$$\det(T-\tau I)=(1+\tau)^{2K}-1=0. \tag{C2}$$

Equation (C2) has the solutions

$$\tau_\nu=\theta_\nu-1, \quad \theta_\nu=\exp(i\pi\nu/K), \quad \nu=0,1,\ldots,2K-1, \tag{C3}$$

and, therefore, $\tau_0=0$ and τ_ν has a negative real part, $\nu=1,\ldots,2K-1$. With $\tau=\tau_\nu$ in Eq. (C1) we find that

$$x_{m\nu}=(2K)^{-1/2}\theta_\nu^{-m}, \quad m=0,1,\ldots,2K-1, \tag{C4}$$

and we have normalized $X_\nu=[x_{0\nu},x_{1\nu},\ldots,x_{2K-1,\nu}]^{\sim}$ so that

$$(X_\nu,X_\nu)\equiv X_\nu^{*\sim}X_\nu=1, \tag{C5}$$

where * denotes complex conjugate. It follows from the relations

$$\sum_{m=0}^{2K-1}\theta_\nu^m=2K\delta_{\nu0}, \quad \theta_\nu^*=\theta_{-\nu}=\theta_\nu^{-1}$$

that

$$(X_\nu,X_\mu)=\delta_{\nu\mu}. \tag{C6}$$

Let us consider the problem $M_1W=sW$. We assume s and W have the expansion given in Eq. (49) and substitute these into Eq. (47). Equating the coefficients of ϵ^σ equal to zero, we obtain

$$(\mathfrak{A}_{s_0}+ik_0I)(\mathfrak{A}_{s_0}-ik_0I)Y_0=0, \tag{C7}$$

$$\begin{aligned} &(\mathfrak{A}_{s_0}+ik_0I)(\mathfrak{A}_{s_0}-ik_0I)Y_{\sigma+1}\\ &\quad+\sum_{\alpha=1}^{\sigma+1}\left(\sum_{\beta+\gamma=\alpha}s_\beta s_\gamma I-2\lambda s_\alpha T\right)Y_{\sigma+1-\alpha}\\ &\quad+BY_\sigma=0, \quad \sigma=0,1,\ldots, \end{aligned} \tag{C8}$$

where $\mathfrak{A}_{s_0}=\lambda T-s_0I$. From Eq. (C7) we obtain the $4K$ values

$$\begin{aligned} s_0=s_{0\nu}^{(\alpha)}&=(-1)^\alpha ik_0+\lambda\tau_\nu,\\ &\alpha=1,2, \ \nu=0,1,\ldots,2K-1. \end{aligned} \tag{C9}$$

The characteristic vector corresponding to $s_{0\nu}^{(\alpha)}$ is $Y_0=X_\nu$. For $\sigma=0$, Eq. (C8) gives

$$\begin{aligned} (\mathfrak{A}_{s_0}+ik_0I)(\mathfrak{A}_{s_0}-ik_0I)Y_1-2s_1(\lambda T-s_0I)Y_0\\ +BY_0=0. \end{aligned} \tag{C10}$$

We set $s_0=s_{0\mu}^{(\alpha)}$, $Y_0=X_\mu$ in Eq. (C10) and assume that Y_1 has the expansion $Y_1=\sum_{\nu=0}^{2K-1}c_{\mu\nu}^{(\alpha)}X_\nu$ in terms of the complete set of vectors X_ν. When these substitutions are made, Eq. (C10) becomes

$$\sum_{\nu=0}^{2K-1} c_{\mu\nu}^{(\alpha)}\lambda(\tau_\nu - \tau_\mu)[\lambda(\tau_\nu - \tau_\mu) - (-1)^\alpha 2ik_0]X_\nu + (-1)^\alpha 2ik_0 s_1 X_\mu + BX_\mu = 0. \quad \text{(C11)}$$

Now, if we form the scalar product of Eq. (C11) with X_μ and use the fact that $(X_\mu, X_\nu) = \delta_{\mu\nu}$, we obtain

$$s_1 = -(-1)^\alpha (2ik_0)^{-1}(X_\mu, BX_\mu) = 0, \quad \text{(C12)}$$

since

$$(X_\mu, BX_\mu) = k_0^2 \sum_{m=0}^{2K-1} b_m x_{m\mu}^* x_{m\mu} = k_0^2(2K)\sum_{m=0}^{2K-1} b_m = 0. \quad \text{(C13)}$$

Thus, the first-order correction to s is zero. For $\nu \neq \mu$, we find

$$c_{\mu\nu}^{(\alpha)} = -\{\lambda(\tau_\nu - \tau_\mu)[\lambda(\tau_\nu - \tau_\mu) - (-1)^\alpha(2ik_0)]\}^{-1}(X_\nu, BX_\mu). \quad \text{(C14)}$$

$c_{\mu\mu}^{(\alpha)}$ is undetermined by this procedure. We turn now to the calculation of the second-order correction to s. Setting $\sigma = 1$ in Eq. (C8), we obtain

$$(\mathfrak{A}_{s_0} + ik_0 I)(\mathfrak{A}_{s_0} - ik_0 I)Y_2 - 2s_1(\lambda T - s_0 I)Y_1 + [s_1^2 I - 2s_2(\lambda T - s_0 I)]Y_0 + BY_1 = 0. \quad \text{(C15)}$$

In this equation we set $s_0 = s_{0\mu}^{(\alpha)}$, $Y_0 = X_\mu$, $s_1 = 0$, $Y_1 = \sum_{\nu=0}^{2K-1} c_{\mu\nu}^{(\alpha)} X_\nu$, and we assume that Y_2 has the expansion $Y_2 = \sum_{\nu=0}^{2K-1} d_{\mu\nu}^{(\alpha)} X_\nu$. Equation (C15) becomes

$$\sum_{\nu=0}^{2K-1} d_{\mu\nu}^{(\alpha)}\lambda(\tau_\nu - \tau_\mu)[\lambda(\tau_\nu - \tau_\mu) - (-1)^\alpha 2ik_0]X_\nu + (-1)^\alpha 2ik_0 s_2 X_\mu + \sum_{\nu=0}^{2K-1} c_{\mu\nu}^{(\alpha)} BX_\nu = 0. \quad \text{(C16)}$$

Forming the scalar product of Eq. (C16) with X_μ, we obtain

$$s_2 = s_{2\mu}^{(\alpha)} = -(2ik_0)^{-1}\sum_{\nu=0}^{2K-1} c_{\mu\nu}^{(\alpha)}(X_\mu, BX_\nu). \quad \text{(C17)}$$

Using Eq. (C14) and the fact that $(X_\mu, BX_\mu) = 0$, we obtain

$$s_{2\mu}^{(\alpha)} = (2ik_0)^{-1}(-1)^\alpha \sum_{\substack{\nu=0 \\ \nu\neq\mu}}^{2K-1} \{\lambda(\tau_\nu - \tau_\mu)[\lambda(\tau_\nu - \tau_\mu] - (-1)^\alpha 2ik_0]\,|(X_\mu, BX_\nu)|^2\}. \quad \text{(C18)}$$

It is not difficult to show that

$$(X_\mu, BX_\nu) = \tfrac{1}{2}[(-1)^{\mu-\nu} - 1]k_0^2(2K)^{-1}\csc^2[(\mu-\nu)\pi/2K]. \quad \text{(C19)}$$

Thus, the summand in Eq. (C18) is nonzero only for $\nu - \mu$ odd, and since $\tau_\nu - \tau_\mu = \theta_\nu - \theta_\mu$, we can express Eq. (C18) in the form shown in Eq. (50).

Finally, let us consider the problem $M_2 W = sW$. Using Eqs. (48) and (49), we obtain the perturbation equations

$$\mathfrak{D}_0 Y_0 = 0, \quad \text{(C20)}$$

$$\sum_{\alpha+\beta=\sigma} \mathfrak{D}_\alpha Y_\beta + 2\lambda(TB + BT)Y_{\sigma-1} - 4\sum_{\alpha+\beta=\sigma-1} s_\alpha B Y_\beta = 0, \quad \text{(C21)}$$

where

$$\mathfrak{D}_0 = \mathfrak{A}_{s_0}(\mathfrak{A}_{s_0} + 2ik_0 I)(\mathfrak{A}_{s_0} - 2ik_0 I) \quad \text{(C22)}$$

$$\mathfrak{D}_\sigma = -3s_\sigma(\mathfrak{A}_{s_0} + 3^{-1/2}2ik_0 I)(\mathfrak{A}_{s_0} - 3^{-1/2}2ik_0 I] + 3(1-\delta_{\sigma 1})\sum_{\alpha+\beta=\sigma-2} s_{\alpha+1}s_{\beta+1}\mathfrak{A}_{s_0} - (1-\delta_{\sigma 1})(1-\delta_{\sigma 2})\sum_{\mu+\nu=\sigma-3} s_{\mu+1}\sum_{\alpha+\beta=\nu} s_{\alpha+1}s_{\beta+1}, \quad \sigma = 1,2,\ldots. \quad \text{(C23)}$$

From Eq. (C20) we find that $s_0 = s_{0\mu}^{(\beta)} = 2ik_0\beta + \lambda\tau_\mu$, $\beta = -1, 0, 1$ with the corresponding eigenvectors $Y_{0\mu}^{(\beta)} = X_\mu$. If we substitute these values into Eq. (C21) with $\sigma = 1$ and assume $Y_1 = \sum_{\nu=0}^{2K-1} c_{\mu\nu}^{(\beta)} X_\nu$, we find as before that $s_1 = 0$ and that for $\mu \neq \nu$

$$c_{\mu\nu}^{(\beta)} = -2[\lambda(\tau_\nu - \tau_\mu) + 2ik_0(1-|\beta|)]^{-1} \times [\lambda(\tau_\mu - \tau_\nu) - 2ik_0(1-|\beta|+\beta)]^{-1}(X_\nu, BX_\mu). \quad \text{(C24)}$$

To determine the second-order correction to s, we set $\sigma = 2$ in Eq. (C21) and evaluate the resulting equation for $s_0 = s_{0\mu}^{(\beta)}$, $s_1 = 0$, $Y_0 = X_\mu$ and $Y_1 = \sum_{\nu=0}^{2K-1} c_{\mu\nu}^{(\beta)} X_\nu$. Assuming the expansion $Y_2 = \sum_{\nu=0}^{2K-1} d_{\mu\nu}^{(\beta)} X_\nu$, we find

$$\sum_{\nu=0}^{2K-1} d_{\mu\nu}^{(\beta)}\lambda(\tau_\nu - \tau_\mu)[\lambda(\tau_\nu - \tau_\mu) + 2ik_0(1-|\beta|-2\beta)] \times [\lambda(\tau_\nu - \tau_\mu) - 2ik_0(1-|\beta|+\beta)]X_\nu + (12|\beta| - 4)k_0^2 s_2 X_\mu + 2\lambda\sum_{\nu=0}^{2K-1} c_{\mu\nu}^{(\beta)}(TB + BT)X_\nu - 4s_{0\mu}^{(\beta)}\sum_{\nu=0}^{2K-1} c_{\mu\nu}^{(\beta)} X_\nu = 0. \quad \text{(C25)}$$

Taking the scalar product of Eq. (C25) with X_μ and using the fact that $X_\mu^{*\sim} T = \tau_\mu X^{*\sim}$, we obtain

$$s_2 = s_{2\mu}^{(\beta)} = -2(12|\beta| - 4)^{-1}k_0^{-2}\sum_{\nu=0}^{2K-1}[\lambda(\tau_\nu - \tau_\mu) - 4\beta ik_0] \times c_{\mu\nu}^{(\beta)}(X_\mu, BX_\nu). \quad \text{(C26)}$$

If we substitute for $c_{\mu\nu}^{(\beta)}$ from Eq. (C24) and use Eq. (C29), we find that $s_{2\nu}^{(\beta)}$ may be expressed in the form shown in Eq. (51).

APPENDIX D

The forward equation for the $p_{ij}(\tau, t)$ of the type II process may be obtained using the techniques of Feller.[10] We find that for $\mu, \nu = 0, \pm 1, \pm 2, \ldots$,

$$\frac{\partial p_{\mu\nu}}{\partial t} = \lambda(p_{\mu,\nu+1} - 2p_{\mu\nu} + p_{\mu,\nu+1}), \quad p_{\mu\nu}(\tau,\tau) = \delta_{\mu\nu}. \quad \text{(D1)}$$

For fixed τ and μ the solution of this set of ordinary differential equations for $\varphi_\nu(t) \equiv p_{\mu\nu}(\tau, t)$ may be found by introducing the function $Q(s,t) = \sum_{\nu=-\infty}^{\infty} \varphi_\nu(t)\times \exp(is\nu)$, $-\pi \leq s \leq \pi$, $i = \sqrt{-1}$. Summing the equations $\varphi_\nu' \exp(is\nu) = \lambda(\varphi_{\nu+1} - 2\varphi_\nu + \varphi_{\nu-1})\exp(is\nu)$ over all ν, we find that

$$\frac{\partial Q}{\partial t} = 2\lambda(\cos(s) - 1)Q, \quad Q(s,\tau) = e^{i\mu s}. \quad \text{(D2)}$$

The solution of (D2) is $Q(s,t) = \exp[i\mu s + 2\lambda(\cos(s) - 1)(t-\tau)]$, and it is clear from the definition of Q that

$$p_{\mu\nu}(\tau, t) = Q_\nu(t) = (2\pi)^{-1}\int_{-\pi}^{\pi} Q(s,t)^{-i\nu s} ds. \quad \text{(D3)}$$

Substituting for $Q(s,t)$ in Eq. (D3) and introducing the

change of variables $z = e^{is}$, we find

$$p_{\mu\nu}(\tau,t) = \exp[-2\lambda(t-\tau)](2\pi i)^{-1} \times \int_\Gamma \exp[\lambda(t-\tau)(z+z^{-1})]z^{-(\nu-\mu+1)}dz \equiv c\exp[-2\lambda(t-\tau)], \quad \text{(D4)}$$

where Γ is the unit circle. The function c in Eq. (D4) is just the coefficient of $w^{\nu-\mu}$ in the Laurent expansion of

$$\exp[\lambda(t-\tau)(w+w^{-1})] = \sum_{n=-\infty}^{\infty} w^n I_n[2\lambda(t-\tau)],$$

where the I_n are modified Bessel functions of index n.[9] Thus,

$$p_{\mu\nu}(\tau,t) = \exp[-2\lambda(t-\tau)]I_{\nu-\mu}[2\lambda(t-\tau)]. \quad \text{(D5)}$$

The characteristic value problem for the matrix $\bar{T}$ associated with the type II process is somewhat more complicated than that associated with T. If we let $D_n(x)$ be the determinant of the $n \times n$ matrix

$$D_n(x) \equiv \det\begin{pmatrix} x & 1 & 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & x & 1 & 0 & 0 & \dots & & 0 \\ 0 & 1 & x & 1 & 0 & \dots & & 0 \\ \vdots & & & & & & & \\ 0 & 0 & \dots & & & 1 & x & 1 \\ 1 & 0 & \dots & & & 0 & 1 & x \end{pmatrix}, \quad \text{(D6)}$$

then the characteristic equation for $\bar{T}$, $\det(\bar{T} - \bar{\tau}I) = 0$, is just $D_{4K'}[-(2+\bar{\tau})] = 0$. In order to find the roots of this equation we shall first find a simple expression for D_n. Expanding the determinant in Eq. (D6) by the elements of the first row and performing similar manipulations with the resulting factors, we find that

$$D_n(x) = x\Delta_{n-1}(x) - 2\Delta_{n-2}(x) + 2(-1)^n, \quad \text{(D7)}$$

where $\Delta_n(x)$ is the determinant of the $n \times n$ tridiagonal matrix

$$\Delta_n(x) \equiv \det\begin{pmatrix} x & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & x & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & x & 1 & 0 & \dots & 0 & 0 \\ \vdots & & & & & & & \\ 0 & & & & & 1 & x & \end{pmatrix}. \quad \text{(D8)}$$

Expanding $\Delta_n(x)$ by the first row, we find the relation

$$\Delta_n(x) = x\Delta_{n-1}(x) - \Delta_{n-2}(x). \quad \text{(D9)}$$

Equations (D7) and (D8) imply that $D_{n+1} = \Delta_{n+1} - \Delta_{n-1} + 2(-1)^n$, while Eq. (D9) gives the relation $\Delta_{n+1} - \Delta_{n-1} = x(\Delta_n - \Delta_{n-2}) - (\Delta_{n-1} - \Delta_{n-3})$. Combining these two recurrence relations and defining

$$C_n(x) \equiv D_n(x) + 2(-1)^n, \quad \text{(D10)}$$

we find that $C_n(x)$ satisfies the recurrance relation

$$C_{n+1} = xC_n(x) - C_{n-1}(x). \quad \text{(D11)}$$

Using Eq. (D6), we obtain, for example, $C_3(x) = x^3 - 3x$, $C_4(x) = x^4 - 4x^2 + 2$. Thus, the C_n are modified Chebyshev polynomials of degree n,[9] and we have

$$D_n(x) = [C_n(x) - 2(-1)^n] = 2[T_n(x/2) - (-1)^n], \quad \text{(D12)}$$

where T_n is the Chebyshev polynomial of degree n. The characteristic equation for $\bar{T}$ is therefore $T_{4K'}[-\frac{1}{2}\bar{\tau} - 1] - 1 = 0$, and this implies that the characteristic values of $\bar{T}$ are

$$\bar{\tau}_\nu = -2[1 - \cos(\nu\pi/2K')], \quad \nu = 0,1,\dots,4K'-1. \quad \text{(D13)}$$

The characteristic vectors corresponding to this set of characteristic values are determined from the equations $\bar{T}\bar{X}_\nu = \bar{\tau}_\nu\bar{X}_\nu$ or

$$\begin{aligned} &\bar{x}_{1\nu} - 2\bar{x}_{0\nu} + \bar{x}_{4K'-1,\nu} = \bar{\tau}_\nu\bar{x}_{0\nu}, \\ &\bar{x}_{4K'-2} - 2\bar{x}_{4K'-1,\nu} + \bar{x}_{0\nu} = \bar{\tau}_\nu\bar{x}_{4K'-1,\nu}, \\ &\bar{x}_{j-1,\nu} - 2\bar{x}_{j\nu} + \bar{x}_{j+1,\nu} = \bar{\tau}_\nu\bar{x}_{j\nu}, \quad j = 1,2,\dots,4K'-2, \end{aligned} \quad \text{(D14)}$$

and it is easy to see that Eq. (D14) leads to the equations $\bar{x}_{j\nu} = S_{j-1}(z_\nu)\bar{x}_{1\nu} - S_{j-2}(z_\nu)\bar{x}_{0\nu}$, $\bar{x}_{4K'-j,\nu} = S_j(z_\nu)\bar{x}_{0\nu} - S_{j-1}(z_\nu)\bar{x}_{1\nu}$, $j = 1,2,\dots,2K'$, where $z_\nu = 2\cos(\nu\pi/2K')$ and S_j are modified Chebyshev polynomials. These equations lead to the following components for the normalized characteristic vectors

$$\begin{aligned} &\bar{x}_{j\nu} = (2K')^{-1/2}\sin(j\nu\pi/2K'), \quad j = 0,1,\dots,4K'-1, \\ &\qquad\qquad \nu = 1,2,\dots,2K'-1, \\ &\bar{x}_{0\nu} = -(2K')^{-1/2}\sin(\nu\pi/2K'), \\ &x_{j\nu} = (2K')^{-1/2}\sin((j-1)\nu\pi/2K'), \\ &\quad j = 1,2,\dots,4K'-1, \ \nu = 2K'+1,\dots,4K'-1, \\ &\bar{x}_{j0} = (4K')^{-1/2}, \quad x_{j,2K'} = (-1)^j(4K')^{-1/2}, \\ &\quad j = 0,1,\dots,4K'-1, \end{aligned} \quad \text{(D15)}$$

where $(\bar{X}_\mu, \bar{X}_\nu) = \delta_{\mu\nu}$.

In order to use the perturbation methods of Appendix C, we need to compute the scalar products $(\bar{X}_\mu, B\bar{X}_\nu) = k_0^2\sum_{j=0}^{4K'-1} b_j\bar{x}_{j\mu}\bar{x}_{j\nu}$, where the $b_j, j = 0,\dots,4K'-1$ are the possible values of the type II process under consideration. In the case $\mu = 0$, for example, we find using Eq. (D15) that for $\nu = 1,2,\dots,2K'-1$,

$$(\bar{X}_0, B\bar{X}_\nu) = [1-(-1)^\nu]\left[K'\sin\left(\frac{\nu\pi}{2}\right) + 2\sum_{j=1}^{K'-1} j\sin\left(\frac{j\nu\pi}{2K'}\right)\right].$$

Since

$$\sum_{j=1}^{n} j\sin j\theta = -\frac{d}{d\theta}\left(\sum_{j=0}^{n}\cos j\theta\right) = -\frac{d}{d\theta}\left(\sum_{j=0}^{n} T_j(\cos\theta)\right),$$

and since the sum of the first $n+1$ Chebyshev polynomials is $\sum_0^n T_j(x) = \frac{1}{2}[1 + U_{n-1}(x) + U_n(x)]$, we find that $(\bar{X}_0, B\bar{X}_\nu) = [1-(-1)^\nu](4\sqrt{2}K')^{-1}k_0^2\sin(\nu\pi/2)\csc^2(\nu\pi/4K')$. Similar elementary but rather tedious calculations show that $(\bar{X}_\mu, B\bar{X}_\nu) = 0$ for $\mu+\nu = 0,2,4,\dots$ and for $1 \le \nu, \mu \le 2K'-1$ and that for $\mu+\nu = 1,3,5,\dots$,

$$(\bar{X}_\mu, B\bar{X}_\nu) = -(\cos(\nu\pi/2K'))^{-1}(\bar{X}_\mu, B\bar{X}_\nu) = (2\sqrt{2}K')^{-1}k_0^2 \times \sin[(\mu+\nu)\pi/2]\csc^2[(\mu+\nu)\pi/4K'],$$
$$\mu = 0, 2K', \quad \nu = 1,3,5,\ldots,2K'-1,$$
$$(\bar{X}_\mu, B\bar{X}_\nu) = -(2K')^{-1}k_0^2\{\sin[(\mu+\nu)\pi/2] \times \cot[(\mu+\nu)\pi/4K'] - \sin[(\nu-\mu)\pi/2] \times \cot[(\nu-\mu)\pi/4K']\}, \quad \text{(D16)}$$
$$2K'+1 \leq \mu,\ \nu \leq 4K'-1,$$
$$(\bar{X}_\mu, B\bar{X}_\nu) = -(4K')^{-1}k_0^2 \sin(\mu\pi/2K')\{\sin[(\mu+\nu)\pi/2] \times \csc^2[(\mu+\nu)\pi/4K'] + \sin[(\nu-\mu)\pi/2] \times \csc^2[(\nu-\mu)\pi/4K']\},$$
$$1 \leq \nu \leq 2K'-1, \quad 2K'+1 \leq \mu \leq 4K'-1.$$

* Research partially supported by National Science Foundation Grant No. GA-29658.

1 J. McKenna and J. A. Morrison, J. Math. Phys. 11, 2348 (1970).
2 J. McKenna and J. A. Morrison (to be published).
3 In Eq. (2) and throughout the rest of the paper, P denotes a probability measure on a Borel field of subsets of a set Ω of points ω. All random quantities are understood to be functions of ω and, for example, $P\{x(0) = a_j\}$ means $P\{\omega: x(\omega,0) = a_j\}$.
4 W. Feller, Trans. Am. Math. Soc. 48, 488 (1940).
5 U. Frisch, "Wave Propagation in Random Media," in *Probabilistic Methods in Applied Mathematics*, edited by A. T. Bharucha-Ried (Academic, New York, 1968).
6 A. I. Markushevich, *Theory of Functions of a Complex Variable* (Prentice-Hall, Englewood Cliffs, N. J., 1965), Vol. II.
7 R. C. Bourret, Can. J. Phys. 44, 2519 (1966).
8 G. Papanicolaou and J. B. Keller (to be published).
9 M. Abramowitz and I. A. Segun, *Handbook of Mathematical Functions* (Dover, New York, 1965).
10 W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1968) Vol. I.

Asymptotic Behavior of Vacuum Space-Times

W. E. Couch*
Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania 15213
and
R. J. Torrence†
Department of Mathematics, University of Calgary, Calgary, Alberta, Canada

(Received 26 July 1971)

Newman and Penrose have given conditions on the asymptotic form of the Weyl tensor in empty space-time that are sufficient to insure that the space-time is asymptotically flat at null infinity and has the peeling property. We give considerably weaker conditions and show them to be sufficient for asymptotic flatness. Under the weaker conditions the asymptotic behavior of the Weyl tensor is more general than the case where the peeling property holds. The asymptotic dependence on a suitably defined radial coordinate is given for the basis null tetrad, the spin coefficients, and the tetrad components of the Weyl tensor.

1. INTRODUCTION

Many attempts to investigate gravitational radiation from bounded sources have been carried out by considering vacuum space-times asymptotically flat at future null infinity.[1-4] In these investigations one wishes to set initial data which results in solutions possessing such asymptotic flatness. Conditions on the initial data necessary for asymptotic flatness are not known; as a result, one simply looks for conditions on the initial data sufficient to insure asymptotic flatness. It will be shown in this paper that the conditions usually assumed are much stronger than needed.

Expressed in terms of the components $(\Psi_0, \Psi_1, \Psi_2, \Psi_3, \Psi_4)$ of the Weyl tensor with respect to a null tetrad $(l^\mu, n^\mu, m^\mu, \bar{m}^\mu)$, the usual assumptions are

$$\Psi_0 = O(r^{-5}), \quad \frac{\partial}{\partial r}\Psi_0 = O(r^{-6}), \quad \frac{\partial}{\partial x^i}\Psi_0 = O(r^{-5}),$$

where l_μ is chosen to be orthogonal to null hypersurfaces, r is an affine parameter along null geodesics to which l^μ is tangent, and the x^i are coordinates[5] that label different null geodesics. The coordinates and tetrad used are more completely defined in Sec. 3. With these assumptions one can, in particular, establish the well-known "peeling" result

$$\Psi_A = O(r^{A-5}), \quad A = 0,\ldots,4.$$

In this paper it will be shown that if we assume

$$\Psi_0 = O(r^{-2-\epsilon_0}), \quad \frac{\partial}{\partial r}\Psi_0 = O(r^{-3-\epsilon_0}), \quad \frac{\partial}{\partial x^i}\Psi_0 = O(r^{-2-\epsilon_0}),$$

where $\epsilon_0 > 0$, then

$$\Psi_0 = O(r^{-2-\epsilon_0}),$$
$$\Psi_1 = O(r^{-2-\epsilon_1}),$$
$$\Psi_2 = O(r^{-2-\epsilon_2}),$$
$$\Psi_3 = O(r^{-2}),$$
$$\Psi_4 = O(r^{-1}),$$

where $\epsilon_0 > \epsilon_1 > \epsilon_2$. Thus under the weaker hypothesis the "peeling" result is modified, but the space is still asymptotically flat in the sense that all the Ψ's go to zero as $r \to \infty$.

Unfortunately it is not known whether the weakening of the conditions is significant in the sense that it will permit a wider class of space-times to be considered; there may not be any interesting asymptotically flat solutions which violate the stronger conditions but are permitted by the new, weaker ones. Penrose[6] has given requirements sufficient for a solution to

$$(\bar{X}_\mu, B\bar{X}_\nu) = -(\cos(\nu\pi/2K'))^{-1}(\bar{X}_\mu, B\bar{X}_\nu) = (2\sqrt{2}K')^{-1}k_0^2 \times \sin[(\mu+\nu)\pi/2]\csc^2[(\mu+\nu)\pi/4K'],$$
$$\mu = 0, 2K', \quad \nu = 1,3,5,\ldots,2K'-1,$$
$$(\bar{X}_\mu, B\bar{X}_\nu) = -(2K')^{-1}k_0^2\{\sin[(\mu+\nu)\pi/2] \times \cot[(\mu+\nu)\pi/4K'] - \sin[(\nu-\mu)\pi/2] \times \cot[(\nu-\mu)\pi/4K']\}, \quad \text{(D16)}$$
$$2K'+1 \leq \mu,\ \nu \leq 4K'-1,$$
$$(\bar{X}_\mu, B\bar{X}_\nu) = -(4K')^{-1}k_0^2 \sin(\mu\pi/2K')\{\sin[(\mu+\nu)\pi/2] \times \csc^2[(\mu+\nu)\pi/4K'] + \sin[(\nu-\mu)\pi/2] \times \csc^2[(\nu-\mu)\pi/4K']\},$$
$$1 \leq \nu \leq 2K'-1, \quad 2K'+1 \leq \mu \leq 4K'-1.$$

* Research partially supported by National Science Foundation Grant No. GA-29658.

1 J. McKenna and J. A. Morrison, J. Math. Phys. 11, 2348 (1970).
2 J. McKenna and J. A. Morrison (to be published).
3 In Eq. (2) and throughout the rest of the paper, P denotes a probability measure on a Borel field of subsets of a set Ω of points ω. All random quantities are understood to be functions of ω and, for example, $P\{x(0) = a_j\}$ means $P\{\omega: x(\omega,0) = a_j\}$.
4 W. Feller, Trans. Am. Math. Soc. 48, 488 (1940).
5 U. Frisch, "Wave Propagation in Random Media," in *Probabilistic Methods in Applied Mathematics*, edited by A. T. Bharucha-Ried (Academic, New York, 1968).
6 A. I. Markushevich, *Theory of Functions of a Complex Variable* (Prentice-Hall, Englewood Cliffs, N. J., 1965), Vol. II.
7 R. C. Bourret, Can. J. Phys. 44, 2519 (1966).
8 G. Papanicolaou and J. B. Keller (to be published).
9 M. Abramowitz and I. A. Segun, *Handbook of Mathematical Functions* (Dover, New York, 1965).
10 W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1968) Vol. I.

Asymptotic Behavior of Vacuum Space-Times

W. E. Couch*
Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania 15213
and
R. J. Torrence†
Department of Mathematics, University of Calgary, Calgary, Alberta, Canada

(Received 26 July 1971)

Newman and Penrose have given conditions on the asymptotic form of the Weyl tensor in empty space-time that are sufficient to insure that the space-time is asymptotically flat at null infinity and has the peeling property. We give considerably weaker conditions and show them to be sufficient for asymptotic flatness. Under the weaker conditions the asymptotic behavior of the Weyl tensor is more general than the case where the peeling property holds. The asymptotic dependence on a suitably defined radial coordinate is given for the basis null tetrad, the spin coefficients, and the tetrad components of the Weyl tensor.

1. INTRODUCTION

Many attempts to investigate gravitational radiation from bounded sources have been carried out by considering vacuum space-times asymptotically flat at future null infinity.[1-4] In these investigations one wishes to set initial data which results in solutions possessing such asymptotic flatness. Conditions on the initial data necessary for asymptotic flatness are not known; as a result, one simply looks for conditions on the initial data sufficient to insure asymptotic flatness. It will be shown in this paper that the conditions usually assumed are much stronger than needed.

Expressed in terms of the components $(\Psi_0, \Psi_1, \Psi_2, \Psi_3, \Psi_4)$ of the Weyl tensor with respect to a null tetrad $(l^\mu, n^\mu, m^\mu, \bar{m}^\mu)$, the usual assumptions are

$$\Psi_0 = O(r^{-5}), \quad \frac{\partial}{\partial r}\Psi_0 = O(r^{-6}), \quad \frac{\partial}{\partial x^i}\Psi_0 = O(r^{-5}),$$

where l_μ is chosen to be orthogonal to null hypersurfaces, r is an affine parameter along null geodesics to which l^μ is tangent, and the x^i are coordinates[5] that label different null geodesics. The coordinates and tetrad used are more completely defined in Sec. 3. With these assumptions one can, in particular, establish the well-known "peeling" result

$$\Psi_A = O(r^{A-5}), \quad A = 0,\ldots,4.$$

In this paper it will be shown that if we assume

$$\Psi_0 = O(r^{-2-\epsilon_0}), \quad \frac{\partial}{\partial r}\Psi_0 = O(r^{-3-\epsilon_0}), \quad \frac{\partial}{\partial x^i}\Psi_0 = O(r^{-2-\epsilon_0}),$$

where $\epsilon_0 > 0$, then

$$\Psi_0 = O(r^{-2-\epsilon_0}),$$
$$\Psi_1 = O(r^{-2-\epsilon_1}),$$
$$\Psi_2 = O(r^{-2-\epsilon_2}),$$
$$\Psi_3 = O(r^{-2}),$$
$$\Psi_4 = O(r^{-1}),$$

where $\epsilon_0 > \epsilon_1 > \epsilon_2$. Thus under the weaker hypothesis the "peeling" result is modified, but the space is still asymptotically flat in the sense that all the Ψ's go to zero as $r \to \infty$.

Unfortunately it is not known whether the weakening of the conditions is significant in the sense that it will permit a wider class of space-times to be considered; there may not be any interesting asymptotically flat solutions which violate the stronger conditions but are permitted by the new, weaker ones. Penrose[6] has given requirements sufficient for a solution to

have the peeling property, and included in these is the requirement that the space–time be conformally compactfiable. Any solution which is asymptotically flat in our weakened sense but does not have the peeling property must violate at least one of Penrose's requirements. Some indication of the possible significance of the weakened conditions can be obtained by considering the linearized theory. In that case it is clear that asymptotic flatness at future null infinity places a restriction on the variety of advanced (incoming) radiation fields which are permitted. Replacing $\Psi_0 = O(r^{-5})$ by $\Psi_0 = O(r^{-2-\epsilon_0})$ permits a wider class of advanced radiation fields (even among those required to have a finite total energy).

2. TWO LEMMAS

In the proof of asymptotic flatness for $\Psi_0 = O(r^{-2-\epsilon_0})$ given in the next section, two results on the asymptotic behavior of solutions of systems of linear ordinary differential equations are needed; these two results will be stated here as Lemma 1 and Lemma 2. The notation $F(x) = O(g(x))$ means that $|F|/g$ is bounded as $x \to \infty$ and $F(x) = o(g(x))$ means that F/g has the limit zero as $x \to \infty$.

Lemma 1: Let the complex $n \times n$ matrix B and the complex column vector f be given functions of x where

$$B = O(x^{-1-\epsilon}) \quad \text{and} \quad f = O(x^{-1-\epsilon}), \qquad \epsilon > 0.$$

Let the $n \times n$ matrix A be independent of x and have no eigenvalues with a positive real part. Suppose also that any eigenvalue of A with vanishing real part is regular (i.e., its multiplicity is equal to the number of linearly independent eigenvectors corresponding to it.) Then all of the solutions of

$$\frac{dy}{dx} = \left(\frac{A}{x} + B\right)y + f$$

are bounded as $x \to \infty$.

A similar form of this lemma is proven by Newman and Penrose.[1]

Lemma 2: If $d^2Y/dx^2 = -QY$, where

$$Y = \begin{pmatrix} y_1 & y_2 \\ \bar{y}_1 & \bar{y}_2 \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} 0 & \Psi_0 \\ \bar{\Psi}_0 & 0 \end{pmatrix}$$

and if $\Psi_0 = O(r^{-2-\epsilon})$, $\epsilon > 0$, then there exists Y such that

$$Y = Y^0 x + O(x)$$

and

$$\frac{dY}{dx} = Y^0 + O(x^{-\epsilon}),$$

where Y^0 is a constant.

Proof: The second-order matrix equation is equivalent to the equations $d^2y_1/dx^2 = -\Psi_0\bar{y}_1$, $d^2y_2/dx^2 = -\Psi_0\bar{y}_2$. If we put $y_1 = c + id$ and $\Psi_0 = a + ib$, the first equation is equivalent to $d^2c/dx^2 = -ac - bd$ and $d^2d/dx^2 = -bc + ad$. Now, by putting $dc/dx = c_1$ and $(d/dx)\, d = d_1$, this last pair of equations is equivalent to the first order system of equations

$$\frac{d}{dx}\begin{pmatrix} c \\ c_1 \\ d \\ d_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} c \\ c_1 \\ d \\ d_1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ -a & 0 & -b & 0 \\ 0 & 0 & 0 & 0 \\ -b & 0 & a & 0 \end{pmatrix}\begin{pmatrix} c \\ c_1 \\ d \\ d_1 \end{pmatrix}.$$

It follows,[7] since $\int^\infty x|\Psi_0(x)|dx > 0$, that $c = o(x)$, $d = o(x)$, $c_1 = o(1)$, and $d_1 = o(1)$. An identical argument with respect to the quantity y_2 allows us to conclude that $Y = o(x)$ and $dy/dx = o(1)$. Then $d^2Y/dx^2 = QY = o(x)O(x^{-2-\epsilon}) = o(x^{-1-\epsilon})$, which implies that $dY/dx = o(x^{-\epsilon}) + Y^0$, where Y^0 is a constant. If we define $\tilde{Y} = Y - xY^0$, then $d\tilde{Y}/dx = o(x^{-\epsilon})$. Thus $d(\tilde{Y}x^{-1})/dx = -(\tilde{Y}/x)/x + o(x^{-1-\epsilon})$ and we can apply Lemma 1 to each column of $\tilde{Y}/x$ to conclude that $\tilde{Y}/x = O(1)$, or that $\tilde{Y} = O(x)$. Thus $Y = Y^0x + O(x)$.[8]

3. ASYMPTOTIC FLATNESS

We use a null tetrad[1] of basis vectors $(l_\mu, n_\mu, m_\mu, \bar{m}_\mu)$ satisfying $l_\mu n^\mu = -m_\mu m^\mu = 1$, $l_\mu m^\mu = n_\mu m^\mu = 0$ The vector l_μ is chosen as the gradient of null hypersurfaces labeled by the coordinate $x^0 = u$ so that $l_\mu = u_{,\mu}$. The vectors n_μ, m_μ, and $\bar{m}_\mu$ are parallelly propagated along the geodesics to which l^μ is tangent. The coordinate x^1 is taken to be the affine parameter r, along l^μ, and the coordinates x^i label different null geodesics in each null hypersurface. Our tetrad then has the form $l^\mu = \delta^\mu_1$, $n^\mu = \delta^\mu_0 + U\delta^\mu_1 + X^i\delta^\mu_i$, and $m^\mu = \omega\delta^\mu_1 + \xi^i\delta^\mu_i$. The gravitational field is given by the tetrad components of the Weyl tensor, $\Psi_0 = -C_{\mu\nu\rho\sigma}l^\mu m^\nu l^\rho m^\sigma$, $\Psi_1 = -C_{\mu\nu\rho\sigma}l^\mu n^\nu l^\rho m^\sigma$, $\Psi_2 = -C_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu l^\rho m^\sigma$, $\Psi_3 = -C_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu l^\rho n^\sigma$, and $\Psi_4 = -C_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu \bar{m}^\rho n^\sigma$. The nonzero spin coefficients formed from the tetrad are defined as

$$\rho = l_{\mu;\nu}m^\mu\bar{m}^\nu,$$
$$\alpha = \tfrac{1}{2}(l_{\mu;\nu}n^\mu\bar{m}^\nu - m_{\mu;\nu}\bar{m}^\mu\bar{m}^\nu),$$
$$\beta = \tfrac{1}{2}(l_{\mu;\nu}n^\mu m^\nu - m_{\mu;\nu}\bar{m}^\mu m^\nu),$$
$$\tau = \bar{\alpha} + \beta,$$
$$\nu = -n_{\mu;\nu}\bar{m}^\mu n^\nu,$$
$$\gamma = \tfrac{1}{2}(l_{\mu;\nu}n^\mu n^\nu - m_{\mu;\nu}\bar{m}^\mu n^\nu),$$
$$\mu = -n_{\mu;\nu}\bar{m}^\mu m^\nu,$$
$$\sigma = l_{\mu;\nu}m^\mu m^\nu,$$
$$\lambda = -n_{\mu;\nu}\bar{m}^\mu\bar{m}^\nu.$$

Intrinsic derivatives with respect to the tetrad $(l^\mu, n^\mu, m^\mu, \bar{m}^\mu)$ are denoted respectively by $(D, \Delta, \delta, \bar{\delta})$. The equations that we will need are the vacuum field equations, the Bianchi identities, and certain equations that arise from the existence of the tetrad. Of all these equations, we need especially two classes: radial equations, which are those that do not contain Δ, and nonradial equations, which are those containing Δ.

The radial field equations are

$$D\rho = \rho^2 + \sigma\bar{\sigma}, \tag{3.1a}$$

$$D\sigma = 2\rho\sigma + \Psi_0, \tag{3.1b}$$

$$D\alpha = \rho\alpha + \bar{\sigma}\beta, \tag{3.1c}$$

$$D\beta = \rho\beta + \sigma\alpha + \Psi_1, \tag{3.1d}$$

$$D\lambda = \rho\lambda + \bar{\sigma}\mu, \tag{3.1e}$$

$$D\mu = \rho\mu + \sigma\lambda + \Psi_2, \tag{3.1f}$$

$$D\gamma = \tau\alpha + \bar{\tau}\beta + \Psi_2, \tag{3.1g}$$

$$D\nu = \tau\lambda + \bar{\tau}\mu + \Psi_3; \tag{3.1h}$$

the radial Bianchi identities are

$$D\Psi_1 - \bar{\delta}\Psi_0 = 4\rho\Psi_1 - 4\alpha\Psi_0, \tag{3.1i}$$

$$D\Psi_2 - \bar{\delta}\Psi_1 = 3\rho\Psi_2 - 2\alpha\Psi_1 - \lambda\Psi_0, \tag{3.1j}$$

$$D\Psi_3 - \bar{\delta}\Psi_2 = 2\rho\Psi_3 - 2\lambda\Psi_1, \tag{3.1k}$$

$$D\Psi_4 - \bar{\delta}\Psi_3 = \rho\Psi_4 + 2\alpha\Psi_3 - 3\lambda\Psi_2; \tag{3.1l}$$

the radial equations satisfied by the tetrad are

$$D\xi^i = \rho\xi^i + \sigma\bar{\xi}^i, \tag{3.1m}$$

$$D\omega = \rho\omega + \sigma\bar{\omega} - \tau, \tag{3.1n}$$

$$DX^i = \bar{\tau}\xi^i + \tau\bar{\xi}^i, \tag{3.1o}$$

$$DU = \bar{\tau}\omega + \tau\bar{\omega} - (\gamma + \bar{\gamma}); \tag{3.1p}$$

The nonradial field equations are

$$\Delta\lambda - \bar{\delta}\nu = 2\alpha\nu + (\bar{\gamma} - 3\gamma - \mu - \bar{\mu})\lambda - \Psi_4, \tag{3.2a}$$

$$\delta\nu - \Delta\mu = \gamma\mu - 2\nu\beta + \bar{\gamma}\mu + \mu^2 + \lambda\bar{\lambda}, \tag{3.2b}$$

$$\delta\gamma - \Delta\beta = \tau\mu - \sigma\nu + (\mu - \gamma + \bar{\gamma})\beta + \bar{\lambda}\alpha, \tag{3.2c}$$

$$\delta\tau - \Delta\sigma = 2\tau\beta + (\bar{\gamma} + \mu - 3\gamma)\sigma + \bar{\lambda}\rho, \tag{3.2d}$$

$$\Delta\rho - \bar{\delta}\tau = (\gamma + \bar{\gamma} - \bar{\mu})\rho - 2\alpha\tau - \lambda\sigma - \Psi_2, \tag{3.2e}$$

$$\Delta\alpha - \bar{\delta}\gamma = \rho\nu - \tau\lambda - \lambda\beta + (\bar{\gamma} - \gamma - \bar{\mu})\alpha - \Psi_3; \tag{3.2f}$$

the nonradial Bianchi identities are

$$\Delta\Psi_0 - \delta\Psi_1 = 4\gamma\Psi_0 - \mu\Psi_0 - 4\tau\Psi_1 - 2\beta\Psi_1 + 3\sigma\Psi_2, \tag{3.2g}$$

$$\Delta\Psi_1 - \delta\Psi_2 = \nu\Psi_0 + 2\gamma\Psi_1 - 2\mu\Psi_1 - 3\tau\Psi_2 + 2\sigma\Psi_3, \tag{3.2h}$$

$$\Delta\Psi_2 - \delta\Psi_3 = 2\nu\Psi_1 - 3\mu\Psi_2 - 2\tau\Psi_3 + 2\beta\Psi_3 + \sigma\Psi_4, \tag{3.2i}$$

$$\Delta\Psi_3 - \delta\Psi_4 = 3\nu\Psi_2 - 2\gamma\Psi_3 - 4\mu\Psi_3 - \tau\Psi_4 + 4\beta\Psi_4; \tag{3.2j}$$

the nonradial equations satisfied by the tetrad are

$$\delta X^i - \Delta\xi^i = (\mu + \bar{\gamma} - \gamma)\xi^i + \bar{\lambda}\bar{\xi}^i, \tag{3.2k}$$

$$\delta U - \Delta\omega = (\mu + \bar{\gamma} - \gamma)\omega + \bar{\lambda}\bar{\omega} - \bar{\nu}. \tag{3.2l}$$

The remainder of the equations are

$$\begin{aligned}
\delta\rho - \bar{\delta}\sigma &= \tau\rho + (\bar{\beta} - 3\alpha)\sigma - \Psi_1,\\
\delta\alpha - \bar{\delta}\beta &= \mu\rho - \lambda\sigma - 2\alpha\beta + \alpha\bar{\alpha} + \beta\bar{\beta} - \Psi_2,\\
\delta\lambda - \bar{\delta}\mu &= \bar{\tau}\mu + (\bar{\alpha} - 3\beta)\lambda - \Psi_3,\\
\delta\bar{\xi}^i - \bar{\delta}\xi^i &= (\bar{\beta} - \alpha)\xi^i + (\bar{\alpha} - \beta)\bar{\xi}^i,\\
\delta\bar{\omega} - \bar{\delta}\omega &= (\bar{\beta} - \alpha)\omega + (\bar{\alpha} - \beta)\bar{\omega} + \mu - \bar{\mu}.
\end{aligned} \tag{3.2'}$$

Consider first Eqs. (3.1a) and (3.1b). Putting

$$P = \begin{pmatrix} \rho & \sigma \\ \bar{\sigma} & \rho \end{pmatrix}, \qquad Q = \begin{pmatrix} 0 & \Psi_0 \\ \bar{\Psi}_0 & 0 \end{pmatrix},$$

we can write those two equations as

$$\frac{dP}{dr} = P^2 + Q.$$

If Y satisfies

$$\frac{dY}{dr} = -PY, \tag{3.3}$$

then

$$\frac{d^2Y}{dr^2} = -QY, \tag{3.4}$$

and, by solving Eq. (3.3) for P,

$$P = -\frac{dY}{dr}\,Y^{-1} \tag{3.5}$$

satisfies $dP/dr = P^2 + Q$ if Y is nonsingular. But we assume that

$$\Psi_0 = O(r^{-2-\epsilon_0}) \tag{3.6}$$

so that Lemma 2 is applicable to Eq. (3.4), and we can conclude that there exists a Y such that

$$Y = Y^0 r + O(r),$$

$$\frac{dY}{dr} = Y^0 + O(r^{-\epsilon_0}). \tag{3.7a}$$

If we now consider $\hat{Y} \equiv r^k(Y - Y^0 r)$ we see that

$$\frac{d}{dr}\,\hat{Y} = \frac{k}{r}\,\hat{Y} + O(r^{k-\epsilon_0}) \tag{3.8}$$

and Lemma 1 is applicable if $k \le 0$ and $k - \epsilon_0 < -1$. This allows us to conclude, setting $k = -1 + \epsilon_\rho$, that

$$Y = Y^0 r + O(r^{1-\epsilon_\rho}) \tag{3.7b}$$

where $\epsilon_\rho < \epsilon_0$ and $\epsilon_\rho \le 1$. Substituting Eqs. (3.7) into Eq. (3.5), we obtain[9]

$$\rho = -r^{-1} + O(r^{-1-\epsilon_\rho}), \qquad \sigma = O(r^{-1-\epsilon_\rho}), \tag{3.9}$$

It follows immediately from Eqs. (3.1a), (3.1b), and (3.9) that

$$D\rho = r^{-2} + O(r^{-2-\epsilon_\rho}),$$

$$D\sigma = O(r^{-2-\epsilon_\rho}).$$

The asymptotic behavior of $\partial\rho/\partial x^i$ and $\partial\sigma/\partial x^i$ is found by differentiating Eqs. (3.1a), (3.1b) with respect to x^i. The resulting equations may be put into

a form where Lemma 1 is applicable to the column vector

$$\left(r^{1+\epsilon_\rho}\frac{\partial\rho}{\partial x^i}, r^{1+\epsilon_\rho}\frac{\partial\sigma}{\partial x^i}, r^{1+\epsilon_\rho}\frac{\partial\bar{\sigma}}{\partial x^i}\right)$$

with $A = (-1+\epsilon_\rho)I$. We have then

$$\frac{\partial\rho}{\partial x^i}, \frac{\partial\sigma}{\partial x^i} = O(r^{-1-\epsilon_\rho}).$$

The asymptotic forms for ρ and σ given by Eq.(3.9) and the assumed form for Ψ_0, $D\Psi_0$, and $(\partial/\partial x^i)\,\Psi_0$ allow the radial equations for ξ^i, α, β, ω, and Ψ_1 to be written in such a way that Lemma 1 is applicable to the column vector $(r\xi^i, r\bar{\xi}^i, r\alpha, r\bar{\alpha}, r\beta, r\bar{\beta}, \omega, \bar{\omega}, r^{2+\epsilon_1}\Psi_1, r^{2+\epsilon_1}\bar{\Psi}_1)$ with

$$A = \begin{pmatrix}
0 & \cdots & & & & & & & & 0 \\
\vdots & & & & & & & & & \vdots \\
0 & \cdots & & & & & & & & 0 \\
0 & 0 & 0 & 0 & -1 & -1 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & \cdots & & & & & & & 0 & -2+\epsilon_1 & 0 \\
0 & \cdots & & & & & & & 0 & 0 & -2+\epsilon_1
\end{pmatrix},$$

where $\epsilon_1 \le 2$ and $\epsilon_1 < \epsilon_0$. Then we conclude

$$\xi^i, \alpha, \beta = O(r^{-1}),$$
$$\omega = O(1),$$
$$\Psi_1 = O(r^{-2-\epsilon_1}).$$

It then follows from Eq. (3.1) that

$$D\xi^i, D\alpha, D\beta = O(r^{-2}),$$
$$D\omega = O(r^{-1}),$$
$$D\Psi_1 = O(r^{-3-\epsilon_1}).$$

After differentiation of Eqs. (3.1c), (3.1d), (3.1i), (3.1m), and (3.1n) with respect to x^i, we can apply Lemma 1 and show that

$$\frac{\partial}{\partial x^i}\xi^j, \frac{\partial}{\partial x^i}\alpha, \frac{\partial\beta}{\partial x^i} = O(r^{-2}),$$
$$\frac{\partial\omega}{\partial x^i} = O(1),$$
$$\frac{\partial\Psi_1}{\partial x^i} = O(r^{-2-\epsilon_1}).$$

The results established thus far allow Eqs. (3.1e), (3.1f), and (3.1j) to be written in such a way that Lemma 1 is applicable to the column vector $(r\lambda, r\mu, r^{2+\epsilon_2}\Psi_2)$ with

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \epsilon_2 - 1 \end{pmatrix}$$

where $\epsilon_2 < \epsilon_1$ and $\epsilon_2 \le 1$. Then we conclude that

$$\mu, \lambda = O(r^{-1}),$$
$$\Psi_2 = O(r^{-2-\epsilon_2}).$$

Equations (3.1e), (3.1f), and (3.1j) then show that

$$D\mu, D\lambda = O(r^{-2}),$$
$$D\Psi_2 = O(r^{-3-\epsilon_2}).$$

Again we can apply Lemma 1 to the equations resulting from taking $\partial/\partial x^i$ of Eqs. (3.1e), (3.1f), and (3.1j) and show that

$$\frac{\partial\mu}{\partial x^i}, \frac{\partial\lambda}{\partial x^i} = O(r^{-1}),$$
$$\frac{\partial\Psi_2}{\partial x^i} = O(r^{-2-\epsilon_2}).$$

We next apply Lemma 1 to each of the remaining radial equations individually and find

$$\Psi_3, \frac{\partial\Psi_3}{\partial x^i} = O(r^{-2}),$$
$$\Psi_4, \frac{\partial\Psi_4}{\partial x^i} = O(r^{-1}),$$
$$\gamma, \nu, X^j, \frac{\partial\gamma}{\partial x^i}, \frac{\partial\nu}{\partial x^i}, \frac{\partial X^j}{\partial x^i} = O(1),$$
$$U, \frac{\partial U}{\partial x^i} = O(r^{1+\delta}), \quad \delta > 0,$$
$$D\Psi_3 = O(r^{-3}),$$
$$D\Psi_4 = O(r^{-2}),$$
$$D\gamma, D\nu, DX^i = O(r^{-1}), \quad DU = O(1).$$

In order to establish that every solution of the field equations that satisfies Ψ_0, $\partial\Psi_0/\partial x^i = O(r^{-2-\epsilon_0})$ and $D\Psi_0 = O(r^{-3-\epsilon_0})$ has an asymptotic form at least as strong as that established by applying Lemma 1 to the radial equations, it is necessary to show that the time derivative $(\partial/\partial u)$ of all the quantities also has such an asymptotic form. To show this, it is necessary to integrate the radial equations for some quantities and obtain explicit r dependence for their leading orders. We give the results for all the explicit r dependence we can obtain without further specifying the assumptions on Ψ_0:

$$\begin{aligned}
\xi^i &= (\xi^{i0}/r) + O(r^{-1-\epsilon_\rho}), \\
\alpha &= (\alpha^0/r) + O(r^{-1-\epsilon_\rho}), \\
\beta &= (\beta^0/r) + O(r^{-1-\epsilon_\rho}) + O(r^{-1-\epsilon_1}), \\
\omega &= -\tau^0 + O(r^{-\epsilon_\rho}) + O(r^{-1-\epsilon_1}), \\
\gamma &= \gamma^0 - [(\tau^0\alpha^0 + \bar{\tau}^0\beta^0)/r] \\
&\quad + O(r^{-1-\epsilon_2}) + O(r^{-1-\epsilon_\rho}), \\
\lambda &= (\lambda^0/r) + O(r^{-1-\epsilon_\rho}), \\
\mu &= (\mu^0/r) + O(r^{-1-\epsilon_2}), \\
\Psi_3 &= (\Psi_3^0/r^2) + O(r^{-2-\epsilon_2}) \qquad (3.10a) \\
\nu &= \nu^0 - [(\tau^0\lambda^0 + \bar{\tau}^0\mu^0 + \Psi_3^0)/r] \\
&\quad + O(r^{-1-\epsilon_\rho}) + O(r^{-1-\epsilon_2}), \\
X^i &= X^{i0} - [(\tau^0\bar{\xi}^{i0} + \bar{\tau}^0\xi^{i0})/r] \\
&\quad + O(r^{-1-\epsilon_\rho}) + O(r^{-1-\epsilon_1}), \\
\Psi_4 &= (\Psi_4^0/r) + O(r^{-1-\epsilon_\rho}), \\
U &= U^0 - (\gamma^0 + \bar{\gamma}^0)r + O(r^{-\epsilon_\rho}) + O(r^{-\epsilon_2}),
\end{aligned}$$

where $\frac{\partial}{\partial r}(\xi^{i0}, \alpha^0, \beta^0, \text{etc.}) = 0$. Note that the use of $\tau^0 = \bar{\alpha}^0 + \beta^0$ and of τ, ω, and γ as given by Eq. (3.10) has enabled us by integration of Eq. (3.1p) to find that U falls off faster than the asymptotic form for U found by application of Lemma 1 to Eq. (3.1p). We still have

$$\begin{aligned} \Psi_0 &= O(r^{-2-\epsilon_0}), \\ \Psi_1 &= O(r^{-2-\epsilon_1}), \\ \Psi_2 &= O(r^{-2-\epsilon_2}), \\ \rho &= -r^{-1} + O(r^{-1-\epsilon_\rho}), \qquad \sigma = O(r^{-1-\epsilon_\rho}). \end{aligned} \tag{3.10b}$$

The requirement that Eqs. (3.2′) hold asymptotically merely imposes some relationships among the functions independent of r that occur in Eq. (3.10). We consider these relations to hold but do not display them. Equations (3.2′) have also been used in establishing ω and U as given by Eqs. (3.10); from the radial equations alone one can establish only that $\omega = -\tau^0 + O(r^{-\epsilon_\omega})$, $\epsilon_\omega < \epsilon_\rho$, $\epsilon_\omega < \epsilon_1$, and $U = U^0 - (\gamma^0 + \bar{\gamma}^0)r + O(r^{-\epsilon_\omega}) + O(r^{-\epsilon_2})$.

We substitute $\Delta = \partial/\partial u + UD + X^i\,\partial/\partial x^i$ into Eqs. (3.2) and solve for the time-differentiated quantities that occur. This results in expressions for the time-differentiated quantities in terms of quantities whose asymptotic behavior is given by Eqs. (3.10). We find then that the asymptotic dependence of $(\dot{\Psi}_0, \dot{\Psi}_1, \dot{\Psi}_2, \dot{\Psi}_3, \dot{\lambda}, \dot{\mu}, \dot{\beta}, \dot{\sigma}, \dot{\rho}, \dot{\alpha}, \dot{\xi}^i, \dot{\omega})$ is the same as that of $(\Psi_0, \Psi_1, \Psi_2, \Psi_3, \lambda, \mu, \beta, \sigma, \rho, \alpha, \xi^i, \omega)$. The quantities $(\dot{\Psi}_4, \dot{U}, \dot{X}^i, \dot{\nu}, \dot{\gamma})$ are found to have the same asymptotic form as $(\Psi_4, U, X^i, \nu, \gamma)$ by differentiation of Eqs. (3.1p), (3.1o), (3.1l), (3.1h), and (3.1g) with respect to u and applying Lemma 1. This establishes that the asymptotic orders for the tetrad, spin coefficients, and tetrad components of the Weyl tensor are those given by Eq. (3.10).

* Research supported in part by the National Science Foundation.
† Research supported in part by the National Research Council of Canada.

1 E. T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).
2 E. T. Newman and T. W. J. Unti, J. Math. Phys. **3**, 891 (1962).
3 H. Bondi, M. Van der Burg, and A. Metzner, Proc. Roy. Soc. (London) **A269**, 21 (1962).
4 W. E. Couch, R. J. Torrence, A. I. Janis, and E. T. Newman, J. Math. Phys. **9**, 484 (1968).
5 Greek indices range and sum from 0 to 3, lower case Latin indices from 2 to 3.
6 R. Penrose, Proc. Roy. Soc. (London) **A284**, 159 (1965).
7 E. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York, 1955), p. 106. Exercise 35 is directly applicable to our system; in our case the r of exercise 35 has the value one and the characteristic roots of A all vanish.
8 This result was also needed for the work presented in Ref. 1. It appears that their justification was inadequate; thus this more elaborate argument is given here.
9 We must assume that Y^0 is a nonsingular matrix. The discussion of this point in Ref. 1, p. 575, applies here.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 13, NUMBER 1 JANUARY 1972

Non-L^2 Solutions of Exactly Soluble Relativistic Wave Equations and the Lee Model

C. A. Nelson and A. K. Rajagopal
Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803
(Received 31 August 1971)

In nonrelativistic potential theory there can exist singularities of the S matrix which are associated with wavefunctions belonging to a non-L^2 class. In this paper the non-L^2 character of these singularities is shown to persist in several relativistic models: in (1) two schemes, including that of the Klein–Gordon equation and the coupling of a classical relativistic field to an external exponential source, in (2) the pair theory, classical and quantized, with separable interactions, and in (3) the Lee model. For (2) and (3), poles of the "proper" Jost function $\bar{f}_-(k, 0)$, i.e., the Fredholm determinant for the outgoing scattering state, correspond to non-L^2 class solutions of the associated dynamical field equation in coordinate space. These non-L^2 solutions will be called "shadow" fields. They are of special importance and bear the same relation as the "shadow" states in potential theory because they also are of dynamical origin and do not appear in the completeness or unitarity relations by virtue of their non-L^2 status.

1. INTRODUCTION

Recently, in nonrelativistic potential theory, it was shown that there can exist singularities of the S matrix which are associated with wavefunctions belonging to a non-L^2 class.[1] To be precise, for S identified from the Jost functions,

$$S(k) = f_-(k, 0)/f_+(k, 0), \tag{1.1}$$

any singularities of the Jost function $f_-(k, 0)$ correspond to a non-L^2 class of solutions of the Schrödinger equation, unlike the zeros of $f_+(k, 0)$ which correspond to the L^2 class, i.e., genuine bound states. Such non-L^2 states (often referred to as "shadow" states for this reason) appear in the same part of the physical k plane as the bound state poles and are "dynamical" in that their analytic properties, such as position and discontinuities, are a function of the potential. That there does exist such a well-defined non-L^2 class of solutions associated with some singularities of the S matrix was proven by construction of the full Green's function for the Schrödinger equation with the boundary conditions of regularity at the origin and spherical outgoing waves at large distance.[1] From it, first, the completeness statement was derived which determines the complete set of L^2 class of solutions. Then, from the usual operator relation between the T matrix and the full Green's function, the S matrix was constructed explicitly and seen to contain additional singularities. Hence, their associated wavefunctions belong to a non-L^2 class.[1]

Historically, almost a quarter of a century ago, Ma[2,3] discovered an example of such singularities in the particular case on an exponential potential. The

where $\frac{\partial}{\partial r}(\xi^{i0}, \alpha^0, \beta^0, \text{etc.}) = 0$. Note that the use of $\tau^0 = \bar{\alpha}^0 + \beta^0$ and of τ, ω, and γ as given by Eq. (3.10) has enabled us by integration of Eq. (3.1p) to find that U falls off faster than the asymptotic form for U found by application of Lemma 1 to Eq. (3.1p). We still have

$$\begin{aligned} \Psi_0 &= O(r^{-2-\epsilon_0}), \\ \Psi_1 &= O(r^{-2-\epsilon_1}), \\ \Psi_2 &= O(r^{-2-\epsilon_2}), \\ \rho &= -r^{-1} + O(r^{-1-\epsilon_\rho}), \qquad \sigma = O(r^{-1-\epsilon_\rho}). \end{aligned} \tag{3.10b}$$

The requirement that Eqs. (3.2′) hold asymptotically merely imposes some relationships among the functions independent of r that occur in Eq. (3.10). We consider these relations to hold but do not display them. Equations (3.2′) have also been used in establishing ω and U as given by Eqs. (3.10); from the radial equations alone one can establish only that $\omega = -\tau^0 + O(r^{-\epsilon_\omega})$, $\epsilon_\omega < \epsilon_\rho$, $\epsilon_\omega < \epsilon_1$, and $U = U^0 - (\gamma^0 + \bar{\gamma}^0)r + O(r^{-\epsilon_\omega}) + O(r^{-\epsilon_2})$.

We substitute $\Delta = \partial/\partial u + UD + X^i\,\partial/\partial x^i$ into Eqs. (3.2) and solve for the time-differentiated quantities that occur. This results in expressions for the time-differentiated quantities in terms of quantities whose asymptotic behavior is given by Eqs. (3.10). We find then that the asymptotic dependence of $(\dot{\Psi}_0, \dot{\Psi}_1, \dot{\Psi}_2, \dot{\Psi}_3, \dot{\lambda}, \dot{\mu}, \dot{\beta}, \dot{\sigma}, \dot{\rho}, \dot{\alpha}, \dot{\xi}^i, \dot{\omega})$ is the same as that of $(\Psi_0, \Psi_1, \Psi_2, \Psi_3, \lambda, \mu, \beta, \sigma, \rho, \alpha, \xi^i, \omega)$. The quantities $(\dot{\Psi}_4, \dot{U}, \dot{X}^i, \dot{\nu}, \dot{\gamma})$ are found to have the same asymptotic form as $(\Psi_4, U, X^i, \nu, \gamma)$ by differentiation of Eqs. (3.1p), (3.1o), (3.1l), (3.1h), and (3.1g) with respect to u and applying Lemma 1. This establishes that the asymptotic orders for the tetrad, spin coefficients, and tetrad components of the Weyl tensor are those given by Eq. (3.10).

* Research supported in part by the National Science Foundation.
† Research supported in part by the National Research Council of Canada.

1 E. T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).
2 E. T. Newman and T. W. J. Unti, J. Math. Phys. **3**, 891 (1962).
3 H. Bondi, M. Van der Burg, and A. Metzner, Proc. Roy. Soc. (London) **A269**, 21 (1962).
4 W. E. Couch, R. J. Torrence, A. I. Janis, and E. T. Newman, J. Math. Phys. **9**, 484 (1968).
5 Greek indices range and sum from 0 to 3, lower case Latin indices from 2 to 3.
6 R. Penrose, Proc. Roy. Soc. (London) **A284**, 159 (1965).
7 E. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York, 1955), p. 106. Exercise 35 is directly applicable to our system; in our case the r of exercise 35 has the value one and the characteristic roots of A all vanish.
8 This result was also needed for the work presented in Ref. 1. It appears that their justification was inadequate; thus this more elaborate argument is given here.
9 We must assume that Y^0 is a nonsingular matrix. The discussion of this point in Ref. 1, p. 575, applies here.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 13, NUMBER 1 JANUARY 1972

Non-L^2 Solutions of Exactly Soluble Relativistic Wave Equations and the Lee Model

C. A. Nelson and A. K. Rajagopal
Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803
(Received 31 August 1971)

In nonrelativistic potential theory there can exist singularities of the S matrix which are associated with wavefunctions belonging to a non-L^2 class. In this paper the non-L^2 character of these singularities is shown to persist in several relativistic models: in (1) two schemes, including that of the Klein–Gordon equation and the coupling of a classical relativistic field to an external exponential source, in (2) the pair theory, classical and quantized, with separable interactions, and in (3) the Lee model. For (2) and (3), poles of the "proper" Jost function $\bar{f}_-(k, 0)$, i.e., the Fredholm determinant for the outgoing scattering state, correspond to non-L^2 class solutions of the associated dynamical field equation in coordinate space. These non-L^2 solutions will be called "shadow" fields. They are of special importance and bear the same relation as the "shadow" states in potential theory because they also are of dynamical origin and do not appear in the completeness or unitarity relations by virtue of their non-L^2 status.

1. INTRODUCTION

Recently, in nonrelativistic potential theory, it was shown that there can exist singularities of the S matrix which are associated with wavefunctions belonging to a non-L^2 class.[1] To be precise, for S identified from the Jost functions,

$$S(k) = f_-(k, 0)/f_+(k, 0), \tag{1.1}$$

any singularities of the Jost function $f_-(k, 0)$ correspond to a non-L^2 class of solutions of the Schrödinger equation, unlike the zeros of $f_+(k, 0)$ which correspond to the L^2 class, i.e., genuine bound states. Such non-L^2 states (often referred to as "shadow" states for this reason) appear in the same part of the physical k plane as the bound state poles and are "dynamical" in that their analytic properties, such as position and discontinuities, are a function of the potential. That there does exist such a well-defined non-L^2 class of solutions associated with some singularities of the S matrix was proven by construction of the full Green's function for the Schrödinger equation with the boundary conditions of regularity at the origin and spherical outgoing waves at large distance.[1] From it, first, the completeness statement was derived which determines the complete set of L^2 class of solutions. Then, from the usual operator relation between the T matrix and the full Green's function, the S matrix was constructed explicitly and seen to contain additional singularities. Hence, their associated wavefunctions belong to a non-L^2 class.[1]

Historically, almost a quarter of a century ago, Ma[2,3] discovered an example of such singularities in the particular case on an exponential potential. The

essential non-L^2 character of such singularities was not recognized until very recently.[1] Ma showed that for this potential in the s wave S matrix there exist poles that do not contribute to the completeness. Because such poles exist in this and other potentials, Sudarshan[4] quite recently has stressed their importance with regard to the concept of "shadow" states. These non-L^2 singularities, indeed, are of intrinsic interest as "shadow" states, first, because of their *dynamical origin* and, second, because by virtue of their non-L^2 status they possess a very elegant reason for not appearing in the completeness statement and hence the *unitarity relation*. Their lack of appearance in the unitarity relation in contrast to other approaches is not due to energy conservation[5] nor to the introduction of standing waves.[6] They do not lead, in the theories studied so far, to any complicated analytic properties or violations of causality.

From recent history of particle physics it will be recalled that Regge poles also originated through work in potential theory. Since that time they have been investigated and proven useful both formally and phenomenologically on more interesting and complex theoretical levels. The question if in a similar way these non-L^2 singularities would appear in domains closer to that of a true interacting quantum field theory becomes of considerable interest. In particular, as with Regge poles, the investigation of these non-L^2 class of solutions when the theory is relativized, i.e., in relativistic wave equations, becomes of a similar importance. Then, as a preface to consideration of these singularities in a nonperturbative quantum field theory model, we consider their behavior in a separable model, which also shows the effects of nonlocality on their properties. In this case, the criteria for the occurance of a non-L^2 singularity is the vanishing of a "proper" Jost function which is the same as the Fredholm determinant. This also enables us to introduce some definitions and notations[7] regarding Jost solutions which we shall use repeatedly in the remainder of the paper. These matters are dealt with in Secs. 2 and 3. While the separable model can be quantized to give a soluble quantum field theory with pairwise interactions in which the non-L^2 singularities persist in the S matrix, we prefer to consider them instead in Sec. 4 in the Lee model. For both this field theoretic model and the pairwise theory, classical or quantized, the non-L^2 character of the wavefunctions associated with these singularities can be demonstrated by general arguments. For each of the various models considered in this paper, we construct specific examples of non-L^2 singularities and show that, as in nonrelativistic potential theory, they correspond to the vanishing of a Wronskian. Thus, the non-L^2 singularities are associated with a breakdown of the linear independence of the two basic solutions employed in construction of the S matrix. By analogy with the nonrelativistic results, we call these "shadow" fields. The last section contains some concluding remarks.

2. EXACT SOLUTIONS OF SCALAR AND VECTOR COUPLING MODELS

We consider the effects of relativity on these non-L^2 singularities in two different schemes for the coupling of a classical field to an external exponential source. These particular relativistic models were employed previously by Guralnik and Hagen[8] to determine the properties of Regge poles for the $1/r$ type of source. With the same spirit, we approach them here from the viewpoint of the results we found for the exponential potential in the nonrelativistic case. In the case of the vector coupling model, in the fixed source limit the wave equation is that for the coupling of a Klein–Gordon field to a $e^{-r/b}$ type source. In both these cases, we find that the non-L^2 character of these singularities persist and are the same as in the Schrödinger theory, except for the obvious change in kinematics.

A. The Scalar Coupling Model

The scalar model is defined by the Lagrangian

$$\begin{aligned} L &\equiv \int dx\, \mathcal{L}(x) \\ &= \int dx \{\varphi^{\mu *}\partial_\mu \varphi - \tfrac{1}{2}[m + \chi(x)]^2 \varphi\varphi^* \\ &\quad + \tfrac{1}{2}\varphi^{\mu *}\varphi_\mu + \text{c.c.}\}, \end{aligned} \tag{2.1}$$

where the metric is $(+,+,+,-)$. This gives the wave equation

$$\{-\partial_\mu^2 + [m + \chi(x)]^2\}\varphi(x) = 0 \tag{2.2}$$

or, for $\chi(x)$ spherically symmetric, the radial equation

$$\begin{aligned} &\frac{E^2 - m^2}{2m}\varphi(x) \\ &\quad = \left(\frac{p_r^2}{2m} + \frac{l(l+1)}{2mr^2} + \frac{1}{2m}[\chi(x)]^2 + \chi(x)\right)\varphi(x). \end{aligned} \tag{2.3}$$

It is useful to compare this equation (2.3) with the usual Schrödinger equation for coupling to a spherically symmetric source

$$E_{NR}\psi(x) = \left(\frac{p_r^2}{2m} + \frac{l(l+1)}{2mr^2} + \chi(x)\right)\psi(x), \tag{2.4}$$

where $E_{NR} = k^2/2m$. Notice that $(E^2 - m^2)/2m$ of (2.3) corresponds to the nonrelativistic energy variable E_{NR}. The physical sheet is defined by $\text{Im}\,[(E^2 - m^2)^{1/2}] > 0$.

It is useful both here and in our subsequent discussion to employ the s wave Jost solutions $f_\pm(k, r)$ of Eq. (2.4).[7] When $f_\pm(k, r)$ are evaluated at $r = 0$, they are called the Jost functions $f_\pm(k, 0)$. If $f_\pm(k, r)$ exist such that

$$\lim_{r\to\infty} e^{\mp ikr} f_\pm(k, r) = 1 \tag{2.5}$$

and are *linearly independent of each other*, then the most general *regular solution* of (2.4) can be constructed as

$$\varphi(k, r) = (1/2ik)[f_-(k, 0)f_+(k, r) - f_+(k, 0)f_-(k, r)], \tag{2.6}$$

such that for $r \to 0$, $\varphi(k, r) \to 0$, and $\varphi'(k, r) \to 1$ (prime means differentiation with respect to r). It should be noted that the Wronskian of the Jost solutions $f_\pm(k, r)$ is

$$W(f_+, f_-) = -2ik \tag{2.7}$$

$[W(\psi_1, \psi_2) = \psi_1\psi_2' - \psi_1'\psi_2]$. The full Green's function $\mathcal{G}^{(+)}$ associated with (2.4) that satisfies the boundary conditions of regularity at the origin and spheri-

cal waves at infinity can also be constructed from $f_\pm(k,r)$. Then via $T = V + V\mathcal{G}^{(+)}V$, the corresponding T matrix, and hence the S matrix, as given by (1.1) can be constructed.

For

$$\chi(x) = -v_0 e^{-r/a} \tag{2.8}$$

both (2.3) and (2.4) are soluble exactly for s waves. Substitution of $z = e^{-r/a}$ into (2.3) yields a form of Whittaker's equation. The unique solutions satisfying (2.5) exist and are

$$\begin{aligned} f_+(k,r) &= z^{(\gamma-1)/2}e^{-dz/2}\,y_1(\alpha,\gamma;dz) \\ &= z^{(\gamma-1)/2}e^{-dz/2}\Phi(\alpha,\gamma;dz), \end{aligned} \tag{2.9a}$$

$$\begin{aligned} f_-(k,r) &= f_+(-k,r) \\ &= z^{(\gamma-1)/2}e^{-dz/2}\,y_2(\alpha,\gamma;dz) \\ &= z^{(1-\gamma)/2}e^{-dz/2}\Phi(\alpha-\gamma+1,2-\gamma;dz), \end{aligned} \tag{2.9b}$$

where the second lines give the standard definitions[9] of y_1 and y_2 in terms of the confluent hypergeometric function of the first kind, Φ. For the scalar coupling model,

$$\begin{aligned} \alpha &= \tfrac{1}{2} - a(m + ik), \\ \gamma &= 1 - i2ak, \\ d &= 2av_0, \end{aligned} \tag{2.10}$$

where both α and γ are simple linear functions of k. Then

$$\begin{aligned} S(k) &= f_-(k,0)/f_+(k,0) \\ &= \Phi(\alpha-\gamma+1,2-\gamma;d)/\Phi(\alpha,\gamma;d) \end{aligned} \tag{2.11}$$

and the bound states correspond to the zeros of $f_+(k,0)$, i.e.,

$$\Phi(\alpha,\gamma;d) = 0. \tag{2.12}$$

Since, for fixed z', $\Phi(\alpha',\gamma',z')$ is an entire function of α' and a meromorphic function of γ' with simple poles at $\gamma' = 0, -1, \cdots, f_-(k,0)$ has poles at

$$\gamma = n + 1$$

or

$$2iak = -n, \quad n \text{ integer } (\neq 0), \tag{2.13}$$

which are identical with the positions of the non-L^2 singularities in the nonrelativistic case for the exponential potential!

Since $y_2/\Gamma(2-\gamma)$ tends to a finite limit at $\gamma = 2, \cdots$, but then is simply a multiple of y_2 there, the Jost functions $f_\pm(k,r)$ are not linearly independent at values of k given by (2.13). However, the regular solution $\varphi(k,r)$ defined by (2.6) and (2.9) can be computed for these values. As y_5 and y_7 provide a fundamental system of solutions under all circumstances, we shall express $\varphi(k,r)$ in terms of them in order to study its behavior. Using

$$\begin{aligned} y_1 &= \frac{\Gamma(\gamma)}{\Gamma(\gamma-\alpha)}\,e^{i\pi\alpha\epsilon}y_5 + \frac{\Gamma(\gamma)}{\Gamma(\alpha)}\,e^{i\pi(\alpha-\gamma)\epsilon}y_7, \\ y_2 &= e^{i\pi(\alpha-\gamma)\epsilon}\left(-\frac{\Gamma(2-\gamma)}{\Gamma(1-\alpha)}\,y_5 + \frac{\Gamma(2-\gamma)}{\Gamma(\alpha-\gamma+1)}\,y_7\right), \end{aligned} \tag{2.14}$$

where $\epsilon = \operatorname{sgn}(\operatorname{Im} z) = \pm 1$ as $\operatorname{Im} z \gtrless 0$, we obtain

$$\varphi(k,r) = N(z)(\gamma-1)\,e^{i\pi(\gamma-\alpha)}[y_5(d)y_7(dz) - y_7(d)y_5(dz)] \tag{2.15}$$

with

$$N(z) = z^{(\gamma-1)/2}\,e^{-d(1+z)/2}/2ik.$$

Computing $\varphi(k,r)$ for $k = in/2a$, we see that for all α

$$\lim_{r\to\infty}\varphi(k,r) = -e^{nr/2a}\left[\frac{e^{-d/2}}{2ikd^n}\,\Psi(\alpha,n+1;d)\right], \tag{2.16}$$

where Ψ is the confluent hypergeometric function of the second kind. Thus, these, too, are non-L^2 singularities!

As a check we compute the nonrelativistic limit of $f_+(k,r)$, as given by Eq. (2.9a), as $m \to \infty$, so $\alpha \to -am$. We use Tricomi's expansion[10]

$$\begin{aligned} \Phi(\alpha,\gamma;y) = \Gamma(\gamma)(-\alpha y)^{(1-\gamma)/2}\,e^{y/2}\sum_{j=0}^{\infty}A_j\left(\frac{y}{2}\right)^{j/2} \\ \times J_{\gamma-1+j}[2\sqrt{-\alpha y}] \end{aligned} \tag{2.17}$$

with

$$A_0 = 1, \quad A_1 = -\gamma/2, \quad A_2 = \gamma(\gamma+1)/2,$$

and

$$(j+1)A_{j+1} = -\tfrac{1}{2}\gamma A_j + \tfrac{1}{2}(\gamma+j-1)A_{j-1} - \tfrac{1}{4}\alpha A_{j-2},$$

and assume it is an asymptotic expansion in m. We obtain as $\alpha \to -am$, m large,

$$f_+(k,r) \to cJ_{-i2ak}(2a\sqrt{2mv_0}\,e^{-r/2a}), \tag{2.18}$$

with c independent of r. This is the same as in the Schrödinger theory.[1]

B. The Vector Coupling Model

This model has been of special interest because of its application in quantum electrodynamics. It is defined by the Lagrangian density

$$\begin{aligned} \mathfrak{L}(x) = -\varphi^{\mu*}\partial_\mu\varphi - \tfrac{1}{2}m^2\varphi^*\varphi - \tfrac{1}{2}\varphi^\mu\varphi_\mu{}^* \\ + i\varphi^{\mu*}\varphi A_\mu + \text{c.c.} \end{aligned} \tag{2.19}$$

In the fixed source limit $\mathbf{A} = 0, A^0 = \chi(r)$, this leads to the wave equation

$$\begin{aligned} \frac{E^2-m}{2m}\,\varphi(x) \\ = \left[\frac{p_r^{\,2}}{2m} + \frac{l(l+1)}{2mr^2} - \frac{1}{2m}[\chi(r)]^2 + \frac{E}{m}\,\chi(r)\right]\varphi(x). \end{aligned} \tag{2.20}$$

For

$$\chi(x) = -v_0 e^{-r/a}, \tag{2.21}$$

this equation (2.20) is exactly soluble for s waves and in fact, by the substitution of $z = e^{-r/a}$, yields again a form of Whittaker's equation. For the vector coupling model all the relativistic results for the previous model then follow but with

$$\begin{aligned} \alpha_v &= \tfrac{1}{2} - ia(E-k), \\ \gamma_v &= 1 - i2ak, \\ d_v &= -i2av_0. \end{aligned} \tag{2.22}$$

Notice that the positions in the k plane of the non-L^2 singularities are unaltered. In the nonrelativistic limit, i.e., $m \to \infty$, now $\alpha_v \to -iam$ which yields again the corresponding Schrödinger result, Eq. (2.18).

We have also examined the case of the coupling of a spin $-\frac{1}{2}$ field to an $e^{-r/b}$ type source. As above, for s wave scattering the Dirac equation is exactly soluble in terms of confluent hypergeometric functions when $j = -\frac{1}{2}$. This value is unphysical, and so we do not discuss it any further here.

3. PROPER JOST FUNCTIONS, COMPLETENESS CONDITION, AND S MATRIX FOR SEPARABLE INTERACTIONS

The effects of nonlocality on these "non-L^2 singularities" of the S matrix can be studied with an external potential of the separable type. Although this separable theory can be quantized to give a soluble quantum field theory with pairwise interactions, we do not "dress it up" but instead treat it here on the classical level where it can be clearly shown that it is the vanishing of the proper Jost function $\bar{f}_-(k,0) = C(k)f_-(k,0)$ where $C(k)$ is real, rather than $f_-(k,0)$, which is the correct criteria for the existence of non-L^2 states. The proper Jost function is the same as the Fredholm determinant of the kernel of the corresponding Lippmann–Schwinger equation for the outgoing scattering state. Notice that $C(k) = 1$ for a local central potential.[11,7] For simplicity, we discuss the s-wave case only and a discussion of the higher angular momentum is straightforward and leads to similar conclusions.

For a separable interaction the Hamiltonian is defined by

$$H = \tfrac{1}{2}\int dx[\dot{\varphi}^2 + (\nabla\varphi)^2 + m^2\varphi^2 - \lambda\rho(\mathbf{x})\varphi(\mathbf{x})\int d\mathbf{x}'\rho(\mathbf{x}')\varphi(\mathbf{x}')] \quad (3.1)$$

where λ is real and $\rho(\mathbf{x})$ spherically symmetric. In the quantized version ρ describes the nucleon density and we could set the nucleon energies equal to zero since this is a static theory. This Hamiltonian leads to the wave equation ($\mathbf{k}^2 = E^2 - m^2$)

$$\mathbf{k}^2\varphi(\mathbf{x}) = -\nabla^2\varphi(\mathbf{x}) - \int d\mathbf{x}'\rho(\mathbf{x})\rho(\mathbf{x}')\varphi(\mathbf{x}'). \quad (3.2)$$

It follows that the Wronskian

$$W(f_+(k,r), f_-(k,r))_{r=0} = W(f_+(k,r), f_-(k,r))_{r=\infty} = -2ik \quad (3.3)$$

and

$$W(f_+(k,r), f_-(k,r)) \neq 0, \quad (3.4)$$

for all r, must hold so that $f_+(k,r)$ and $f_-(k,r)$ are linearly independent for any r. In momentum space for the outgoing scattering state, the corresponding wave equation is

$$\psi^{(+)}(\mathbf{k};\mathbf{p}) = \delta(\mathbf{k}-\mathbf{p}) - \int d\mathbf{p}'\,\frac{1}{k^2 - p^2 + i\epsilon}\,\frac{2\lambda}{\pi} \times \rho(p)\rho(p')\psi^{(+)}(\mathbf{k},\mathbf{p}'), \quad (3.5)$$

where

$$\rho(p) \equiv \frac{1}{4\pi}\int d\mathbf{x}\, e^{i\mathbf{p}\cdot\mathbf{x}}\rho(\mathbf{x}) \quad (3.6)$$

and $\rho(\mathbf{x}) = \rho(|\mathbf{x}|)$, i.e., s wave interaction. The solution is

$$\psi^{(+)}(\mathbf{k};\mathbf{p}) = \delta(\mathbf{k}-\mathbf{p}) - \frac{2\lambda}{\pi}\,\frac{\rho(k)\rho(p)}{D_+(k)(k^2 - p^2 + i\epsilon)}, \quad (3.7)$$

where

$$D_+(k) = 1 + \frac{2\lambda}{\pi}\int d\mathbf{p}\,\frac{\rho(p)^2}{k^2 - p^2 + i\epsilon} \quad (3.8)$$

is the Fredholm determinant. There can be at most one bound state, $\psi^{(n)}(\mathbf{p})$, and if it exists, its characterization is given by $D(k = iK_n) = 0$ for $K_n > 0$, where $D(k)$ is the analytic continuation of $D_+(k)$ from $k + i\epsilon$ into theUHP.

The completeness of $\psi^{(+)}(\mathbf{k};\mathbf{p})$ and $\psi^{(n)}(\mathbf{p})$ is arrived at by a suitable contour integration in the UHP.[7,12] This also determines the complete set of the L^2 class of solutions of Eq. (3.2).[13] From (3.7) it is seen that only the zeros of $D_+(k)$ contribute to the completeness of the L^2 class of solutions of Eq. (3.2). Notice from (3.8) that $D(k)$ has no complex zeros in the k UHP.[12]

The S matrix for the outgoing scattering states is

$$S(k) = \int d\mathbf{p}\psi^{(-)}(\mathbf{k},\mathbf{p})^*\psi^{(+)}(\mathbf{k},\mathbf{p}) = D_-(k)/D_+(k), \quad (3.9)$$

where $\psi^{(-)}(\mathbf{k},\mathbf{p})$ is the ingoing scattering solution of (3.3) and $D_-(k) = D(k - i\epsilon)$. Notice that $D_-(k)$ appears directly in $S(k)$. Thus, by the preceeding discussion the eigenfunctions associated with the singularities of $D_-(k)$ in $S(k)$ do not appear in the completeness statement and hence by exclusion do not belong to the L^2 class of solutions of (3.2).

As for a central local potential, the Jost function $f_+(k,0)$ can be defined either by

$$\psi^{(+)}_{l=0}(k;r)'\Big|_{r=0} = \frac{k}{f_+(k,0)}\,\varphi(k;r)'\Big|_{r=0} = \frac{k}{f_+(k,0)}, \quad (3.10)$$

where $\psi^{(+)}_{l=0}(k,r)$ is the s-wave Fourier transform of $\psi^{(+)}(\mathbf{k};\mathbf{p})$ and $\varphi(k;r)$ the regular solution of (3.2), or directly by evaluation at $r = 0$ of $f_+(k,r)$ as determined by its integral equation.[14] Primes here denote differentiation. The result is the same:

$$f_\pm(k,0) = \bar{f}_\pm(k,0)/C(k) \quad (3.11)$$

with

$$\bar{f}_+(k,0) = 1 + \frac{2\lambda}{\pi}\int dp\,\frac{\rho(p)^2}{k^2 - p^2 + i\epsilon}, \quad (3.12a)$$

$$C(k) = 1 + \frac{2\lambda}{\pi}\,\mathcal{P}\int dp\,\frac{\rho(p)^2}{k^2 - p^2} + 4\pi\lambda\rho(k) \times\int_0^\infty dr\, r\,\rho(r)\cos kr, \quad (3.12b)$$

where $\bar{f}_\pm(k,0)$ are called "proper Jost functions" and $C(k)$ is real. Notice that only the proper solutions contribute to the full Green's function associated with the s wave radial equation for (3.2) which satisfies

$$\left(\frac{\partial^2}{\partial r^2} - k^2 - 4\pi\lambda\int_0^\infty dr'rr'\rho(r)\rho(r')\right)\mathcal{G}^{(+)}(k,r,r') = -\delta(r-r') \quad (3.13)$$

with the boundary conditions that it is regular at $r = 0$ and contains only outgoing waves for $r \to \infty$. Expli-

citly in terms of the proper solutions, it is

$$\mathcal{G}^{(+)}(k;r,r') = -\frac{1}{\bar{f}_+(k,0)}\left[\bar{f}_+(k,r)\varphi(k,r')\eta_+(r-r') + \bar{f}_+(k,r')\varphi(k,r)\eta_+(r'-r)\right], \quad (3.14)$$

where $\eta_+(r-r') = 1$ for $r > r'$, and is zero otherwise. Hence, by derivations of the type for the local potential only the proper Jost solutions contribute to the completeness statement and S matrix.[15]

It is useful to construct specific examples displaying the non-L^2 character of the wavefunctions associated with the singularities of $\bar{f}_-(k,0)$. We construct two examples, one where the S matrix has a non-L^2 pole of order two and, for the other, a simple non-L^2 pole.

A. Non-L^2 Pole of Order Two

Here

$$\rho(k) = b^2/(1 + k^2b^2), \quad (3.15)$$

and so $\rho(r) = e^{-r/b}/r$ which has the characteristic exponential tail at large distances. Then,

$$\bar{f}_+(k,r) = \left(1 + \frac{2\pi\lambda b^3}{1 + k^2b^2}\right)e^{ikr} - \frac{4\pi\lambda b^3}{(1 + k^2b^2}(1 + ikb)e^{-r/b} \quad (3.16)$$

$$\bar{f}_-(k,r) = \bar{f}_+(-k,r),$$

and

$$S(k) = \bar{f}_-(k,0)/\bar{f}_+(k,0),$$

where

$$\bar{f}_+(k,0) = [(1 - ikb)^2 - 2\pi\lambda b^3]/(1 - ikb)^2 \to 1 - \tfrac{1}{2}\pi\lambda b^3 \quad \text{as } kb \to i,$$

$$\bar{f}_-(k,0) = [(1 + ikb)^2 - 2\pi\lambda b^3]/(1 + ikb)^2 \to -\frac{2\pi\lambda b^3}{(1 + ikb)^2} \quad \text{as } kb \to i \quad (3.17)$$

so that the proper Jost function $\bar{f}_-(k,0)$ has a pole of order two. Defining $\tilde{f}_-(k,r) = (1 + ikb)f_-(k,r)$, $\tilde{f}_+(k,r) = \tilde{f}_-(-k,r)$ to be the solutions so that the divergent factor is removed from the $f_-(k,r)$ Jost solution, we have that the Wronskian of the two basic solutions is

$$W(\tilde{f}_+(k,r),\tilde{f}_-(k,r)) = -i2(1 + k^2b^2)\left(k - \frac{4\pi\lambda b^2}{C(k)}e^{-r/b}\sin kr\right) \quad (3.18)$$

with

$$C(k) = 1 + 2\pi\lambda b^3/(1 + k^2b^2).$$

It is equal at $r = 0$ and $r = \infty$, and vanishes for all r for $kb = i$. Thus, the starting solutions (3.16) for constructing the S matrix where $\bar{f}_+(k,r)$ and $\bar{f}_-(k,r)$ are assumed to be linearly independent are not so at $kb = i$. If one tries to solve the differential equation (3.2) at $kb = i$ for the regular solution $\varphi(k,r)$, it is found that $\varphi(k = ib^{-1}, r)$ does not exist and, hence, is not of the L^2 class of solutions. For instance, the integral equation

$$\varphi(k,r) = \frac{\sin kr}{k} - \frac{4\pi\lambda}{k}\int_0^r dr'\, r'\rho(r')\sin k(r-r')z(k) \quad (3.19)$$

with

$$z(k) = \int_0^\infty dr''\, r''\rho(r'')\varphi(k,r'') \quad (3.20)$$

leads to

$$\varphi(k,r) = -\tfrac{1}{2}b(e^{-r/b} - e^{r/b}) + 2\pi\lambda b\, z(k)[(r + \tfrac{1}{2}b)e^{-r/b} - \tfrac{1}{2}be^{r/b}]. \quad (3.21)$$

That $z(k)$ is undefined for $kb = i$ can be seen by the use of (3.21) in (3.20).

B. Simple Non-L^2 Pole

Here

$$\rho(k) = \left[c^2/(1 + k^2c^2)\right]^{1/2} \quad (3.22)$$

so that

$$\rho(r) = \frac{2}{\pi c}\frac{1}{r}K_1\left(\frac{r}{c}\right) \quad (3.23)$$

which has an exponential tail as $r \to \infty$ and a logrithmic singularity as $r \to 0$. Hence, from (3.12b) we see that $C(k)$ is divergent for all k. However, $\bar{f}_+(k,0)$ are still defined and

$$\bar{f}_+(k,0) = [(1 - ikc) - 4\pi\lambda c]/(1 - ikc) \quad (3.24)$$

so that

$$S(k) = \{(1 - ikc)[(1 - ikc) - 4\pi\lambda c]\}/\{(1 + ikc)[(1 + ikc) - 4\pi\lambda]\}. \quad (3.25)$$

A pole of $\bar{f}_-(k,0)$ is found at $kc = i$. Again from the differential equation (3.2) for (3.23), $\varphi(k = ic^{-1}, r)$ does not exist and thus belongs to the non-L^2 class of solutions.

4. THE $N\theta$ SECTOR OF THE LEE MODEL

The completeness statement and S matrix for this quantum field theory model are derived from the "state functions" by essentially identical arguments to those used in the preceding section for the pairwise theory.[12] Hence, the singularities of $\bar{f}_-(k,0) = D_-(\omega(\mathbf{k}))$ are associated with non-L^2 solutions of the corresponding wave equation in coordinate space, that is, the dynamical equation analogous to the Schrödinger equation in nonrelativistic potential theory. We construct here a specific example of this.

The Hamiltonian for the Lee model[16] is ($E_N = 0$, which is all right for a static theory)

$$H = m_0V^\dagger V + \int \omega(\mathbf{q})a^\dagger(\mathbf{q})a(\mathbf{q})d\mathbf{q} + \int f(\mathbf{q})N^\dagger Va^\dagger(\mathbf{q})d\mathbf{q} + \int f^*(\mathbf{q})V^\dagger Na(\mathbf{q})d\mathbf{q} \quad (4.1)$$

with $[N,N^\dagger] = [V,V^\dagger] = 1$, $[a(\mathbf{q})a^\dagger(\mathbf{q})] = \delta(\mathbf{q} - \mathbf{q}')$, and all others zero. $f(\mathbf{q}) = f(|\mathbf{q}|)$ and so we have only s wave interactions. Applying the Hamiltonian to the master state

$$|\Phi_\lambda\rangle \equiv \varphi_\lambda|V\rangle + \int d\mathbf{q}'\varphi_\lambda(\mathbf{q}')|N,\mathbf{q}'\rangle$$

with $|V\rangle \equiv V^\dagger|0\rangle$, $|N,\mathbf{q}'\rangle = N^\dagger a^\dagger(\mathbf{q}')|0\rangle$, we obtain the physical outgoing wave no scattering state with energy $\lambda = \omega(k)$,

$$\psi^{(+)}_{\omega(\mathbf{k})} = \begin{bmatrix} \dfrac{f^*(\mathbf{k})}{D_+(\omega(\mathbf{k}))} \\ \delta(\mathbf{q}-\mathbf{k}) + \dfrac{f^*(\mathbf{k})f(\mathbf{q})}{D_+(\omega(\mathbf{k}))[\omega(\mathbf{k})-\omega(\mathbf{q})+i\epsilon]} \end{bmatrix} \quad (4.2)$$

with

$$D(z) = [z-M]\left(1 + \int \frac{|f(\mathbf{k})|^2}{[M-\omega(\mathbf{k})][z-\omega(\mathbf{k})]}\right), \quad (4.3)$$

where M is the physical mass of the stable V particle. The S matrix determined by taking the inner product of $\psi^{(+)}_{\omega(\mathbf{k})}$ with the corresponding ingoing wave state of energy $\omega(\mathbf{k})$ is

$$S(k) = \langle \text{in}, \mathbf{k} | \text{out}, \mathbf{k} \rangle = D_-(\omega(\mathbf{k}))/D_+(\omega(\mathbf{k})). \quad (4.4)$$

Briefly then, we now take nonrelativistic kinematics, $\omega(\mathbf{k}) = \mathbf{k}^2$, and choose a particular $f(\mathbf{k})$:

$$|f(\mathbf{k})|^2 = \frac{2\lambda}{\pi}[M-\omega(\mathbf{k})]\left(\frac{b^2}{1+k^2b^2}\right)^2.$$

Thus $D(z)$, $S(k)$, and the continuum components of $\psi^{(+)}_{\omega(k)}$ are essentially the same as the corresponding quantities in first example we considered in the pairwise theory. Thus, the Fourier transform of $\psi^{(+)}_{\omega(\mathbf{k})}$ into coordinate space becomes non-L^2 as $kb \to i$, in the same way as in the previous example.

5. CONCLUSIONS

The analysis presented in this paper concerning the singularities of the S matrix demonstrates that the non-L^2 character of the poles of $\bar{f}_-(k, 0)$ persists in domains closer to that of an interacting quantum field theory. From the scalar and vector coupling models it is seen that relativization does not destroy the non-L^2 poles of the exponential potential. For the pair theory, classical and quantized, and for the Lee model it is now possible to make the definite statement that poles of the proper Jost function $\bar{f}_-(k, 0)$ correspond to non-L^2 class solutions of the associated dynamical field equation in coordinate space and that zeros of $\bar{f}_+(k, 0)$ to the L^2 class of solutions. A corollary of these conclusions is that non-L^2 states, as "shadow" states, possess inherently a very elegant reason for not appearing in the unitarity relation, and may be called "shadow" fields in complete correspondence with its nonrelativistic counterpart.

[1] C. A. Nelson, A. K. Rajagopal, and C. S. Shastry, J. Math. Phys. **12**, 737 (1971). The present paper is its sequel.

[2] S. T. Ma, Phys. Rev. **69**, 668 (1946); **71**, 195 (1947). See Ref. 1, where more recent work on such singularities are cited.

[3] R. Jost, Helv. Phys. Acta **20**, 256 (1947).

[4] S. N. Biswas, T. Pradhan, and E. C. G. Sudarshan, University of Texas at Austin preprint (September 1970).

[5] T. D. Lee and G. C. Wick, Nucl. Phys. **B9**, 209 (1969); **B10**, 1 (1969); T. D. Lee, CERN Report No. Th. 914 (unpublished). A. M. Gleeson and E. C. G. Sudarshan, Phys. Rev. D **1**, 474 (1970).

[6] E. C. G. Sudarshan, "Action at a Distance", Report to the Austin Symposium on the Past Decade in Particle Theory, April 1970; *Fields and Quanta*, Vol. II (in press). C. A. Nelson and E. C. G. Sudarshan, University of Texas at Austin, Preprint CPT-94, AEC-44 (March 1970); Austin preprint (April 1970).

[7] The relativistic wave equations (2. 3) and (2. 20) have the same differential equation structure as that of the Schrödinger equation and so the approach of Jost may be used. The definitions and notations employed here is that of R. G. Newton, *Scattering Theory of Waves and Particles*, (McGraw-Hill, New York, 1966).

[8] G. S. Guralnik and C. R. Hagen, Phys. Rev. **130**, 1259 (1963).

[9] See A. Erdélyi (Ed.), *Bateman Manuscript Project: Higher Transcendental Functions* (McGraw Hill, New York, 1953) Vol. I, Chap. VI. Volumes in this series are hereafter referred to as HTFI, II, III.

[10] HTFI-6-13-(6).

[11] R. Jost and A. Pais, Phys. Rev. **82**, 840 (1951).

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

[13] B. Friedman, *Principles and Techniques of Applied Mathematics*, (Wiley, New York, 1956).

[14] The boundary conditions of $\varphi(k; r)$ and $f_\pm(k, r)$ at $r = 0$ and $r \to \infty$, respectively, are the same as for a local potential.

[15] An alternative procedure is to introduce a proper regular solution $\bar{\varphi}(k, r) = C(k)\varphi(k, r)$ and a proper regular Green's function $\bar{\mathcal{G}}^{(+)} = C(k)\mathcal{G}^{(+)}$, i.e.,

$$\lim_{r\to\infty} e^{\mp ikr}\bar{f}_\pm(k, r) = C(k),$$

$$\lim_{r\to 0} \bar{\varphi}(k, r)' = C(k)$$

and to work with these proper solutions as for a local theory, e.g., $\psi^{(+)}(k, r)'|_{r=0} = k\bar{\varphi}(k, r)'|_{r=0}/\bar{f}_+(k, 0)$. It is permissible to modify the boundary conditions for $f_+(k, r)$ and $\varphi(k, r)$ as they are mathematical constructs; however, see the discussion in the text concerning the Wronskian. Note also C. S. Warke and R. K. Bhaduri, Nucl. Phys. **A162**, 289 (1971).

[16] T. D. Lee, Phys. Rev. **95**, 1329 (1954).

Electromagnetic Radiation in a Uniformly Moving, Homogeneous Medium

Günther Johannsen

Danish Engineering Academy, Electrical Department, The Technical University of Denmark, Aalborg, Denmark

(Received 12 July 1971)

A new method of treating radiation problems in a uniformly moving, homogeneous medium is presented. A certain transformation technique in connection with the four-dimensional Green's function method makes it possible to elaborate the Green's functions of the governing differential equations in the rest system of the medium, whereas the final integrals determining the field may be calculated in the rest system of the source.

I. INTRODUCTION

In recent years several papers have been concerned with electromagnetic radiation in a uniformly moving medium with a more or less complicated structure. The usual approach is to formulate the governing differential equation in the observer system (rest system of the source) K and then evaluate the solution for this equation.[1-11]

In two articles[12,13] the author showed that advantage may be derived from the fact that the governing differential equations are simplest in the rest frame K' of the medium, the medium being simple (i.e.,

$$\psi^{(+)}_{\omega(\mathbf{k})} = \begin{bmatrix} \dfrac{f^*(\mathbf{k})}{D_+(\omega(\mathbf{k}))} \\ \delta(\mathbf{q}-\mathbf{k}) + \dfrac{f^*(\mathbf{k})f(\mathbf{q})}{D_+(\omega(\mathbf{k}))[\omega(\mathbf{k})-\omega(\mathbf{q})+i\epsilon]} \end{bmatrix} \quad (4.2)$$

with

$$D(z) = [z-M]\left(1 + \int \frac{|f(\mathbf{k})|^2}{[M-\omega(\mathbf{k})][z-\omega(\mathbf{k})]}\right), \quad (4.3)$$

where M is the physical mass of the stable V particle. The S matrix determined by taking the inner product of $\psi^{(+)}_{\omega(\mathbf{k})}$ with the corresponding ingoing wave state of energy $\omega(\mathbf{k})$ is

$$S(k) = \langle \text{in}, \mathbf{k} | \text{out}, \mathbf{k} \rangle = D_-(\omega(\mathbf{k}))/D_+(\omega(\mathbf{k})). \quad (4.4)$$

Briefly then, we now take nonrelativistic kinematics, $\omega(\mathbf{k}) = \mathbf{k}^2$, and choose a particular $f(\mathbf{k})$:

$$|f(\mathbf{k})|^2 = \frac{2\lambda}{\pi}[M-\omega(\mathbf{k})]\left(\frac{b^2}{1+k^2b^2}\right)^2.$$

Thus $D(z)$, $S(k)$, and the continuum components of $\psi^{(+)}_{\omega(k)}$ are essentially the same as the corresponding quantities in first example we considered in the pairwise theory. Thus, the Fourier transform of $\psi^{(+)}_{\omega(\mathbf{k})}$ into coordinate space becomes non-L^2 as $kb \to i$, in the same way as in the previous example.

5. CONCLUSIONS

The analysis presented in this paper concerning the singularities of the S matrix demonstrates that the non-L^2 character of the poles of $\bar{f}_-(k, 0)$ persists in domains closer to that of an interacting quantum field theory. From the scalar and vector coupling models it is seen that relativization does not destroy the non-L^2 poles of the exponential potential. For the pair theory, classical and quantized, and for the Lee model it is now possible to make the definite statement that poles of the proper Jost function $\bar{f}_-(k, 0)$ correspond to non-L^2 class solutions of the associated dynamical field equation in coordinate space and that zeros of $\bar{f}_+(k, 0)$ to the L^2 class of solutions. A corollary of these conclusions is that non-L^2 states, as "shadow" states, possess inherently a very elegant reason for not appearing in the unitarity relation, and may be called "shadow" fields in complete correspondence with its nonrelativistic counterpart.

[1] C. A. Nelson, A. K. Rajagopal, and C. S. Shastry, J. Math. Phys. **12**, 737 (1971). The present paper is its sequel.

[2] S. T. Ma, Phys. Rev. **69**, 668 (1946); **71**, 195 (1947). See Ref. 1, where more recent work on such singularities are cited.

[3] R. Jost, Helv. Phys. Acta **20**, 256 (1947).

[4] S. N. Biswas, T. Pradhan, and E. C. G. Sudarshan, University of Texas at Austin preprint (September 1970).

[5] T. D. Lee and G. C. Wick, Nucl. Phys. **B9**, 209 (1969); **B10**, 1 (1969); T. D. Lee, CERN Report No. Th. 914 (unpublished). A. M. Gleeson and E. C. G. Sudarshan, Phys. Rev. D **1**, 474 (1970).

[6] E. C. G. Sudarshan, "Action at a Distance", Report to the Austin Symposium on the Past Decade in Particle Theory, April 1970; *Fields and Quanta*, Vol. II (in press). C. A. Nelson and E. C. G. Sudarshan, University of Texas at Austin, Preprint CPT-94, AEC-44 (March 1970); Austin preprint (April 1970).

[7] The relativistic wave equations (2. 3) and (2. 20) have the same differential equation structure as that of the Schrödinger equation and so the approach of Jost may be used. The definitions and notations employed here is that of R. G. Newton, *Scattering Theory of Waves and Particles*, (McGraw-Hill, New York, 1966).

[8] G. S. Guralnik and C. R. Hagen, Phys. Rev. **130**, 1259 (1963).

[9] See A. Erdélyi (Ed.), *Bateman Manuscript Project: Higher Transcendental Functions* (McGraw Hill, New York, 1953) Vol. I, Chap. VI. Volumes in this series are hereafter referred to as HTFI, II, III.

[10] HTFI-6-13-(6).

[11] R. Jost and A. Pais, Phys. Rev. **82**, 840 (1951).

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

[13] B. Friedman, *Principles and Techniques of Applied Mathematics*, (Wiley, New York, 1956).

[14] The boundary conditions of $\varphi(k; r)$ and $f_\pm(k, r)$ at $r = 0$ and $r \to \infty$, respectively, are the same as for a local potential.

[15] An alternative procedure is to introduce a proper regular solution $\bar{\varphi}(k, r) = C(k)\varphi(k, r)$ and a proper regular Green's function $\bar{\mathcal{G}}^{(+)} = C(k)\mathcal{G}^{(+)}$, i.e.,

$$\lim_{r\to\infty} e^{\mp ikr}\bar{f}_\pm(k, r) = C(k),$$

$$\lim_{r\to 0} \bar{\varphi}(k, r)' = C(k)$$

and to work with these proper solutions as for a local theory, e.g., $\psi^{(+)}(k, r)'|_{r=0} = k\bar{\varphi}(k, r)'|_{r=0}/\bar{f}_+(k, 0)$. It is permissible to modify the boundary conditions for $f_+(k, r)$ and $\varphi(k, r)$ as they are mathematical constructs; however, see the discussion in the text concerning the Wronskian. Note also C. S. Warke and R. K. Bhaduri, Nucl. Phys. **A162**, 289 (1971).

[16] T. D. Lee, Phys. Rev. **95**, 1329 (1954).

Electromagnetic Radiation in a Uniformly Moving, Homogeneous Medium

Günther Johannsen

Danish Engineering Academy, Electrical Department, The Technical University of Denmark, Aalborg, Denmark

(Received 12 July 1971)

A new method of treating radiation problems in a uniformly moving, homogeneous medium is presented. A certain transformation technique in connection with the four-dimensional Green's function method makes it possible to elaborate the Green's functions of the governing differential equations in the rest system of the medium, whereas the final integrals determining the field may be calculated in the rest system of the source.

I. INTRODUCTION

In recent years several papers have been concerned with electromagnetic radiation in a uniformly moving medium with a more or less complicated structure. The usual approach is to formulate the governing differential equation in the observer system (rest system of the source) K and then evaluate the solution for this equation.[1-11]

In two articles[12,13] the author showed that advantage may be derived from the fact that the governing differential equations are simplest in the rest frame K' of the medium, the medium being simple (i.e.,

homogeneous, isotropic, lossless and non dispersive) and conducting (but otherwise simple), respectively.

It is the purpose of the present paper to show that a certain transformation technique in connection with the four-dimensional Green's function method makes it possible to derive advantage from both the "simplicity" of the governing differential equation in K' and the "simplicity" of the source distribution in K; the structure of the medium may be rather complicated.

The basic idea of the method can be described in this way: The first step in the Green's function technique is to find a fundamental solution corresponding to a source which is represented by a four-dimensional δ function. Such a space–time impulse is bound to a single space–time point (world point). As far as we are concerned with a single world point, all inertial systems are equally suitable, and therefore, in the first step, the rest system of the medium should be used because the governing differential equations are simplest in this system.

The next step consists of a summation over the world points of the source distribution. This step is in principle independent of the first one, and, therefore, the most convenient inertial system should be used for the second step. Usually, the rest system of the source (if such a system exists, of course) is the most suitable one for the summation procedure.

Mathematically, the formalism provides the definition of a Green's tensor function which can be evaluated in an arbitrary inertial frame. (In K' this function may even be well known from the corresponding radiation problem, where the medium is at rest relatively to the source). This Green's function can be transformed to another arbitrary inertial frame (e.g., the rest frame K of the source), where the final integrals which represent the field may be worked out.

The general formalism is elaborated for a homogeneous medium and then applied to some examples: An isotropic medium and a lossless, electrically, uniaxially, anisotropic medium are considered. The first case has been treated before by Besieris and Compton,[9] Chen and Yen,[10] and the author,[13] the second case by S. W. Lee and Lo.[6] Case one includes a simple medium (Refs. 1, 5, 8, 12). Application of the method will also be made to a dispersive but otherwise simple medium (ionized gas), a case which has been treated before by K. S. H. Lee and Papas.[4]

II. FORMAL SOLUTION IN THE REST SYSTEM OF THE MEDIUM

In this section all quantities refer to the system of inertia K' in which the medium is at rest. The transformation of the results into four-dimensional tensor language will be done in the next section. In preparation for this transformation we will use three-dimensional, Cartesian tensor notation (cf. Ref. 14) in this section. Latin subscripts run from 1 to 4; Greek subscripts run from 1 to 3. The coordinate $x_4 \equiv ict$, where t is the time and c the speed of light in vacuum. Repeated subscripts obey the summation convention, and commas in subscripts denote partial differentiation with respect to coordinates.

The constitutive equations are given by

$$D'_\kappa = \epsilon_{\kappa\lambda} E'_\lambda, \tag{1a}$$

$$B'_\kappa = \mu_{\kappa\lambda} H'_\lambda, \tag{1b}$$

$$J^{0\prime}_\kappa = \sigma_{\kappa\lambda} E'_\lambda, \tag{1c}$$

where the dyadics $\epsilon_{\kappa\lambda}, \mu_{\kappa\lambda}$, and $\sigma_{\kappa\lambda}$ are assumed to be independent of the space–time coordinates x'_γ (i.e., the medium is homogeneous). $J^{0\prime}_\kappa$ denotes the free current density, whereas the free charge density $\rho^{0\prime}$ is assumed to be vanishing.

A vector potential A'_κ and a scalar potential $A'_4 \equiv (i/c)\Phi'$ can be introduced in the usual way. The field is given by

$$\mu_{\kappa\lambda} H'_\lambda = \epsilon_{\kappa\lambda\nu} A'_{\nu,\lambda}, \tag{2a}$$

$$E'_\kappa = ic\,(A'_{4,\kappa} - A'_{\kappa,4}), \tag{2b}$$

where $\epsilon_{\kappa\lambda\nu}$ is the three-dimensional permutation symbol, and the differential equations for the potentials can be written as

$$\epsilon_{\kappa\lambda\nu}\epsilon_{\rho\sigma\tau}\,\mu^{-1}_{\nu\tau} A_{\sigma,\lambda\rho} + c^2\epsilon_{\kappa\lambda}(A'_{4,4\lambda} - A'_{\lambda,44}) - ic\sigma_{\kappa\lambda}(A'_{4,\lambda} - A'_{\lambda,4}) = J'_\kappa, \tag{3a}$$

$$c^2\epsilon_{\lambda\nu}(A'_{\nu,\lambda 4} - A'_{4,\lambda\nu}) = J'_4. \tag{3b}$$

J'_κ is the current density of the source and $J'_4 \equiv ic\rho'$, where ρ' is the charge density of the source. Finally $\mu^{-1}_{\nu\tau}\mu_{\tau\lambda} = \delta_{\nu\lambda}$ which is the three-dimensional Kronecker symbol. [The matrices representing the constitutive dyadics in (1) are assumed to be regular.]

Introducing the definitions $n^2 \equiv c^2\mu\epsilon$ and

$$\epsilon^\star_{\kappa\lambda} \equiv \epsilon^{-1}\epsilon_{\kappa\lambda} - \delta_{\kappa\lambda}, \tag{4a}$$

$$(\mu^{-1}_{\kappa\nu})^\star \equiv \mu\mu^{-1}_{\kappa\lambda} - \delta_{\kappa\lambda}, \tag{4b}$$

$$\sigma^\star_{\kappa\lambda} \equiv \sigma^{-1}\sigma_{\kappa\lambda} - \delta_{\kappa\lambda}, \tag{4c}$$

where ϵ, μ, σ are constants (different from zero), we may write the differential equations (3) in the form

$$\epsilon_{\kappa\lambda\nu}\epsilon_{\rho\sigma\tau}(\mu^{-1}_{\nu\tau})^\star A'_{\sigma,\rho\lambda} + n^2[\epsilon^\star_{\kappa\lambda}(A'_{4,\lambda 4} - A'_{\lambda,44}) - (i/c)(\sigma/\epsilon)\sigma^\star_{\kappa\lambda}(A'_{4,\lambda} - A'_{\kappa,4})] - A'_{\kappa,\lambda\lambda} - n^2 \times [A'_{\kappa,44} - (i/c)(\sigma/\epsilon)A'_{\kappa,4}] = \mu J'_\kappa, \tag{5a}$$

$$\epsilon^\star_{\lambda\nu}(A'_{\nu,4\lambda} - A'_{4,\nu\lambda}) - A'_{4,\lambda\lambda} - n^2[A'_{4,44} - (i/c)(\sigma/\epsilon) \times A'_{4,4}] = \mu J'_4/n^2. \tag{5b}$$

In (5) we also have applied the Gauge condition

$$A'_{\lambda,\lambda} + n^2[A'_{4,4} - (i/c)(\sigma/\epsilon)A'_4] = 0 \tag{6}$$

in the usual way[15] to the isotropic terms of the equations.

Applying fourfold space–time Fourier transforms [cf. (12), Sec. III] to the differential equations (5), we obtain the matrix equation

$$(M'_{rs} + V'\delta_{rs})a'_s = s'_r, \tag{7}$$

where

$$M'_{\kappa\lambda} \equiv -\epsilon_{\kappa\sigma\nu}\epsilon_{\rho\lambda\tau}(\mu_\nu^{-1})^* k'_\rho k'_\sigma + n^2[k'^2_4\epsilon^*_{\kappa\lambda} - (i/c)(\sigma/\epsilon)k'_4\sigma^*_{\kappa\lambda}], \quad (8a)$$

$$M'_{\kappa 4} \equiv -n^2[k'_4\epsilon^*_{\kappa\lambda} - (i/c)(\sigma/\epsilon)\sigma^*_{\kappa\lambda}]k'_\lambda, \quad (8b)$$

$$M'_{4\kappa} \equiv -\epsilon^*_{\lambda\kappa}k'_\lambda k'_4, \quad (8c)$$

$$M'_{44} \equiv +\epsilon^*_{\kappa\lambda}k'_\kappa k'_\lambda, \quad (8d)$$

$$V' \equiv k'_\lambda k'_\lambda + n^2[k'^2_4 - (1/c)(\sigma/\epsilon)k'_4], \quad (8e)$$

and

$$s'_k \equiv \mu(j'_\kappa, j'_4/n^2). \quad (9)$$

It is noticed that M'_{rs} vanishes when the medium is isotropic.

Defining

$$g'_{rs} \equiv (M'_{rs} + V'\delta_{rs})^{-1}, \quad (10)$$

we finally get

$$a'_r = g'_{rt}s'_t \quad (11)$$

which establishes the formal solution of the problem in the transformed domain.

III. TRANSLATION INTO FOUR-DIMENSIONAL TENSOR FORMALISM

Consider a tensor field $T\ldots(x_r)$ of arbitrary order (we use four-dimensional tensors as in Ref. 12) and a fourfold Fourier integral

$$T\ldots(x_r) = \int_{-\infty}^{\infty}\int_{-i\infty}^{i\infty} t\ldots(k_r)e^{ik_s x_s}d(k_r) = \int_{M_4} t\ldots(k_r)e^{ik_s x_s}d(k_r), \quad (12)$$

where the integration is to be taken over the whole Minkowski 4-space M_4.

A proper Lorentz transformation can be applied to the integration variables k_r as well as to the space-time variables x_r. The inner tensor product $k_r x_r$ is invariant, and we still have to integrate over the whole space M_4 because a Lorentz transformation is a one-to-one mapping of M_4 on itself. Therefore, the Fourier amplitudes $t\ldots$ transform like a tensor if $T\ldots$ does, and vice versa (because the transformation matrix does not depend on x_r).

The four-current density J_r transforms like a tensor (cf. Ref. 16) and so does the four-potential A_r (cf. below). Therefore, the Fourier amplitudes j_r and a_r are tensors of order one (4-vectors).

We make also use of the 4-vector

$$S_r \equiv S_{rs}J_s, \quad S_{rs} \equiv \mu(\delta_{rs} + \kappa/n^2 U_s U_r), \quad (13)$$

where $\kappa \equiv (n^2 - 1)/c^2$ and U_r is the 4-velocity, i.e., $U'_r = (0, 0, 0, ic)$ and $S'_r = (J'_\rho, J'_4/n^2)$. It is seen that the Fourier amplitudes of S'_r are given by (9).

Furthermore, we define a second-order tensor g_{rs} by means of the components in K' as given by (10). The tensor character of Eq. (11) now ensures that

$$a_r = g_{rt}s_t \quad (14)$$

is valid in an arbitrary inertial frame, and, referring to the statements made in the beginning of this section, we are able to give a covariant, formal solution of the problem:

$$A_n(x_r) = \frac{1}{(2\pi)^4}\int_{M_4} G_{ns}(u_r)J_s(z_r)d(z_r), \quad (15)$$

where

$$G_{ns}(u_r) \equiv \int_{M_4} g_{nm}(k_r)S_{ms}e^{ik_t u_t}d(k_r), \quad u_t \equiv x_t - z_t. \quad (16)$$

For a nondispersive medium, n (and therefore also S_{rs}) is independent of the variables k_r. In that case we use a modified definition of the Green's tensor function:

$$A_n(x_r) = \frac{1}{(2\pi)^4}\int_{M_4} G_{ns}(u_r)S_s(z_r)d(z_r), \quad (15')$$

where

$$G_{ns}(u_r) \equiv \int_{M_4} g_{ns}(k_r)e^{ik_t u_t}d(k_r). \quad (16')$$

Since the Green's tensor function (16) is simplest in K' (which is a consequence of the fact that the governing differential equations are simplest in K'), an evaluation of the integral in (16) should be sought (may be well known) in the rest frame of the medium (cf. the following sections).

In Minkowski's theory the field quantities $E_\kappa, B_\kappa, H_\kappa, D_\kappa$ transform in a certain manner from one inertial frame to another, which is called the Lorentz transformation of the field. This can be expressed very briefly by saying that the field quantities constitute the components of two second-order field tensors F_{rs} and H_{rs} (consult Ref. 17 for the definition).

Define a 4-vector A_r in such a way that the A'_r are identical with the potentials entering Eqs. (2), then the tensor equation

$$F_{rs} = c(A_{s,r} - A_{r,s}) \quad (17)$$

is valid because (17) is equivalent to (2) in the rest system K', which in turn shows that the primes may be omitted in (2b), whereas the spatial part of (17) is equivalent to the equation $B_\kappa = \epsilon_{\kappa\lambda\nu}A_{\nu,\lambda}$.

In order to express the field quantities H_κ and D_κ by means of the potential, this can only be done via the constitutive relations. Therefore, it might be convenient to have covariant forms of these relations to our disposal.

Following Marx[18] in a slightly deviating way, we define

$$F_r \equiv (1/c)F_{rs}U_s, \quad (18a)$$

$$\tilde{F}_r \equiv (1/ic^2)F^*_{rs}U_s, \quad (18b)$$

$$K_r \equiv (1/c^2)H_{rs}U_s, \quad (18c)$$

$$\tilde{K}_r \equiv (1/ic)H^*_{rs}U_s, \quad (18d)$$

where F^*_{rs}, H^*_{rs} are dual tensors, i.e., $F^*_{rs} \equiv (1/2!)\epsilon_{rsmn}F_{mn}$, say. It is readily seen that $F'_r = (E'_\rho, 0)$, $\tilde{F}'_r = (B'_\rho, 0)$, $K'_r = (D'_\rho, 0)$, and $\tilde{K}'_r = (H'_\rho, 0)$.

By defining tensors e_{rs}, m_{rs} by means of $e'_{\rho\sigma} \equiv \epsilon_{\rho\sigma}$, $m'_{\rho\sigma} \equiv \mu_{\rho\sigma}$, $e'_{4\sigma} \equiv m'_{4\sigma} \equiv 0$ (the remaining components can be chosen arbitrarily), it follows from (1) that

$$K_r = e_{rs} F_s, \tag{19a}$$

$$F_r = m_{rs} \tilde{K}_s. \tag{19b}$$

Using tensor relations inverse to (18) (expressing second-order field tensors in terms of first order field tensors), one can express H_{rs} in terms of F_{rs}, say (cf. Ref. 18). Marx[18] proposed that the components e'_{r4}, m'_{r4} are vanishing. In order to get regular matrices, we choose $e'_{\rho 4} \equiv m'_{\rho 4} \equiv 0$, $\epsilon'_{44} \equiv \epsilon$, and $m'_{44} \equiv \mu$.

If wanted, even a covariant, four-dimensional formulation of the "wave equations" can be obtained from (5) by means of a simple translation technique. (The usual way is to introduce 4-potentials into the 4-tensor formulation of Maxwell–Minkowski equations, as was done by Viglin[7] and earlier by Jauch and Watson[19] for the isotropic case.) Defining m^{-1}_{rs} such that $m^{-1}_{rs} m_{st} = \delta_{rt}$, we obviously have $(m^{-1}_{\rho\sigma})' = \mu^{-1}_{\rho\sigma}$, $(m^{-1}_{44})' = 1/\mu$. Furthermore, the definitions $e^{\star}_{rs} \equiv (1/\epsilon) e_{rs} - \delta_{rs}$, $(m^{-1}_{rs})^{\star} \equiv \mu (m^{-1}_{rs}) - \delta_{rs}$ imply that the spatial parts of these tensors in K' are given By (4) and that the remaining components in K' are equal to zero.

Consider the first term of Eq. (5a); we get

$$\epsilon_{\kappa\lambda\nu}\epsilon_{\rho\sigma\tau}(\mu^{-1}_{\nu\tau})^{\star} A'_{\sigma,\rho\lambda} = \epsilon_{\kappa l n 4}\epsilon_{rst4}(m^{-1}_{nt})^{\star\prime} A'_{s,rl} = -(1/c^2)\epsilon_{\kappa l n u}\epsilon_{rstv}(m^{-1}_{nt})^{\star\prime} A'_{s,rl} U'_u U'_v.$$

It is observed that the "fourth component" (κ is replaced by the integer 4 in the final expression) vanishes.

The second term in Eq. (5a) is translated in this way:

$$n^2 \epsilon^{\star}_{\kappa\lambda}(A'_{4,\lambda 4} - A'_{\lambda,44}) = -\frac{n^2}{c^2} e^{\star\prime}_{\kappa l}(A'_{r,ls} - A'_{l,rs}) U'_r U'_s.$$

Again, it is noticed that the fourth component of the final expression vanishes.

In order to translate the third term of (5a), we define tensors s_{rs} and $s^{\star}_{rs}$ by means of $s'_{\rho\sigma} \equiv \sigma_{\rho\sigma}$, $s'_{4\rho} \equiv s'_{\rho 4} \equiv 0$, $s'_{44} \equiv \sigma$, $s^{\star}_{rs} \equiv s_{rs}/\sigma - \delta_{rs}$. Since $s^{\star\prime}_{\rho\sigma} = \sigma^{\star}_{\kappa\lambda}$ [cf. (4)], we conclude that

$$-(i\sigma/c\epsilon)\sigma^{\star}_{\kappa\lambda}(A'_{4,\lambda} - A'_{\lambda,4}) = -(\sigma/c^2\epsilon) s^{\star}_{\kappa l}(A'_{ul} - A'_{lu}) U'_u$$

and that the fourth component of the final expression is zero.

The first term of (5b) is equal to the fourth component of $-(1/c^2)\, e^{\star\prime}_{ln}(A'_{n,rl} - A'_{r,nl}) U'_r U'_k$, whereas the first three components are equal to zero.

It is not difficult to translate the remaining, isotropic terms of (5a) and (5b) into the four-dimensional form of the isotropic wave equation.

Omitting primes, we finally get a 4-tensor formulation of Eq. (5):

$$(1/c^2)\{[\epsilon_{klnu}\epsilon_{rstv}(m^{-1}_{nt})^{\star} A_{s,rl} + n^2 e^{\star}_{kl}(A_{u,lv} - A_{l,uv})] U_u V_v - (\sigma/\epsilon) s^{\star}_{kl}(A_{u,l} - A_{l,u}) U_u + e^{\star}_{ln}(A_{n,rl} - A_{r,nl}) U_r U_k\} + A_{k,nn} - [(n^2 - 1)/c^2] A_{k,rs} U_r U_s = -S_k. \tag{20}$$

Finally, we want to make some remarks on dispersive media. In K' the constitutive parameters of a dispersive medium are assumed to be functions of the frequency $\omega' = ck'_4/i$ corresponding to a Fourier component of the field.

Since the formalism provides a decomposition of the field into Fourier components in all inertial frames, we may define a "frequency" referring to an arbitrary inertial system K:

$$\omega \equiv ck_4/i. \tag{21}$$

In order to get covariant forms for the functions which express the dependency of the constitutive parameters on ω', we observe that

$$\omega' = -k'_r U'_r = -k_r U_r. \tag{22}$$

Clearly, (22) shows that in K (different from K') the constitutive parameters depend not only on k_4 (i.e., ω) but also on the spatial part k_ρ of the "propagation tensor" k_r.

IV. ISOTROPIC, NONDISPERSIVE MEDIUM

For an isotropic, nondispersive medium it is seen from (4), (8), and (10) that $\epsilon^{\star}_{\kappa\lambda} = 0$, etc., $M'_{rs} = 0$, and $g'_{rs} = (1/V')\delta_{rs}$. Defining a Green's function $G'(u'_t)$ such that $G'_{rs}(u'_t) = G'(u'_t)\delta_{rs}$ we find that this function is given by the integral [cf. (16a)]

$$G'(u'_r) = \int_{M_4} \frac{1}{V'(k'_r)} e^{ik'_t u'_t} d(k'_r), \tag{23}$$

where $V'(k'_r)$ is given by (8e).

The substitution $k'_4 + il'_4 \to k'_4$, $l'_4 \equiv (i/2c)(\sigma/\epsilon)$ leads to

$$G'(u'_r) = e^{il'_4 u'_4} \iiint_{-\infty}^{\infty} \int_{\alpha - i\infty}^{\alpha + i\infty} \times \frac{e^{ik'_t u'_t}}{k'_\lambda k'_\lambda + n^2(k'^2_4 + l'^2_4)} d(k'_r), \alpha \equiv -|l'_4|. \tag{24}$$

The integral in (21) is the Fourier integral of the time-dependent Green's function of the Klein–Gordon differential equation. Since G is a tensor of zero order (Lorentz invariant) the Lorentz transformation of this well-known function[20] to an arbitrary inertial frame can be achieved by translating the dependency of this function on the variables u_r into covariant tensor language, as was done by the author in Ref. 13. This is easier than evaluating the integrals in the covariant edition of (24) which is given by

$$G(u_r) = e^{l_t u_t} \int_{M_4} \frac{e^{ik_t u_t}}{k_t k_t - \kappa(k_t U_t)^2 - [(n/c) l_t U_t]^2} d(k_r), \tag{25}$$

where we have defined a 4-vector l_r such that $l'_r = (0, 0, 0, (i/2c)(\sigma/\epsilon))$.

The "difficult" way of evaluating (25) was used by Chen and Yen[10] and, for the lossless case, by Compton.[5] Earlier K.S.H. Lee and Papas[1] evaluated these integrals for a time-harmonic source in a simple medium. [Harmonic time dependency implies that the k_4-integration in (25) "disappears". This is readily seen from (15a) in connection with (25):

Assuming that S_s in (15a) is time-harmonic, the z_4-integration introduces a time-dependent δ function in (24)]

V. IONIZED GAS

Consider a medium which is dispersive but otherwise simple (i.e., isotropic and lossless). Furthermore, assume that $\mu = \mu_0$ and

$$\epsilon(\omega') = \epsilon(-k_r U_r) = \epsilon_0[1 + (ck_{(p)}/k_r U_r)^2], \quad (26)$$

where μ_0, ϵ_0 are the permeability and the dielectric constant of vacuum, $k_{(p)} \equiv i\omega_{(p)}/c$, and $\omega_{(p)}$ is a real number.

Following Ref. 4, we wish to evaluate the Green's function for the unbounded space in the form of a one-dimensional integral, but our approach is different from that in Ref. 4: A well-known result referring to the rest frame K' of the medium is transformed to K. In the case which is under consideration Eq. (16) reduces to

$$G'_{ns}(u'_t) = \mu_0 \int_{M_4} \frac{\delta_{ns} - \kappa/n^2 U'_n U'_s}{k'_\lambda k'_\lambda + n^2 k'^2_4} e^{ik'_t u'_t} d(k'_r), \quad (27)$$

where we have made use of (13).

It is well known from the corresponding radiation problem in a nonmoving and nondispersive medium that the threefold space integral in (27) can be reduced to a single integral (of course, this result can be obtained by exactly the same steps which are used in Ref. 4 but with the simplification that v is zero).

We have

$$G'_{ns}(u'_t) = i\pi^2\mu_0 \int_{-i\infty}^{i\infty} \int_{\infty}^{\infty} + [\delta_{ns} + (\kappa/n^2) U'_n U'_s] H_0^1[\rho' \sqrt{-(k'^2_3 + n^2 k'^2_4)}] \times e^{i(k'_3 u'_3 + k'_4 u'_4)} dk'_3 dk'_4, \quad (28)$$

where $\rho' \equiv (u'^2_1 + u'^2_2)$ and H_0^1 is the Hankel function of zero order and first kind.

By means of (26), the argument of the Hankel function can be written as $\rho'[k^2_{(p)} - (k'^2_3 + k'^2_4)]^{1/2}$.

As in Ref. 4 we choose (without loss of generality) the velocity of K relative to K' to be directed along the $x_3(x'_3)$ axis. This implies that the first and the second component of a 4-vector are uneffected by the Lorentz transformation so that $\rho' = \rho$. Furthermore, $k'^2_3 + k'^2_4 = k^2_3 + k^2_4$ and $k'_3 u'_3 + k'_4 u'_4 = k_3 u_3 + k_4 u_4$ (because these expressions are tensors of zero order, i.e., invariants).

The transformation of (27) to an arbitrary inertial frame K is now almost trivial. Eliminating κ/n^2 by means of (26), we obtain

$$G_{ns}(u_t) = i\pi^2\mu_0 \int_{-i\infty}^{i\infty} \int_{\infty}^{\infty} \left(\delta_{ns} + \frac{k^2_{(p)} U_n U_s}{c^2 k^2_{(p)} + (k_r U_r)^2}\right) \times H_0^1[\rho \sqrt{k^2_{(p)} - (k^2_3 + k^2_4)}] e^{i(k_3 u_3 + k_4 u_4)} dk_3 dk_4. \quad (29)$$

The first term of this expression (containing δ_{ns}) may be integrated with respect to k_3, leading to the result of Ref. 4.

On the other hand, it can be done in this way: Since $k'_\lambda k'_\lambda + n^2 k'^2_4 = k'_r k'_r - k^2_{(p)} = k_r k_r - k^2_{(p)}$, the first term of (27) can be written as

$$\mu_0 \delta_{ns} \int_{-i\infty}^{i\infty} e^{ik_4 u_4} \iiint_{-\infty}^{\infty} \frac{e^{ik_\tau u_\tau}}{k_\nu k_\nu - (k^2_{(p)} - k^2_4)} d(k_\rho) dk_4.$$

This expression is covariant (!), and the spatial integral of this term is well known (from the simple wave equation) and is equal to $\exp[ir(k^2_{(p)} - k^2_4)^{1/2}]/4\pi r$ in agreement with Ref. 4.[21]

VI. LOSSLESS, ELECTRICALLY, UNIAXIALLY ANISOTROPIC MEDIUM

Consider a medium for which $\sigma, \sigma^*_{\kappa\lambda}$, and $(\mu^{-1}_{\kappa\lambda})$ are vanishing and

$$\epsilon^*_{\kappa\lambda} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \tau - 1 \end{pmatrix}, \quad (30)$$

where τ is a given positive constant [cf. (4)].

M_{kl} as given by (8) reduces considerably, and g'_{rs} is readily found from (10). We get

$$g'_{rs}(k'_t) = \begin{pmatrix} 1/V' & & 0 & 0 \\ 0 & 1/V' & 0 & 0 \\ 0 & 0 & \dfrac{X'}{V'Z'} & \dfrac{(\tau-1)n^2 k'_3 k'_4}{V'Z'} \\ 0 & 0 & \dfrac{(\tau-1)k'_3 k'_4}{V'Z'} & \dfrac{Y'}{V'Z'} \end{pmatrix}, \quad (31)$$

where V' is given by (8e) and

$$X' \equiv k'^2_1 + k'^2_2 + \tau k'^2_3 + (nk'_4)^2, \quad (32a)$$

$$Y' \equiv k'^2_1 + k'^2_2 + k'^2_3 + \tau(nk'_4)^2, \quad (32b)$$

$$Z' \equiv k'^2_1 + k'^2_2 + \tau k'^2_3 + \tau(nk'_4)^2. \quad (32c)$$

If we restrict the discussion to inertial frames moving on the axis of symmetry (i.e., the x'_3 direction) it is consistent (i.e., independent of the inertial frame) to define two 4-vectors s^o_r and s^e_r in such a way that $s_r = s^o_r + s^e_r$ and $s^o_3 = s^o_4 = s^e_1 = s^e_2 = 0$.

Case I: Consider a source for which $s^e_r = 0$, e.g., a longitudinally orientated current loop (magnetic dipole). Obviously, $s_1 = s'_1$, $s_2 = s'_2$, and it is immediately seen from (31) that the only mode which is excited is identical with the isotropic case. This mode is called the ordinary or the first mode (cf. Ref. 11).

Case II: Next, consider a source for which $s^o_r = 0$, e.g., a longitudinally orientated, electrical dipole. Using the continuity equation $s'_3 k'_3 + s'_4 n^2 k'_4 = 0$, we derive, by means of (14) and (31), the equation

$$g_{rt}(k_n) s_t = g^e(k_n) \delta_{rt} s_t, \quad g^e(k_n) \equiv 1/Z(k_n). \quad (33)$$

Obviously, only one mode is excited. The mode is of a

different form than in Case I and is called the extraordinary or second mode (cf. Ref. 11).

It is not difficult to find the corresponding extraordinary Green's function G^e from the ordinary Green's function $G^o \equiv G$ as given by (23), (24) (with $\sigma = 0$).

The connection between the Green's functions is given by

$$G^e(u_r) = \tau^{-1} G^o(u_1, u_2, u_3/\sqrt{\tau}, u_4/\sqrt{\tau}). \tag{34}$$

In K', (34) is verified just by inspection of the corresponding fourfold Fourier integrals, and it is easily seen that the relation is covariant under those special Lorentz transformations which we are dealing with in this section.

$G(u_r)$ may be found in Ref. 13. Applying the extraordinary Green's function $G^e(u_r)$ to the case of a time-harmonic source, we readily derive the extraordinary Green's function as given by Lee and Lo.[6]

Case III: Finally, consider an arbitrary source distribution. From the continuity equation $s'_\rho k'_\rho + n^2 s'_4 k'_4 = 0$ in connection with (14) and (31), we get

$$a'_r = \tilde{g}'_{rt} s'_t, \tag{35}$$

where we have defined a tensor $\tilde{g}_{rt}$ by means of the components in K':

$$\tilde{g}'_{rt} \equiv \begin{pmatrix} 1/V' & 0 & 0 & 0 \\ 0 & 1/V' & 0 & 0 \\ 0 & 0 & 1/Z' & 0 \\ 0 & 0 & 0 & 1/Z' \end{pmatrix}$$

$$+ \frac{1-\tau}{V'Z'} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ k'_1 k'_3 & k'_2 k'_3 & 0 & 0 \\ k'_1 k'_4 & k'_2 k'_4 & 0 & 0 \end{pmatrix} \tag{36}$$

The contribution to the Green's function $\tilde{G}'_{rt}$ corresponding to the first matrix term in (36) can immediately be obtained from Cases I and II. As to the second matrix term, the evaluation of the Fourier integral is more complicated. But, if $\tilde{G}'_{rt}$ can be found in K', the transformation to another inertial system K is simple. In fact, the second matrix in (27) is not affected at all by Lorentz transformations between inertial frames moving on the axis of symmetry. (Of course, the way in which the matrix depends on k_r is affected.)

Though there is no need for working in K at all (as was done in Ref. 6), it is easily verified (in K') that

$$V = k_r k_r + (1 - n^2)(k_r U_r/c)^2, \tag{37a}$$

$$Z = k_r k_r + (\tau - 1)(k_r T_r)^2 + (1 - \tau n^2)(k_r U_r/c)^2, \tag{37b}$$

where we have defined a 4-vector T_r such that $T'_r = (0, 0, 1, 0)$.

If the frequency ω is introduced by means of (21), the equations $V = 0$ and $Z = 0$ represent surfaces called the ordinary and extraordinary dispersion surface, respectively. These surfaces are of fundamental significance when asymptotic expressions of the Fourier integrals are to be evaluated (as was done in Ref. 6), but this can more easily be done in K' (cf. Refs. 22, 23), whereupon the result can be translated to K as outlined here.

[1] K. S. H. Lee and C. H. Papas, J. Math. Phys. 5, 1668 (1964).
[2] C. T. Tai, Trans. IEEE Antennas Propagation 13, 322 (1965).
[3] R. T. Compton, Jr. and C. T. Tai, Trans. IEEE Antennas Propagation 13, 574 (1965).
[4] K. S. H. Lee and C. H. Papas, Trans. IEEE Antennas Propagation 13, 799 (1965).
[5] R. T. Compton, J. Math. Phys. 7, 2145 (1966).
[6] S. W. Lee and Y. T. Lo, Radio Sci. 1, 313 (1966).
[7] A. S. Viglin, Zh. Eksp. Teor. Fiz. 50, 85 (1966) [Sov. Phys. JETP 23, 59 (1966)].
[8] C. T. Tai, J. Math. Phys. 8, 646 (1967).
[9] I. M. Besieris and R. T. Compton, Jr., J. Math. Phys. 8, 2445 (1967).
[10] K. C. Chen and J. L. Yen, J. Math. Phys. 9, 2081 (1968).
[11] I. M. Besieris, J. Math. Phys. 10, 1156 (1969).
[12] Günther Johannsen, J. Math. Phys. 11, 3251 (1970).
[13] Günther Johannsen, J. Math. Phys. 12, 1349 (1971).
[14] J. L. Synge and A. Schild, *Tensor Calculus* (University of Toronto Press, Toronto, 1966).
[15] D. S. Jones, *The Theory of Electromagnetism* (Macmillan, New York, 1964), pp. 14–15.
[16] C. Møller, *The Theory of Relativity* (Oxford U.P., London, 1952), pp. 196–98.
[17] A. Sommerfeld, *Electrodynamics* (Academic, New York, 1952), pp. 216–17.
[18] G. Marx, Acta Phys. Acad. Sci. Hung. 3, 75 (1953).
[19] J. M. Jauch and K. M. Watson, Phys. Rev. 74, 951 (1948).
[20] P. M. Porse and H. Feshback, *Methods of Theoretical Physics* (McGraw-Hill, New York) Vol. I., p. 856.
[21] Attention is drawn to the fact that in K' Eq. (26) (i.e., $\epsilon/\epsilon_0 = 1 - \omega'^2_{(p)}/\omega'^2$) is assumed to hold for all ω'. This is true even if we assume harmonic time dependency in K (as was done in Ref. 4), because a spatial Fourier transformation in K does not correspond to a pure spatial Fourier transformation in K'.
[22] G. A. Deschamps and O. Kesler, Trans. IEEE Antennas Propagation 11, 783 (1964).
[23] M. J. Lighthill, Phil. Trans. Roy. Soc. London, Ser. A, 252, 397 (1960).

Structure of the Wigner 9j Coefficients in the Bargmann Approach

Alfred C. T. Wu
Department of Physics, University of Michigan, Ann Arbor, Michigan 48104
(Received 19 June 1971)

Bargmann's treatment of the Clebsch–Gordan (3j) and Racah (6j) coefficients is here extended to the case of Wigner 9j coefficients. The generating function for the 9j coefficient is computed by the analytic method. The result is compared to the Schwinger's expression derived with the algebraic (boson operator) method. The full symmetry of the Wigner 9j coefficients is manifest and transparent in the Bargmann's formalism. A new explicit expression for the Wigner 9j coefficient is derived as a sixfold sum which may be regarded as the analog of the Racah's formula for the Racah coefficient.

I. INTRODUCTION

The practical implication in the use of the rotation group machinery in handling the composition of angular momentum states obviously warrants a detailed analysis of the coupling and recoupling coefficients such as the Clebsch–Gordan, Racah, and Wigner 9j coefficients. Bargmann's beautiful exposition of the representations of the rotation group[1] in the Wigner *Festschrift* issue of the *Reviews of Modern Physics* (October, 1962) is presumably well known. It is quite remarkable that results on the representations and the coupling and recoupling coefficients all come out in such a coherent, transparent and effortless way. We have in mind here, in particular, the following features of the coupling and recoupling coefficients: (i) derivation of the generating functions; (ii) symmetry properties; and (iii) explicit summation expressions. For the Clebsch–Gordan (3j) and Racah (6j) coefficients, these features are all worked out in detail in the Bargmann article. For the Wigner 9j coefficients,[2] the standard reference for the generating function is the celebrated unpublished report by Schwinger,[3] where everything is obtained by the algebraic (boson operator) method. The 72-element symmetry for the 9j symbol is discussed by Jahn and Hope.[4] For an easy access to many classic references as well as a highly readable introductory account of the quantum theory of angular momentum, the reader is referred to the volume edited by Biedenharn and Van Dam.[5]

It would seem desirable, at least for methodological interest, to push the Bargmann scheme to treat the case of the Wigner 9j coefficients. Indeed we find that the analytic approach is sufficiently powerful to render the treatment feasible.

Thus the generating function for the 9j symbol is computed (Sec. II). The result is essentially equivalent to that of Schwinger.[6] In Sec. III, the 9j coefficient is extracted by expansion from the generating function. In Sec. IV, the symmetry of the 9j coefficient is easily read off in the Bargmann formalism. Since the Regge symmetry,[7] which boosts the previously known symmetry of Clebsch–Gordan and Racah coefficients by a factor of six (from 12 to 72 elements for the former and from 24 to 144 elements for the latter), is entirely contained in the generating function approach,[8] there does not seem to be an obvious addition to the known 72 element symmetry for the 9j symbols. The fact that the 9j coefficients possess less symmetry than the Racah coefficients may be due to the tighter structure of the former.

In Sec. V, the fifteenfold sum expression for the 9j coefficient in Sec. III is reduced, after proper consideration of the constraint conditions, to a sixfold sum. This new explicit form [Eq. (65) or (65a) below] which manifests full symmetry of the 9j coefficient may be regarded as the analog of the well-known single-sum expression for the Racah coefficient. As one of the consistency checks, it is explicitly verified that when one of the 9j values approaches zero, the sixfold sum for the 9j coefficient collapses to a single sum for the Racah coefficient with the correct phase as well as normalization factors. Finally, the following question is posed. Inasmuch as the Clebsch–Gordan coefficient may be regarded as a ${}_3F_2$ function at $x=-1$, and the Racah coefficient as a Saalschutzian ${}_4F_3$ function at $x=1$, the seeming inference that the Wigner 9j coefficient might also belong to some ${}_pF_q$ function turns out to be unwarranted.

II. GENERATING FUNCTION FOR THE WIGNER 9j COEFFICIENT

As is well known, the Wigner 9j symbol is the recoupling coefficient which connects two different schemes of adding up four angular momenta (such as from ls to jj couplings).[9] In parallel to Bargmann's treatment of the Racah coefficient where the 6j symbol is defined through a sixfold sum of the product of four 3j symbols,[10] our starting point here is that the Wigner 9j symbol is defined through a ninefold sum of the product of six 3j symbols[11]:

$$\begin{Bmatrix} j_{11} & j_{12} & j_{13} \\ j_{21} & j_{22} & j_{23} \\ j_{31} & j_{32} & j_{33} \end{Bmatrix} = \sum_{\text{all } m_{ij}} \begin{pmatrix} j_{11} & j_{12} & j_{13} \\ m_{11} & m_{12} & m_{13} \end{pmatrix} \begin{pmatrix} j_{21} & j_{22} & j_{23} \\ m_{21} & m_{22} & m_{23} \end{pmatrix} \begin{pmatrix} j_{31} & j_{32} & j_{33} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \times \begin{pmatrix} m_{11} & m_{21} & m_{31} \\ j_{11} & j_{21} & j_{31} \end{pmatrix} \begin{pmatrix} m_{12} & m_{22} & m_{32} \\ j_{12} & j_{22} & j_{32} \end{pmatrix} \begin{pmatrix} m_{13} & m_{23} & m_{33} \\ j_{13} & j_{23} & j_{33} \end{pmatrix} . \tag{1}$$

It is clear that there are six triangular relations, one for each triplets of j belonging to the rows and columns on the left-hand side of Eq. (1). The following short-hand notations are found convenient:

$$J_p \equiv \sum_{q=1}^{3} j_{pq}, \qquad p = 1, 2, 3, \tag{2a}$$

$$K_p \equiv \sum_{q=1}^{3} j_{qp}, \tag{2b}$$

$$k_{pq} \equiv J_p - 2j_{pq}, \quad p, q = 1, 2, 3, \tag{3a}$$

$$k'_{pq} \equiv K_q - 2j_{pq}. \tag{3b}$$

There are a number of constraints on the 18 k and k', namely

$$\sum_q k_{pq} = J_p, \tag{4a}$$

$$\sum_p k'_{pq} = K_q, \tag{4b}$$

$$k_{pq} - k'_{pq} = J_p - K_q, \tag{5}$$

$$\sum_p J_p = \sum_p K_p \equiv J. \tag{6}$$

The generating function $\rho(\tau, \tau')$ for the 9j symbol will be written as

$$\rho(\tau, \tau') = N \sum_{k_{pq}} \sum_{k'_{pq}} \{9j\} \prod_{p,q} \tau_{pq}^{k_{pq}} \tau'^{k'_{pq}}_{pq}, \tag{7}$$

where the normalization constant

$$N \equiv \left[\frac{\prod_p (J_p + 1)!\ (K_p + 1)!}{\prod_{p,q} k_{pq}!\ k'_{pq}!} \right]^{1/2} \tag{8}$$

comes from the six sets of triangular coefficients. On account of Eq. (1), the generating function can be readily cast into the integral representation

$$\rho(\tau, \tau') = \int d\mu_{18}(\zeta_{pq}) \prod_{p=1}^{3} \Phi_p \prod_{q=1}^{3} \tilde{\Phi}_q$$
$$= \int d\mu_{18}(\zeta_{pq}) \exp \sum_{p=1}^{3} (D_p + \tilde{D}_p), \tag{9}$$

where the six Φ's are the generating functions for the 3j symbols, one for each triplet in Eq. (1). Explicitly, we have[12] ($p, q = 1, 2, 3$)

$$\Phi_p \equiv \Phi(\tau_{p1}, \tau_{p2}, \tau_{p3}; \zeta_{p1}, \zeta_{p2}, \zeta_{p3}) = \exp D_p, \tag{10a}$$

$$\tilde{\Phi}_q \equiv \Phi(\tau'_{1q}, \tau'_{2q}, \tau'_{3q}; \Gamma^T \bar{\zeta}_{1q}, \Gamma^T \bar{\zeta}_{2q}, \Gamma^T \bar{\zeta}_{3q}) = \exp \tilde{D}_q, \tag{10b}$$

where ζ_{pq} is a shorthand notation for the two-component complex

$$\zeta_{pq} \equiv \begin{pmatrix} \xi_{pq} \\ \eta_{pq} \end{pmatrix}.$$

Recall that the ξ and η are the variables that enter in the basis function

$$v_m^j(\zeta) = \frac{\xi^{j+m} \eta^{j-m}}{\sqrt{(j+m)!\,(j-m)!}}. \tag{11}$$

In Eq. (10b), Γ^T is the transpose of $\Gamma \equiv \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, and $\bar{\zeta}$ is the complex conjugate of ζ. The D are 3 × 3 determinants, namely

$$D_p = \begin{vmatrix} \tau_{p1} & \tau_{p2} & \tau_{p3} \\ \xi_{p1} & \xi_{p2} & \xi_{p3} \\ \eta_{p1} & \eta_{p2} & \eta_{p3} \end{vmatrix}, \tag{12a}$$

$$\tilde{D}_q = \begin{vmatrix} \tau'_{1q} & \tau'_{2q} & \tau'_{3q} \\ \bar{\eta}_{1q} & \bar{\eta}_{2q} & \bar{\eta}_{3q} \\ -\bar{\xi}_{1q} & -\bar{\xi}_{2q} & -\bar{\xi}_{3q} \end{vmatrix}. \tag{12b}$$

Finally by $d\mu_N(z)$ is meant the "Gaussian" measure in the space of N-dimensional complex variables, namely

$$d\mu_N(z) = \pi^{-N} \exp(-\bar{z} \cdot z)\, d^N z. \tag{13}$$

The proof for Eq. (9) is obvious and will be omitted here.[13]

The computation of the generating function for the 9j symbol is thus reduced to the evaluation of the right-hand side of Eq. (9). This can be done in three steps.

Step 1: Integrate over the variables ζ_{1q}. Only four of the six D's, namely D_1, $\tilde{D}_q$, $q = 1, 2, 3$, contribute here. Write

$$\sum_q \tilde{D}_q = \sum_q (c_q \bar{\xi}_{1q} + d_q \bar{\eta}_{1q}) + E, \tag{14}$$

where

$$c_q \equiv \tau'_{3q} \bar{\eta}_{2q} - \tau'_{2q} \bar{\eta}_{3q}, \tag{15}$$

$$d_q \equiv -\tau'_{3q} \bar{\xi}_{2q} + \tau'_{2q} \bar{\xi}_{3q}, \tag{16}$$

$$E \equiv \sum_q \tau'_{1q} \bar{\delta}_{2q,3q}, \tag{17}$$

with

$$\delta_{pq,p'q'} \equiv \begin{vmatrix} \xi_{pq} & \xi_{p'q'} \\ \eta_{pq} & \eta_{p'q'} \end{vmatrix}; \tag{18}$$

then[14]

$$\int d\mu_6(\zeta_{11}, \zeta_{12}, \zeta_{13}) \exp(c \cdot \bar{\xi}_1 + d \cdot \bar{\eta}_1) \times \exp(D_1 + E) = \exp f, \tag{19}$$

$$f = \begin{vmatrix} \tau_{11} & \tau_{12} & \tau_{13} \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{vmatrix} + E. \tag{20}$$

Step 2: Integrate over the variables ζ_{2q}. Rewrite

$$f = \begin{vmatrix} \sigma_1 & \sigma_2 & \sigma_3 \\ \bar{\xi}_{21} & \bar{\xi}_{22} & \bar{\xi}_{23} \\ \bar{\eta}_{21} & \bar{\eta}_{22} & \bar{\eta}_{23} \end{vmatrix} + (c' \cdot \bar{\xi}_2 + d' \cdot \bar{\eta}_2) + E', \tag{21}$$

where

$$\sigma \equiv \begin{pmatrix} \tau_{11}\tau'_{32}\tau'_{33} \\ \tau_{12}\tau'_{31}\tau'_{33} \\ \tau_{13}\tau'_{31}\tau'_{32} \end{pmatrix}, \tag{22}$$

$$c'_p \equiv \sum_q H_{pq} \bar{\eta}_{3q}, \tag{23}$$

$$d'_p \equiv -\sum_q H_{pq} \bar{\xi}_{3q}, \tag{24}$$

$$H \equiv \begin{pmatrix} \tau'_{11} & -\tau_{13}\tau'_{22}\tau'_{31} & \tau_{12}\tau'_{23}\tau'_{31} \\ \tau_{13}\tau'_{21}\tau'_{32} & \tau'_{12} & -\tau_{11}\tau'_{23}\tau'_{32} \\ -\tau_{12}\tau'_{21}\tau'_{33} & \tau_{11}\tau'_{22}\tau'_{33} & \tau'_{13} \end{pmatrix}, \tag{25}$$

$$E' \equiv (\lambda \times \bar{\xi}_3) \cdot \bar{\eta}_3, \tag{26}$$

with

$$\lambda \equiv \begin{pmatrix} \tau_{11}\tau'_{22}\tau'_{23} \\ \tau_{12}\tau'_{21}\tau'_{23} \\ \tau_{13}\tau'_{21}\tau'_{22} \end{pmatrix}. \tag{27}$$

Since E' and D_3 are independent of the variables ζ_{2q}, the relevant expression is

$$\exp(E' + D_3) \int \prod_q d\mu_3(\xi_{2q}) \exp(c'\cdot\bar{\xi}_2) \times \int \prod_q d\mu_3(\eta_{2q}) \exp(d'\cdot\bar{\eta}_2) \exp[(\sigma \times \bar{\xi}_2)\cdot\bar{\eta}_2] \exp D_2 = \exp(E' + D_3) \int \prod_q d\mu_3(\xi_{2q}) \exp(c'\cdot\bar{\xi}_2) \exp X, \tag{28}$$

where

$$X \equiv d''\cdot\xi_2 + \bar{\xi}_2^T A \xi_2 \tag{29}$$

with

$$d''_q \equiv (d' \times \tau_2)_q, \tag{30}$$

$$A = \begin{pmatrix} \sigma_2\tau_{22} + \sigma_3\tau_{23} & -\sigma_2\tau_{21} & -\sigma_3\tau_{21} \\ -\sigma_1\tau_{22} & \sigma_1\tau_{21} + \sigma_3\tau_{23} & -\sigma_3\tau_{22} \\ -\sigma_1\tau_{23} & -\sigma_2\tau_{23} & \sigma_1\tau_{21} + \sigma_2\tau_{22} \end{pmatrix}. \tag{31}$$

The last integral in Eq. (28) can be performed with the aid of Bargmann's theorems on the Laplacian integrals.[15] The result is

$$\int \Pi d\xi_{2q} \exp[-\bar{\xi}_2^T(1 - A)\xi_2 + c'\cdot\bar{\xi}_2 + d''\cdot\xi_2] = \det(1 - A)^{-1} \exp d''^T (1 - A)^{-1} c' \tag{32}$$

with

$$\det(1 - A) = (1 - \sigma_1\tau_{21} - \sigma_2\tau_{22} - \sigma_3\tau_{23})^2. \tag{33}$$

Last Step: Integrate over the remaining variables ζ_{3q}. The remaining expression is

$$\rho \equiv \det(1 - A)^{-1} \int d\mu_3(\xi_{3q}) d\mu_3(\eta_{3q}) \times \exp[E' + D_3 + d''^T (1 - A)^{-1} c']. \tag{34}$$

The last term in Eq. (34) can be rewritten as

$$d''^T (1 - A)^{-1} c' = \bar{\xi}_3^T P \eta_3 \tag{35}$$

in which c' is given by Eq. (23), d'' by Eq. (30) and

$$(1 - A)^{-1} = [\det(1 - A)]^{-1/2} \times \begin{pmatrix} 1 - \sigma_1\tau_{21} & -\sigma_2\tau_{21} & -\sigma_3\tau_{21} \\ -\sigma_1\tau_{22} & 1 - \sigma_2\tau_{22} & -\sigma_3\tau_{22} \\ -\sigma_1\tau_{23} & -\sigma_2\tau_{23} & 1 - \sigma_3\tau_{23} \end{pmatrix}, \tag{36}$$

$$P \equiv [\det(1 - A)]^{-1/2} H^T \begin{pmatrix} 0 & \tau_{23} - \tau_{22} \\ -\tau_{23} & 0 & -\tau_{21} \\ \tau_{22} & -\tau_{21} & 0 \end{pmatrix} H. \tag{37}$$

Equation (34) then becomes (dropping now the subscripts 3 on the ζ's)

$$\rho = \det(1 - A)^{-1} \int d\mu_3(\xi)\, d\mu_3(\eta) \times \exp[(\lambda \times \bar{\xi})\cdot\bar{\eta} + (\tau_3 \times \xi)\cdot\eta + \bar{\xi}^T P \bar{\eta}] = \det(1 - A)^{-1} \int d\mu_3(\xi) \exp[(\lambda \times \bar{\xi} + \bar{\xi}^T P)\cdot(\tau_3 \times \xi)] = \det(1 - A)^{-1} \int d\mu_3(\xi) \exp \bar{\xi}^T \Omega \xi = \det(1 - A)^{-1} \cdot \det(1 - \Omega)^{-1}, \tag{38}$$

where

$$\Omega_{pq} = (\lambda\cdot\tau_3)\delta_{pq} - \tau_{3p}\lambda_q + Q_{pq} \tag{39}$$

with

$$Q \equiv P \begin{pmatrix} 0 & -\tau_{33} & \tau_{32} \\ \tau_{33} & 0 & -\tau_{31} \\ -\tau_{32} & \tau_{31} & 0 \end{pmatrix}. \tag{40}$$

The final result is remarkably simple. The generating function for the $9j$ symbol reads

$$\rho(\tau, \tau') = N \sum_{k,k'} \{9j\} \prod_{p,q} \tau_{pq}^{k_{pq}} \tau'^{k'_{pq}} = \det(1 - A)^{-1} \cdot \det(1 - \Omega)^{-1} \equiv [G(\tau, \tau')]^{-2} \tag{41}$$

in which $\det(1 - A)^{-1}$ is given by (33) and

$$\det(1 - \Omega) = [1 - \tau_{31}(\lambda_1 + P_{23}) - \tau_{32}(\lambda_2 + P_{31}) - \tau_{33}(\lambda_3 + P_{12})]^2. \tag{42}$$

With the substitution of the matrix elements P_{ij} from (37), we get

$$G(\tau, \tau') = 1 - \sum_{p,q=1}^{3} a_{pq} - \sum_{\alpha=1}^{6} b_\alpha, \tag{43}$$

where

$$(a_{pq}) \equiv \begin{pmatrix} \tau_{21}\tau_{31}\tau'_{12}\tau'_{13} & \tau_{22}\tau_{32}\tau'_{11}\tau'_{13} & \tau_{23}\tau_{33}\tau'_{11}\tau'_{12} \\ \tau_{11}\tau_{31}\tau'_{22}\tau'_{23} & \tau_{12}\tau_{32}\tau'_{21}\tau'_{23} & \tau_{13}\tau_{33}\tau'_{21}\tau'_{22} \\ \tau_{11}\tau_{21}\tau'_{32}\tau'_{33} & \tau_{12}\tau_{22}\tau'_{31}\tau'_{33} & \tau_{13}\tau_{23}\tau'_{31}\tau'_{32} \end{pmatrix}, \tag{44}$$

$$(b_\alpha) \equiv \begin{pmatrix} \tau_{11}\tau'_{11}\tau_{32}\tau'_{32}\tau_{23}\tau'_{23} \\ \tau_{31}\tau'_{31}\tau_{22}\tau'_{22}\tau_{13}\tau'_{13} \\ \tau_{21}\tau'_{21}\tau_{12}\tau'_{12}\tau_{33}\tau'_{33} \\ -\tau_{11}\tau'_{11}\tau_{22}\tau'_{22}\tau_{33}\tau'_{33} \\ -\tau_{21}\tau'_{21}\tau_{32}\tau'_{32}\tau_{13}\tau'_{13} \\ -\tau_{31}\tau'_{31}\tau_{12}\tau'_{12}\tau_{23}\tau'_{23} \end{pmatrix}. \tag{45}$$

Note that the sum of the six b terms may be written as $\det|\tau_{pq}\tau'_{pq}|$, namely,

$$\sum_{\alpha=1}^{6} b_\alpha = -\begin{vmatrix} \tau_{11}\tau'_{11} & \tau_{12}\tau'_{12} & \tau_{13}\tau'_{13} \\ \tau_{21}\tau'_{21} & \tau_{22}\tau'_{22} & \tau_{23}\tau'_{23} \\ \tau_{31}\tau'_{31} & \tau_{32}\tau'_{32} & \tau_{33}\tau'_{33} \end{vmatrix}, \tag{46}$$

where the indices of the entries in (46) may be regarded as complementary to those in (44).

Comparison with Schwinger's result shows that the Schwinger's expression [Eq. (4.37) of Ref. 3] differs from our Eq. (43) in a few changes of signs and some shifts in the primes. Since our expression [Eqs. (44) and (45)] cannot tolerate such changes without ruining the symmetry of the problem, we believe that these discrepancies are most likely typographical in origin.

III. EXPANSION FORMULA FOR THE 9j COEFFICIENT

Expanding $[G(\tau, \tau')]^{-2}$ in powers of τ and τ', we have

$$N \sum_{k,k'} \{9j\} \prod_{p,q} \tau_{pq}^{k_{pq}} \tau'^{k'_{pq}} = \sum_{n=0} (n+1) \Big(\sum_{p,q=1} a_{pq} + \sum_{\alpha=1}^{6} b_\alpha \Big)^n = \sum_{n=0} (n+1)! \prod_{p,q=1}^{3} \prod_{\alpha=1}^{6} \frac{a_{pq}^{\nu_{pq}}}{\nu_{pq}!} \frac{b_\alpha^{\omega_\alpha}}{\omega_\alpha!} . \quad (47)$$

Comparing the coefficients of τ_{pq} and τ'_{pq} on both sides of this equation, we get

$$\{9j\} = N^{-1} \sum_{\nu,\omega} \frac{(n+1)!}{\prod_{p,q=1}^{3} \prod_{\alpha=1}^{6} \nu_{pq}! \omega_\alpha!} (-1)^{\omega_4+\omega_5+\omega_6}, \quad (48)$$

where the summations on the right-hand side are subject to the following set of matrix component-wise constraints:

$$k_{pq} = \hat{\nu}_{pq} + \omega_{pq}, \quad (49a)$$

$$k'_{pq} = (\hat{\nu}_{qp})^T + \omega_{pq}, \quad (49b)$$

$$n = \sum_{p,q=1}^{3} \nu_{pq} + \sum_{a=1}^{6} \omega_\alpha, \quad (50)$$

where

$$(\omega_{pq}) \equiv \begin{pmatrix} \omega_1 + \omega_4 & \omega_3 + \omega_6 & \omega_2 + \omega_5 \\ \omega_3 + \omega_5 & \omega_2 + \omega_4 & \omega_1 + \omega_6 \\ \omega_2 + \omega_6 & \omega_1 + \omega_5 & \omega_3 + \omega_4 \end{pmatrix}. \quad (51)$$

The caret operation on any matrix element z_{pq} is defined here as

$$\hat{z}_{pq} \equiv \sum_{p=1}^{3} z_{pq} - z_{pq} = z_{lq} + z_{mq}, \quad l \neq m \neq p. \quad (52)$$

Note that $k_{pq}, k'_{pq}, \nu_{pq}$, and ω_α are all nonnegative integers. With the aid of the identities (4) and (6), we have

$$\sum_{q=1}^{3} \nu_{pq} = n - J_p, \quad (53a)$$

$$\sum_{q=1}^{3} \nu_{qp} = n - K_p. \quad (53b)$$

Thus Eq. (49) reads

$$k_{pq} = n - K_q - \nu_{pq} + \omega_{pq} \equiv \mathfrak{N}_{pq}, \quad (49a')$$

$$k'_{pq} = n - J_p - \nu_{pq} + \omega_{pq} \equiv \mathfrak{N}'_{pq}. \quad (49b')$$

From Eqs. (3a) and (3b), we get the same expression for j_{pq}, namely

$$2j_{pq} = -n + J_p + K_q + \nu_{pq} - \omega_{pq}. \quad (54)$$

The implication of this statement will be taken up in Sec. V. Before we eliminate those redundant summation variables, let us dispose of the symmetry properties of the 9j coefficient.

IV. SYMMETRY OF THE 9j COEFFICIENT

The symmetry of the 9j coefficient is embodied in those operations which leave the generating function formally invariant. By formally invariant we include those cases where the determinant [Eq. (46)] may undergo a change of sign, thus resulting in an overall phase factor for the 9j symbol. It will be convenient to speak of the operations on the various sets of nine quantities $j_{pq}, k_{pq}, k'_{pq}, \tau_{pq}, \tau'_{pq}, \nu_{pq}$, and a_{pq} in terms of their respective 3×3 matrix arrays. Thus one can easily read off the symmetry from Eqs. (48)–(54).

(i) A permutation of given two rows or two columns in the j matrix implies a corresponding permutation in the k and k' matrices simultaneously. This induces a corresponding permutation in the ν and ω matrices on account of the constraint conditions $\mathcal{P}k = \mathcal{P}\mathfrak{N}$ and $\mathcal{P}k' = \mathcal{P}\mathfrak{N}'$. The effect on the six ω_α are such that the set $(\omega_1, \omega_2, \omega_3)$ is mapped onto $(\omega_4, \omega_5, \omega_6)$ and *vice versa*. Thus for an odd number of permutations of rows and columns, there is a net change of phase equal to $(-1)^{\Sigma\omega_\alpha}$ which is $(-1)^J \equiv (-1)^{\Sigma j_{pq}}$ by virtue of Eqs. (4), (6), (49), and (50).

(ii) Transposition of the j matrix implies $k \leftrightarrow k'^T$ and hence $\mathfrak{N} \leftrightarrow \mathfrak{N}'^T$, or $\nu \leftrightarrow \nu^T$ and $\omega \leftrightarrow \omega^T$, (i.e., $\omega_5 \leftrightarrow \omega_6$). This clearly leaves Eq. (48) invariant.

The symmetry (i) and (ii) corresponds to precisely the 72 element symmetry discussed by Jahn and Hope.[4] As was in the case of 3j and 6j symbols, the Bargmann approach enables one to read off the underlying symmetry in a quite transparent way.

V. EXPLICIT EXPRESSION FOR THE 9j COEFFICIENT

The expansion formula (48) calls for a sum over 15 variables $(\nu_{pq}, \omega_\alpha)$ [if we disregard n as being fixed by (50)] subject to the constraint conditions (49). Half of the 18 constraints in (49) are actually redundant. The simplest way to see this is to note that the j matrix (54) reconstructed from the k and k' matrices turns out to be identical for both (49a) and (49b). Thus the number of independent constraints equals the number of the given j_{pq}, which is nine in this problem.

We find that, consistent with all the constraint conditions, the 15 summation variables are expressible in terms of a basic set of six independent variables z_α. The solutions to (49) can be described in general as follows.

We let

$$z_{pq} \equiv \nu_{pq} + j_{pq}, \quad (55)$$

$$\omega'_{pq} \equiv \omega_{pq} + h_{pq}, \quad (56)$$

where

$$(h_{pq}) \equiv \begin{pmatrix} h_1 + h_4 & h_3 + h_6 & h_2 + h_5 \\ h_3 + h_5 & h_2 + h_4 & h_1 + h_6 \\ h_2 + h_6 & h_1 + h_5 & h_3 + h_4 \end{pmatrix} \quad (57)$$

with[16]

$$h_1 \equiv j_{11} + j_{23} + j_{32}, \quad h_2 \equiv j_{31} + j_{22} + j_{13},$$
$$h_3 \equiv j_{21} + j_{12} + j_{33}, \quad h_4 \equiv j_{11} + j_{22} + j_{33},$$
$$h_5 \equiv j_{21} + j_{32} + j_{13}, \quad h_6 \equiv j_{31} + j_{12} + j_{23}. \quad (58)$$

Then Eqs. (49a) and (49b) are exhausted by

$$\hat{z}_{pq} + \omega'_{pq} = J \quad \text{for all} \quad p, q = 1, 2, 3 \tag{59}$$

by virtue of the identities

$$J - h_{pq} = k_{pq} + \hat{j}_{pq} = k'_{pq} + (\hat{j}_{qp})^T \tag{60}$$

and

$$\hat{z}_{pq} = (\hat{z}_{qp})^T . \tag{61}$$

An explicit solution then calls for a particular choice of (i) z_{pq} in terms of a basic set of six z_α and (ii) ω_α in terms of ω_{pq}. Among many essentially equivalent sets of such solutions, one sufficiently symmetric set is obtained by taking the structure of z_{pq} in exactly the same form as that of the ω_{pq} of (51), namely

$$z_{pq} = \begin{pmatrix} z_1 + z_4 & z_3 + z_6 & z_2 + z_5 \\ z_3 + z_5 & z_2 + z_4 & z_1 + z_6 \\ z_2 + z_6 & z_1 + z_5 & z_3 + z_4 \end{pmatrix}. \tag{62}$$

Then the ω_α may be expressed as

$$\omega_\alpha = \tfrac{1}{2}(J - n) + z_\alpha - h_\alpha, \qquad \alpha = 1, \ldots, 6 \tag{63}$$

with

$$n = \sum_{\alpha=1}^{6} z_\alpha . \tag{64}$$

The last expression follows from (53) and (55).

We thus get an explicit expression for the $9j$ coefficient from Eqs. (48), (55), and (63):

$$\begin{Bmatrix} j_{11} & j_{12} & j_{13} \\ j_{21} & j_{22} & j_{23} \\ j_{31} & j_{32} & j_{33} \end{Bmatrix} = N^{-1} \sum_{z_\alpha} \frac{(-1)^{[(J+\Sigma_{\beta=1}^{6} z_\beta)/2 + z_4 + z_5 + z_6]} \left(1 + \sum_{\beta=1}^{6} z_\beta\right)!}{\prod_{p,q=1}^{3} (z_{pq} - j_{pq})! \prod_{\alpha=1}^{6} [\tfrac{1}{2}(J - \sum_{\beta=1}^{6} z_\beta) + z_\alpha - h_\alpha]!}, \tag{65}$$

where z_{pq} are given by (62) and the summation is over all $z_\alpha \geq 0$, such that all the factorial quantities are nonnegative integers.

The symmetry discussed in Sec. IV is of course *manifest* here. An odd number of permutations of rows and columns of the j matrix which can be compensated by the interchange of the set (z_1, z_2, z_3) with (z_4, z_5, z_6) yields a net change of phase equal to

$$(-1)^{z_4 + z_5 + z_6 - z_1 - z_2 - z_3} = (-1)^{\Sigma \omega_\alpha} = (-1)^J$$

by virtue of (63).

We record below two alternate forms which, while manifesting slightly less symmetry than (65), have other redeeming values:

(a) The unpleasant factor of $\frac{1}{2}$ in (65) may be removed by taking in place of (63),

$$\begin{aligned} \omega_p &= J - n + z_p - h_p , \\ \omega_{p+3} &= z_{p+3} - h_{p+3}, \end{aligned} \tag{63'}$$

then

$$\begin{Bmatrix} j_{11} & j_{12} & j_{13} \\ j_{21} & j_{22} & j_{23} \\ j_{31} & j_{32} & j_{33} \end{Bmatrix} = N^{-1}(-1)^J \sum_{z_\alpha} \frac{(-1)^{z_4 + z_5 + z_6} (1 + \sum_{\beta=1}^{6} z_\beta)!}{\prod_{p,q=1}^{3} (z_{pq} - j_{pq})! \prod_{p=1}^{3} (J - \sum_\beta z_\beta + z_p - h_p)! (z_{p+3} - h_{p+3})!} . \tag{65'}$$

(b) The summation variables z's in (65) and (65') may be integers or half-integers (depending on j_{pq}). By a simple change of basis, the summation variables can be chosen to take on nonnegative *integers* only. From (65a), let $(p = 1, 2, 3)$

$$\begin{aligned} x_p &\equiv z_p + z_4 - j_{pp} = \nu_{pp} , \\ x_{p+3} &\equiv z_{p+3} - x_4(1 - \delta_{p1}) - h_{p+3}; \end{aligned}$$

then

$$\begin{Bmatrix} j_{11} & j_{12} & j_{13} \\ j_{21} & j_{22} & j_{23} \\ j_{31} & j_{32} & j_{33} \end{Bmatrix} = N^{-1} \sum_{x_\alpha \geq 0} \frac{(-1)^{x_4 + x_5 + x_6} (1 + t_o + \sum_{\beta=1}^{6} x_\beta - x_4)!}{\prod_{p=1}^{3} x_p! x_4! (x_5 + x_4)! (x_6 + x_4)! \prod_{p \neq q=1}^{3} (x_{pq} + t_{pq})! \prod_{p=1}^{3} (s_p + x_p - \sum_{\beta=1}^{6} x_\beta)!} , \tag{65''}$$

where the matrix (x_{pq}) is formed in terms of x_α analogously to the structure of the matrices ω_{pq}, z_{pq} and h_{pq} before except that x_4 is now absent along the diagonals. The parameters are given as follows $(p \neq q \neq r = 1, 2, 3)$:

$$\begin{aligned} t_{pq} &\equiv j_{qr} + j_{rp} - j_{pp} - j_{qq}, \\ t_o &\equiv J - 2\,\mathrm{tr} j = \mathrm{tr} k = \mathrm{tr} k', \\ s_p &\equiv \mathrm{tr} j - j_{qr} - j_{rq} . \end{aligned}$$

The following remarks seem in order.

(i) As already pointed out in the beginning of this section, the number of outstanding summation variables equals the number of terms (a's and b's beside the identity term) in the generating function less the number of the j's in the problem. We recall that, for the case of Racah coefficient, such a reduction from a sevenfold to a single sum, as shown by Bargmann, readily yields the well-known Racah's formula.[17] In this spirit, Eq.(65) may be regarded as the corresponding explicit expression for the 9j coefficient.

(ii) As one of the consistency checks on the expression (65), it is perhaps instructive to verify explicitly the following well known identity[18]:

$$\begin{Bmatrix} j_{11} & j_{12} & j_{13} \\ j_{21} & j_{22} & j_{13} \\ j_{31} & j_{31} & 0 \end{Bmatrix} = \frac{(-1)^{j_{12}+j_{21}+j_{13}+j_{31}}}{\sqrt{(2j_{13}+1)(2j_{31}+1)}} \begin{Bmatrix} j_{11} & j_{12} & j_{13} \\ j_{22} & j_{21} & j_{31} \end{Bmatrix}. \tag{66}$$

In establishing (66) directly from (65), we observe that in the limit $j_{33} \to 0$, the six triangle relations of the 9j symbol collapse into four triangle relations for the 6j symbol, namely Φ_3 and $\tilde{\Phi}_3$ of Eq. (10) both tend to 1. This implies that the following limit is to be taken[19]: $(\tau_{31}, \tau_{32}, \tau'_{13}, \tau'_{23}) \to 0$, $(\tau_{33}, \tau'_{33}) \to 1$, and $\tau'_{32} \to \tau_{31}$, $\tau_{23} \to \tau'_{13}$. The net effect is that in Eqs. (44) and (45), $(a_{11}, a_{12}, a_{21}, a_{22}) \to 0$ and $(b_1, b_2, b_5, b_6) \to 0$. This necessitates $(\nu_{11}, \nu_{12}, \nu_{21}, \nu_{22}) \to 0$ and $(\omega_1, \omega_2, \omega_5, \omega_6) \to 0$. The rest of the variables are now expressible in terms of one independent variable, say $x = k_{23} - 2z_4$, namely

$$(z_{pq} - j_{pq}) = (\nu_{pq}) = \begin{pmatrix} 0 & 0 & x \\ 0 & 0 & J_1 - J_2 + x \\ J_1 - K_1 + x & J_1 - K_2 + x & J_2 - K_3 - x \end{pmatrix},$$
$$n = J_1 + x,$$
$$\omega_3 = k'_{12} - x,$$
$$\omega_4 = k'_{11} - x. \tag{67}$$

In this way, the sixfold sum in (65) is simply reduced to a single sum, and we have

$$\{9j\}_{j_{33}=0} = N^{-1}(-1)^{k'_{11}} \sum_x \frac{(-1)^x (1+J_1+x)!}{x!\,(k'_{11}-x)!\,(k'_{12}-x)!\,(k_{23}-x)!\,(J_1-J_2+x)!} \frac{1}{(J_1-K_1+x)!\,(J_1-K_2+x)!}$$
$$= N^{-1}(-1)^{k'_{11}} \frac{(1+J_1)!}{k'_{11}!\,k'_{12}!\,k_{23}!\,(J_1-J_2)!\,(J_1-K_1)!\,(J_1-K_2)!}$$
$$\times {}_4F_3(2+J_1, -k'_{11}, -k'_{12}-k_{23}; 1+J_1-J_2, 1+J_1-K_1, 1+J_1-K_2; 1). \tag{68}$$

Apart from a normalization constant, the right-hand side is precisely the Racah coefficient, since in our notation we have [17, 20]

$$\begin{Bmatrix} j_{11} & j_{12} & j_{13} \\ j_{22} & j_{21} & j_{31} \end{Bmatrix}$$
$$= \left(\frac{\prod k_{ij}!}{\prod_i (J_i+1)!\,(K_i+1)!} \right)^{1/2}$$
$$\times \frac{(-1)^{J_1}(1+J_1)!}{(J_1-J_2)!\,\prod_i (J_1-K_i)!\,k'_{11}!\,k'_{12}!\,k_{23}!}$$
$$\times {}_4F_3(2+J_1, -k'_{11}, -k'_{12}, -k_{23}; 1+J_1-J_2, 1+J_1-K_1, 1+J_1-k_2; 1). \tag{69}$$

Combining (68) with (69) gives immediately (66) with the correct constant and phase factors as expected.

We note in passing that this ${}_4F_3$ function structure for the Racah coefficient *in general* satisfies a criterion called Saalschutzian[21] [i.e., ${}_AF_B(a_i; b_j; x)$ with $\sum b_j = 1 + \sum a_i$]. A contrary statement in the literature[22] is traceable to an incorrect transcription of the Racah's formula.

(iii) In the present approach, we have no immediate contact [other than through Eq. (1)] with the other standard formula, namely[18]

$$\begin{Bmatrix} j_{11} & j_{12} & j_{13} \\ j_{21} & j_{22} & j_{23} \\ j_{31} & j_{32} & j_{33} \end{Bmatrix}$$
$$= \sum_j (-1)^{2j}(2j+1) \begin{Bmatrix} j_{11} & j_{21} & j_{31} \\ j_{32} & j_{33} & j \end{Bmatrix} \begin{Bmatrix} j_{12} & j_{22} & j_{32} \\ j_{21} & j & j_{23} \end{Bmatrix} \begin{Bmatrix} j_{13} & j_{23} & j_{33} \\ j & j_{11} & j_{12} \end{Bmatrix}. \tag{70}$$

We leave it as an open question here whether (65) could yield (70) directly.

(iv) Finally, it seems desirable to examine the functional structure of the 9j coefficient in the light of known facts on the lower hierarchy. While the Clebsch–Gordan coefficient is a ${}_3F_2(a_1, a_2, a_3; b_1, b_2; x)$ function at $x = -1$, and the Racah coefficient is a ${}_4F_3(a_1, a_2, a_3, a_4; b_1, b_2, b_3; x)$ function at $x = 1$, a naive conjecture that the 9j coefficient might also belong to some hypergeometric ${}_pF_q$ family turns out to be false. The best that can be said in this regard is that the 9j symbol is a folded products of either ${}_3F_2$ or ${}_4F_3$ functions. Stated otherwise, (65b) can be transcribed into integral representations (of dimensions either six or nine) which can be shown to be quite different from the known representations of a single ${}_pF_q$ function.

[1] V. Bargmann, Rev. Mod. Phys. 34, 829 (1962). The notation of this paper is followed when possible.

[2] E. P. Wigner, famous unpublished 1940 manuscript, reprinted in Ref. 5.

[3] J. Schwinger, famous unpublished 1952 AEC Report, reprinted in Ref. 5.
[4] H. A. Jahn and J. Hope, Phys. Rev. **93**, 318 (1954).
[5] L. C. Biedenharn and H. Van Dam, Eds., *Quantum Theory of Angular Momentum* (Academic, New York, 1968). A very extensive bibliography on the subject may be found.
[6] Apart from some possible misprints in Eq. (4.37) of Ref. 3. See remarks following Eq. (46) of the present paper.
[7] T. Regge, Nuovo Cimento **10**, 544 (1958); **11**, 116 (1959).
[8] See Ref. 1.
[9] See, e.g., A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U. P., Princeton, N.J., 1957). D. M. Brink and G. R. Satchler, *Angular Momentum* (Oxford U. P., London, 1968), 2nd ed. Other references may be found in Ref. 5.
[10] Ref. 1, Eq. (4.1).
[11] Eq. (1) is quite well known. See, e.g., Ref. 2, Eq. (78b), Ref. 4, Eq. (3), and many standard texts such as Ref. 9.
[12] The particular combination of $\Gamma^T\bar{\zeta}$ in Eq. (10b) as compared to ζ in Eq. (10a) in essence accounts for the difference in writing the contravariant and covariant indices in the corresponding state vectors. For details, see, e.g., Refs. 1 and 2.
[13] The proof is exactly parallel to that leading to Eq. (4.13) of Ref. 1.
[14] The integral in Eq. (19) is done exactly the same way as the steps leading to Eq. (4.16) of Ref. 1.
[15] See the Appendix of Ref. 1.
[16] It is obvious that the index labeling of h_α is in accordance with that of Eq. (45).
[17] G. Racah, Phys. Rev. **62**, 146 (1942). Cf. Ref. 1, esp. equation following (4.20); and Eq. (69) of the present paper.
[18] See any standard textbook, such as Ref. 9.
[19] Note that $(k_{31}, k_{32}, k'_{13}, k'_{23}) \to 0$ as $j_{33} \to 0$.
[20] It is trivial to transcribe the summation expression of the Racah's formula (Ref. 17) into a hypergeometric function of the ${}_4F_3$ type (with argument $x = 1$).
[21] See, e.g., L. J. Slater, *Generalized Hypergeometric Functions* (Cambridge U. P., Cambridge, 1966), esp. p. 42.
[22] D. A. Akyeampong and M. A. Rashid, Trieste Preprint IC/70/134. See Appendix. In their starting point, Eq. (A.1), b should read d in the next to the last factorial. Hence their last condition (if $b = d$) is spurious. *Note added in proof:* Dr. Akyeampong (private communication) informs the author that this point has been revised in a later version (J. Math. Phys., to appear).

A Class of Matrix Ensembles*

Freeman J. Dyson
Institute for Advanced Study, Princeton, New Jersey 08540
(Received 23 August 1971)

A class of random matrix ensembles is defined, with the purpose of providing a realistic statistical description of the Hamiltonian of a complicated quantum-mechanical system (such as a heavy nucleus) for which an approximate model Hamiltonian is known. An ensemble of the class is specified by the model Hamiltonian H_0, an observed eigenvalue distribution-function $r(E)$, and a parameter τ which may be considered to be a fictitious "time." Each of H_0, $r(E)$, and τ may be chosen independently. The ensemble consists of matrices M which are obtained from H_0 by an invariant random Brownian-motion process, lasting for a time τ and tending to pull the eigenvalues of M toward the distribution $r(E)$. For small τ the ensemble allows only small perturbations of H_0. As $\tau \to \infty$, the ensemble tends to a stationary limit independent of H_0 and depending on $r(E)$ alone. The following quantitative results are obtained. (1) It is proved that the global eigenvalue distribution in the limit $\tau \to \infty$ becomes identical with the observed distribution $r(E)$. (2) A nonlinear partial differential equation is obtained for the global eigenvalue distribution $\rho(E, \tau)$ as a function of E and τ. Solution of this equation will show how the distribution changes from the initial form specified by H_0 at $\tau = 0$ to the final form $r(E)$ at $\tau = \infty$. Approximate solution shows that deviations of $\rho(E, \tau)$ from $r(E)$ extending over an interval containing m eigenvalues will disappear exponentially as soon as τ is of the order of mD^2, where D is the local mean level spacing. (3) Exact analytic expressions are obtained for the correlation functions representing the probabilities for finding n eigenvalues at assigned positions $(E_1, \ldots, E_n)$ in the ensemble in the limit $\tau \to \infty$, irrespective of the positions of the remaining $(N\text{-}n)$ eigenvalues. It is made plausible, but not proved, that these correlation functions tend to limits as $N \to \infty$, which are universal functions independent of $r(E)$. If proved, this statement would imply that the local statistical properties (spacing distributions, etc.) of eigenvalues in the ensemble become, when τ and N are both large, universal properties independent of the global eigenvalue distributions. In particular, the spacing distributions would be identical with those calculated for more special ensembles by Wigner, Gaudin, and Mehta.

1. THE WIGNER ENSEMBLE

Wigner[1] proposed, as a mathematical tool rather than as a physical model for the description of complicated nuclei, the "Gaussian Ensemble" of random matrices. The Gaussian ensemble E_G is defined as the set of all real symmetric $(N \times N)$ matrices M with the probability-distribution

$$p(M) = c \exp(-\mathrm{Tr}M^2/a^2). \tag{1.1}$$

Here N is any integer, a is a real number, and c is a normalization constant depending on N and a. Wigner suggested that in some respects the statistical behavior of the eigenvalues of a matrix M chosen at random in the ensemble E_G would mimic the behavior of highly-excited energy levels of a complex nucleus. Since the number of levels of a real nucleus is infinite, the representation of the levels by a finite matrix cannot be complete. It was Wigner's suggestion that the levels of the nucleus and of the random matrix in E_G should behave in the same way *locally*, that is to say, so long as attention is confined to a group of n consecutive levels, where n is a number very small compared to N.

One of the first results of the theory of random matrices was Wigner's Semicircle Law. This law[2] states that the density of eigenvalues per unit energy E of a matrix in the ensemble E_G tends to the limit

$$r(E) = (2/\pi a^2)(Na^2 - E^2)^{1/2}, \quad |E| < N^{1/2}a,$$
$$= 0, \quad |E| > N^{1/2}a, \tag{1.2}$$

as $N \to \infty$. The law was proved by Wigner for a large class of matrix ensembles of which E_G is a special case. The essential requirement upon which the validity of (1.2) depends is that the matrix elements of M be independent random variables.

In spite of its mathematical elegance, the semicircle law is in violent contradiction with the facts of nuclear physics. The level density in real nuclei has roughly the form

[3] J. Schwinger, famous unpublished 1952 AEC Report, reprinted in Ref. 5.
[4] H. A. Jahn and J. Hope, Phys. Rev. **93**, 318 (1954).
[5] L. C. Biedenharn and H. Van Dam, Eds., *Quantum Theory of Angular Momentum* (Academic, New York, 1968). A very extensive bibliography on the subject may be found.
[6] Apart from some possible misprints in Eq. (4.37) of Ref. 3. See remarks following Eq. (46) of the present paper.
[7] T. Regge, Nuovo Cimento **10**, 544 (1958); **11**, 116 (1959).
[8] See Ref. 1.
[9] See, e.g., A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U. P., Princeton, N.J., 1957). D. M. Brink and G. R. Satchler, *Angular Momentum* (Oxford U. P., London, 1968), 2nd ed. Other references may be found in Ref. 5.
[10] Ref. 1, Eq. (4.1).
[11] Eq. (1) is quite well known. See, e.g., Ref. 2, Eq. (78b), Ref. 4, Eq. (3), and many standard texts such as Ref. 9.
[12] The particular combination of $\Gamma^T\bar{\zeta}$ in Eq. (10b) as compared to ζ in Eq. (10a) in essence accounts for the difference in writing the contravariant and covariant indices in the corresponding state vectors. For details, see, e.g., Refs. 1 and 2.
[13] The proof is exactly parallel to that leading to Eq. (4.13) of Ref. 1.
[14] The integral in Eq. (19) is done exactly the same way as the steps leading to Eq. (4.16) of Ref. 1.
[15] See the Appendix of Ref. 1.
[16] It is obvious that the index labeling of h_α is in accordance with that of Eq. (45).
[17] G. Racah, Phys. Rev. **62**, 146 (1942). Cf. Ref. 1, esp. equation following (4.20); and Eq. (69) of the present paper.
[18] See any standard textbook, such as Ref. 9.
[19] Note that $(k_{31}, k_{32}, k'_{13}, k'_{23}) \to 0$ as $j_{33} \to 0$.
[20] It is trivial to transcribe the summation expression of the Racah's formula (Ref. 17) into a hypergeometric function of the ${}_4F_3$ type (with argument $x = 1$).
[21] See, e.g., L. J. Slater, *Generalized Hypergeometric Functions* (Cambridge U. P., Cambridge, 1966), esp. p. 42.
[22] D. A. Akyeampong and M. A. Rashid, Trieste Preprint IC/70/134. See Appendix. In their starting point, Eq. (A.1), b should read d in the next to the last factorial. Hence their last condition (if $b = d$) is spurious. *Note added in proof:* Dr. Akyeampong (private communication) informs the author that this point has been revised in a later version (J. Math. Phys., to appear).

A Class of Matrix Ensembles*

Freeman J. Dyson
Institute for Advanced Study, Princeton, New Jersey 08540
(Received 23 August 1971)

A class of random matrix ensembles is defined, with the purpose of providing a realistic statistical description of the Hamiltonian of a complicated quantum-mechanical system (such as a heavy nucleus) for which an approximate model Hamiltonian is known. An ensemble of the class is specified by the model Hamiltonian H_0, an observed eigenvalue distribution-function $r(E)$, and a parameter τ which may be considered to be a fictitious "time." Each of H_0, $r(E)$, and τ may be chosen independently. The ensemble consists of matrices M which are obtained from H_0 by an invariant random Brownian-motion process, lasting for a time τ and tending to pull the eigenvalues of M toward the distribution $r(E)$. For small τ the ensemble allows only small perturbations of H_0. As $\tau \to \infty$, the ensemble tends to a stationary limit independent of H_0 and depending on $r(E)$ alone. The following quantitative results are obtained. (1) It is proved that the global eigenvalue distribution in the limit $\tau \to \infty$ becomes identical with the observed distribution $r(E)$. (2) A nonlinear partial differential equation is obtained for the global eigenvalue distribution $\rho(E, \tau)$ as a function of E and τ. Solution of this equation will show how the distribution changes from the initial form specified by H_0 at $\tau = 0$ to the final form $r(E)$ at $\tau = \infty$. Approximate solution shows that deviations of $\rho(E, \tau)$ from $r(E)$ extending over an interval containing m eigenvalues will disappear exponentially as soon as τ is of the order of mD^2, where D is the local mean level spacing. (3) Exact analytic expressions are obtained for the correlation functions representing the probabilities for finding n eigenvalues at assigned positions $(E_1, \ldots, E_n)$ in the ensemble in the limit $\tau \to \infty$, irrespective of the positions of the remaining $(N\text{-}n)$ eigenvalues. It is made plausible, but not proved, that these correlation functions tend to limits as $N \to \infty$, which are universal functions independent of $r(E)$. If proved, this statement would imply that the local statistical properties (spacing distributions, etc.) of eigenvalues in the ensemble become, when τ and N are both large, universal properties independent of the global eigenvalue distributions. In particular, the spacing distributions would be identical with those calculated for more special ensembles by Wigner, Gaudin, and Mehta.

1. THE WIGNER ENSEMBLE

Wigner[1] proposed, as a mathematical tool rather than as a physical model for the description of complicated nuclei, the "Gaussian Ensemble" of random matrices. The Gaussian ensemble E_G is defined as the set of all real symmetric $(N \times N)$ matrices M with the probability-distribution

$$p(M) = c \exp(-\mathrm{Tr}M^2/a^2). \tag{1.1}$$

Here N is any integer, a is a real number, and c is a normalization constant depending on N and a. Wigner suggested that in some respects the statistical behavior of the eigenvalues of a matrix M chosen at random in the ensemble E_G would mimic the behavior of highly-excited energy levels of a complex nucleus. Since the number of levels of a real nucleus is infinite, the representation of the levels by a finite matrix cannot be complete. It was Wigner's suggestion that the levels of the nucleus and of the random matrix in E_G should behave in the same way *locally*, that is to say, so long as attention is confined to a group of n consecutive levels, where n is a number very small compared to N.

One of the first results of the theory of random matrices was Wigner's Semicircle Law. This law[2] states that the density of eigenvalues per unit energy E of a matrix in the ensemble E_G tends to the limit

$$r(E) = (2/\pi a^2)(Na^2 - E^2)^{1/2}, \quad |E| < N^{1/2}a,$$
$$= 0, \quad |E| > N^{1/2}a, \tag{1.2}$$

as $N \to \infty$. The law was proved by Wigner for a large class of matrix ensembles of which E_G is a special case. The essential requirement upon which the validity of (1.2) depends is that the matrix elements of M be independent random variables.

In spite of its mathematical elegance, the semicircle law is in violent contradiction with the facts of nuclear physics. The level density in real nuclei has roughly the form

$$r(E) = c\,\exp(bE^{1/2}) \tag{1.3}$$

and bears not the slightest resemblance to a semicircle. The disagreement between (1.2) and (1.3) does not contradict Wigner's hypothesis that the eigenvalues of a matrix in E_G and the levels of a nucleus should have the same local behavior. Equation (1.2) describes the global behavior of eigenvalues in E_G, and Eq. (1.3) the global behavior of nuclear levels. Their local behavior could well be the same even when their global behavior is different. Nevertheless the semicircle law is a stumbling block for anyone who wishes to believe that the Wigner ensemble gives a useful insight into the behavior of real nuclei. Professor G. E. Uhlenbeck[3] expressed the problem succinctly with the words: "If you admit that the Wigner ensemble gives a completely wrong answer for the level density, why do you believe any of the other predictions of random-matrix theory?"

The only effective way to answer Uhlenbeck's criticism is to exhibit a new ensemble E_B of random matrices which satisfies four requirements: (1) It predicts a global level-density distribution in agreement with observation. (2) It predicts a local statistical behavior of energy levels in agreement with the Wigner ensemble. (3) Its definition is physically plausible. (4) Its consequences are amenable to mathematical treatment. The purpose of this paper is to construct such an ensemble. Section 2 is occupied with the definition of E_B, Sec. 3 with the verification that it has property (1), and Sec. 4 with an incomplete verification of property (2). Whether E_B also satisfies requirements (3) and (4) is a question of taste which the reader must judge for himself.

This is not the first proposal of a new ensemble to describe nuclear levels while avoiding the unrealistic features of the Wigner ensemble. The most successful attempts in this direction were made recently by Bohigas and Flores[4,5] and independently by French and Wong.[6,7] The Bohigas–Flores and French–Wong ensembles take into account from the beginning the fact that the nuclear Hamiltonian is not truly random but is composed of one-particle energies and two-particle interactions whose general structure is known from the empirical validity of the nuclear shell model. Study of these ensembles is still at an early stage. They will certainly be further refined and improved, and it is much too soon to pronounce them inadequate because they have not yet overcome all difficulties. In their present form, the Bohigas–Flores and French–Wong ensembles do not satisfy items (1) and (4) in our list of requirements. They predict a level density which is roughly Gaussian instead of semicircular, but still quite different from the empirical density (1.3). They are mathematically so intractable that all theoretical analysis so far has been based on numerical Monte–Carlo computations of the behavior of matrices of comparatively low order. We do not yet have any firm evidence of the behavior of these ensembles in the limit $N \to \infty$. On the other hand, they seem to satisfy well the requirements (2) and (3). Further study and development of these ensembles will certainly be helpful in bringing random-matrix theory into closer contact with physical reality.

Another previous proposal of a new type of random-matrix ensemble was made by Leff[8–10] and by Fox and Kahn[11]; see also Sec. 17.3 of Mehta's book.[12] The Leff–Fox–Kahn proposal was in essence identical with a limiting case of the ensemble E_B proposed in this paper. Only their physical motivation was different, and they did not go very far in calculating the consequences of their proposal.

The suffix B in the symbol E_B stands for "Brown," since the most natural definition of the new ensemble arises from the Brownian-motion model of random matrices introduced by Dyson.[13]

2. THE BROWNIAN ENSEMBLE

Let H_0 be the Hamiltonian of an approximate model of a nucleus. For example, H_0 might be

$$H_0 = \sum n_j \epsilon_j, \tag{2.1}$$

where the ϵ_j are the energies of independent-particle orbits in a nuclear shell model, and the n_j are Fermion occupation numbers taking the values 0 and 1. Alternatively, H_0 might be a collective model Hamiltonian or a hybrid single-particle-collective model. We assume that H_0 has one unrealistic feature; there is a high-energy cutoff, so that the total number N of nuclear states is finite. H_0 is then a real symmetric $(N \times N)$ matrix.

Let $r(E)$ be a smooth function representing the density of levels of an actual nucleus around the energy E. For example, $r(E)$ might have the empirical form (1.3). We do not assume that the eigenvalues of H_0 follow the distribution $r(E)$, but we suppose the high-energy cutoff applied to $r(E)$ so as to make the total number of levels

$$\int r(E)dE = N, \tag{2.2}$$

the same for $r(E)$ as for the model Hamiltonian H_0.

Consider a classical electrostatic model in which $r(E)$ is the density of negative electric charge distributed along a straight line. According to the laws of two-dimensional electrostatics, this charge generates a potential

$$u(E) = \int r(\lambda)\,\log|\lambda - E|\,d\lambda \tag{2.3}$$

and an electric field

$$f(E) = -\frac{\partial u}{\partial E} = \int r(\lambda)(\lambda - E)^{-1}d\lambda. \tag{2.4}$$

Let M be any real symmetric $(N \times N)$ matrix. We can define a generalized electric field

$$f(M) = \int r(\lambda)(\lambda I_N - M)^{-1}d\lambda. \tag{2.5}$$

In the space of matrices M, this field is the gradient of a potential

$$f(M) = -\,\mathrm{grad}\,u(M), \tag{2.6}$$

$$u(M) = \int r(\lambda)\,\log|\,\mathrm{Det}(\lambda I_N - M)|\,d\lambda. \tag{2.7}$$

Although in the application to nuclear physics we are interested in real symmetric M, it is mathematically convenient to consider at the same time the cases in which H_0 and M are either complex Hermitian or quaternion-real self-adjoint matrices. The definitions

(2.5)–(2.7) are meaningful in all three cases. As usual (see Dyson[14]) we introduce the parameter β to denote the number of independent components

$$M_{ij}^{\alpha}, \quad \alpha = 1,\ldots,\beta \quad i,j = 1,\ldots,N, \tag{2.8}$$

of each matrix element of M. Thus $\beta = 1,2,4$ for M real, complex or quaternionic, respectively.

Let the matrix M undergo a Brownian motion constrained by the "electric field" $\beta f(M)$. The change of M during an infinitesimal time-interval δt is an infinitesimal random matrix δM whose first moment is given by

$$\langle \delta M \rangle = \beta f(M)\delta t, \tag{2.9}$$

while the second moments are

$$\langle \delta M_{ij}^{\alpha} \delta M_{kl}^{\gamma} \rangle = \delta_{\alpha\gamma}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\delta t. \tag{2.10}$$

The form of the right-hand side of Eq. (2.10) is determined by the requirement that the Brownian process be independent of the representation of M. The form of Eq. (2.9) is imposed by the physical requirement that the eigenvalues of M should be "pulled" toward the distribution $r(E)$ during the Brownian process. If $p(M,t)$ is the probability distribution of matrices M at time t, the moment equations (2.9) and (2.10) imply the Ornstein–Uhlenbeck equation

$$(\partial p/\partial t) = \operatorname{div}[\operatorname{grad} p - \beta p f], \tag{2.11}$$

which fixes the development of p with time. Here the operations div and grad are defined in the space of matrices M. The parameter t is a fictitious time having nothing to do with real physical time.

We are now ready to define the Brownian ensemble E_B. Given the model Hamiltonian H_0, the level-density $r(E)$, and the positive real number τ, *E_B is defined as the set of matrices M with probability-density $p(M,\tau)$, where $p(M,t)$ is the solution of* (2.11) *with the initial condition*

$$p(M,0) = \delta(M - H_0). \tag{2.12}$$

Thus E_B is the ensemble obtained by starting with $M = H_0$ and allowing M to follow the Brownian process (2.9), (2.10) for a finite time τ.

What is the meaning of the parameter τ? If L is the length of the energy interval over which the distribution $r(E)$ extends, then

$$D = LN^{-1} \tag{2.13}$$

is an average spacing between levels. We have the rough estimate

$$\operatorname{Tr}(H_0^2) \sim NL^2, \tag{2.14}$$

while for small τ, Eq. (2.10) gives

$$\langle \operatorname{Tr}(M - H_0)^2 \rangle \sim N^2\tau. \tag{2.15}$$

Equations (2.14) and (2.15) are comparable when

$$\tau \sim ND^2. \tag{2.16}$$

For $\tau \ll ND^2$, we are in the domain of "small perturbations," and the matrices M in E_B are still concentrated around the starting Hamiltonian H_0. For $\tau \gg ND^2$, the memory of the initial condition is rapidly lost, and the ensemble E_B approaches the unique stationary solution of Eq. (2.11),

$$q(M) = p(M,\infty) = c\,\exp[-\beta u(M)], \tag{2.17}$$

with $u(M)$ defined by Eq. (2.7). We denote by E_S the stationary ensemble of matrices with probability-distribution $q(M)$. E_S is independent of H_0. The ensemble studied earlier by Leff[8,9,10] and Fox and Kahn[11] was essentially identical with E_S.

To describe more precisely the approach of E_B to the limit E_S, we observe that the Ornstein–Uhlenbeck equation (2.11) is self-adjoint with respect to the metric

$$(p,p') = \int pq^{-1}p'\,dM, \tag{2.18}$$

where the integration extends over the space of matrices M. For if R denotes the operator

$$Rp = \operatorname{div}[\operatorname{grad} p + \beta p \operatorname{grad} u], \tag{2.19}$$

then

$$(p,Rp') = -\int[(\operatorname{grad}(q^{-1}p)\cdot\operatorname{grad}(q^{-1}p')]q\,dM = (p',Rp). \tag{2.20}$$

The eigenstates p_j and eigenvalues λ_j of R are related by

$$\lambda_j = -\left(\int |\operatorname{grad}(q^{-1}p_j)|^2 q\,dM\right), \tag{2.21}$$

with the normalization condition

$$(p_j,p_k) = \delta_{jk}. \tag{2.22a}$$

Equation (2.21) implies that all λ_j are strictly negative except for the unique zero eigenvalue

$$\lambda_0 = 0, \quad p_0 = q. \tag{2.22b}$$

The probability distribution of the ensemble E_B is explicitly

$$p(M,\tau) = q^{-1}(H_0)\sum_j \exp(\lambda_j\tau)p_j(H_0)p_j(M). \tag{2.23}$$

We can estimate roughly the size of the smallest nonzero λ_j by taking in Eq. (2.21) a linear ansatz

$$p_j(M) = q(M)\,\operatorname{Tr}(AM) \tag{2.24}$$

with A a constant matrix. Then

$$\lambda_j = -\operatorname{Tr}(A^2), \tag{2.25}$$

while the normalization condition (2.22a) gives

$$1 \approx N^{-2}(\operatorname{Tr}A^2)\int(\operatorname{Tr}M^2)q\,dM \approx ND^2(\operatorname{Tr}A^2). \tag{2.26}$$

Hence the nonzero λ_j are all of magnitude

$$\lambda_j \sim -(ND^2)^{-1} \tag{2.27}$$

or greater. Equation (2.23) then implies that the ensemble E_B tends to the limiting distribution (2.17) with exponential speed as soon as $\tau \gg ND^2$.

The physical meaning of the definition of E_B may be summarized as follows. The true Hamiltonian M of a nucleus is derived from the model Hamiltonian H_0 by an unknown perturbation. The ensemble E_B is obtained by regarding all perturbations of H_0 as equally likely, subject to two restrictions, (i) the overall strength of the perturbation is determined by the parameter τ, and (ii) the perturbation is biassed in such a way as to pull the eigenvalues of M toward the observed distribution-function $r(E)$. As $\tau \to \infty$, E_B tends rapidly to the ensemble E_S which is independent of the starting-model H_0 and depends only on $r(E)$.

3. EIGENVALUE DISTRIBUTIONS

Let $P(E_1, \ldots, E_N, \tau)$ be the joint probability distribution of the eigenvalues of a matrix M in the ensemble E_B. The Brownian process (2.9), (2.10) induces a Brownian motion of the eigenvalues E_j with

$$\langle \delta E_j \rangle = \beta\left[f(E_j) - \sum_{k \neq j} (E_k - E_j)^{-1}\right]\delta t = -\beta \frac{\partial w}{\partial E_j}\delta t, \tag{3.1}$$

$$w = \sum_j u(E_j) - \sum\sum_{j<k} \log|E_k - E_j|, \tag{3.2}$$

$$\langle \delta E_j \delta E_k \rangle = 2\delta_{jk}\delta t. \tag{3.3}$$

Equations (3.1)–(3.3) imply the Ornstein–Uhlenbeck equation

$$\frac{\partial P}{\partial \tau} = \sum_j \frac{\partial}{\partial E_j}\left(\frac{\partial P}{\partial E_j} + \beta P \frac{\partial w}{\partial E_j}\right) \tag{3.4}$$

for the joint eigenvalue distribution. As $\tau \to \infty$ and $E_B \to E_S$, the eigenvalue distribution tends to the limiting form

$$P(E_1, \ldots, E_N, \infty) = c\ \exp(-\beta w) = c\left(\prod_{j<k} |E_k - E_j|^{\beta}\right) \exp\left(-\beta \sum_j u(E_j)\right), \tag{3.5}$$

which has been studied in detail by Leff.[10]

It is convenient to normalize the joint probability distribution so that

$$\int \cdots \int P(E_1, \ldots, E_N, \tau) dE_1 \ldots dE_N = N! \tag{3.6}$$

The one- and two-level distribution functions in E_B are

$$\rho(E, \tau) = [1/(N-1)!] \int \cdots \times \int P(E, E_2, \ldots, E_N, \tau) dE_2 \cdots dE_N, \tag{3.7}$$

$$\rho_2(E, E', \tau) = [1/(N-2)!] \int \cdots \times \int P(E, E', E_3, \ldots, E_N, \tau) dE_3 \cdots dE_N, \tag{3.8}$$

with the normalizations

$$\int \rho(E, \tau) dE = N, \tag{3.9}$$

$$\int \rho_2(E, E', \tau) dE' = (N-1)\rho(E, \tau). \tag{3.10}$$

When Eq. (3.4) is integrated over $E_2, \ldots, E_N$, the result is

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial E}\left[\frac{\partial \rho}{\partial E} + \beta\rho\left(\frac{\partial u}{\partial E}\right) + \beta \int \rho_2(E, E', \tau)(E' - E)^{-1} dE'\right]. \tag{3.11}$$

We write

$$\rho_2(E, E') = \rho(E)\rho(E')[1 - y(E, E')], \tag{3.12}$$

the dependence of each factor on τ being understood. Equations (3.9) and (3.10) then imply

$$\int \rho(E') y(E, E') dE' = 1. \tag{3.13}$$

Equation (3.11) becomes

$$\partial \rho / \partial t = \frac{\partial}{\partial E}\left(\frac{\partial \rho}{\partial E} + \beta\rho \frac{\partial (u - v)}{\partial E} - \beta\rho \int \rho(E') y(E, E')(E' - E)^{-1} dE'\right), \tag{3.14}$$

with

$$u - v = \int [r(\lambda) - \rho(\lambda)] \log|\lambda - E| d\lambda. \tag{3.15}$$

All equations up to this point are exact. We now make the approximation of assuming that the function $y(E, E')$, which according to Eq. (3.12) describes the statistical correlation between eigenvalues at the positions E and E', is a short-range quantity. Specifically, we shall neglect the second and higher moments of $y(E, E')$ with respect to the difference variable $(E' - E)$. Physically, the range within which $y(E, E')$ is significant may be expected to be of the order of the mean level-spacing D. If $\rho(E)$ varies appreciably only over intervals of the order of $L = ND$, then the errors involved in neglecting the second moment of $y(E, E')$ should be of the order N^{-2}. We call the neglect of the second moment the *hydrodynamical approximation*, since the same approximation is made in deducing the equations of classical hydrodynamics from the Boltzmann equation.

It is convenient to transform $y(E, E')$ to a function of sum and difference variables

$$y(E, E') = Y(s, z), \tag{3.16}$$

$$s = \tfrac{1}{2}(E + E'), \quad z = E' - E. \tag{3.17}$$

The first moment vanishes,

$$\int z Y(s, z) dz = 0, \tag{3.18}$$

because $Y(s, z)$ is even in z. Equation (3.13) becomes

$$\int \rho(E + z) Y(E + \tfrac{1}{2}z, z) dz = 1. \tag{3.19}$$

When $\rho(E + z)$ and $Y(E + \frac{1}{2}z, z)$ are expanded in a Taylor series around the values $\rho(E)$ and $Y(E, z)$, the terms in (3.19) linear in z vanish by (3.18), and the higher terms are assumed negligible. Therefore Eq. (3.19) reduces to

$$\rho(E) \int Y(E, z) dz = 1. \tag{3.20}$$

We now apply the same procedure to the integral which appears in Eq. (3.14). The integral is

$$\int \rho(E + z) Y(E + \tfrac{1}{2}z, z) z^{-1} dz = \int \left(\rho(E) + z \frac{\partial \rho}{\partial E} + \ldots\right)\left(Y(E, z) + \tfrac{1}{2}z \frac{\partial Y}{\partial E} + \ldots\right) z^{-1} dz = (\partial \rho / \partial E) \int Y(E, z) dz + \tfrac{1}{2}\rho(\partial / \partial E) \int Y(E, z) dz, \tag{3.21}$$

the terms in z^{-1} and z again vanishing by symmetry, and the higher terms by hypothesis. By virtue of Eq. (3.20), (3.21) reduces to

$$\tfrac{1}{2}\rho^{-1}\frac{\partial\rho}{\partial E}. \tag{3.22}$$

Hence the "hydrodynamic" equation (3.14) becomes

$$\frac{\partial\rho}{\partial\tau} = \frac{\partial}{\partial E}\left((1-\tfrac{1}{2}\beta)\frac{\partial\rho}{\partial E} + \beta\rho\frac{\partial(u-v)}{\partial E}\right) = \frac{\partial}{\partial E}\left(\beta\rho\frac{\partial\psi}{\partial E}\right), \tag{3.23}$$

with

$$\psi = u - v + T\log\rho, \tag{3.24}$$

where $(u-v)$ is given by Eq. (3.15), and

$$T = \beta^{-1} - \tfrac{1}{2}. \tag{3.25}$$

Instead of a linear equation (3.4) for the N-level distribution-function P, we have a nonlinear equation (3.23) for the single-level distribution-function ρ. The potential ψ has the form of a "free energy," being equal to the electrostatic energy generated by the charge distribution $(r-\rho)$, plus an entropy term $(-TS)$, with

$$S = -\log\rho, \tag{3.26}$$

and the fictitious temperature T given by Eq. (3.25). The case of complex Hermitian matrices ($\beta = 2$) is particularly simple, since then $T = 0$ and ψ is linear in ρ. In the application to nuclear physics, we have $\beta = 1$ and $T = \frac{1}{2}$.

Equation (3.23) takes a simpler appearance when expressed in terms of Lagrangian rather than Eulerian variables. Let $E(n,\tau)$ be defined as a function of the continuous variable n by

$$\int_{-\infty}^{E(n,\tau)}\rho dE = n. \tag{3.27}$$

Thus, for integer n, $E(n,\tau)$ is roughly the mean position in the ensemble E_B of the nth smallest eigenvalue of M. The Lagrangian form of Eq. (3.23) is

$$\frac{dE}{d\tau} = -\beta\frac{\partial\psi}{\partial E}. \tag{3.28}$$

The derivative on the left of Eq. (3.28) is taken at fixed n, that on the right at fixed τ. In terms of Lagrangian variables,

$$\psi(E) = u(E) - \int\log|E(n)-E|dn - T\log\left(\frac{\partial E}{\partial n}\right). \tag{3.29}$$

From Eq. (3.23) or (3.28) it follows that the level-density ρ for the stationary ensemble E_S is characterized by

$$\psi(E) = \text{const}. \tag{3.30}$$

The stationary ρ thus satisfies the equation

$$\rho(E) = c\,\exp\{-T^{-1}\int[r(\lambda)-\rho(\lambda)]\log|\lambda-E|d\lambda\}. \tag{3.31a}$$

This equation has appeared before in the calculation of the asymptotic form of the level-spacing distribution for large spacings (see Dyson[15]). Recently Mehta[16] and Widom[17] have checked the accuracy of Eq. (3.31a) in the level-spacing application by a more exact analysis. They found that the error is actually not, as claimed by Dyson,[15] of the order N^{-2}, but is proportional to $(N^{-2}\log N)$. It is likely that the hydrodynamical approximation which led to Eq. (3.23) will also in general introduce errors of order $(N^{-2}\log N)$.

For the application to nuclear physics, a precision of order $(N^{-2}\log N)$ goes far beyond what is required. Statistical fluctuations and imperfections in the experimental data make it meaningless to retain even terms of order $(N^{-1}\log N)$ in the theory. But the troublesome nonlinear term $(T\log\rho)$ in Eq. (3.23) is in fact of order $(N^{-1}\log N)$ relative to u and v. It is therefore sufficiently accurate to drop this term and to take for the equation defining the level-density ρ in the ensemble E_B,

$$\frac{\partial\rho}{\partial\tau} = \frac{\partial}{\partial E}\{\beta\rho\int[r(\lambda)-\rho(\lambda)](E-\lambda)^{-1}d\lambda\}. \tag{3.31b}$$

In this approximation the level-density of the stationary ensemble E_S becomes simply

$$\rho(E) = r(E). \tag{3.32}$$

The ensemble E_S thus satisfies requirement (1) of Sec. 1 by predicting a global level-density $\rho(E)$ identical with the observed distribution $r(E)$.

From Eq. (3.31b) it is easy to see how the level-density $\rho(E,\tau)$ of the ensemble E_B tends to the limit $r(E)$ as $\tau\to\infty$. First of all, if we choose the starting Hamiltonian H_0 to have precisely the observed level-density $r(E)$, then Eq. (3.31b) implies

$$\rho(E,\tau) = r(E) \tag{3.33}$$

for all τ. In this case, requirement (1) is fulfilled exactly by E_B. In practice we shall generally choose a model H_0 which has a level density not identical with $r(E)$ but reasonably close to it. Then the deviation

$$\delta(E,\tau) = \rho(E,\tau) - r(E) \tag{3.34}$$

satisfies approximately the linear equation

$$\frac{\partial\delta}{\partial\tau} = \frac{\partial}{\partial E}[\beta r(E)\int\delta(\lambda)(\lambda-E)^{-1}d\lambda], \tag{3.35}$$

which can be solved by an eigenfunction expansion similar to Eq. (2.23). Instead of solving Eq. (3.35) exactly, we here examine the qualitative behavior of $\delta(E,\tau)$. Consider a range of energy E small compared with L, so that $r(E)$ can be considered constant, say $r(E) = D^{-1}$ where D is the local mean level spacing. Within this range, the local behavior of $\delta(E,\tau)$ can be represented as a superposition of Fourier components

$$\delta(E,\tau) \sim \exp(ikE - \gamma\tau), \tag{3.36}$$

with wavelengths mD, where

$$m = (2\pi/|k|D). \tag{3.37}$$

Equation (3.35) then gives the dispersion relation

$$\gamma = (\beta\pi|k|/D) = (2\pi^2\beta/mD^2). \tag{3.38}$$

So the component of δ with wavelength mD dies away exponentially with the characteristic lifetime

$$t = (mD^2/2\pi^2\beta). \tag{3.39}$$

From Eq. (3.39) we see in detail how the level-density $\rho(E,\tau)$ conforms locally to the prescribed density $r(E)$ on various time scales. After a time $\tau = mD^2$, ρ has adjusted itself to r everywhere over energy intervals of length mD. Beginning at $\tau = D^2$, the correlations between neighbouring levels are established. Successively longer-range adjustments are made as τ increases, until for $\tau > ND^2$ the global level distribution approaches $r(E)$ over the whole range of E, and the ensemble E_B satisfies requirement (1) everywhere.

It was proved in Sec. 2 that not only the one-level distribution ρ but the entire ensemble E_B tends to its limiting form E_S when $\tau > ND^2$. It is therefore reasonable to make the conjecture that *all the local many-level correlations within any group of m consecutive levels of a matrix in E_B tend to the corresponding correlations in the stationary ensemble E_S as soon as* $\tau > mD^2$. We have proved this statement for the one-level distribution $\rho(E,\tau)$ for any value of m, and for all the multiple correlations when $m = N$. Physical intuition suggests that the statement should be true in general, since the process of "thermal equilibrization" of the one-level distribution ρ within a given energy interval must go hand-in-hand with the equilibrization of the higher-order correlations within the same interval. The precise form of the many-level correlations in E_S will be determined in the following section.

4. EIGENVALUE CORRELATIONS

Mehta[18] has calculated analytically the correlation functions of all orders for the eigenvalues of a matrix in the Wigner ensemble E_G. His results take a simpler form in the limit as the total number of eigenvalues N tends to infinity. For any integer $n \leq N$, let

$$P_n(E_1,\ldots,E_n)dE_1\cdots dE_n \tag{4.1}$$

be the probability for finding an eigenvalue in each of the infinitesimal intervals $(E_1, E_1 + dE_1)\cdots(E_n, E_n + dE_n)$, irrespective of the positions of the remaining levels. This P_n is the nth-order correlation function for the ensemble in question. Suppose now that the mean level spacing at some place E is

$$D = [P_1(E)]^{-1}. \tag{4.2}$$

As $N \to \infty$, D will generally tend to zero. We suppose that the n variables $(E_1,\ldots,E_n)$ are all chosen in the neighbourhood of E, so that

$$E_j = E + Dx_j, \quad j = 1,\ldots,n. \tag{4.3}$$

We keep the quantities x_j fixed as $N \to \infty$, while E may vary and D tends to zero according to Eq. (4.2). Then Mehta proved that the limit

$$Y_n(x_1,\ldots,x_n) = \lim_{N\to\infty}[D^n P_n(E_1,\ldots,E_n)] \tag{4.4}$$

exists and is independent of E. The functions Y_n are called the *local* correlation functions for the infinite Wigner ensemble. If we know all the Y_n for a particular ensemble of matrices, then we have a complete description of the local statistical behavior of the sequence of eigenvalues of matrices in the ensemble as $N \to \infty$.

Mehta[18] found the explicit form of Y_n for the Wigner ensemble. Y_n is an nth order Pfaffian or quaternion determinant,

$$Y_n(x_1,\ldots,x_n) = Q\ \det(\sigma_1(x_i - x_j)), \tag{4.5}$$

where σ_1 is a quaternion with the $[2\times 2]$ matrix representation

$$\sigma_1(r) = \begin{bmatrix} s(r) & Ds(r) \\ Is(r) - \epsilon(r) & s(r) \end{bmatrix} \tag{4.6}$$

and

$$s(r) = (\sin(\pi r)/\pi r), \tag{4.7}$$

$$Ds(r) = \frac{ds(r)}{dr}, \tag{4.8}$$

$$Is(r) = \int_0^r s(r')dr', \tag{4.9}$$

$$\epsilon(r) = \tfrac{1}{2}\frac{r}{|r|}. \tag{4.10}$$

The meaning of the notation Q det is fully explained in Dyson,[19] where Dyson had shown that the identical set of local correlation-functions Y_n appeared in a different ensemble (the so-called circular orthogonal ensemble of unitary matrices), in the limit when the order of the matrices tends to infinity.

It is plausible that the local statistical behavior of an eigenvalue sequence should become independent of the global eigenvalue distribution when the total number of eigenvalues becomes large. We therefore conjecture that the behavior which Mehta found for the Wigner ensemble and Dyson found for the circular ensemble holds equally for the ensembles E_B and E_S, independently of the level-density function $r(E)$. Specifically, we conjecture that in the stationary ensemble E_S with $\beta = 1$, the limit (4.4) exists and is equal to (4.5) for any choice of $r(E)$. We conjecture that in the Brownian ensemble E_B with $\beta = 1$, the same limit exists with the same value, provided that the parameter τ defining E_B varies with N in such a way that

$$(\tau/D^2) \to \infty \text{ as } N \to \infty. \tag{4.11}$$

The condition (4.11) is weak and is easily satisfied, for example by taking τ constant. According to Eq. (3.39), the condition (4.11) should be sufficient to ensure that eigenvalue correlations come to statistical equilibrium over distances larger than mD for any finite value of m.

The conjectured local correlation functions in E_B are independent of H_0, $r(E)$, and τ, but they are not independent of β. For $\beta = 2$ and 4, a different set of functions replaces (4.5). Namely, for $\beta = 2$,

$$Y_n(x_1,\ldots,x_n) = \det(s(x_i - x_j)) \tag{4.12}$$

is an ordinary $(n\times n)$ determinant, while for $\beta = 4$,

$$Y_n(x_1,\ldots,x_n) = Q\ \det(\sigma_4(x_i - x_j)) \tag{4.13}$$

is a quaternion determinant with

$$\sigma_4(r) = \begin{bmatrix} s(2r) & Ds(2r) \\ Is(2r) & s(2r) \end{bmatrix}. \tag{4.14}$$

Equations (4.12) and (4.14) were proved by Mehta[18]

for the ensembles which are analogous to the Wigner ensemble E_0 with $\beta = 2$ and 4, namely the Gaussian ensembles of matrices whose elements are complex or quaternion real, respectively. The same Y_n are presumably valid for the ensembles E_β and E_S, when $\beta = 2, 4$.

The mathematical treatment of eigenvalue correlations is always simplest in the case $\beta = 2$. For this case, working with the ensemble E_S and an arbitrary distribution-function $r(E)$, Fox and Kahn[11] proved that the limit (1.4) exists and has the value (1.12). They obtained an explicit formula for the correlation functions in the ensemble E_S for finite N when $\beta = 2$, namely

$$P_n(E_1, \ldots, E_n) = \det[K_N(E_i, E_j)], \tag{4.15}$$

where

$$K_N(x, y) = f(x)\sum_{j=1}^{N} p_j(x)p_j(y), \tag{4.16}$$

and the $p_j(x)$ are orthogonal polynomials defined by

$$\int p_j(x)p_k(x)f(x)dx = \delta_{jk}, \tag{4.17}$$

with the weight-function

$$f(x) = \exp[-\beta u(x)] \tag{4.18}$$

and $u(x)$ given by Eq. (2.3). The existence and value of the limit (4.4) then follows from known theorems on the behavior of the kernel $K_N(x, y)$ for large N.

Exact expressions for the correlation functions, analogous to Eq. (4.15), exist also for the ensemble E_S in the cases $\beta = 1, \beta = 4$. Consider first the case $\beta = 1$. We introduce the family of skew-orthogonal polynomials $q_j(x)$ defined by

$$\int\int q_j(x)q_k(y)f(x)f(y)\epsilon(x - y)dxdy = Z_{jk}, \tag{4.19}$$

where Z_{jk} is the canonical antisymmetric matrix

$$Z_{jk} = \delta_{j+1,k} - \delta_{k+1,j}, \quad j + k + 1 \equiv 0 \ (\text{mod}4) \text{ only}. \tag{4.20}$$

We write

$$\varphi_j(x) = f(x)q_j(x), \quad \psi_j(x) = \int \epsilon(x - y)\varphi_j(y)dy, \tag{4.21}$$

and define the kernels

$$L(x, y) = \sum Z_{jk}\varphi_j(x)\psi_k(y), \tag{4.22}$$

$$L^\dagger(x, y) = -\sum Z_{jk}\psi_j(x)\varphi_k(y), \tag{4.23}$$

$$\lambda(x, y) = \sum Z_{jk}\varphi_j(x)\varphi_k(y), \tag{4.24}$$

$$\Lambda(x, y) = \sum Z_{jk}\psi_j(x)\psi_k(y), \tag{4.25}$$

$$\epsilon(x, y) = \epsilon(x - y), \tag{4.26}$$

the sums over j, k in Eqs. (4.22)–(4.25) going from 1 to N. When N is even, we define a quaternion kernel by the (2×2) matrix representation

$$Q(x, y) = \begin{bmatrix} L & -\lambda \\ \Lambda - \epsilon & L^\dagger \end{bmatrix}, \tag{4.27}$$

and the correlation-functions in E_S for $\beta = 1$ are given by

$$P_n(E_1, \ldots, E_n) = Q \det[Q(E_i, E_j)]. \tag{4.28}$$

When N is odd, the last skew-orthogonal polynomial $q_N(x)$ is unpaired, and its normalization is left arbitrary by Eq. (4.19). In this case we supplement Eq. (4.19) by the extra condition

$$\int q_j(x)f(x)dx = \delta_{jN}, \tag{4.29}$$

and introduce the additional kernels

$$M(x, y) = \varphi_N(x), \quad M^\dagger(x, y) = \varphi_N(y), \tag{4.30}$$

$$\mu(x, y) = \psi_N(x), \quad \mu^\dagger(x, y) = \psi_N(y). \tag{4.31}$$

The result (4.28) then still holds if we define Q for $\beta = 1, N$ odd, by

$$Q(x, y) = \begin{bmatrix} L + M & -\lambda \\ \Lambda - \epsilon + \mu - \mu^\dagger & L^\dagger + M^\dagger \end{bmatrix}, \tag{4.32}$$

instead of by (4.27).

When $\beta = 4$, we require a different family of skew-orthogonal polynomials $r_j(x)$ defined by

$$\int\int f(x)\epsilon(x - y)[r_j(x)r_k(y) - r_k(x)r_j(y)]dxdy = Z_{jk}, \tag{4.33}$$

instead of Eq. (4.19). Instead of Eq. (4.21), we take

$$\varphi_j(x) = f(x)r_j(x), \quad \psi_j(x) = \int \epsilon(x - y)r_j(y)dy. \tag{4.34}$$

The kernels $L, L^\dagger, \lambda$, and Λ are defined as before by Eqs. (4.22)–(4.25), but with the summations now running from 1 to $2N$. The correlation functions for $\beta = 4$ are then given by Eq. (4.28) with

$$Q(x, y) = \begin{bmatrix} L & -\lambda \\ \Lambda & L^\dagger \end{bmatrix}. \tag{4.35}$$

We omit here the proofs of the results (4.27), (4.28), (4.32), and (4.35). The proofs follow precisely the lines laid down in Dyson[19] and Mehta,[18] where the corresponding results were proved for the circular and Gaussian ensembles. The results are exhibited here in order to demonstrate that the ensemble E_S is mathematically tractable, in the sense that all eigenvalue correlation functions can be precisely calculated in analytic form. This is to justify our claim that the ensemble E_S satisfies item (4) in the list of requirements stated in Sec. 1.

It remains to be proved that the correlation functions (4.28) tend as $N \to \infty$ to the forms (4.5) and (4.14) which are valid for the circular and Gaussian ensembles with $\beta = 1, 4$. To prove this, we require an extension of some standard results concerning orthogonal polynomials to the case of skew-orthogonal polynomials. More generally, in order to make the explicit formulas of this section useful, we need to develop the theory of skew-orthogonal polynomials until it becomes a working tool as handy as the existing theory of orthogonal polynomials. We intend to come back to this subject in a future publication.

We have also left unproven our assertion that the en-

semble E_B in the limit (4.11) has the same universal local correlation functions given by Eq. (4.5), (4.12), and (4.13) in the cases $\beta = 1, 2, 4$, respectively. This assertion was strictly proved only for the stationary ensemble E_S with $\beta = 2$. For E_S in the cases $\beta = 1, 4$, the assertion is made mathematically plausible by the existence of the explicit formula (4.28) for the correlation functions. For the general Brownian ensemble E_B, it remains to be seen whether any similarly explicit formula for the correlation functions can be found.

Failing an exact formula, one could probably develop a sufficiently accurate theory of the behavior of the correlation functions of the Brownian ensemble for large N by following the "hydrodynamical" approximation scheme which was used in Sec. 3 to study the behavior of the one-level distribution function $\rho(E, \tau)$.

Note added in proof: The author has received a preprint, "Exact Results for a Quantum Many-body Problem in One Dimension. II," from Dr. B. Sutherland. Sutherland has found the exact eigenvalues and eigenfunctions of the operator (2.19) for the Gaussian and circular ensembles. It remains to be seen whether Sutherland's method also works for the ensemble E_S.

ACKNOWLEDGMENTS

This work was begun while the author was enjoying the hospitality of the Rockefeller University as a visiting lecturer. He is grateful to Professor George Uhlenbeck for supplying the initial stimulus and to Dr. Madan Lal Mehta for helpful conversations and for communication of unpublished results.

* Dedicated to Professor George Uhlenbeck on the occasion of his seventieth birthday.
[1] E. P. Wigner, *Proceedings of the Fourth Canadian Mathematical Congress* (U. of Toronto Press, Toronto, 1957), p. 174.
[2] E. P. Wigner, Ann. Math. **67**, 325 (1958).
[3] G. E. Uhlenbeck (private communication, 1971).
[4] O. Bohigas and J. Flores, Phys. Letters **34B**, 261 (1971).
[5] O. Bohigas and J. Flores, Phys. Letters **35B**, 383 (1971).
[6] J. B. French and S. S. M. Wong, Phys. Letters **33B**, 449 (1970).
[7] J. B. French and S. S. M. Wong, Phys. Letters **35B**, 5 (1971).
[8] H. S. Leff, Phys. Rev. **136**, A355 (1964).
[9] H. S. Leff, J. Math. Phys. **5**, 756 (1964).
[10] H. S. Leff, J. Math. Phys. **5**, 763 (1964).
[11] D. Fox and P. B. Kahn, Phys. Rev. **134**, B1151 (1964).
[12] M. L. Mehta, *Random Matrices* (Academic, New York, 1967).
[13] F. J. Dyson, J. Math. Phys. **3**, 1191 (1962).
[14] F. J. Dyson, J. Math. Phys. **3**, 1199 (1962).
[15] F. J. Dyson, J. Math. Phys. **3**, 157 (1962).
[16] M. L. Mehta, "On Spacing Distributions for Large Spacings," Saclay, Preprint DPh-T/71/12 (1971).
[17] H. Widom, "The Strong Szegö Limit Theorem for Circular Arcs," University of California, Santa Cruz, Preprint (1971).
[18] M. L. Mehta, Commun. Math. Phys. **20**, 245 (1971).
[19] F. J. Dyson, Commun. Math. Phys. **19**, 235 (1970). This definition of quaternion determinants is due to E. H. Moore, "General Analysis, Part I," Memoirs of the American Philosophical Society, (Philadelphia, 1935), p. 116.

Currents in Classical Field Theories

Robert Hermann*
Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903
(Received 13 November 1970)

The current vector field associated with a one-parameter group of transformations of a classical field is defined in a coordinate-free way.

1. INTRODUCTION

The aim of this paper is to provide a coordinate-free definition in the language of manifold and differential-operator theory[1,2,3] for the "current" vector-field associated with a one-parameter group of transformations of a "classical" field. Our approach will be somewhat different from that used by Trautman[4]. Such a formula should be useful for proving general results about the relations between "fields" and "currents," in both classical and quantum field theories, although in this paper the goal is the more modest one of showing that such a formula exists.

We will now describe the differential-geometric setting for the formula. Let E and M be manifolds, with $\pi: E \to M$ a map that defines E as a local product fiber space over M. (In the physical applications, M will be R^4, space-time, and the fibers of E will represent the "values" of the field, i.e., a cross-section map $\gamma: M \to E$ will represent a "field".) Let $J^1(E)$ be the bundle of one-jets[3] of cross sections of E, and let $L: J^1(E) \to R$ be a real-valued function on $J^1(E)$. (L should be regarded as the "Lagrangian" of the field[2,3].) Let X be a vector field (in the sense of manifold theory[1]) on E that is "projectable" under π, i.e., there is a vector field X_M on M such that:

$$\pi^*(X_M(f)) = X(\pi^*(f)) \tag{1.1}$$

for all $f \epsilon F(M)$.

[We adopt the differential-geometric notations of Ref. 1. In particular, $F(M)$ denotes the C^∞, real-valued functions on M and π^*: $F(M) \to F(E)$ denotes the pull-back map on functions defined by π.] Such an X generates a one-parameter group of diffeomorphisms of E that maps fibers into fibers, hence, acts on the space of cross sections of the fiber space E, i.e., physically, the group acts on the "classical fields." Condition (1.1) can also be interpreted group theoretically. It is the condition that π intertwine the one-parameter group of diffeomorphisms generated by X and X_M on E and M.

Suppose now that L, a volume element form "dp" on M, and a cross section map $\gamma: M \to E$, are fixed. We will then show that to each vector field X on E that satisfies (1.1) there is a vector field Y on M that may be regarded as the "current" associated with X.

semble E_B in the limit (4.11) has the same universal local correlation functions given by Eq. (4.5), (4.12), and (4.13) in the cases $\beta = 1, 2, 4$, respectively. This assertion was strictly proved only for the stationary ensemble E_S with $\beta = 2$. For E_S in the cases $\beta = 1, 4$, the assertion is made mathematically plausible by the existence of the explicit formula (4.28) for the correlation functions. For the general Brownian ensemble E_B, it remains to be seen whether any similarly explicit formula for the correlation functions can be found.

Failing an exact formula, one could probably develop a sufficiently accurate theory of the behavior of the correlation functions of the Brownian ensemble for large N by following the "hydrodynamical" approximation scheme which was used in Sec. 3 to study the behavior of the one-level distribution function $\rho(E, \tau)$.

Note added in proof: The author has received a preprint, "Exact Results for a Quantum Many-body Problem in One Dimension. II," from Dr. B. Sutherland. Sutherland has found the exact eigenvalues and eigenfunctions of the operator (2.19) for the Gaussian and circular ensembles. It remains to be seen whether Sutherland's method also works for the ensemble E_S.

ACKNOWLEDGMENTS

This work was begun while the author was enjoying the hospitality of the Rockefeller University as a visiting lecturer. He is grateful to Professor George Uhlenbeck for supplying the initial stimulus and to Dr. Madan Lal Mehta for helpful conversations and for communication of unpublished results.

* Dedicated to Professor George Uhlenbeck on the occasion of his seventieth birthday.
[1] E. P. Wigner, *Proceedings of the Fourth Canadian Mathematical Congress* (U. of Toronto Press, Toronto, 1957), p. 174.
[2] E. P. Wigner, Ann. Math. **67**, 325 (1958).
[3] G. E. Uhlenbeck (private communication, 1971).
[4] O. Bohigas and J. Flores, Phys. Letters **34B**, 261 (1971).
[5] O. Bohigas and J. Flores, Phys. Letters **35B**, 383 (1971).
[6] J. B. French and S. S. M. Wong, Phys. Letters **33B**, 449 (1970).
[7] J. B. French and S. S. M. Wong, Phys. Letters **35B**, 5 (1971).
[8] H. S. Leff, Phys. Rev. **136**, A355 (1964).
[9] H. S. Leff, J. Math. Phys. **5**, 756 (1964).
[10] H. S. Leff, J. Math. Phys. **5**, 763 (1964).
[11] D. Fox and P. B. Kahn, Phys. Rev. **134**, B1151 (1964).
[12] M. L. Mehta, *Random Matrices* (Academic, New York, 1967).
[13] F. J. Dyson, J. Math. Phys. **3**, 1191 (1962).
[14] F. J. Dyson, J. Math. Phys. **3**, 1199 (1962).
[15] F. J. Dyson, J. Math. Phys. **3**, 157 (1962).
[16] M. L. Mehta, "On Spacing Distributions for Large Spacings," Saclay, Preprint DPh-T/71/12 (1971).
[17] H. Widom, "The Strong Szegö Limit Theorem for Circular Arcs," University of California, Santa Cruz, Preprint (1971).
[18] M. L. Mehta, Commun. Math. Phys. **20**, 245 (1971).
[19] F. J. Dyson, Commun. Math. Phys. **19**, 235 (1970). This definition of quaternion determinants is due to E. H. Moore, "General Analysis, Part I," Memoirs of the American Philosophical Society, (Philadelphia, 1935), p. 116.

Currents in Classical Field Theories

Robert Hermann*
Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903
(Received 13 November 1970)

The current vector field associated with a one-parameter group of transformations of a classical field is defined in a coordinate-free way.

1. INTRODUCTION

The aim of this paper is to provide a coordinate-free definition in the language of manifold and differential-operator theory[1,2,3] for the "current" vector-field associated with a one-parameter group of transformations of a "classical" field. Our approach will be somewhat different from that used by Trautman[4]. Such a formula should be useful for proving general results about the relations between "fields" and "currents," in both classical and quantum field theories, although in this paper the goal is the more modest one of showing that such a formula exists.

We will now describe the differential-geometric setting for the formula. Let E and M be manifolds, with $\pi: E \to M$ a map that defines E as a local product fiber space over M. (In the physical applications, M will be R^4, space-time, and the fibers of E will represent the "values" of the field, i.e., a cross-section map $\gamma: M \to E$ will represent a "field".) Let $J^1(E)$ be the bundle of one-jets[3] of cross sections of E, and let $L: J^1(E) \to R$ be a real-valued function on $J^1(E)$. (L should be regarded as the "Lagrangian" of the field[2,3].) Let X be a vector field (in the sense of manifold theory[1]) on E that is "projectable" under π, i.e., there is a vector field X_M on M such that:

$$\pi^*(X_M(f)) = X(\pi^*(f)) \tag{1.1}$$

for all $f \epsilon F(M)$.

[We adopt the differential-geometric notations of Ref. 1. In particular, $F(M)$ denotes the C^∞, real-valued functions on M and π^*: $F(M) \to F(E)$ denotes the pull-back map on functions defined by π.] Such an X generates a one-parameter group of diffeomorphisms of E that maps fibers into fibers, hence, acts on the space of cross sections of the fiber space E, i.e., physically, the group acts on the "classical fields." Condition (1.1) can also be interpreted group theoretically. It is the condition that π intertwine the one-parameter group of diffeomorphisms generated by X and X_M on E and M.

Suppose now that L, a volume element form "dp" on M, and a cross section map $\gamma: M \to E$, are fixed. We will then show that to each vector field X on E that satisfies (1.1) there is a vector field Y on M that may be regarded as the "current" associated with X.

We will give an explicit but coordinate free formula for Y that will facilitate the understanding of its mathematical and physical properties.

2. THE CURRENT ASSOCIATED WITH A VECTOR FIELD

Let $V(E,M)$ denote the vector fields on E that are projectable under π, i.e., satisfy condition (1.1). Let $\Gamma(E)$ denote the space of all cross-section maps $M \to E$. Let $J^1(E)$ denote the space of 1-jets of cross sections.[3] Then, denoting by π^1 the projection map $J^1(E) \to M$, it defines $J^1(E)$ as a fiber space over M. Given $\gamma \in \Gamma(E)$, one can define[3] a cross-section map $j^1(\gamma): M \to J^1(E)$ called the 1-jet of γ.

There is a Lie algebra homomorphism $V(E,M) \to V(J^1(E),M)$, denoted by $X \to X^1$; the vector field X^1 on $J^1(E)$ is called the *first order prolongation* of X. We will give an explicit formula for X^1 later on. Let "dp" denote a fixed volume element differential form on M, i.e., "dp" is an everywhere nonzero m-differential form on M $(m = \dim M)$. Let $L: J^1(E) \to R$ be a Lagrangian, and let ω be the following M form on $J^1(E)$:

$$\omega = L\pi^{1*}(dp). \tag{2.1}$$

Let $X \in V(E,M)$, and $\gamma \in \Gamma(E)$. Given $f \in F(M)$, let fX denote the vector field $\pi^*(f)X$ in $V(E,M)$. Let $(fX)^1$ be the first-order prolongation of this vector field to be a vector field on $J^1(E)$. Now, define Y by the following formula:

$$Y(f)dp = j^1(\gamma)^*((fX)^1(\omega)) - fj^1(\gamma)^*(X^1(\omega)). \tag{2.2}$$

[$(fX)^1(\omega)$ denotes the Lie derivative[1] of the form ω on $J^1(E)$ by the vector field $(fX)^1$.]

It should be clear that Y defined by 2.2 is a linear map $F(M) \to F(M)$. We will show that it is indeed a vector field by calculating its value in a local coordinate system. Suppose (x_μ), $0 \le \mu, \nu \le m-1$, are functions on M that form a coordinate system for M. We suppose that E is a product $M \times N$ of M with a manifold N. Let (ϕ_a), $1 \le a, b \le n$, be coordinates of N. Then, (x_μ, ϕ_a) defines a coordinate system for E. Suppose a cross section $\gamma: M \to E$ is determined by functions $\phi_a(x)$. There are functions $\phi_{a\mu}$ on $J^1(E)$ such that

$$J^1(\gamma)^*(\phi_{a\mu}) = \frac{\partial}{\partial x_\mu}(\phi_a(x)).$$

Then, the functions $(x_\mu, \phi_a, \phi_{a\mu})$ form a coordinate system for $J^1(E)$. The Lagrangian L becomes a function $L(x, \phi, \phi_{a\mu})$ of these variables. Let:

$$dp = dx_0 \wedge \cdots \wedge dx_{m-1},$$

$$L_\mu = \frac{\partial L}{\partial x_\mu} \qquad L_a = \frac{\partial L}{\partial \phi_a},$$

$$L_{a\mu} = \frac{\partial L}{\partial \phi_{a\mu}},$$

$$\partial_a = \frac{\partial}{\partial \phi_a}, \quad \partial_\mu = \frac{\partial}{\partial x_\mu}, \quad \partial_{a\mu} = \frac{\partial}{\partial \phi_{a\mu}}.$$

Suppose $X \in V(E,M)$ is of the following form:

$$X = A_\mu(x)\partial_\mu + A_a(x,\phi)\partial_a. \tag{2.3}$$

(It is readily seen that such an X admits a description of this form.) Then, one can prove that X^1, its prolongation, is of the following form:

$$X^1 = A_\mu\partial_\mu + A_a\partial_a + [\partial_\mu(A_a) - \phi_{a\nu}\partial_\mu(A_\nu) + [\partial_b(A_a)\phi_{b\mu}]\partial_{a\mu} \tag{2.4}$$

We can now work out Y, given by (2.2) in these coordinates. If $X \in V(E,M)$ is of form (2.3), then

$$X^1(\omega) = X^1(L)dp + LX^1(dp)$$
$$= [A_\mu L_\mu + A_a L_a + (\partial_\mu A_a - \phi_{a\nu}\partial_\mu(A_\nu) + \partial_b(A_a)\phi_{b\mu})L_{a\mu}]dp + L\partial_\mu(A_\mu)dp.$$

Thus, for $f \in F(M)$,

$$(fX)^1(\omega) - f(X^1(\omega)) = [(\partial_\mu(f)A_a - \phi_{a\nu}\partial_\mu(f)A_\nu)L_{a\mu} + L\partial_\mu(f)A_a]dp.$$

Then, we see finally that (2.2) takes the following form:

$$Y = [j^1(\gamma)^*(A_a L_{a\mu} - \phi_{a\nu}A_\nu L_{a\mu} + LA_\mu)]\partial_\mu. \tag{2.5}$$

In particular, we see that Y, considered as a map: $F(M) \to F(M)$ is a first order linear differential operator, i.e., a vector field in $V(M)$. The reader will readily verify that, modulo a changed notation and interpretation, (2.4) is indeed the formula for a "current" generated by a group of transformations on the fields that is to be found in all quantum field theory books in one form or another.

3. CONSERVATION PROPERTIES OF CURRENTS

Recall[1] how the divergence of a vector field Y on M is defined:

$$Y(dp) = \operatorname{div}(Y)dp.$$

Let us continue with the notations and coordinate systems used in Sec. 2. Then, one readily obtains the following formula:

$$\text{If } Y = B_\mu(x)\partial_\mu, \text{ then } \operatorname{div} Y = \partial_\mu(B_\mu). \tag{3.1}$$

Our goal now is to calculate the divergence of the current vector field Y, given by formula (2.5) in case the cross-section map $\gamma: M \to E$ is an *extremal* of the Lagrangian L, i.e., satisfies the following Euler equations:

$$\partial_\mu(j^1(\gamma)^*(L_{a\mu})) = j^1(\gamma)^*(L_a). \tag{3.2}$$

From 2.4, we see that

$$B_\mu = j^1(\gamma)^*(A_a L_{a\mu} - \phi_{a\nu}A_\nu L_{a\mu} + LA_\mu). \tag{3.3}$$

With X defined by (2.3) and X_M by (1.1), we have

$$X_M = A_\mu\partial_\mu; \tag{3.4}$$

hence

$$\operatorname{div} X = \partial_\mu(A_\mu). \tag{3.5}$$

Also, using (2.4),

$$X^1(L) = A_\mu L_\mu + A_a L_a + [\partial_\mu(A_a) - \phi_{a\nu}\partial_\mu(A_\nu) + \partial_b(A_a)\phi_{b\mu}]L_{a\mu};$$

hence

$$j^1(\gamma)^*(X^1(L)) = j^1(\gamma)^*(A_\mu L_\mu + A_a L_a) + \partial_\mu(j^1(\gamma)^*(A_a))j^1(\gamma)^*(L_{a\mu}) - \partial_\nu(\phi_a)\partial_\mu(A_\nu). \quad (3.6)$$

Combining (3.2), (3.3), and (3.6), we have

$$\begin{aligned}\partial_\mu(B_\mu) = {} & \partial_\mu(j^1(\gamma)^*(A_a))j^1(\gamma)^*(L_{a\mu}) + j^1(\gamma)^*(A_a L_a) \\ & - \partial_\mu\partial_\nu(\phi_a)A_\nu j^1(\gamma)^*(L_{a\mu}) - \partial_\nu(\phi_a)\partial_\mu(A_\nu) \\ & \times j^1(\gamma)^*(L_{a\mu}) - \partial_\nu(\phi_a)A_\nu j^1(\gamma)^*(L_a) \\ & + [j^1(\gamma)^*(L_\mu) + j^1(\gamma)^*(L_a)\partial_\mu\phi_a \\ & + j^1(\gamma)^*(L_{a\mu})\partial_\mu\partial_\nu\phi_a]A_\mu \\ & + j^1(\gamma)^*(L)\partial_\mu(A_\mu) = j^1(\gamma)^*[X^1(L) + \partial_\mu(A_\mu)L].\end{aligned}$$

Hence, we have

$$\mathrm{div}Y = j^1(\gamma)^*[X^1(L) + \mathrm{div}X_M L] \quad (3.7)$$

or, since $d(Y \lrcorner p) = Y(dp) = \mathrm{div}(Y)dp$,

$$d(Y \lrcorner dp) = j^1(\gamma)^*(X^1(\omega)). \quad (3.8)$$

(3.8) is the main formula for the study of the connection between "conserved currents" and "symmetries" of L. Let us say that X is an *infinitesimal symmetry of the Lagrangian* L if

$$0 = X^1(\omega) = X^1(L\pi^{1*}(dp)). \quad (3.9)$$

Let us say that a vector field Y on M is a *conserved current* if

$$d(Y \lrcorner dp) = 0. \quad (3.10)$$

If (3.10) is satisfied, and if N is a submanifold of M of one lower dimension, one can set

$$f(\gamma, N, Y) = \int_N Y \lrcorner dp. \quad (3.11)$$

Then, the conservation condition (3.10) guarantees using Stokes' formula[1] that (3.11) really does not depend on the choice of the "Cauchy data" submanifold N, hence a define a real valued function $\gamma \to f(\gamma, N, Y)$ in the space of all extremal cross sections γ; this function then defines what one might call a "classical conserved observable." (In the quantum field theoretic version of these ideas, (3.11) is an operator in Hilbert space which commutes with the Hamiltonian, i.e., a "conserved charge.") Let us then restate formula (3.8) in the following way:

Theorem 1: If X is a vector field on E that generates an infinitesimal symmetry of the Lagrangian L, then formula (2.2) defines a vector field Y on M that generates a conserved current.

---

* Supported by the Office of Naval Research under contract NR-041-378. Reproduction in whole or part is permitted for any purpose of the United States Government.

[1] R. Hermann, *Differential Geometry and the Calculus of Variations* (Academic, New York, 1968).
[2] R. Hermann, *Lie Algebras and Quantum Mechanics* (Benjamin, New York, 1970).
[3] R. Hermann, *Vector Bundles in Mathematical Physics* (Benjamin, New York, 1970), Vol. I.
[4] A Trautman, Commun. Math. Phys. 6, 248 (1967).

Velocity-Dominated Singularities in Irrotational Dust Cosmologies

D. Eardley, E. Liang,* and Rainer Sachs*
University of California at Berkeley, Berkeley, California 94720
(Received 12 March 1971)

We consider irrotational dust solutions of the Einstein equations. We define "velocity-dominated" singularities of these solutions. We show that a velocity-dominated singularity can be considered as a three-dimensional manifold with an invariantly and uniquely defined inner metric tensor, extrinsic curvature tensor, and scalar bang time function. We compute this structure for a variety of known exact models. The structure of the singularity uniquely determines the solution in a certain class of spatially inhomogeneous models. We briefly discuss the b boundary (Schmidt boundary). In an appendix we generalize conformal transformations to "stretch" transformations and calculate the curvature form of a stretched metric.

I. INTRODUCTION

A. Motivation

If one thinks gravity may play a dominant role even near the big bang, one would like an invariant characterization of the structure of the cosmological singularity for at least two reasons.

First, the present lumps in the universe may be due to conditions holding so early that we should attribute these conditions to the singularity itself in our models. Then one could perhaps see what conditions on the singularity give reasonable lumps. Perhaps one could even make initial conditions part of physics by postulating field equations for the intrinsic structure.

Second, it is at least conceivable that the big bang is still going on in isolated parts of the universe.[1] Let us consider an irrotational, hydrodynamic, general relativistic model, so that cosmological time is well defined. The exact spherically symmetric Tolman–Bondi dust models[2] and linearized perturbations of Robertson–Walker models both indicate that it would be very artificial to require that the big bang went off simultaneously in the past of each matter world line. We might reasonably require that the singularity be in some sense spacelike, but it turns out that this is a very mild restriction. In fact there are Tolman–Bondi models which are homogeneous (throughout this paper "homogeneous" means spatially homogeneous)

$$X^1(L) = A_\mu L_\mu + A_a L_a + [\partial_\mu(A_a) - \phi_{a\nu}\partial_\mu(A_\nu) + \partial_b(A_a)\phi_{b\mu}]L_{a\mu};$$

hence

$$j^1(\gamma)^*(X^1(L)) = j^1(\gamma)^*(A_\mu L_\mu + A_a L_a) + \partial_\mu(j^1(\gamma)^*(A_a))j^1(\gamma)^*(L_{a\mu}) - \partial_\nu(\phi_a)\partial_\mu(A_\nu). \quad (3.6)$$

Combining (3.2), (3.3), and (3.6), we have

$$\begin{aligned}\partial_\mu(B_\mu) = {} & \partial_\mu(j^1(\gamma)^*(A_a))j^1(\gamma)^*(L_{a\mu}) + j^1(\gamma)^*(A_a L_a) \\ & - \partial_\mu\partial_\nu(\phi_a)A_\nu j^1(\gamma)^*(L_{a\mu}) - \partial_\nu(\phi_a)\partial_\mu(A_\nu) \\ & \times j^1(\gamma)^*(L_{a\mu}) - \partial_\nu(\phi_a)A_\nu j^1(\gamma)^*(L_a) \\ & + [j^1(\gamma)^*(L_\mu) + j^1(\gamma)^*(L_a)\partial_\mu\phi_a \\ & + j^1(\gamma)^*(L_{a\mu})\partial_\mu\partial_\nu\phi_a]A_\mu \\ & + j^1(\gamma)^*(L)\partial_\mu(A_\mu) = j^1(\gamma)^*[X^1(L) + \partial_\mu(A_\mu)L].\end{aligned}$$

Hence, we have

$$\mathrm{div}Y = j^1(\gamma)^*[X^1(L) + \mathrm{div}X_M L] \quad (3.7)$$

or, since $d(Y \lrcorner p) = Y(dp) = \mathrm{div}(Y)dp$,

$$d(Y \lrcorner dp) = j^1(\gamma)^*(X^1(\omega)). \quad (3.8)$$

(3.8) is the main formula for the study of the connection between "conserved currents" and "symmetries" of L. Let us say that X is an *infinitesimal symmetry of the Lagrangian* L if

$$0 = X^1(\omega) = X^1(L\pi^{1*}(dp)). \quad (3.9)$$

Let us say that a vector field Y on M is a *conserved current* if

$$d(Y \lrcorner dp) = 0. \quad (3.10)$$

If (3.10) is satisfied, and if N is a submanifold of M of one lower dimension, one can set

$$f(\gamma, N, Y) = \int_N Y \lrcorner dp. \quad (3.11)$$

Then, the conservation condition (3.10) guarantees using Stokes' formula[1] that (3.11) really does not depend on the choice of the "Cauchy data" submanifold N, hence a define a real valued function $\gamma \to f(\gamma, N, Y)$ in the space of all extremal cross sections γ; this function then defines what one might call a "classical conserved observable." (In the quantum field theoretic version of these ideas, (3.11) is an operator in Hilbert space which commutes with the Hamiltonian, i.e., a "conserved charge.") Let us then restate formula (3.8) in the following way:

Theorem 1: If X is a vector field on E that generates an infinitesimal symmetry of the Lagrangian L, then formula (2.2) defines a vector field Y on M that generates a conserved current.

---

* Supported by the Office of Naval Research under contract NR-041-378. Reproduction in whole or part is permitted for any purpose of the United States Government.

[1] R. Hermann, *Differential Geometry and the Calculus of Variations* (Academic, New York, 1968).
[2] R. Hermann, *Lie Algebras and Quantum Mechanics* (Benjamin, New York, 1970).
[3] R. Hermann, *Vector Bundles in Mathematical Physics* (Benjamin, New York, 1970), Vol. I.
[4] A Trautman, Commun. Math. Phys. 6, 248 (1967).

Velocity-Dominated Singularities in Irrotational Dust Cosmologies

D. Eardley, E. Liang,* and Rainer Sachs*
University of California at Berkeley, Berkeley, California 94720
(Received 12 March 1971)

We consider irrotational dust solutions of the Einstein equations. We define "velocity-dominated" singularities of these solutions. We show that a velocity-dominated singularity can be considered as a three-dimensional manifold with an invariantly and uniquely defined inner metric tensor, extrinsic curvature tensor, and scalar bang time function. We compute this structure for a variety of known exact models. The structure of the singularity uniquely determines the solution in a certain class of spatially inhomogeneous models. We briefly discuss the b boundary (Schmidt boundary). In an appendix we generalize conformal transformations to "stretch" transformations and calculate the curvature form of a stretched metric.

I. INTRODUCTION

A. Motivation

If one thinks gravity may play a dominant role even near the big bang, one would like an invariant characterization of the structure of the cosmological singularity for at least two reasons.

First, the present lumps in the universe may be due to conditions holding so early that we should attribute these conditions to the singularity itself in our models. Then one could perhaps see what conditions on the singularity give reasonable lumps. Perhaps one could even make initial conditions part of physics by postulating field equations for the intrinsic structure.

Second, it is at least conceivable that the big bang is still going on in isolated parts of the universe.[1] Let us consider an irrotational, hydrodynamic, general relativistic model, so that cosmological time is well defined. The exact spherically symmetric Tolman–Bondi dust models[2] and linearized perturbations of Robertson–Walker models both indicate that it would be very artificial to require that the big bang went off simultaneously in the past of each matter world line. We might reasonably require that the singularity be in some sense spacelike, but it turns out that this is a very mild restriction. In fact there are Tolman–Bondi models which are homogeneous (throughout this paper "homogeneous" means spatially homogeneous)

and isotropic far from the "center," have a singularity not reachable by future pointing timelike geodesics, but have a big bang which near the center is still going off (Fig. 1). The central singularities, assuming a reasonable opacity law, appear very small from the outside.[3] If these models can be generalized to more realistic equations of state, the result would be possible models for quasars or perhaps even galactic centers.

To try to analyze either possibility one really needs to assign an invariant structure to the singularity. The Tolman–Bondi singularities discussed do have two obvious kinds of structure: the rate at which baryons are being fed into the nonsingular parts of the universe, and the time at which a singularity appears on a particular matter line. How about other models?

The main purpose of this paper is to obtain the structure for one special case: irrotational dust with a velocity-dominated singularity.

B. Outline

In Sec. II we define a velocity-dominated singularity. Friedmann models are velocity-dominated because the effects of spatial curvature, which are determined by $k = \pm 1$ or 0, are negligible at early times. One can explicitly integrate the Einstein evolution equations to find the general form of the metric near a velocity-dominated singularity. Some functions obtained by the integration are identified as the positive definite metric, extrinsic curvature tensor, and scalar bang time of a three-dimensional manifold identified as the singularity. We show that this structure is uniquely and invariantly defined. Finally we discuss the constraints placed on this structure by the constraint equations and the requirement that our approximation be self-consistent.

In Sec. III we work out the metric, extrinsic curvature, and bang time for a variety of exact solutions. The largest known class of exact solutions obeys our velocity-dominated assumption. In this class the structure of the singularity uniquely determines the model. In Sec. IV we give some preliminary results on the b boundary.[4] In Sec. V we give a short summary. In an appendix we analyze metrics obtained from some other metric by stretching along N orthonormal directions with scalar functions. We use the Cartan formalism to calculate the curvature form. The relations derived include those for conformally related metrics and those for diagonalizable metrics as special cases.

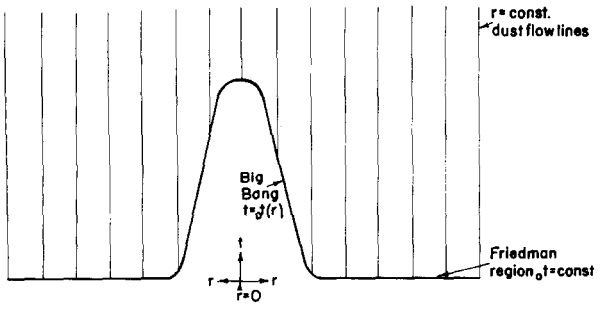

Fig. 1. "Bump" singularity at the central region of a loosely bound ($\beta = 0$) Tolman–Bondi model. The singularity is of the $(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3})$ type in the ${}_0t, {}_3 \neq 0$ region. As ${}_0t, {}_3 \to 0$, the singularity becomes Friedman-like. There should in general be a (1, 0, 0) singularity beneath the bump (see Fig. 2). But this does not influence our region of interest (above the singularity) here.

Many of our methods are essentially due to Misner.[5] Formally, our results are similar to those given in a classic paper by Lifshitz and Khalatnikov.[6]

II. THE STRUCTURE OF THE SINGULARITY

We work in a coordinate patch of a C^∞ manifold throughout. A solution of the Einstein dust equations $G^\mu_\nu = -\rho u^\mu u_\nu$ ($\mu, \nu = 0 \cdots 3$, $u^\mu u_\mu = -1$) is called irrotational if $u_{\mu;\nu} = u_{\nu;\mu}$. In such models there exist "comoving geodesic normal" coordinates characterized by[7]

$$ds^2 = -dt^2 + g_{ab}(x^c, t)dx^a dx^b, \qquad a, b = 1 \cdots 3,$$
$$u^\mu = \delta^\mu_0. \tag{1}$$

The only allowed transformations are

$$x'^a = x'^a(x^b), \tag{2a}$$

$$t' = t + \text{const}. \tag{2b}$$

We shall denote covariant differentiation within the hypersurfaces $t = \text{const}$ with respect to $g_{ab}(x^c, \text{const})$ by a vertical stroke, the three-dimensional Einstein tensor by ${}^3G^a_b$, and the second fundamental form (the field "velocity") by K^a_b:

$$\frac{1}{2}\frac{\partial g_{ab}}{\partial t} = K^c_b g_{ac}. \tag{3}$$

We let $\alpha = +[\det(g_{ab})]^{1/2}$. Then the "evolution equations" are (3) and[7]

$$-G^a_b = -{}^3G^a_b + \left(\frac{\partial}{\partial t} + K\right)L^a_b - \delta^a_b\left[\frac{2}{3}\left(\frac{\partial}{\partial t} + \frac{K}{2}\right)K + \tfrac{1}{2}L^2\right] = 0. \tag{4}$$

Here

$$K = K^a_a, \qquad L^a_b = K^a_b - \tfrac{1}{3}\delta^a_b K, \qquad L^2 = L^b_a L^a_b. \tag{5}$$

Equation (5) implies $L^2 \geq 0$, $L^a_a = 0$, $K = \partial \ln\alpha/\partial t$. The "constraint equations" are

$$G^0_b = K^a_{b/a} - K_{,b} = 0; \tag{6}$$

$G_{00} = \frac{1}{2}(2\,{}^3G + L^2 - \frac{2}{3}K^2)$ determines ρ. Assuming (3), we find that the evolution equations (4) are the Euler–Lagrange equations of $I = \int \alpha d^4x(L^2 - \frac{2}{3}K^2 - 2\,{}^3G)$. Suppose a metric obeys just the evolution equations (3) and (4). Then the Bianchi identities $G^\mu_{\nu;\mu} = 0$ give, after a short calculation,

$$G^0_b = {}_0G^0_b\alpha^{-1}. \tag{7}$$

Here ${}_0G^0_b(x^c)$ is a function of integration, independent of t, but in general a function of x^c. We shall use the left subscript 0 in this manner throughout. From (7) we see that (3), (4), and (6) are involutive. Similarly, (3), (4), and (6) imply

$$G^0_0 = {}_0G^0_0\alpha^{-1}. \tag{8}$$

In the approximations which follow, it will be convenient to replace the evolution equations by integro-differential equations, a trick familiar in linear theories.

Suppose for a moment ${}^3G^a_b$ and α are, miraculously, known. Then the trace-free part of (4) can be integrated immediately:

$$-L^a_b = \alpha^{-1}\left({}_0M^a_b + \int {}^3H^a_b\alpha\right). \tag{9}$$

Here $\int f$ means

$$\int_{A(x^c)}^{t} dt' f(x^c, t'),$$

with x^c fixed: ${}_0M^b_a$ transforms as a tensor density under (2a).

If we now consider L^a_b as known and if

$$L^a_b(t', x^c)L^b_c(t, x^c) = L^a_b(t, x^c)L^b_c(t', x^c),$$

the trace-free part of (3) can likewise be integrated:

$$g_{ab} = \alpha^{2/3}{}_0h_{ac}\exp\left(2\int L^c_b\right), \tag{10}$$

where $A_{ac}\exp(B^c_b)$ means $A_{ac}[\delta^c_d + B^c_dB^d_b(2!)^{-1} + \cdots]$. In general, the exponential term in (10) must be replaced by an ordered product. Finally, combining the traces of (3) and (4) yields the following integro-differential equation for α:

$$\alpha\frac{\partial^2\alpha}{\partial t^2} - \frac{1}{2}\frac{\partial\alpha^2}{\partial t} - \frac{\alpha^2}{2}{}^3G = -\tfrac{3}{4}\left({}_0M^a_b + \int {}^3H^a_b\alpha\right)\left({}_0M^b_a + \int {}^3H^b_a\alpha\right). \tag{11}$$

A solution of (11) substituted into (9) and (10) then converts (10) into a covariant, but very messy, functional equation for g_{ab}, with Einstein tensor ${}^3G^a_b$.

By Raychaudhuri's theorem $\rho(x^c, t) > 0$ implies $\rho = \infty$ for some $t = {}_0t(x^c)$. We call ${}_0t$ the bang time and assume throughout that ${}_0t$ is a C^∞ function of x^c; ${}_0t$ transforms as a scalar under (2a). Thus $t = {}_0t(x^c)$ becomes a three-dimensional C^∞ differential manifold which we call the "singularity." Now the essential idea suggested by the Friedmann models is to drop the ${}^3G^a_b$ terms from (9)–(11) and use the result as a first approximation to g_{ab} near the singularity. To make this idea somewhat more precise, we use the following definition.

Let $\hat{g}_{ab}$, with corresponding $\hat{K}^a_b$ and bang-time function ${}_0\hat{t} \equiv {}_0t$ be an exact dust solution. Suppose there is a real g_{ab} which near ${}_0t$ obeys

(a) $g_{ab} = g_{ba}$, signature $(g_{ab}) = +3$.

(b) g_{ab} obeys (3) and

$$\left(\frac{\partial}{\partial t} + K\right)L^a_b = \delta^a_b\left[\frac{2}{3}\left(\frac{\partial}{\partial t} + \frac{K}{2}\right)K + \tfrac{1}{2}L^2\right]. \tag{4'}$$

(c) There is some component, say u, of g_{ab} such that for x^c fixed and $t \to {}_0t$,

$$(\hat{K}^a_b - K^a_b)(K^c_dK^d_c)^{-1/2} \to 0,$$
$$(\hat{g}_{ab} - g_{ab})u^{-1} \to 0. \tag{12}$$

Then we call g_{ab} and its singularity "velocity dominated" and call g_{ab} the first approximation. (4′) implies that both sides are separately zero and that we can set ${}^3G^a_b = 0$ in (9)–(11). Condition (a) may or may not be independent of (b) and (c). Not every solution of (3) and (4′) is a first approximation since (6), $\rho > 0$, and the consistency condition

$${}^3G^a_b(K^c_dK^d_c)^{-1} \to 0 \tag{13}$$

place restrictions on g_{ab}. The definition is invariant under the transformations (2).

To find a general first approximation, we must now set ${}^3G^a_b = 0$ in (9)–(11), analyze uniqueness from (12), analyze (6) and (13), and finally estimate the order of magnitude of the correction terms. With ${}^3G^a_b = 0$ in (9)–(11) various cases arise. In each case there is an exact vacuum or dust solution which coincides with our first approximation for suitably adjusted integration functions, and we name the various cases by the corresponding exact solution.

The Heckmann–Schücking-like[8] solution of (9)–(11), with ${}^3G^a_b = 0$, is

$$g_{ab} = {}_0g_{ac}\exp2\{{}_0K^c_b[\ln(t - {}_0t) - \ln(t - {}_0t')] + \tfrac{2}{3}\delta^c_b\ln(t - {}_0t')\}, \quad t > {}_0t > {}_0t' > -\infty,$$
$$\leftrightarrow K^a_b = {}_0K^a_b\left(\frac{1}{t - {}_0t} - \frac{1}{t - {}_0t'}\right) + \tfrac{2}{3}\delta^a_b(t - {}_0t')^{-1},$$
$$\alpha = {}_0\alpha(t - {}_0t)(t - {}_0t'). \tag{14}$$

Here ${}_0K^a_b$ is restricted by the famous condition[6]

$${}_0K^a_b\,{}_0K^b_a = 1 = {}_0K^a_a. \tag{15}$$

All other cases can be generated from (14). If ${}_0t = {}_0t'$ in (15) we get the Friedmann-like solution

$$g_{ab} = {}_0g_{ab}(t - {}_0t)^{4/3} \leftrightarrow K^a_b = \tfrac{2}{3}\delta^a_b(t - {}_0t)^{-1}, \quad \alpha = {}_0\alpha(t - {}_0t)^2. \tag{16}$$

If we write

$${}_0g_{ab} = {}_cg'_{ac}\exp\{({}_0K^c_b - \tfrac{2}{3}\delta^c_b)\ln(-{}_0t')\}$$

in (15) and take the limit ${}_0t' \to -\infty$, with ${}_0g'_{ab}, {}_0K^a_b$ fixed, we get (dropping the prime) the Kasner-like solution:

$$g_{ab} = {}_0g_{ab}\exp[{}_0K^c_b\ln(t - {}_0t)], \quad K^a_b = {}_0K^a_b(t - t_0)^{-1}, \quad \alpha = {}_0\alpha(t - {}_0t), \tag{17}$$

with ${}_0K^a_b$ still subject to (15). Near $t = {}_0t$, (17) and (14) are very similar, but (14) contains the extra arbitrary function ${}_0t'$. Finally, a similar limit gives the flatlike case

$$g_{ab} = {}_0g_{ab}, \quad K^a_b = 0. \tag{18}$$

Equation (15) or its specializations and condition (a) of our definition imply that the Jordan normal form of ${}_0K^a_b$ is diagonal and real; that is, ${}_0K^a_b$ can be diagonalized at any one point of the singularity by a real coordinate transformation. We then find that ${}_0g_{ab} = {}_0g_{ba}$, $\mathrm{sgn}\,{}_0g_{ab} = +3$, ${}_0K_{ab} \equiv {}_0g_{ac}\,{}_0K^c_b = {}_0K_{ba}$ and that ${}_0g_{ab}$ and ${}_0K_{ab}$ transform as tensors under (2a). Equations (12) and (14)–(18) now imply that two first approximations to the same exact metric can differ at most by having different functions ${}_0t'$ in the range ${}_0t > {}_0t' \geq -\infty$. We expect, but have not proved, that ${}_0t'$ is unique as well. In any case, ${}_0g_{ab}$ and ${}_0K_{ab}$ are uniquely and invariantly defined. We call them the metric and extrinsic curvature of the singularity.

We next analyze (6) and (13), then estimate correction terms. There are exact solutions which have one behavior on open regions of the singularity and another on the boundaries of these regions. However, for simplicity we confine our attention to an open region such that only one of the forms (14) and (16)–(18) holds everywhere.

In the flatlike case we then get from (13) ${}^3G^a_b = 0$ and that (6) is an identity. Then g_{ab} is regular. Thus, this case is not of interest, except perhaps on submanifolds of the singularity.

In the Friedmannlike case (16) we have

$$K^a_{b/a} - K_{,b} = \frac{2}{3(t - {}_0t)^2}\, {}_0t_{,b}, \tag{19}$$

in agreement with (7). If we assume $(\Gamma^a_{bc} - \hat{\Gamma}^a_{bc})$ $(t - {}_0t) \to 0$, then (6), (12), (16), and (19) give

$${}_0t_{,b} = 0. \tag{20}$$

With ${}_0t = \text{const}$ and ${}_0G^a_b$ the Einstein tensor of the singularity, one now gets

$${}^3G^a_b = {}_0G^a_b (t - {}_0t)^{-4/3}. \tag{21}$$

Thus (13) is an identity. Substituting (21) back into (9)–(11) then gives a second approximation. A short calculation shows that the metric g_{ab} correct to second order obeys (12). Thus, one has a systematic approximation scheme. Whether the approximations always converge is not known; examples show that they do converge for some models more general than the Friedmann models. Since ${}_0t' = {}_0t = \text{const}$, this Friedmann-like case is completely equivalent to the isotropic case of Lifshitz and Khalatnikov.[6]

The analysis in the Heckmann–Schücking-like case is more cumbersome. Let $\tau^a_A(x^b)$ be the unit eigenvectors of ${}_0K^a_b$ ($A, B \cdots = 1 \cdots 3$; capital indices are not subjected to the summation convention):

$${}_0K^a_b \tau^b_Y = P_Y \tau^a_Y, \qquad \tau^a_A \tau^b_B\, {}_0g_{ab} = \delta_{AB}. \tag{22}$$

Then (22) implies

$${}_0K^a_b = \textstyle\sum_A P_A \tau_{Aa} \tau^b_A \tag{23}$$

and (16) reads

$$\textstyle\sum_A P_A = 1 = \sum_A P_A^2. \tag{24}$$

(24) implies that we can order the P_A by the convention $1 \geq P_1 \geq P_2 \geq 0 \geq P_3 \geq -\frac{1}{3}$. Also, let $\|$ denote covariant differentiation within the singularity with respect to ${}_0g_{ab}$. We can use (14), (A1), and (A3) to analyze the constraint equation (6). If we ignore (24) for the moment, the dominant term is

$$\tau^b_B(K^a_{b/a} - K_{,b}) = (t - {}_0t)^{-2}\, {}_0t_{,a} \times \tau^a_A[P_A - 1 + \textstyle\sum_B P_B(P_B - P_A)] + \cdots . \tag{25a}$$

In contrast to the Friedmann-like case, this inverse square term vanishes by (24). Assuming now (24), we get

$$\tau^b_B(K^a_{b/a} - K_{,b}) = (t - {}_0t)^{-1}[{}_0K^a_{b\|a}\tau^b_B + (P_B - 1)(t - {}_0t')\, {}_0t_{,c}\tau^c_B] + \cdots . \tag{25b}$$

Here the prime denotes the two terms obtained by the substitutions ${}_0t \leftrightarrow {}_0t'$, ${}_0K'^a_b = \frac{2}{3}\delta^a_b - {}_0K^a_b$, $P_A \leftrightarrow P'_A = \frac{2}{3} - P_A$. The $(t - {}_0t)^{-1}$ dependence near $t = {}_0t$ agrees with (7) since g_{ab} obeys (3) and (4) (but only to first order in the approximation). Since the dominant term (24) cancels out, we cannot set (25a) equal to zero directly—in general, correction terms from the second order can contribute to (25a). However, (7) shows that $K^a_{b|a} - K^a_{,b} = 0$ is in any case three conditions within the singularity. The Friedmann-like expression (19) can be obtained directly from (25a) by letting ${}_0t' \to {}_0t$. Similarly, in the Kasner-like case we find that (6) is given by the appropriate limit ${}_0t' \to -\infty$ of (25a):

$$K^a_{b/a} - K_{,b} = (t - {}_0t)^{-1}({}_0K^a_{b\|a} - {}_0K^c_{c,b}) = (t - {}_0t)^{-1}\, {}_0K^a_{b\|a}. \tag{26}$$

We turn finally to (13) for the Heckmann–Schücking-like case. If ${}_0t = \text{constant}$, the analysis is not hard. Using (13) and (A5) gives, after considerable calculation the following two results: If $\mathbf{P} \equiv (P_1, P_2, P_3) \neq (1, 0, 0)$, then the Lifshitz–Khalatnikov surface orthogonal condition for τ_{3a},

$$\tau_{31,2} - \tau_{32,1} = 0 \tag{27}$$

holds. Second, if (27) holds, then ${}^3G^a_b$ substituted back into (9)–(11) gives a unique, consistent second approximation. The Kasner-like case behaves similarly. Thus the extra function ${}_0t'$ does not modify the analysis of Ref. 6 in an essential way.

If ${}_0t \neq \text{const}$ many special subcases arise. However for general values of $\mathbf{P}$ one finds from (13), (14), and (A5) not only (27) but also other conditions, for example,

$${}_0t_{,a}\tau^a_A = 0, \qquad A = 1, 2. \tag{28}$$

We omit a detailed discussion of the various special cases and turn instead to the examples.

III. EXAMPLES

In the following we indicate the singularity structure for some known exact irrotational dust models. The models are velocity dominated unless otherwise indicated. In the following, we do not give ${}_0t'$ for all cases; in fact, in many cases it is not known whether or not a Heckmann–Schücking-like solution is perhaps actually Kasner-like (${}_0t' = -\infty$). We consider only dust models with $\rho > 0$.

A. Homogeneous Models

1. Friedmann Models

$$ds^2 = -dt^2 + a^2(t)d\sigma^2, \tag{29}$$

where $d\sigma^2$ is the metric of a three-space of constant curvature. Then near $t = a = 0$, $a \cong \epsilon t^{2/3}$, with $\epsilon = \frac{1}{6}\alpha\rho = \text{const}$. Thus the singularity is Friedmann-like, with

$${}_0ds^2 = {}_0g_{ab}dx^a dx^b = \epsilon^2 d\sigma^2. \tag{30}$$

2. Heckmann-Schücking (Bianchi I) Models[8]

$$ds^2 = -dt^2 + \textstyle\sum_a l_a^2 (dx^a)^2,$$
$$l_a = \epsilon^{2/3 - P_a}\, t^{P_a}(t - {}_0t')^{2/3 - P_a},$$

$$P_a = \text{const}, \quad \epsilon = \tfrac{3}{4}\alpha\rho = \text{const}, \quad {}_0t' = \text{const},$$
$$|{}_0\bar{t}| = \epsilon^{-1}, \quad \textstyle\sum_a P_a = 1 = \sum_a P_a^2. \tag{31}$$

This is the prototype of the Heckmann–Schücking-like singularities, with

$${}_0ds^2 = \textstyle\sum_a \epsilon^{4/3-2P_a}(dx^a)^2, \quad {}_0K^a_b = \text{diag}(P_1, P_2, P_3),$$
$${}_0t = 0, \quad |{}_0t'| = \epsilon^{-1}. \tag{32}$$

3. Other Models

In the Kantowski–Sachs models[9] the singularities are Heckmann–Schücking-like. In general, each dust line in each model hits a $p = [1, 0, 0]$ singularity and a $\mathbf{P} = [\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}]$ singularity. However, in "open solution a"[9] there is only a $[1, 0, 0]$ singularity. The Bianchi V models[8] have a Heckmann–Schücking-like $\mathbf{p} = [\frac{1}{2}(1+\sqrt{3}), \frac{1}{3}, \frac{1}{2}(1-\sqrt{3})]$ singularity hypersurface. The Ellis–McCallum models[10] contain one example which is *not* velocity dominated. A structure can be assigned to the singularity, but then $\mathbf{p} = (\frac{3}{4}, \frac{3}{4}, \frac{1}{2})$. The mixmaster models[11] are apparently not velocity dominated, but the structure of the singularity is not well understood.

B. Inhomogeneous Models

We study in detail two classes of exact solutions of inhomogeneous dust models. Both admit a three-parameter group and have very similar structures. Because of the complexity involved, we discuss the general solutions in a little detail. All possible singularities that may arise are listed. [There may arise other $(1, 0, 0)$ singularities besides those already listed here, both in ${}_0t,_3 = 0$ and ${}_0t,_3 \neq 0$ regions.]

1. Plane Symmetric Models

$$ds^2 = -dt^2 + \phi^2(z, t)(dx^2 + dy^2) + \psi^2(z, t)dz^2. \tag{33}$$

The metric in this form still allows the arbitrary coordinate transformations: $\tau \to t + c$, $z \to f(z)$. The Einstein equations reduce to two essential equations in the generic case ($\kappa \neq 0$):

$$\phi_{,3} = \kappa\psi, \quad \kappa = \kappa(z) \neq 0 \tag{34}$$

and

$$2\phi\ddot{\phi} + \dot{\phi}^2 - \kappa^2 = 0. \tag{35}$$

Here and in the following the subscript denotes $\partial/\partial z$, overhead dots denote $\partial/\partial t$. Equation (35) has a first integral

$$\phi(\dot{\phi}^2 - \kappa^2) = \lambda(z). \tag{36}$$

The Bianchi identities then require $\lambda_{,3} = \kappa\epsilon$, where $\epsilon = \alpha\rho > 0$. λ and either κ or ϵ are two arbitrary real functions of z to be specified as initial data. Depending on the values of λ and κ, a variety of cases arise. All the different cases *may* (but need not) occur in one single solution. Different spatial coordinate neighborhoods $\{z\,|\,z$ satisfies the conditions of the specific case$\}$ may have different time developments, and thus different types of singularities.

We give the exact solution for the various cases.

Case (a): $\kappa = 0$. Equation (34) does not define ψ. Then using another member of the Einstein equations together with Eq. (35) gives

$$\phi = {}_0\phi(t - {}_0t)^{2/3}, \quad \psi = {}_0\psi(t - {}_0t)^{2/3},$$
$${}_0\phi, {}_0\psi, {}_0t = \text{const}. \tag{37}$$

The special case $\kappa \equiv 0$ (identically for all z) reduces to the Einstein–De Sitter solution of the Friedmann models.

Case (b): $\kappa \neq 0$, $\lambda > 0$. Equations (34) and (36) give

$$\phi = \frac{\lambda}{\kappa^2}\sinh^2\left(\frac{\eta}{2}\right),$$
$$\psi = \frac{1}{\kappa}\left(\frac{\lambda}{\kappa^2}\right)_{,3}\sinh^2\left(\frac{\eta}{2}\right) - \coth\left(\frac{\eta}{2}\right) \times \left[\xi_{,3} + \frac{1}{2}\left(\frac{\lambda}{\kappa^3}\right)_{,3}(\sinh\eta - \eta)\right],$$
$$t - \xi(z) = (\lambda/2\kappa^3)(\sinh\eta - \eta),$$
$$\rho = \frac{\kappa^4\epsilon}{2\lambda^2}\left(\sinh^6\frac{\eta}{2}\left\{\frac{1}{2\kappa}\left(\frac{\lambda}{\kappa^2}\right)_{,3} - \frac{\coth\eta/2}{\sinh^2\eta/2} \times \left[\xi_{,3} + \frac{1}{2}\left(\frac{\lambda}{\kappa^3}\right)_{,3}(\sinh\eta - \eta)\right]\right\}\right)^{-1}. \tag{38}$$

Case (c): $\kappa \neq 0$, $\lambda < 0$, put $\mu = -\lambda > 0$:

$$\phi = \frac{\mu}{\kappa^2}\cosh^2\frac{\eta}{2},$$
$$\psi = \frac{1}{\kappa}\left(\frac{\mu}{\kappa^2}\right)_{,3}\cosh^2\frac{\eta}{2} - \tanh\frac{\eta}{2} \times \left[\xi' + \frac{1}{2}\left(\frac{\mu}{\kappa^3}\right)_{,3}(\sinh\eta + \eta)\right],$$
$$t - \xi = \frac{\mu}{2\kappa^3}(\sinh\eta + \eta), \ \xi \text{ function of } z \text{ only},$$
$$\rho = \frac{\kappa^4\epsilon}{2\mu^2}\left(\cosh^6\frac{\eta}{2}\left\{\frac{1}{2\kappa}\left(\frac{\mu}{\kappa^2}\right)_{,3} - \frac{\tanh\eta/2}{\cosh^2\eta/2} \times \left[\xi' + \frac{1}{2}\left(\frac{\mu}{\kappa^3}\right)_{,3}(\sinh\eta + \eta)\right]\right\}\right). \tag{39}$$

Case (d): $\kappa \neq 0, \lambda = 0$. Since $\rho \neq 0$ implies $\lambda' \neq 0$, this can only happen at isolated values of z, say z_0. Then at z_0:

$$\phi = \kappa_0(t - c_0),$$
$$\psi = \frac{\epsilon_0}{2\kappa_0{}^2}\ln(t - c) + \frac{(\kappa_{,3})_0}{\kappa_0}(t - c_0), \tag{40}$$

everything being evaluated at z_0.

We now give the singularity structures (Fig. 3).

In Case (a), the singularity is Friedmann-like with

$${}_0ds^2 = {}_0g_{ab}dx^a dx^b = {}_0\phi^2(dx^2 + dy^2) + {}_0\psi^2 dz^2, {}_0t, {}_0\phi, {}_0\psi \text{ constants}. \tag{41}$$

In case (b), there are two possibilities:

(i) If $\xi_{,3} = 0$, we have only one singularity. Near $\eta = 0$, $\phi \to 0$, $\psi \to 0$ with

$$\phi \cong f(z)(t - {}_0t)^{2/3}, \quad \psi \cong g(z)(t - {}_0t)^{2/3}, \quad {}_0t = \text{const}. \tag{42}$$

Thus the singularity is inhomogeneous but is Friedmann-like, with

$${}_0ds^2 = {}_0g_{ab}dx^a dx^b = f^2(z)(dx^2 - dy^2) + g^2(z)dz^2. \tag{43}$$

(ii) If $\zeta_{,3} \neq 0$, we have two singularities in general. They fall in the Heckmann–Schücking pattern. The one at $\eta = 0$ gives

$${}_0K^a_b = \text{diag}(\tfrac{2}{3}, \tfrac{2}{3}, -\tfrac{1}{3}), \quad {}_0ds^2 = f^2(z)(dx^2 + dy^2) + h^2(z)dz^2, \quad {}_0t = \zeta(z), \tag{44}$$

where $f(z)$ is the same function as in (43). The other one happens at

$$\sinh^2\frac{\eta_0}{2}\left(\frac{1}{\kappa}\right)\left(\frac{\lambda}{\kappa^2}\right)_{,3} = \coth\frac{\eta_0}{2}\zeta_{,3} + \tfrac{1}{2}\left(\frac{\lambda}{\kappa^3}\right)_{,3}(\sinh\eta_0 - \eta_0),$$

near which we have $\phi \cong F(z)$, $\psi \cong G(z)[t - {}_0\bar{t}(z)]$, ${}_0\bar{t} \equiv t(\eta_0)$.

We thus have

$${}_0\bar{K}^a_b = \text{diag}(0, 0, 1), \quad {}_0d\bar{s}^2 = F^2(z)(dx^2 + dy^2) + G^2(z)dz^2. \tag{45}$$

In Case (c) the generic solution has only one singularity at

$$\cosh^2\frac{\eta_0}{2}\frac{1}{\kappa}\left(\frac{\mu}{\kappa^2}\right)_{,3} = \tanh\frac{\eta_0}{2}\left[\zeta_{,3} + \frac{1}{2}\left(\frac{\mu}{\kappa^3}\right)_{,3}(\sinh\eta_0 + \eta_0)\right].$$

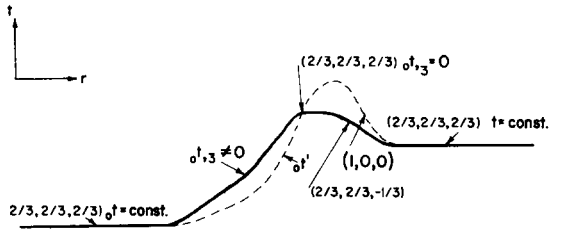

Fig. 2. Possible singularity configuration of a loosely bound ($\beta = 0$) Tolman–Bondi model showing the splitting of a Friedman-like $(\frac{2}{3}, \frac{2}{3}, \frac{2}{3})$ singularity (${}_0t = \text{const}$) into a $(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3})$ ${}_0t_{,3} \neq 0$ singularity and a (1, 0, 0) singularity with bang time ${}_0t'$.

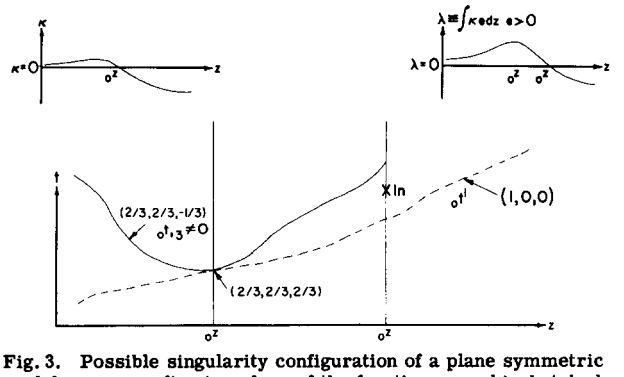

Fig. 3. Possible singularity configuration of a plane symmetric model corresponding to values of the functions κ and λ sketched in the upper corners, showing the complexity of the structure. Note the local Friedman behavior at ${}_0z$ ($\kappa = 0$) and the non-velocity-dominated log behavior (×) at ${}_1z$ ($\lambda = 0$). In the $\lambda < 0$ region there can only be a (1, 0, 0) singularity. Thus, the metric must be discontinuous across ${}_1z$. Such defect can be removed if we restrict the function κ to be nonnegative for all z.

Near η_0 we have

$$\phi \cong F(z), \quad \psi \cong G(z)[t - {}_0t(z)], \quad {}_0t \equiv t(\eta_0). \tag{46}$$

Thus ${}_0K^a_b = \text{diag}(1, 0, 0)$.

In Case (d) the solution has two singularities:

(i) Near $t = c_0$, we have

$$\phi \cong A(t - {}_0t), \quad \psi = B\ln(t - {}_0t). \tag{47}$$

This solution fits into *neither* (14), (16), or (17). However, a closer look reveals that since $\lambda = 0$, Eq. (36) implies $\dot{\phi}^2 = \kappa^2$ always; thus the singularity *need not* be velocity dominated. This serves as a good counter example.

(ii) Near

$$\frac{(\kappa_{,3})_0}{\kappa_0}({}_0\bar{t} - c_0) = \frac{\epsilon_0}{2\kappa_0^2}\ln({}_0\bar{t} - c_0), \tag{48}$$

we have $\phi \cong D$, $\psi \cong E(t - {}_0\bar{t})$. This is again a Heckmann–Schücking-like singularity, with ${}_0\bar{K}^a_b = \text{diag}(0, 0, 1)$.

2. Spherical Symmetric Models[2]

$$ds^2 = -dt^2 + \phi^2(r, t)d\Omega^2 + \psi^2(r, t)dr^2, \quad d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2. \tag{49}$$

Allowed arbitrary coordinate transformations are $t \to t + c$ and $r \to f(r)$. The Einstein equations again reduce to two equations:

$$\phi_{,3} = [1 + \beta(r)]\psi, \quad \beta(r) > -1, \tag{50}$$

$$2\phi\ddot{\phi} + \dot{\phi}^2 - \beta = 0. \tag{51}$$

Here and in the following $_{,3}$ denotes $\partial/\partial r$, a raised dot denotes $\partial/\partial t$. Equation (51) has a first integral

$$\phi(\dot{\phi}^2 - \beta) = \lambda(r). \tag{52}$$

The Bianchi identities require $\lambda_{,3} = \epsilon(1 + \beta)^{1/2}$, where $\epsilon = \alpha\rho > 0$. Again, λ, β, or ϵ are two arbitrary real functions of r specified as initial data. The general solution consists of the following cases.

Case (α): $\beta > 0$. We can put $\kappa^2 = \beta$. Then everything is the same as in the plane symmetric cases, and the singularity analysis becomes completely parallel. We omit the repetition here.

Case (β): $\beta = 0$.

$$\phi = {}_0\phi(r)[t - {}_0t(r)]^{2/3},$$

where ${}_0\phi(r) = \lambda^{1/3}(\frac{3}{2})^{2/3}$;

$$\psi = {}_0\phi_{,3}[t - {}_0t(r)]^{2/3} - \tfrac{2}{3}\,{}_0t_{,3}\,{}_0\phi[t - {}_0t(r)]^{-1/3} = {}_0\phi_{,3}[t - {}_0t(r)]^{1/3}[t - {}_0\bar{t}(r)], \tag{53}$$

$${}_0\bar{t} \equiv {}_0t(r) + \tfrac{2}{3}\,{}_0\phi\,{}_0t_{,3}/{}_0\phi_{,3}.$$

Case (γ): $-1 < \beta < 0$. We put $\kappa^2 = -\beta > 0$. Since now Eq. (52) requires $\lambda > 0$, we get only one possibility:

$$\phi = \frac{\lambda}{\kappa^2}\sin^2\frac{\eta}{2},$$

$$\psi = \frac{1}{(1-\kappa^2)^{1/2}}\left(\frac{\lambda}{\kappa^2}\right)_{,3}\sin^2\frac{\eta}{2} - \frac{\kappa}{(1-\kappa^2)^{1/2}} \times \cot\frac{\eta}{2}\left[\gamma_{,3} + \frac{1}{2}\left(\frac{\lambda}{\kappa^3}\right)_{,3}(\eta - \sin\eta)\right], \tag{54}$$

$$t - \gamma(r) = (\lambda/2\kappa^3)(\eta - \sin\eta).$$

We now consider the structure of the singularities. In Case (β), we have two possibilities (the singularity is flat in this case):

(i) ${}_0t_{,3} = 0$. Then the singularity is Friedmann-like with

$${}_0ds^2 = {}_0\phi^2(r)d\Omega^2 + [{}_0\phi_{,3}(r)]^2dr^2, \quad {}_0t = \text{const}; \tag{55}$$

(ii) ${}_0t_{,3} \neq 0$. Then the singularity is Heckmann-Schücking-like, with

$${}_0K^a_b = \text{diag}(\tfrac{2}{3}, \tfrac{2}{3}, -\tfrac{1}{3}), \quad \text{diag}(0,0,1). \tag{56}$$

Equation (53) shows the smooth transition from Heckmann–Schücking-like to Friedman-like behavior as ${}_0t_{,3}$ approaches zero.

In Case (γ) we again have two possibilities.

(i) $\gamma_{,3} = 0$. Then the singularity near $\eta = 0, \pm 2\pi n$ (n integers) is Friedmann-like with

$${}_0ds^2 = f^2(r)d\Omega^2 + g^2(r)dr^2, \quad {}_0t = \text{const}. \tag{57}$$

(ii) $\gamma_{,3} \neq 0$. We have two singularities of the Heckmann–Schücking type.

For the first singularity, near $\eta = 0, \pm 2\pi n$ (n integers), we have

$$\phi \cong f(r)[t - {}_0t(r)]^{2/3}, \quad {}_0t(r) = \gamma(r),$$
$$\psi \cong h(r)(t - {}_0t(r))^{-1/3}, \quad f \text{ same as in (57)}, \tag{58}$$
$$\therefore {}_0K^a_b = \text{diag}(\tfrac{2}{3}, \tfrac{2}{3}, -\tfrac{1}{3}), \quad {}_0ds^2 = f^2d\Omega^2 + h^2dr^2.$$

For the second singularity, near

$$\frac{1}{\kappa}\left(\frac{\lambda}{\kappa^2}\right)_{,3}\sin^2\frac{\eta_0}{2} = \cot\frac{\eta_0}{2}\left[\gamma_{,3} + \frac{1}{2}\left(\frac{\lambda}{\kappa^3}\right)_{,3}(\eta_0 - \sin\eta_0)\right],$$

we have

$$\phi \cong F(r), \quad \psi \cong G(r)[t - {}_0\bar{t}(r)], \quad {}_0\bar{t} \equiv t(\eta_0),$$
$$\therefore {}_0K^a_b = \text{diag}(0,0,1), \quad {}_0ds^2 = F^2(r)d\Omega^2 + G^2(r)dr^2. \tag{59}$$

We note some features of the singularity structure. All cases except (47), which occurs only on submanifolds of the singularity, are velocity dominated. The assumed symmetry restricts the possibilities for p to $(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3})$, $(1,0,0)$, or $(\frac{2}{3}, \frac{2}{3}, \frac{2}{3})$. The latter, Friedmann-like case occurs only where ${}_0t_{,3} = 0$ in agreement with (20). On the other hand, $(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3})$ occurs only where ${}_0t_{,3} \neq 0$ and is then always accompanied by a $(1,0,0)$ singularity occurring along the same dust lines at a different t. One can visualize the Friedmann-like singularity "splitting" into two different singularities (Fig. 2). A $(1,0,0)$ singularity can occur along a dust line without any other kind of singularity occurring along the same dust line.

The structure of the singularity uniquely determines the full metric. For example, one can obtain (38) from (44) by noting $\lambda = 4f^3/9$, $\kappa = (\frac{2}{3})^{1/3}(f^3/h)\,{}_0t_{,3}$, $\zeta = {}_0t$. In such examples one sees that the space–time metric cannot be C^∞ unless the whole singularity structure is C^∞.

IV. THE b BOUNDARY

The assignment of a differentiable structure, metric, and extrinsic curvature to the singularity can thus far be carried out systematically only for the velocity-dominated irrotational dust models. To try to generalize one should probably work with the b-Boundary.[4] We now give some fragmentary results on the b boundary of a first approximation to a velocity-dominated solution.

For present purposes we can define the b boundary of space-time M as follows[12]: Let $O(M)$ be the orthonormal frame bundle over M. The points of $O(M)$ can be locally coordinated by x^μ, X^μ_α ($\alpha, \beta = 1 \cdots 4$ label tetrad vectors; $\mu, \nu = 1 \cdots 4$ vector indices) with $X^\mu_\alpha X^\nu_\beta g_{\mu\nu}(x) = \eta_{\alpha\beta} \equiv \text{diag}(-1,1,1,1)$. Let C be a curve, with parameter q, in $O(M)$, whose tangent is nowhere vertical, $dx^\mu/dq \neq 0$. Introduce along C the line element

$$\left(\frac{dS}{dq}\right)^2 = \sum_{\alpha=1}^{4}\left[\left(\frac{dx^\mu}{dq}\right)X_{\alpha\mu}\right]^2 + \sum_{\alpha<\beta=1}^{4}\left(X^\mu_{\alpha;\nu}\frac{dx^\nu}{dq}X_{\beta\mu}\right)^2. \tag{60}$$

This corresponds to giving $O(M)$ a positive definite Riemmanian structure. One can then complete $O(M)$ in the standard way. Identifying points and Cauchy sequences in $O(M)$ which are equivalent under the homogeneous Lorentz group ($\hat{x}^\mu = x^\mu$, $\hat{X}^\mu_\beta = L^\alpha_\beta X^\mu_\alpha$, $L^\alpha_\beta = \text{const}$, $L^\alpha_\beta L^\gamma_\delta \eta_{\alpha\gamma} = \eta_{\beta\delta}$) gives a boundary, the b boundary, of M.

For a first approximation g_{ab} to a velocity-dominated irrotational dust model the points $({}_0t(x^a), x^a)$ are points of the b boundary. In fact, consider the line in $O(M)$ given by $t = q$, $x^a = \text{const}$, $X^\mu_0 = \delta^\mu_0$, X^μ_a parallel displaced along x^μ. This obviously contains Cauchy sequences terminating at x^a, ${}_0t$. Moreover, for two points $(x^a, {}_0t)$, $(x^a + dx^a, {}_0t + {}_0t_{,a}dx^a)$ the "horizontal" distance element $\sum_\alpha (dx^\mu/dq X_{\alpha\mu})$ may go to zero, but then the "vertical" distance element [the second sum in (60)] contains a term of the form

$$K^a_b K_{ac}\frac{dx^b}{dq}\frac{dx^c}{dq} \to \infty.$$

Thus two neighboring points $({}_0t, x^a)$, with different x^a, are presumably different points of the b boundary. Whether the topology of the b boundary is consistent with the manifold structure we have assumed has not been checked in general.

V. CONCLUSIONS

To characterize and restrict cosmological models by analyzing conditions at the bang will presumably be possible only if we can assign a detailed structure to the bang in all sensible models. We have shown that in velocity-dominated irrotational dust models the bang does have a rather natural and elegant structure which seems to be consistent with the structure of the b boundary. There are many other models, not discussed here, where a very similar structure can be

assigned *(ad hoc)* to the singularity. However, we have not yet been able to assign such a structure to the mixmaster model[5] or the mixmasterlike models.[13] Thus it is possible, though in our opinion not likely, that there are sensible models in which the b boundary cannot usefully be given even a differentiable, much less a Riemannian, structure.

ACKNOWLEDGMENTS

The authors are grateful to A. Ali and R. Tabensky for many discussions. One of us (R.S.) gratefully acknowledges the hospitality of the University of Pittsburgh during part of the time this research was being done.

APPENDIX: STRETCHED METRICS

Let M be an N-dimensional manifold with metric $dS^2 = g_{ab}dX^a dX^b$ $(a, b = 1 \cdots N)$. Let τ_{Aa} $(A, B = 1 \cdots N)$ be an orthonormal frame, $\tau_{Aa}\tau_{Bb}g^{ab} = \delta_{AB}$. We shall use the usual rules of the Cartan formalism: $\tau_A = \tau_{Aa}dx^a$, $\tau_A \wedge \tau_B = \frac{1}{2}(\tau_A \otimes \tau_B - \tau_B \otimes \tau_A)$, $d\tau_A = \sum_B \tau_{AB} \wedge \tau_B$, $\tau_{AB} = -\tau_{BA}$, etc. Thus $ds^2 = \sum_A \tau_A \otimes \tau_A$, the curvature form is $\rho_{AB} = d\tau_{AB} - \sum_C \tau_{AC} \wedge \tau_{CB}$, and the rule for covariant differentiation is

$$V_{a;b}\tau_A^a\tau_B^b = V_{AB} + \sum_C V_C \tau_{CAB}, \tag{A1}$$

where V_A, V_{AB}, and τ_{ABC} are defined by $V_A = V_a\tau_A^a$, $dV_A = \sum_B V_{AB}\tau_B$, and $\tau_{AB} = \sum_C \tau_{ABC}\tau_C$. It will be convenient to define $F_{ABC} = \tau_{ABC} - \tau_{ACB}$ and to let $dF_{ABC} = \sum_D F_{ABCD}\tau_D$.

Now suppose $\omega_Z = e^{A_Z}\tau_Z$ (no sum of course), with the A_Z scalar functions on M. Then we call

$$dl^2 = \sum_A \omega_A \otimes \omega_A \tag{A2}$$

the stretch of ds^2 along τ_A with factors A_A. If we do not specialize τ_A or A_A, this notion is very general. If ds^2 is positive-definite, then locally every positive-definite dl^2 is some stretch of a given ds^2 since both metrics can be simultaneously diagonalized at a point. However, appropriate specializations of τ_A and A_A lead to several useful subcases, as discussed below.

Let $dA_A = \sum_B A_{AB}\tau_B$ and $d\omega_A = \sum_B \omega_{AB} \wedge \omega_B$ with $\omega_{AB} = -\omega_{BA}$. Then a moderate calculation gives

$$2\omega_{XY} = \sum_B \tau_B[e^{A_X - A_Y}(F_{XYB} - 2A_{XY}\delta_{XB}) - \tfrac{1}{2}e^{2A_B - A_X - A_Y}F_{BXY}] - (X, Y), \tag{A3}$$

where (X, Y) denotes interchange of indices. Let $dA_{AB} = \sum_C A_{ABC}\tau_C$, $B_{ABC} = \frac{1}{2}(A_{ABC} + A_{ACB})$. Then a dreary calculation gives for the curvature form $\theta_{AB} = d\omega_{AB} - \sum_C \omega_{AC} \wedge \omega_{CB}$ the following:

$$-2\sum_{A,E}(\omega_A \wedge \omega_E) \otimes \theta_{AE} = \sum_{ABED}(\tau_A \wedge \tau_E) \otimes (\tau_B \wedge \tau_D) \times (M_{AEBD} + N_{AEBD} + N_{BDAE}),$$

$$M_{XYRS} = \sum_C [e^{2A_R + 2A_X - 2A_C}F_{XSC}F_{RYC} - 2e^{2A_X + 2A_Y - 2A_C}A_{XC}A_{YC}\delta_{XR}\delta_{YS} - \tfrac{1}{2}e^{2A_C}(F_{CXR}F_{CYS} + F_{CXY}F_{CRS})] - 4\delta_{XR}e^{2A_X}[B_{XYS} + (A_{XS} - A_{YS})A_{XY} - A_{XS}A_{SY}], \tag{A4}$$

$$N_{XYRS} = \sum_C [-e^{2A_R + 2A_S - 2A_C}(2\delta_{YS}A_{YC}F_{RXC} + \tfrac{1}{2}F_{RXC}F_{SYC}) + e^{2A_X}F_{XSC}F_{CRY}] + e^{2A_X}(F_{XSRY} + 2A_{XS}F_{XYR} - 2A_{YS}F_{XYR} - 2A_{XY}F_{XRS} + 2A_{SY}F_{XRS}).$$

To check (A4) one can let $A_1 = A_2 = \cdots = A_N$; then (A4) yields, after a moderate calculation, the usual expression for conformally related metrics. If we take the appropriate trace of the curvature tensor, we find, after some calculations, the Ricci tensor of dl^2:

$$-4e^{A_S + A_Y}R_{ab}\omega_Y^a\omega_S^b = \sum\{e^{2A_S - 2A_C}[-2F_{SYCC} + 2F_{SYC} \times (-2A_{SC} + 2A_{YC} - F_{AAC} - A_{AC} + 2A_{CC}) - 2\delta_{SY}(B_{YCC} - 2A_{CC}A_{YC} + A_{YC}(A_{AC} + F_{AAC}))] + 2(2B_{YYS} - B_{AYS} - A_{AS}A_{AY} - 2A_{YS}A_{SY} + 2A_{AY}A_{YS} - F_{AASY} - 2A_{CY}F_{CCS} + 2A_{SY}F_{CCS} + A_{YC}F_{CYS} - \tfrac{1}{2}F_{AYC}F_{CSA}) - e^{2A_A - 2A_C} \times F_{ASC}F_{AYC} + \tfrac{1}{2}e^{2A_Y + 2A_S - 2A_C - 2A_A}F_{YAC}F_{SAC}\} + (Y, S), \tag{A5}$$

where $\sum$ means to multiply out all brackets and then sum over A and/or C if they appear as subscripts in a given term. As a check, we note that (A5) reduces to the standard equations for a diagonalizable metric if $d\tau_A = 0$.

Equations (A1), (A3), and (A5) are those needed in the main text, (A3)–(A5) have a variety of useful applications in addition to those already mentioned; for example, if space–time permits an isometry group, it is often convenient to regard the inner metric of each orbit as an (orbit-dependent) stretch of a fixed simple metric. It is conceivable that useful techniques for solving the field equations could be developed by some clever choice of ds^2, τ_A, and A_A.

---

* Partially supported by the National Science Foundation, under Grant No. GP-21495.

1 Compare Y. Ne'eman, Astrophys. J. **141**, 1303 (1965); G. R. Burbidge, E. M. Burbidge, and A. R. Sandage Rev. Mod. Phys. **35**, 947, (1963).
2 H. Bondi, Monthly Notices Roy. Astron. Soc. **410**, 107 (1947).
3 We are grateful to D. Brown for discussions on this point.
4 B. G. Schmidt, "A New Definition of Singularities in General Relativity" (Hamburg, 1970) (unpublished).
5 C. Misner, Astrophys. J. **151**, 431 (1968).
6 E. M. Lifshitz and I. M. Khalatnikov, Usp. Fiz Nauk **80**, 391 (1963) [Sov. Phys. Usp. **6**, 495 (1964)]; see also E. M. Lifshitz and I. M. Khalatnikov, Advan. Phys. **12**, 185 (1963).
7 J. L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1964).
8 G. Heckmann and E. Schücking, "Relativistic Cosmology" in *Gravitation, An Introduction to Current Research* (Wiley, New York, 1962).
9 R. Kantowski and R. K. Sachs, J. Math. Phys. **7**, 443 (1966).
10 G. F. R. Ellis and M. A. H. MacCallum, Commun. Math. Phys. **12**, 108 (1969).
11 C. W. Misner, Phys. Rev. Letters **22**, 1071 (1969). Very useful surveys of homogeneous models, including the mixmaster, are in M. A. H. MacCallum, "A Class of Homogeneous Cosmological Models III" (Cambridge Univeristy, 1969) (unpublished) and K. C.

Jacobs and L. P. Hughston, "More Quantum Cosmology" (unpublished).

[12] The essential idea of the following calculation was suggested to us by R. Geroch and H. Ploss (private communication).

[13] V. A. Belinski and I. M. Khalatnikov, Zh. Eksp. Teor. Fiz. **57**, 2163 (1970) [Sov. Phys. JETP **30**, 1174 (1970)]; and references cited therein.

Statistically Exact Kinematic Dynamo Action in a Box

I. Lerche
Enrico Fermi Institute and Department of Physics, University of Chicago, Chicago, Illinois
(Received 1 July 1971)

A set of statistically *exact* equations is set up to describe dynamo action brought about by velocity turbulence confined to a finite spatial domain (a "box"). Using a variation of a limit-theorem employed elsewhere by Kac, we demonstrate that in large enough boxes a large-scale magnetic field is regenerated by such velocity turbulence.

I. INTRODUCTION

Over the past decade the advancement of turbulent kinematic dynamo theory has been considerable based on the original investigations by Parker.[1] In Parker's original paper it was shown that very rapid cyclonic velocity turbulence in conjunction with a sheared large-scale velocity field was capable of generating dynamo action. Later Braginskii[2] showed that slow turbulence in conjunction with a sheared velocity field would also produce dynamo action.

In 1966, Steenbeck, Krause, and Radler demonstrated that *helical* velocity turbulence on its own gave rise to kinematic dynamo action in an infinite medium. More recently still, it has been shown[3,4] that *isotropic* velocity turbulence on its own gives rise to dynamo action. It has also been shown[3] that both Parker's dynamo equations and those of Braginskii are limiting forms of a master dynamo equation which includes both sheared velocity multiplied by turbulence terms and turbulence terms on their own, either of which (or both) will give dynamo action.

In view of the complexity of the general turbulent kinematic dynamo equations, recourse is normally made to some approximation scheme in order to obtain the basic physical behavior of particular terms in the equations under particular boundary and/or initial value conditions. We are aware of only two cases[5,6] in which some *exact* statistical properties of the turbulent kinematic dynamo equations are set forth—and both of these deal with velocity turbulence in infinite media.

In the first of these papers,[5] a discussion is given of the irregular *generation* of magnetic field brought about by fluctuations in the *level* of dynamo activity (as opposed to irregular generation of magnetic field brought about by random velocity turbulence with a *fixed* level of activity). In the second of these papers[6] the singular eigenvalue equations are set up which describe the behavior of the large-scale field (an equation of the Dyson type) and the small-scale turbulent field energy (an equation of the Bethe-Salpeter type).

In the above two papers, all correlated quantities are taken to be homogeneous in space and time. In the present paper where the turbulence is confined to a "box" the correlated quantities can be neihter homogeneous nor isotropic. Further, as we will demonstrate directly, the finite (but large) size of the box enters directly into the normal modes and their growth rates [*vide* Eq. (58)]. As the size of the box increases both the level of turbulence [given through Eq. (3)] and the attendant normal mode growth rate [given through Eq. (58)] decrease. Thus the generation of magnetic field brought about by turbulence in a "box" is rather different than that obtaining in an infinite medium.

Now most of the astrophysical situations (the Earth, Sun, galaxy, etc.) in which a dynamo mechanism is invoked to account for the continued presence of a large-scale magnetic field are objects of finite size. The question then naturally arises: Are there any statistically *exact* situations which give kinematic dynamo activity when the velocity turbulence is confined to a spatial domain of finite size (a 'box')?

The purpose of the present paper is to answer this question affirmatively by considering a simple form of velocity turbulence confined to a cubical box.

II. EQUATIONS OF MOTION AND THE EVOLUTION OF PROBABILITY

Consider then a finite medium, of constant resistivity η, which is not undergoing either bulk convection or shear, so that only a turbulent velocity, $\mathbf{V}$, with zero mean is present. Then the magnetohydrodynamic equations for the vector potential $\mathbf{A}$ are

$$\left(\frac{\partial}{\partial t} - \eta\nabla^2\right)A_i = \epsilon_{ijk}V_j(\mathbf{x}, t)B_k(\mathbf{x}, t), \tag{1}$$

with the magnetic field B given by

$$B_i(\mathbf{x}, t) = \epsilon_{ijk}\partial A_k/\partial x_j. \tag{2}$$

For random velocities $\mathbf{V}$ which are arbitrary functions of both space and time, Eqs. (1) and (2) are difficult to solve.

Elsewhere,[3,4,6-8] we have investigated some of the exact statistical kinematic dynamo properties of Eqs. (1) and (2) in an *infinite* medium when the random velocity $\mathbf{V}(\mathbf{x}, t)$ was taken to depend on only one coordinate (either spatial or temporal). In the present paper we consider the possibility of kinematic dynamo action when the turbulent velocity is confined to a cubical 'box' of side L. We describe the turbulent velocity field $\mathbf{V}$ by

$$V_x = V_1(t)\,\sin(N\pi x/L)\,\cos(N\pi y/L)\,\cos(N\pi z/L), \tag{3a}$$

Jacobs and L. P. Hughston, "More Quantum Cosmology" (unpublished).

[12] The essential idea of the following calculation was suggested to us by R. Geroch and H. Ploss (private communication).

[13] V. A. Belinski and I. M. Khalatnikov, Zh. Eksp. Teor. Fiz. **57**, 2163 (1970) [Sov. Phys. JETP **30**, 1174 (1970)]; and references cited therein.

Statistically Exact Kinematic Dynamo Action in a Box

I. Lerche
Enrico Fermi Institute and Department of Physics, University of Chicago, Chicago, Illinois
(Received 1 July 1971)

A set of statistically *exact* equations is set up to describe dynamo action brought about by velocity turbulence confined to a finite spatial domain (a "box"). Using a variation of a limit-theorem employed elsewhere by Kac, we demonstrate that in large enough boxes a large-scale magnetic field is regenerated by such velocity turbulence.

I. INTRODUCTION

Over the past decade the advancement of turbulent kinematic dynamo theory has been considerable based on the original investigations by Parker.[1] In Parker's original paper it was shown that very rapid cyclonic velocity turbulence in conjunction with a sheared large-scale velocity field was capable of generating dynamo action. Later Braginskii[2] showed that slow turbulence in conjunction with a sheared velocity field would also produce dynamo action.

In 1966, Steenbeck, Krause, and Radler demonstrated that *helical* velocity turbulence on its own gave rise to kinematic dynamo action in an infinite medium. More recently still, it has been shown[3,4] that *isotropic* velocity turbulence on its own gives rise to dynamo action. It has also been shown[3] that both Parker's dynamo equations and those of Braginskii are limiting forms of a master dynamo equation which includes both sheared velocity multiplied by turbulence terms and turbulence terms on their own, either of which (or both) will give dynamo action.

In view of the complexity of the general turbulent kinematic dynamo equations, recourse is normally made to some approximation scheme in order to obtain the basic physical behavior of particular terms in the equations under particular boundary and/or initial value conditions. We are aware of only two cases[5,6] in which some *exact* statistical properties of the turbulent kinematic dynamo equations are set forth—and both of these deal with velocity turbulence in infinite media.

In the first of these papers,[5] a discussion is given of the irregular *generation* of magnetic field brought about by fluctuations in the *level* of dynamo activity (as opposed to irregular generation of magnetic field brought about by random velocity turbulence with a *fixed* level of activity). In the second of these papers[6] the singular eigenvalue equations are set up which describe the behavior of the large-scale field (an equation of the Dyson type) and the small-scale turbulent field energy (an equation of the Bethe-Salpeter type).

In the above two papers, all correlated quantities are taken to be homogeneous in space and time. In the present paper where the turbulence is confined to a "box" the correlated quantities can be neihter homogeneous nor isotropic. Further, as we will demonstrate directly, the finite (but large) size of the box enters directly into the normal modes and their growth rates [*vide* Eq. (58)]. As the size of the box increases both the level of turbulence [given through Eq. (3)] and the attendant normal mode growth rate [given through Eq. (58)] decrease. Thus the generation of magnetic field brought about by turbulence in a "box" is rather different than that obtaining in an infinite medium.

Now most of the astrophysical situations (the Earth, Sun, galaxy, etc.) in which a dynamo mechanism is invoked to account for the continued presence of a large-scale magnetic field are objects of finite size. The question then naturally arises: Are there any statistically *exact* situations which give kinematic dynamo activity when the velocity turbulence is confined to a spatial domain of finite size (a 'box')?

The purpose of the present paper is to answer this question affirmatively by considering a simple form of velocity turbulence confined to a cubical box.

II. EQUATIONS OF MOTION AND THE EVOLUTION OF PROBABILITY

Consider then a finite medium, of constant resistivity η, which is not undergoing either bulk convection or shear, so that only a turbulent velocity, $\mathbf{V}$, with zero mean is present. Then the magnetohydrodynamic equations for the vector potential $\mathbf{A}$ are

$$\left(\frac{\partial}{\partial t} - \eta\nabla^2\right)A_i = \epsilon_{ijk}V_j(\mathbf{x}, t)B_k(\mathbf{x}, t), \tag{1}$$

with the magnetic field B given by

$$B_i(\mathbf{x}, t) = \epsilon_{ijk}\partial A_k/\partial x_j. \tag{2}$$

For random velocities $\mathbf{V}$ which are arbitrary functions of both space and time, Eqs. (1) and (2) are difficult to solve.

Elsewhere,[3,4,6-8] we have investigated some of the exact statistical kinematic dynamo properties of Eqs. (1) and (2) in an *infinite* medium when the random velocity $\mathbf{V}(\mathbf{x}, t)$ was taken to depend on only one coordinate (either spatial or temporal). In the present paper we consider the possibility of kinematic dynamo action when the turbulent velocity is confined to a cubical 'box' of side L. We describe the turbulent velocity field $\mathbf{V}$ by

$$V_x = V_1(t)\,\sin(N\pi x/L)\,\cos(N\pi y/L)\,\cos(N\pi z/L), \tag{3a}$$

$$V_y = V_2(t)\cos(N\pi x/L)\sin(N\pi y/L)\cos(N\pi z/L), \quad (3b)$$

$$V_z = V_3(t)\cos(N\pi x/L)\cos(N\pi y/L)\sin(N\pi z/L), \quad (3c)$$

where N is a given integer and where the normal components of $\mathbf{V}$ vanishes on the box walls. We point out here that the method to be employed in investigating the kinematic dynamo activity of Eq. (1) works equally well if (i) the box is rectangular rather than cubical, (ii) if the fluid is compressible rather than incompressible, and (iii) if the fluid can flow out of the box rather than being confined to the box. However the extra "bookkeeping" necessary when any one, or all, of the conditions, cubical, incompressible or confined, is relaxed is both unwieldy and, quite frankly, uninviting. It also obscures, in a plethora of notation, the basic structure of the equations. For these reasons we confine our attention to the simple situation given in the text.] The amplitudes $V_1(t), V_2(t), V_3(t)$ are random functions of time with zero mean. For the remainder of this paper we take the velocity turbulence to be incompressible, $\nabla\cdot\mathbf{V} = 0$, so that, from Eqs. (2), we have

$$V_1(t) + V_2(t) + V_3(t) = 0. \quad (4)$$

Therefore only two of the random velocity amplitudes, say V_1 and V_2, are independently assignable.

Now write

$$A_i(x,y,z,t) = \sum_{n,m,l=-\infty}^{\infty} \exp[i\pi L^{-1}(nx + my + lz)] \times A^{(i)}(n,m,l,t). \quad (5)$$

Use Eqs. (3)–(5) in Eq. (1). Then equate the coefficients of like factors $\exp[i\pi L^{-1}(nx + my + lz)]$ to obtain, after a little algebra,

$$\frac{\partial A^{(x)}}{\partial\tau}(n,m,l) + T\pi^2 L^{-2}(n^2+m^2+l^2)A^{(x)}(n,m,l) = T\pi\epsilon(8L)^{-1}[V_2(① - ②) - V_1②], \quad (6a)$$

$$\frac{\partial A^{(y)}}{\partial\tau}(n,m,l) + T\pi^2 L^{-2}(n^2+m^2+l^2)A^{(y)}(n,m,l) = T\pi\epsilon(8L)^{-1}[V_1(③ - ④) - V_2④], \quad (6b)$$

$$\frac{\partial A^{(z)}}{\partial\tau}(n,m,l) + T\pi^2 L^{-2}(n^2+m^2+l^2)A^{(z)}(n,m,l) = T\pi\epsilon(8L)^{-1}[V_1⑤ + V_2⑥], \quad (6c)$$

where

$$① = (n-N)[A^{(y)}(n-N,m-N,l-N) + A^{(y)}(n-N,m-N,l+N) - A^{(y)}(n-N,m+N,l-N) - A^{(y)}(n-N,m+N,l+N)] + (n+N)[A^{(y)}(n+N,m-N,l+N) + A^{(y)}(n+N,m-N,l-N) - A^{(y)}(n+N,m+N,l+N) - A^{(y)}(n+N,m+N,l-N)] - (m-N)[A^{(x)}(n-N,m-N,l-N) + A^{(x)}(n-N,m-N,l+N) + A^{(x)}(n+N,m-N,l+N) + A^{(x)}(n+N,m-N,l-N)] + (m+N)[A^{(x)}(n+N,m+N,l+N) + A^{(x)}(n-N,m+N,l-N) + A^{(x)}(n+N,m+N,l-N) + A^{(x)}(n-N,m+N,l+N)], \quad (7)$$

$$② = (n-N)[A^{(z)}(n-N,m-N,l-N) + A^{(z)}(n-N,m+N,l-N) - A^{(z)}(n-N,m-N,l+N) - A^{(z)}(n-N,m+N,l+N)] + (n+N)[A^{(z)}(n+N,m-N,l-N) + A^{(z)}(n+N,m+N,l-N) - A^{(z)}(n+N,m-N,l+N) - A^{(z)}(n+N,m+N,l+N)] - (l-N)[A^{(x)}(n-N,m-N,l-N) + A^{(x)}(n-N,m+N,l-N) + A^{(x)}(n+N,m-N,l-N) + A^{(x)}(n+N,m+N,l-N)] + (l+N)[A^{(x)}(n+N,m+N,l+N) + A^{(x)}(n-N,m-N,l+N) + A^{(x)}(n+N,m-N,l+N) + A^{(x)}(n-N,m+N,l+N)], \quad (8)$$

$$③ = (m-N)[A^{(x)}(n-N,m-N,l-N) + A^{(x)}(n-N,m-N,l+N) - A^{(x)}(n+N,m-N,l+N) - A^{(x)}(n+N,m-N,l-N)] + (m+N)[A^{(x)}(n-N,m+N,l+N) + A^{(x)}(n-N,m+N,l-N) - A^{(x)}(n+N,m+N,l+N) - A^{(x)}(n+N,m+N,l-N)] - (n-N)[A^{(y)}(n-N,m-N,l-N) + A^{(y)}(n-N,m-N,l+N) + A^{(y)}(n-N,m+N,l+N) + A^{(y)}(n-N,m+N,l-N)] + (n+N)[A^{(y)}(n+N,m+N,l+N) + A^{(y)}(n+N,m-N,l+N) + A^{(y)}(n+N,m+N,l-N) + A^{(x)}(n+N,m-N,l-N)], \quad (9)$$

$$④ = (m-N)[A^{(z)}(n-N,m-N,l-N) + A^{(z)}(n+N,m-N,l-N) - A^{(z)}(n-N,m-N,l+N) - A^{(z)}(n+N,m-N,l+N)] + (m+N)[A^{(z)}(n-N,m+N,l-N) + A^{(z)}(n+N,m+N,l-N) - A^{(z)}(n-N,m+N,l+N) - A^{(z)}(n+N,m+N,l+N)] - (l-N)[A^{(y)}(n-N,m-N,l-N) + A^{(y)}(n+N,m-N,l-N) + A^{(y)}(n-N,m+N,l-N) + A^{(y)}(n+N,m+N,l-N)] + (l+N)[A^{(y)}(n-N,m-N,l+N) + A^{(y)}(n+N,m-N,l+N) + A^{(y)}(n-N,m+N,l+N) + A^{(y)}(n+N,m+N,l+N)], \quad (10)$$

$$⑤ = (l-N)[A^{(x)}(n-N,m-N,l-N) + A^{(x)}(n-N,m+N,l-N) - A^{(x)}(n+N,m-N,l-N) - A^{(x)}(n+N,m+N,l-N)] + (l+N)[A^{(x)}(n-N,m-N,l+N) + A^{(x)}(n-N,m+N,l+N)$$

$$- A^{(x)}(n+N, m-N, l+N) - A^{(x)}(n+N, m+N, l+N)] - (n-N)[A^{(z)}(n-N, m-N, l-N)$$
$$+ A^{(z)}(n-N, m+N, l-N) + A^{(z)}(n-N, m-N, l+N) + A^{(z)}(n-N, m+N, l+N)]$$
$$+ (n+N)[A^{(z)}(n+N, m-N, l-N) + A^{(z)}(n+N, m+N, l-N) + A^{(z)}(n+N, m-N, l+N)$$
$$+ A^{(z)}(n+N, m+N, l+N)], \tag{11}$$

$$⑥ = (l-N)[A^{(y)}(n-N, m-N, l-N) + A^{(y)}(n+N, m-N, l-N) - A^{(y)}(n-N, m+N, l-N)$$
$$- A^{(y)}(n+N, m+N, l-N)] + (l+N)[A^{(y)}(n-N, m-N, l+N) + A^{(y)}(n+N, m-N, l+N)$$
$$- A^{(y)}(n-N, m+N, l+N) - A^{(y)}(n+N, m+N, l+N)] - (m-N)[A^{(z)}(n-N, m-N, l-N)$$
$$+ A^{(z)}(n+N, m-N, l-N) + A^{(z)}(n-N, m-N, l+N) + A^{(z)}(n+N, m-N, l+N)]$$
$$+ (m+N)[A^{(z)}(n-N, m+N, l-N) + A^{(z)}(n+N, m+N, l-N) + A^{(z)}(n-N, m+N, l+N)$$
$$+ A^{(z)}(n+N, m+N, l+N)]. \tag{12}$$

We have suppressed the variable τ in each $A(n, m, l)$ with the understanding that it is present. Further $\tau = t/T$, where T is the correlation time for $\mathbf{V}$. We have also written $V_1 = \epsilon v_1$, $v_2 = \epsilon v_2$, so that $\langle V_1^2 \rangle = \epsilon^2 = \langle V_2^2 \rangle$.

Now consider the probability P of finding the combination of values v_1, v_2, $\{A^{(x)}(n, m, l)\}$, $\{A^{(y)}(n, m, l)\}$, and $\{A^{(z)}(n, m, l)\}$ at time τ. Let the probability of finding v_1 and v_2 *on their own* be described by the operator field $\mathcal{L}(v_1, v_2)$.

Then $P(\tau, v_1, v_2, \{A^{(x)}(n, m, l)\}, \{A^{(y)}(n, m, l)\}, \{A^{(z)}(n, m, l)\})$ satisfies the spatially homogeneous equation

$$\frac{\partial P}{\partial \tau} = \mathcal{L}(v_1, v_2)P - \sum_{\substack{n,m,l=-\infty \\ i=x,y,z}}^{\infty} \frac{\partial}{\partial A^{(i)}(n, m, l)} \times \left(\frac{\partial A^{(i)}(n, m, l)}{\partial \tau} P \right), \tag{13}$$

where the $\partial A^{(i)}(n, m, l)/\partial \tau$ are given by Eqs. (6).

While Eq. (13) is both general and exact, it is difficult to proceed further until the statistical distribution of v_1 and v_2 on their own is specified. For the remainder of this paper we shall take $\mathcal{L}(v_1, v_2)$ to represent a Gaussian velocity distribution in both v_1 and v_2, with the same correlation time and intensity in each component. Then[9]

$$\mathcal{L}(v_1, v_2)P \equiv \frac{\partial}{\partial v_1}(v_1 P) + \frac{\partial}{\partial v_2}(v_2 P) + \frac{\partial^2 P}{\partial v_1^2} + \frac{\partial^2 P}{\partial v_2^2}, \tag{14}$$

and Eq. (13) becomes

$$\frac{\partial P}{\partial \tau} = \frac{\partial}{\partial v_1}(v_1 P) + \frac{\partial}{\partial v_2}(v_2 P) + \frac{\partial^2 P}{\partial v_1^2} + \frac{\partial^2 P}{\partial v_2^2} - \sum_{\substack{n,m,l=-\infty, \\ i=x,y,z}}^{\infty} \frac{\partial}{\partial A^{(i)}(n, m, l)} \left(\frac{\partial A^{(i)}(n, m, l)}{\partial \tau} P \right). \tag{15}$$

The first four terms on the right-hand side of Eq. (14) represent the assumption that the stationary probability distribution over v_1, v_2 alone is the Gaussian $\exp(-\frac{1}{2}(v_1^2 + v_2^2))$.

The initial values of $A^{(i)}(n, m, l)$ are sufficient to determine the solution of (15). Denote them by $A^{(i)}(n, m, l; 0)$, so that at $\tau = 0$ the probability distribution is

$$P(\tau = 0) = (2\pi)^{-1} \exp(-\tfrac{1}{2}(v_1^2 + v_2^2)) \times \prod_{\substack{n,m,l=-\infty \\ i=x,y,z}}^{\infty} \delta[A^{(i)}(n, m, l; \tau) - A^{(i)}(n, m, l; 0)]. \tag{16}$$

To proceed with the solution of Eqs. (15) define the quantities

$$R^j(n', m', l'; v_1, v_2; \tau) = \int A^{(j)}(n', m', l') P \prod_{\substack{n,m,l=-\infty \\ i=x,y,z}}^{\infty} \times dA^{(i)}(n, m, l); \tag{17}$$

so that the amplitude of the component of the ensemble average vector potential $\langle \mathbf{A} \rangle$ varying as $\exp[i\pi L^{-1}(nx + my + lz)]$ is

$$\langle A^{(i)}(n, m, l; \tau) \rangle = \int R^i(n, m, l; v_1, v_2; \tau) dv_1 \, dv_2, \tag{18}$$

with the associated ensemble average magnetic field amplitude

$$\langle B^{(i)}(n, m, l; \tau) \rangle = \exp[-i\pi L^{-1}(nx + my + lz)] \times \epsilon_{ijk} \langle A^{(k)}(n, m, l; \tau) \rangle \frac{\partial}{\partial x_j} \exp[i\pi L^{-1}(nx + my + lz)]. \tag{19}$$

From Eq. (15) we obtain

$$\frac{\partial R^x(n, m, l)}{\partial \tau} = \mathcal{L}(v_1, v_2)R^x(n, m, l) - \pi^2 \eta T L^{-2}(n^2 + m^2 + l^2)R^x(n, m, l) + \pi \epsilon T (8L)^{-1}$$
$$\times (v_2\{(n-N)[R^y(n-N, m-N, l-N) + R^y(n-N, m-N, l+N) - R^y(n-N, m+N, l-N)$$
$$- R^y(n-N, m+N, l+N) - R^z(n-N, m-N, l-N) - R^z(n-N, m+N, l-N)$$
$$+ R^z(n-N, m-N, l+N) + R^z(n-N, m+N, l+N)] + (n+N)[R^y(n+N, m-N, l+N)$$
$$+ R^y(n+N, m-N, l-N) - R^{(y)}(n+N, m+N, l+N) - R^y(n+N, m+N, l-N)$$
$$- R^z(n+N, m-N, l-N) - R^{(z)}(n+N, m+N, l-N) + R^z(n+N, m+N, l+N)$$
$$+ R^z(n+N, m-N, l+N)] + (m+l)[R^x(n+N, m+N, l-N) + R^x(n-N, m+N, l-N)$$

$$- R^x(n-N,m-N,l+N) - R^x(n+N,m-N,l+N)] + (m-l)[R^x(n+N,m+N,l+N)$$
$$+ R^x(n-N,m+N,l+N) - R^x(n+N,m-N,l-N) - R^x(n-N,m-N,l-N)]\}$$
$$- v_1\{(n-N)[R^z(n-N,m-N,l-N) + R^z(n-N,m+N,l-N) - R^z(n-N,m-N,l+N)$$
$$- R^z(n-N,m+N,l+N)] + (n+N)[R^z(n+N,m-N,l-N) + R^z(n+N,m+N,l-N)$$
$$- R^z(n+N,m-N,l+N) - R^z(n+N,m+N,l+N)] - (l-N)[R^x(n-N,m-N,l-N)$$
$$+ R^x(n-N,m+N,l-N) + R^x(n+N,m-N,l-N) + R^x(n+N,m+N,l-N)] + (l+N)$$
$$\times [R^x(n+N,m+N,l+N) + R^x(n-N,m-N,l-N) + R^x(n+N,m-N,l+N)$$
$$+ R^x(n-N,m+N,l+N)]\}, \tag{20}$$

$$\frac{\partial R^y(n,m,l)}{\partial \tau} = \mathcal{L}(v_1,v_2)R^y(n,m,l) - \pi^2\eta TL^{-2}(n^2+m^2+l^2)R^y(n,m,l) + \pi\epsilon T(8L)^{-1}$$
$$\times (v_1\{(m-N)[R^x(n-N,m-N,l-N) + R^x(n-N,m-N,l+N) - R^x(n+N,m-N,l+N)$$
$$- R^x(n+N,m-N,l+N) - R^z(n-N,m-N,l-N) - R^z(n+N,m-N,l-N)$$
$$+ R^z(n-N,m-N,l+N) + R^z(n+N,m-N,l+N)] + (m+N)[R^x(n-N,m+N,l+N)$$
$$+ R^x(n-N,m+N,l-N) - R^x(n+N,m+N,l+N) - R^x(n+N,m+N,l-N)$$
$$- R^z(n-N,m+N,l-N) - R^z(n+N,m+N,l-N) + R^z(n-N,m+N,l+N)$$
$$+ R^z(n+N,m+N,l+N)] + (l-n)[R^y(n-N,m-N,l-N) + R^y(n-N,m+N,l-N)$$
$$- R^y(n+N,m+N,l+N) - R^y(n+N,m-N,l+N)] + (l+N)[R^y(n+N,m+N,l-N)$$
$$+ R^y(n+N,m-N,l-N) - R^y(n-N,m-N,l+N) - R^y(n-N,m+N,l+N)]\}$$
$$- v_2\{(m-N)[R^z(n-N,m-N,l-N) + R^z(n+N,m-N,l-N) - R^z(n-N,m-N,l+N)$$
$$- R^z(n+N,m-N,l+N)] + (m+N)[R^z(n-N,m+N,l-N) + R^z(n+N,m+N,l-N)$$
$$- R^z(n-N,m+N,l+N) - R^z(n+N,m+N,l+N)] - (l-N)[R^y(n-N,m-N,l-N)$$
$$+ R^y(n+N,m-N,l-N) + R^y(n-N,m+N,l-N) + R^y(n+N,m+N,l-N)] + (l+N)$$
$$\times [R^y(n-N,m-N,l+N) + R^y(n+N,m-N,l+N) + R^y(n-N,m+N,l+N)$$
$$+ R^y(n+N,m+N,l+N)]\}) \tag{21}$$

$$\frac{\partial R^z(n,m,l)}{\partial \tau} = \mathcal{L}(v_1,v_2)R^z(n,m,l) - \pi^2\eta TL^{-2}(n^2+m^2+l^2)R^z(n,m,l) + \pi\epsilon T(8L)^{-1}$$
$$\times (v_1\{(l-N)[R^x(n-N,m-N,l-N) + R^x(n-N,m+N,l-N) - R^x(n+N,m-N,l-N)$$
$$- R^x(n+N,m+N,l-N)] + (l+N)[R^x(n-N,m-N,l+N) + R^x(n-N,m+N,l+N)$$
$$- R^x(n+N,m-N,l+N)] - R^x(n+N,m+N,l+N)] - (n-N)[R^z(n-N,m-N,l-N)$$
$$+ R^z(n-N,m+N,l-N) + R^z(n-N,m-N,l+N) + R^z(n-N,m+N,l+N)] + (n+N)$$
$$\times [R^z(n+N,m+N,l-N) + R^z(n+N,m-N,l-N) + R^z(n+N,m-N,l+N)$$
$$+ R^z(n+N,m+N,l+N)]\} + v_2\{(l-N)[R^y(n-N,m-N,l-N) + R^y(n+N,m-N,l-N)$$
$$- R^y(n-N,m+N,l-N) - R^y(n+N,m+N,l-N)] + (l+N)[R^y(n-N,m-N,l+N)$$
$$+ R^y(n+N,m-N,l+N) - R^y(n-N,m+N,l+N) - R^y(n+N,m+N,l+N)] - (m-N)$$
$$\times [R^z(n-N,m-N,l-N) + R^z(n+N,m-N,l-N) + R^z(n-N,m-N,l+N)$$
$$+ R^z(n+N,m-N,l+N)] + (m+N)[R^z(n-N,m+N,l-N) + R^z(n+N,m+N,l-N)$$
$$+ R^z(n-N,m+N,l+N) + R^z(n+N,m+N,l+N)]\}). \tag{22}$$

In writing Eqs. (20)–(22) the dependence of each $R^{(i)}$ on v_1, v_2, and τ is implied rather than spelled out explicitly. By inspection of Eqs. (20)–(22), we see that they represent a linear threefold infinite set of finite difference equations in n, m, and l; that they are second order differential equations in v_1 and v_2; and that the coefficients are independent of time τ. So each and every $R^{(i)}$ has normal modes with the dependence

$$R^{(i)}(n,m,l;v_1,v_2;\tau) = e^{\sigma\tau}R^{(i)}(n,m,l;v_1,v_2), \tag{23}$$

where σ is independent of n, m, l, v_1, and v_2.

In order to demonstrate dynamo activity it is both sufficient and necessary to show that at least one of the allowed "eigenvalues" σ exists with $\mathrm{Re}(\sigma) > 0$.

Before proceeding with the discussion of Eqs. (20)–(22), it is opportune here to write down some auxiliary expressions for we will make use of them in Sec. III.

It will prove convenient to expand later in terms of the eigenfunctions, ψ_r of the homogeneous equation

$$\frac{d^2\psi_r}{dx^2} + \frac{d}{dx}(x\psi_r) + r\psi_r = 0, \tag{24}$$

which are

$$\psi_r(x) = \exp(-\tfrac{1}{2}x^2)H_r(x2^{-1/2}), \tag{25}$$

where H_r is the rth Hermite polynomial. From the recurrence relation for the Hermite polynomials, we obtain

$$2^{1/2}x\psi_r(x) = \psi_{r+1}(x) + 2r\psi_{r-1}(x). \tag{26}$$

It is clear by inspection of Eqs. (20)–(22) that by writing

$$R^{(i)}(n,m,l;v_1,v_2) = \sum_{r,s=0}^{\infty} R^{(i)}_{rs}(n,m,l)\psi_r(v_1)\psi_s(v_2), \tag{27}$$

it is possible to reduce the equations to algebraic finite difference equations in the five-integer space of n, m, l, r, and s. It is also clear that such a reduction will then yield an infinite determinant as a consistency condition for the homogeneous solutions to Eqs. (20)–(22). The zeros of the determinant then yield the allowed σ values. This techniques has been used elsewhere[8] to obtain those eigenvalues giving rise to kinematic dynamo action in an infinite medium, when the random velocity field is a function of only one spatial coordinate.

In the present case where the random velocity field possesses *two* independent amplitudes, which are functions of time, and also depends on all three spatial coordinates, we have been unable to obtain (much less solve!) the infinite determinant. We, therefore, resort to a different technique to investigate the behavior and signature of the eigenvalues σ of Eqs. (20)–(22).

III. REDUCTION OF THE FINITE-DIFFERENCE EQUATIONS USING A VARIATION OF KAC'S "LIMIT-THEOREM" APPROACH

So far the finite-difference Eqs. (20)–(22) [or their equivalent using (27)] are statistically *exact* equations to the problem of kinematic dynamo activity in a box. They are quite difficult to reduce any further exactly.

However, when the box is large enough, we can use a variation of a technique employed by Kac[10] which he ascribes to Smoluchowski.[11] In the form suitable for our equations we write

$$R^{(i)}(n-N,\cdots) = R^{(i)}(n/L - N/L,\cdots).$$

For large values of L (N fixed), we have

$$R^{(i)}(n/L - N/L,\cdots) = R^{(i)}(n/L) - \frac{N}{L}\frac{\partial R^{(i)}(n/L)}{\partial(n/L)} + \cdots,$$

since in the limit $L \to \infty$ (N fixed) only those n such that $n/L \to$ finite contribute. In order to anticipate the transformation to continuous variables we write $N = \Delta L$,

$$L^{-1}(n,m,l) = (\zeta,\nu,\xi) \equiv \mathbf{k},$$

with $|\mathbf{k}| \gg \Delta$.

Then for large enough (but finite) L, Eqs. (20)–(22) reduce to

$$\begin{aligned}\sigma R^x = {} & \mathfrak{L}(v_1,v_2)R^x - \pi^2\eta T(\zeta^2+\nu^2+\xi^2)R^x \\ & + \pi\epsilon T\Delta\Big\{v_2\Big[\zeta\Big(\frac{\partial R^z}{\partial\xi} - \frac{\partial R^y}{\partial\nu}\Big) + \frac{\nu\partial R^x}{\partial\nu} - \xi\frac{\partial R^x}{\partial\xi} \\ & + v_1\Big[\zeta\frac{\partial R^z}{\partial\xi} - R^x - \nu\frac{\partial R^x}{\partial\xi}\Big]\Big\},\end{aligned} \tag{28}$$

$$\begin{aligned}\sigma R^y = {} & \mathfrak{L}(v_1,v_2)R^y - \pi^2\eta T(\zeta^2+\nu^2+\xi^2)R^y \\ & + \pi\epsilon T\Delta\Big\{v_1\Big[\nu\frac{\partial R^z}{\partial\xi} + \zeta\frac{\partial R^y}{\partial\zeta} - \xi\frac{\partial R^y}{\partial\xi} - \nu\frac{\partial R^x}{\partial\zeta}\Big] \\ & - v_2\Big[R^y + \xi\frac{\partial R^y}{\partial\xi} - \nu\frac{\partial R^z}{\partial\zeta}\Big]\Big\},\end{aligned} \tag{29}$$

$$\begin{aligned}\sigma R^z = {} & \mathfrak{L}(v_1,v_2)R^z - \pi^2\eta T(\zeta^2+\nu^2+\xi^2)R^z \\ & + \pi\epsilon T\Delta\Big\{v_1\Big[\zeta\frac{\partial R^z}{\partial\zeta} + R^z - \xi\frac{\partial R^x}{\partial\zeta}\Big] \\ & + v_2\Big[R^z + \nu\frac{\partial R^z}{\partial\nu} - \xi\frac{\partial R^y}{\partial\nu}\Big]\Big\},\end{aligned} \tag{30}$$

where $R^i \equiv R^i(\zeta,\nu,\xi,v_1,v_2)$.

Now write

$$R^i = \sum_{n,m=0} R^i_{nm}(\zeta,\nu,\xi)\psi_n(v_1)\psi_m(v_2), \tag{31}$$

and use the recurrence relation (26) to obtain from Eqs. (28)–(30) the equations

$$\begin{aligned}\sigma R^x_{nm} = {} & -(n+m)R^x_{nm} - \eta\pi^2T(\zeta^2+\nu^2+\xi^2)R^x_{nm} \\ & + \pi\epsilon T\Delta 2^{-1/2}\Big\{\zeta\Big[\frac{\partial R^z_{n,m-1}}{\partial\xi} + 2(m+1)\frac{\partial R^z_{n,m+1}}{\partial\xi} \\ & - \frac{\partial R^y_{n,m-1}}{\partial\nu} - 2(m+1)\frac{\partial R^y_{n,m+1}}{\partial\nu}\Big] \\ & + \nu\frac{\partial}{\partial\nu}[R^x_{n,m-1} + 2(m+1)R^x_{n,m+1}] \\ & - \xi\frac{\partial}{\partial\xi}[R^x_{n,m-1} + 2(m+1)R^x_{n,m+1}] \\ & + \zeta\frac{\partial}{\partial\xi}[R^z_{n-1,m} + 2(n+1)R^z_{n+1,m}] \\ & - \Big(1 + \nu\frac{\partial}{\partial\xi}\Big)[R^x_{n-1,m} + 2(n+1)R^x_{n+1,m}]\Big\},\end{aligned} \tag{32}$$

$$\begin{aligned}\sigma R^y_{nm} = {} & -(n+m)R^y_{nm} - \pi^2\eta T(\zeta^2+\nu^2+\xi^2)R^y_{nm} \\ & + \pi\epsilon T\Delta 2^{-1/2}\Big\{\nu\frac{\partial}{\partial\xi}[R^z_{n-1,m} + 2(n+1)R^z_{n+1,m}] \\ & - \nu\frac{\partial}{\partial\zeta}[R^x_{n-1,m} + 2(n+1)R^x_{n+1,m}] + \nu\frac{\partial}{\partial\xi} \\ & \times [R^z_{n,m-1} + 2(m+1)R^z_{n,m+1}] \\ & + \Big(\zeta\frac{\partial}{\partial\zeta} - \xi\frac{\partial}{\partial\xi}\Big)[R^y_{n-1,m} + 2(n+1)R^y_{n+1,m}] \\ & - \Big(1 + \xi\frac{\partial}{\partial\xi}\Big)[R^y_{n,m-1} + 2(m+1)R^y_{n,m+1}],\end{aligned} \tag{33}$$

$$\begin{aligned}\sigma R^z_{nm} = {} & -(n+m)R^z_{nm} - \pi^2\eta T(\zeta^2+\nu^2+\xi^2)R^z_{nm} \\ & + \pi\epsilon T\Delta 2^{-1/2}\Big\{-\xi\frac{\partial}{\partial\zeta}[R^x_{n-1,m} + 2(n+1)R^x_{n+1,m}] \\ & + \Big(1 + \zeta\frac{\partial}{\partial\zeta}\Big)[R^z_{n-1,m} + 2(n+1)R^z_{n+1,m}] \\ & - \xi\frac{\partial}{\partial\nu}[R^y_{n,m-1} + 2(m+1)R^y_{n,m+1}] \\ & + \Big(1 + \nu\frac{\partial}{\partial\nu}\Big)[R^z_{n,m-1} + 2(m+1)R^z_{n,m+1}]\Big\}.\end{aligned} \tag{34}$$

The finite difference equations (32)–(34) can be converted into exact differential equations by employing a technique due to Kac.[10] Write

$$Q^i(\alpha, \beta; \zeta, \nu, \xi) = \sum_{n,m=0}^{\infty} R^i_{nm}(\zeta, \nu, \xi)\alpha^n\beta^m. \tag{35}$$

We shall refer to the Q^i as 'generator fields' for the R^i_{nm}. Then it is obvious by inspection of Eqs. (32)–(34) that the generator fields satisfy the coupled first-order equations

$$\begin{aligned}\sigma Q^x = &- \alpha Q^x_\alpha - \beta Q^x_\beta - \pi^2\eta T(\zeta^2 + \nu^2 + \xi^2)Q^x \\ &+ \pi\epsilon T\Delta 2^{-1/2}[\zeta(\beta Q^z_\xi + 2Q^z_{\xi\beta} - \beta Q^y_\nu - 2Q^y_{\nu\beta}) \\ &+ \nu(\beta Q^x_\nu + 2Q^x_{\nu\beta}) - \xi(\beta Q^x_\xi + 2Q^x_{\xi\beta}) \\ &+ \zeta(\alpha Q^z_\xi + 2Q^z_{\alpha\xi}) - (\alpha Q^x + 2Q^x_\alpha) \\ &- \nu(\alpha Q^x_\xi + 2Q^x_{\alpha\xi})],\end{aligned} \tag{36}$$

$$\begin{aligned}\sigma Q^y = &- \alpha Q^y_\alpha - \beta Q^y_\beta - \pi^2\eta T(\zeta^2 + \nu^2 + \xi^2)Q^y \\ &+ \pi\epsilon T\Delta 2^{-1/2}[-\nu(\alpha Q^x_\zeta + 2Q^x_{\zeta\alpha} - \alpha Q^z_\xi - 2Q^z_{\alpha\xi}) \\ &+ \zeta(\alpha Q^y_\zeta + 2Q^y_{\alpha\zeta}) - \xi(\alpha Q^y_\xi + 2Q^y_{\alpha\xi}) \\ &+ \nu(\beta Q^z_\xi + 2Q^z_{\beta\xi}) - (\beta Q^y + 2Q^y_\beta) \\ &- \xi(\beta Q^y_\xi + 2Q^y_{\beta\xi})],\end{aligned} \tag{37}$$

$$\begin{aligned}\sigma Q^z = &- \alpha Q^z_\alpha - \beta Q^z_\beta - \pi^2\eta T(\zeta^2 + \nu^2 + \xi^2)Q^z \\ &+ \pi\epsilon T\Delta 2^{-1/2}[-\xi(\alpha Q^x_\zeta + 2Q^x_{\alpha\zeta}) + \alpha Q^z + 2Q^z_\alpha \\ &+ \zeta(\alpha Q^z_\zeta + 2Q^z_{\alpha\zeta}) - \xi(\beta Q^y_\nu + 2Q^y_{\nu\beta}) \\ &+ \beta Q^z + 2Q^z_\beta + \nu(\beta Q^z_\nu + 2Q^z_{\nu\beta})],\end{aligned} \tag{38}$$

where

$$Q^i_{abc\ldots} = \left[\frac{\partial}{\partial a}\cdot\frac{\partial}{\partial b}\cdot\frac{\partial}{\partial c}\cdots\right]Q^i.$$

The advantage of expanding $R^i(\zeta, \nu, \xi, v_1, v_2)$ in the normal modes of $\mathfrak{L}(v_1, v_2)$ and then resumming the coefficients is fairly clear. By so doing we convert the second-order equations (28)–(30) (in v_1 and v_2) into first-order equations (36)–(39) (in α and β) and, as is usual, discussion of first order equations with nonconstant coefficients is somewhat simpler than discussion of second-order equations with nonconstant coefficients.

Equations (36)–(38) can be simplified by the transformation

$$Q^i = f^i \exp[-\tfrac{1}{4}(\alpha^2 + \beta^2)], \tag{39}$$

to yield

$$\begin{aligned}\sigma f^x = &- \alpha f^x_\alpha + \tfrac{1}{2}\alpha^2 f^x - \beta f^x_\beta + \tfrac{1}{2}\beta^2 f^x \\ &- \eta\pi^2 T(\zeta^2 + \nu^2 + \xi^2)f^x \\ &+ \pi\epsilon T\Delta 2^{1/2}[\zeta(f^z_{\xi\beta} - f^y_{\nu\beta}) + \nu(f^x_{\nu\beta} - f^x_{\xi\alpha}) \\ &- \xi f^x_{\xi\beta} + \zeta f^z_{\xi\alpha} - f^x_\alpha],\end{aligned} \tag{40}$$

$$\begin{aligned}\sigma f^y = &- \alpha f^y_\alpha + \tfrac{1}{2}\alpha^2 f^y - \beta f^y_\beta + \tfrac{1}{2}\beta^2 f^y \\ &- \eta\pi^2 T(\zeta^2 + \nu^2 + \xi^2)f^y \\ &+ \pi\epsilon T\Delta 2^{1/2}[\nu(f^z_{\zeta\alpha} - f^x_{\zeta\alpha} + f^z_{\xi\beta}) + \zeta f^y_{\zeta\alpha} - f^y_\beta \\ &- \xi(f^y_{\xi\alpha} + f^y_{\xi\beta})],\end{aligned} \tag{41}$$

$$\begin{aligned}\sigma f^z = &- \alpha f^z_\alpha + \tfrac{1}{2}\alpha^2 f^z - \beta f^z_\beta + \tfrac{1}{2}\beta^2 f^z \\ &- \eta\pi^2 T(\zeta^2 + \nu^2 + \xi^2)f^z \\ &+ \pi\epsilon T\Delta 2^{1/2}[-\xi(f^x_{\zeta\alpha} + f^y_{\nu\beta}) + \zeta f^z_{\zeta\alpha} + \nu f^z_{\nu\beta} \\ &+ f^z_\alpha + f^z_\beta].\end{aligned} \tag{42}$$

Note from Eqs. (40)–(42) that the generator fields f^i are either centrally symmetric or antisymmetric with respect to (ζ, ν, ξ), i.e.,

$$f_S(-\zeta, -\nu, -\xi) = f_S(\zeta, \nu, \xi) \tag{43a}$$

and

$$f_A(-\zeta, -\nu, -\xi) = -f_A(\zeta, \nu, \xi). \tag{43b}$$

IV. LONG-WAVELENGTH SYMMETRIC MODES

We have not succeeded in obtaining the general symmetric solution to Eqs. (40)–(42). But in demonstrating that Eqs. (40)–(42) generate kinematic dynamo activity the general solution of the equations is not essential, as we now demonstrate directly.

Consider the long-wavelength situation $(\zeta^2 + \nu^2 + \xi^2)\eta T \ll 1$.

Write

$$f^i_S = a^i(\alpha, \beta) + b^i_{jl}(\alpha, \beta)k_j k_l + O(k^4) + \cdots,$$

where $\mathbf{k} = (\zeta, \nu, \xi)$. Then to $O(k^0)$, Eqs. (40)–(42) give

$$\sigma a^x = -\alpha a^x_\alpha - \beta a^x_\beta + \tfrac{1}{2}(\alpha^2 + \beta^2)a^x - \pi\epsilon T\Delta 2^{1/2}a^x_\alpha, \tag{44}$$

$$\sigma a^y = -\alpha a^y_\alpha - \beta a^y_\beta + \tfrac{1}{2}(\alpha^2 + \beta^2)a^y - \pi\epsilon T\Delta 2^{1/2}a^y_\beta, \tag{45}$$

$$\begin{aligned}\sigma a^z = &- \alpha a^z_\alpha - \beta a^z_\beta + \tfrac{1}{2}(\alpha^2 + \beta^2)a^z \\ &+ \pi\epsilon T\Delta 2^{1/2}(a^z_\alpha + a^z_\beta).\end{aligned} \tag{46}$$

Note from Eqs. (44)–(46) that the amplitudes a^x, a^y, and a^z are linearly independent. If we had gone to $O(k^2)$ then twenty seven inhomogeneous differential equations for the $b^i_{jl}(\alpha, \beta)$ would have resulted with the a^i acting as source terms for the b's. So the a's form the basis for computing the series expansion of f^i_S in ascending powers of $\mathbf{k}$.

Consider the solution to Eqs. (44)–(46). From Eq. (44) we see that $a^x(\alpha, \beta)$ is separable. Write $a^x(\alpha, \beta) = R(\alpha)S(\beta)$.

Then from Eq. (44) we obtain

$$-\beta S_\beta + \tfrac{1}{2}\beta^2 S = cS, \tag{47}$$

$$(\sigma - \tfrac{1}{2}\alpha^2 - c)R + (\alpha + \pi\epsilon T\Delta 2^{1/2})R_\alpha = 0, \tag{48}$$

where c is the separation constant. The general solution to Eq. (47) is

$$S(\beta) = S\beta^{-c}\exp(\tfrac{1}{4}\beta^2), \tag{49}$$

where S is a constant.

But from Eq. (35) we see that, as $\beta \to 0$, we have

$$f^x(\alpha, \beta \to 0) = \exp\tfrac{1}{4}\alpha^2 \sum_{n=0}^{\infty}\alpha^n R_{no}(\mathbf{k} = 0),$$

so that $c = 0$ and then $S(\beta) = S\exp\tfrac{1}{4}\beta^2$.

With $c \equiv 0$ the general solution to Eq. (48) is

$$R(\alpha) = R(\alpha + \Gamma)^{(\Gamma^2/2 - \sigma)}\exp[\tfrac{1}{4}(\alpha + \Gamma)^2 - \Gamma(\alpha + \Gamma)], \tag{50}$$

where $\Gamma = \pi\epsilon T\Delta 2^{1/2}$.

Now as $\alpha \to 0$, we have

$$f^x_S(\alpha \to 0, \beta \to 0) = R^x_{00}(\mathbf{k} = 0), \tag{51}$$

so that

$$RS = R^x_{00}(\mathbf{k}=0)\Gamma^{(\sigma-\Gamma^2/2)}\exp(\tfrac{3}{4}\Gamma^2), \tag{52}$$

and then

$$Q^x(\alpha,\beta;\mathbf{k}=0) \equiv \sum_{n,m=0}^{\infty} R^x_{nm}(\mathbf{k}=0)\alpha^n\beta^m$$
$$= R^x_{00}(\mathbf{k}=0)(1+\alpha/\Gamma)^{(\Gamma^2/2-\sigma)}\exp(-\tfrac{1}{2}\alpha\Gamma). \tag{53}$$

Note that this is independent of β so that $Q^x(\alpha,\beta,\mathbf{k}=0)$ is synonymous with

$$Q^x(\alpha,\beta=0=\mathbf{k}) = \sum_{n=0}^{\infty} R^x_{n0}(\mathbf{k}=0)\alpha^n. \tag{54}$$

Now replace α by $e^{i\theta}$ and integrate $0 \leq \theta \leq 2\pi$ in Eqs. (53) and (54). From Eq. (54) we have

$$2\pi R^x_{00}(\mathbf{k}=0) = \int_0^{2\pi} Q^x(e^{i\theta},\beta=0=\mathbf{k})d\theta.$$

So there exists a symmetric solution for Q^x if, and only if,

$$2\pi = \int_0^{2\pi} d\theta(1+e^{i\theta}/\Gamma)^{(\Gamma^2/2-\sigma)}\exp(-\tfrac{1}{2}\Gamma e^{i\theta}). \tag{55}$$

Equation (55) gives the allowed σ values as functions of Γ in order that a symmetric solution exist. With $z = e^{i\theta}$ we can convert the integral in Eq. (55) into an integral around the unit circle in the complex z plane centered on the origin. Then Eq. (55) becomes

$$2\pi i = \oint \frac{dz}{z}(1+z/\Gamma)^{(\Gamma^2/2-\sigma)}\exp(-\tfrac{1}{2}\Gamma z). \tag{56}$$

The immediate problem before us is to find the σ values for which the integral on the right-hand side of Eq. (56) is $2\pi i$.

Some of the roots are obvious. If $\frac{1}{2}\Gamma^2 - \sigma = n$ (n a *positive* integer including zero), then Eq. (56) is satisfied. If $\frac{1}{2}\Gamma^2 - \sigma = -m$ (m a positive integer excluding zero), then Eq. (56) is satisfied if also $\Gamma > 1$. If $\Gamma < 1$ and $\frac{1}{2}\Gamma^2 - \sigma = -m$ (m a positive integer), then Eq. (56) is satisfied if and only if

$$\frac{\partial^{m-1}}{\partial z^{m-1}}[z^{-1}\exp(-\tfrac{1}{2}\Gamma z)]_{z=-\Gamma} = 0. \tag{57}$$

If $\frac{1}{2}\Gamma^2 - \sigma$ is not an integer (either positive or negative) then we have so far been unable to obtain the remaining (if any) roots to Eq. (56). However in order to demonstrate dynamo activity we do not require all the roots. For example, the root ($n = 0$), i.e., $\sigma = \frac{1}{2}\Gamma^2 \equiv \pi^2\epsilon^2T^2N^2/L^2$ gives a growing mode with e-folding time

$$\tau_1 = L^2/(N^2\pi^2\epsilon^2). \tag{58}$$

Consider next the y component of field obtained from Eq. (45). We see by inspection that this has the same structure as Eq. (44) if we replace $\alpha \to \beta$ and $\beta \to \alpha$. So

$$Q^y(\alpha,\beta,\mathbf{k}=0)$$
$$= R^y_{00}(\mathbf{k}=0)(1+\beta/\Gamma)^{(\Gamma^2/2-\sigma)}\exp(-\tfrac{1}{2}\Gamma\beta) \tag{59}$$

and the dispersion relation satisfied by Q^y is *exactly* Eq. (55) with the same roots, i.e., the same σ values.

Consider finally the z component of field from Eq. (46). It is again clear by inspection that $a^z(\alpha,\beta)$ is separable.

Using an identical argument to that for Eqs. (44) and (45), it is easy to show that

$$Q^z(\alpha,\beta,\mathbf{k}=0) = R^z_{00}(\mathbf{k}=0)\exp[\tfrac{1}{2}\Gamma(\alpha+\beta)]$$
$$\times(1-\beta/\Gamma)^{(\Gamma^2/2-c)}(1-\alpha/\Gamma)^{(\Gamma^2/2-\sigma+c)}, \tag{60}$$

where c is the separation constant for Eq. (46). Then with $\alpha = e^{i\theta}$, $\beta = e^{i\varphi}$, we obtain the dispersion relation

$$\oint \frac{dz}{z}\exp\tfrac{1}{2}\Gamma z(1-z/\Gamma)^{(\Gamma^2/2-c)}$$
$$\times \oint \frac{dz'}{z'}\exp\tfrac{1}{2}\Gamma z'(1-z'/\Gamma)^{(\Gamma^2/2-\sigma+c)} = -4\pi^2, \tag{61}$$

where the integrals over z and z' are to be taken around the unit circles centered on the origin in the complex z and z' planes, respectively. The alternative is $R^z_{00} \equiv 0$.

But we originally assumed that each and every R^i varied as $\exp(\sigma\tau)$, so the roots of Eq. (61) have to be *identical* to the roots of Eq. (56) if $R^z_{00} \neq 0 \neq R^y_{00} \neq R^z_{00}$.

One set of roots to Eq. (61) is obtained by choosing

$$\tfrac{1}{2}\Gamma^2 - c = L, \quad L \text{ a positive integer including zero,} \tag{62a}$$

$$\tfrac{1}{2}\Gamma^2 - \sigma + c = M, \quad M \text{ a positive integer including zero.} \tag{62b}$$

From Eqs. (62) we obtain

$$\sigma = \Gamma^2 - (L+M). \tag{63}$$

But from Eq. (56) we have

$$\sigma = \tfrac{1}{2}\Gamma^2 - n. \tag{64}$$

Equations (63) and (64) are compatible if, and only if,

$$\tfrac{1}{2}\Gamma^2 = L + M - n, \tag{65}$$

which requires $L + M \geq n$.

So symmetric long-wavelength dynamo modes exist in a box with the growth rate σ and the product $\pi^2\epsilon^2T^2N^2/L^2$ being positive, discrete, integers. So the e-folding time of any such disturbance is $\tau_m = T/m$, $m = 1, 2, \ldots$, where T is the correlation time of the turbulent fluid velocity. And the unstable modes exist when $\pi^2\epsilon^2T^2N^2 = L^2r$, $r = 1, 2, \ldots$. In making these remarks we are assuming that none of R^x_{00}, R^y_{00}, and R^z_{00} are zero.

In the event that R^z_{00} is zero, the modes are

$$\sigma = \tfrac{1}{2}\Gamma^2 - n$$

with no restriction that $\frac{1}{2}\Gamma^2$ must be an integer.

Consider now the antisymmetric modes.

V. LONG-WAVELENGTH ANTISYMMETRIC MODES

We again consider the long-wavelength situation $(\zeta^2 + \nu^2 + \xi^2)\eta T \ll 1$.

Write

$$f^i_A = b^i_1(\alpha,\beta)\zeta + b^i_2(\alpha,\beta)\nu + b^i_3(\alpha,\beta)\xi + O(k^3) + \cdots.$$

Then to lowest order Eqs. (40)–(42) give

$$\sigma b_1^x = \tfrac{1}{2}(\alpha^2+\beta^2)b_1^x - (\alpha\partial_\alpha + \beta\partial_\beta)b_1^x + \Gamma[\partial_\beta b_3^z + \partial_\alpha(b_3^z - b_1^x)], \quad (66)$$

$$\sigma b_2^x = \tfrac{1}{2}(\alpha^2+\beta^2)b_2^x - (\alpha\partial_\alpha + \beta\partial_\beta)b_2^x + \Gamma[\partial_\beta(b_2^x - b_3^x) - \partial_\alpha b_2^x], \quad (67)$$

$$\sigma b_3^x = \tfrac{1}{2}(\alpha^2+\beta^2)b_3^x - (\alpha\partial_\alpha + \beta\partial_\beta)b_3^x - \Gamma(\partial_\alpha + \partial_\beta)b_3^x, \quad (68)$$

$$\sigma b_1^y = \tfrac{1}{2}(\alpha^2+\beta^2)b_1^y - (\alpha\partial_\alpha + \beta\partial_\beta)b_1^y + \Gamma(\partial_\alpha - \partial_\beta)b_1^y, \quad (69)$$

$$\sigma b_2^y = \tfrac{1}{2}(\alpha^2+\beta^2)b_2^y - (\alpha\partial_\alpha + \beta\partial_\beta)b_2^y + \Gamma[\partial_\alpha(b_1^z - b_1^x) + \partial_\beta(b_3^z - b_2^y)], \quad (70)$$

$$\sigma b_3^y = \tfrac{1}{2}(\alpha^2+\beta^2)b_3^y - (\alpha\partial_\alpha + \beta\partial_\beta)b_3^y - \Gamma(\partial_\alpha + 2\partial_\beta)b_3^y, \quad (71)$$

$$\sigma b_1^z = \tfrac{1}{2}(\alpha^2+\beta^2)b_1^z - (\alpha\partial_\alpha + \beta\partial_\beta)b_1^z + \Gamma(2\partial_\alpha + \partial_\beta)b_1^z, \quad (72)$$

$$\sigma b_2^z = \tfrac{1}{2}(\alpha^2+\beta^2)b_2^z - (\alpha\partial_\alpha + \beta\partial_\beta)b_2^z + \Gamma(\partial_\alpha + 2\partial_\beta)b_2^z, \quad (73)$$

$$\sigma b_3^z = \tfrac{1}{2}(\alpha^2+\beta^2)b_3^z - (\alpha\partial_\alpha + \beta\partial_\beta)b_3^z + \Gamma[\partial_\alpha(b_3^z - b_1^x) + \partial_\beta(b_3^z - b_2^y)], \quad (74)$$

where $\partial_\mu \equiv \partial/\partial_\mu$.

Note that Eqs. (68), (69), and (71)–(73) are homogeneous and independent of each other and independent of the remaining four equations (66), (67), (70), and (74) in the sense that no knowledge of the latter is required in order to obtain the general solution to the former. But, of course, the converse is not true.

It is also clear that if we can obtain the general solution to Eq. (71) as a function of α, β, and Γ then we can write down by inspection the general solutions to Eqs. (72) and (73). For if we denote the general solution to Eq. (71) by $F(\alpha, \beta, \Gamma)$, then to within a multiplicative constant $b_1^z = F(\beta, \alpha, -\Gamma)$ and $b_2^z = F(\beta, \alpha, \Gamma)$.

Likewise if we denote the general solution to Eq. (68) by $G(\alpha, \beta, \Gamma)$ then to within a multiplicative constant $b_1^y = G(-\alpha, \beta, \Gamma)$. Consider Eqs. (68) and (71). The general solution to Eq. (68) is

$$b_3^x = T_3^x(\beta+\Gamma)^{(\Gamma^2/2-c)}(\alpha+\Gamma)^{(c-\sigma+\Gamma^2/2)} \times \exp\{\tfrac{1}{4}[(\alpha+\Gamma)^2 + (\beta+\Gamma)^2] - \Gamma(\alpha+\beta+2\Gamma)\}, \quad (75)$$

where c is the separation constant. So

$$b_1^y = T_1^y(\beta+\Gamma)^{(\Gamma^2/2-c^1)}(\alpha+\Gamma)^{(c^1-\sigma+\Gamma^2/2)} \times \exp\{\tfrac{1}{4}[(\Gamma-\alpha)^2 + (\beta+\Gamma)^2] - \Gamma(\beta-\alpha+2\Gamma)\}, \quad (76)$$

where, in general, c^1 is not the same as c.

The general solution to Eq. (71) is

$$b_3^y = T_3^y(\beta+2\Gamma)^{(2\Gamma^2-q)}(\Gamma+\alpha)^{(q-\sigma+\Gamma^2/2)} \times \exp\{\tfrac{1}{4}[(\alpha+\Gamma)^2 + (\beta+2\Gamma)^2] - \Gamma(2\beta+\alpha+5\Gamma)\}. \quad (77)$$

So

$$b_1^z = T_1^z(\alpha-2\Gamma)^{(2\Gamma^2-q')}(\beta-\Gamma)^{(q'-\sigma+\Gamma^2/2)} \times \exp\{\tfrac{1}{4}[(\beta-\Gamma)^2 + (\alpha-2\Gamma)^2] + \Gamma(2\alpha+\beta-5\Gamma)\}, \quad (78)$$

$$b_2^z = T_2^z(\alpha+2\Gamma)^{(2\Gamma^2-q'')}(\beta+\Gamma)^{(q''-\sigma+\Gamma^2/2)} \times \exp\{\tfrac{1}{4}[(\alpha+2\Gamma)^2 + (\beta+\Gamma)^2] - \Gamma(2\alpha+\beta+5\Gamma)\}, \quad (79)$$

and, in general, $q \neq q' \neq q''$.

From Eqs. (35) and (39) we obtain

$$b_1^i(\alpha, \beta, \mathbf{k}=0) = \exp[\tfrac{1}{4}(\alpha^2+\beta^2)] \times \sum_{n,m=0}^{\infty} \alpha^n\beta^m \frac{\partial R_{nm}^i(\zeta, 0, 0)}{\partial\zeta}\Big|_{\zeta=0}, \quad (80a)$$

$$b_2^i(\alpha, \beta, \mathbf{k}=0) = \exp[\tfrac{1}{4}(\alpha^2+\beta^2)] \times \sum_{n,m=0}^{\infty} \alpha^n\beta^m \frac{\partial R_{nm}^i(0, \nu, 0)}{\partial\nu}\Big|_{\nu=0}, \quad (80b)$$

$$b_3^i(\alpha, \beta, \mathbf{k}=0) = \exp[\tfrac{1}{4}(\alpha^2+\beta^2)] \times \sum_{n,m=0}^{\infty} \alpha^n\beta^m \frac{\partial R_{nm}^i(0, 0, \xi)}{\partial\xi}\Big|_{\xi=0}, \quad (80c)$$

So that

$$b_1^i(0,0,0) = \frac{\partial R_{00}^i}{\partial\zeta}\Big|_{\mathbf{k}=0}, \quad b_2^i(0,0,0) = \frac{\partial R_{00}^i}{\partial\nu}\Big|_{\mathbf{k}=0}, \quad b_3^i(0,0,0) = \frac{\partial R_{00}^i}{\partial\xi}\Big|_{\mathbf{k}=0}.$$

Proceeding as for the symmetric modes, we set $\alpha = e^{i\theta}$, $\beta = e^{i\varphi}$, and integrate Eqs. (75) and (80) over $0 \leq \theta, \varphi \leq 2\pi$ to obtain dispersion relations similar to those obtained for the symmetric modes. One set of solutions gives

$$\tfrac{1}{2}\Gamma^2 - c' = i, \quad \tfrac{1}{2}\Gamma^2 - \sigma + c' = j, \quad (81a)$$

$$2\Gamma^2 - q = I, \quad \tfrac{1}{2}\Gamma^2 - \sigma + q = J, \quad (81b)$$

$$2\Gamma^2 - q' = K, \quad \tfrac{1}{2}\Gamma^2 - \sigma + q' = L, \quad (81c)$$

$$2\Gamma^2 - q'' = M, \quad \tfrac{1}{2}\Gamma^2 - \sigma + q'' = U, \quad (81d)$$

where i, j, I, J, K, L, M, and U are positive integers including zero. It is clear that an unstable solution exists if $\frac{1}{2}\Gamma^2$ is itself a positive integer. Or if there is only one nonzero derivative of R_{00}^i then only one pair of Eqs. (81) needs to be satisfied. For example, consider that Eq. (81c) survives. Then

$$q' = 2\Gamma^2 - K, \quad \sigma = \tfrac{5}{2}\Gamma^2 - (K+L). \quad (82)$$

One set of solutions is $K = 0 = L$ giving $\sigma = \frac{5}{2}\Gamma^2$, so that in this case Γ^2 can be a continuous variable and $\sigma > 0$.

Alternatively we could consider the remaining four equations (66), (67), (70), and (74). They become an independent set if we choose $b_3^x = b_1^y = b_3^y = b_1^z = b_2^z = 0$.

And once again we obtain dispersion relations giving rise to unstable solutions, i.e., regenerative kinematic dynamo action. We see no point in including these

modes here. The mathematical analysis is similar to that already given in this section and the dispersion relations are encompassed in one or more of Eqs. (81).

VI. CONCLUSION AND DISCUSSION

In this paper we have demonstrated for the first time that turbulent velocity fields confined to a 'box' give rise to kinematic dynamo action using an *exact* statistical set of kinematic dynamo equations. The analysis carried through to the end of Sec. II is *exact* and we also outline there how to obtain the *exact* infinite determinant which describes the normal modes (and their eigenvalues σ) pertaining to the ensemble average magnetic field.

In view of the complexity of the general equations, we analyzed them only partially using a limit-theorem approach previously employed by Kac.[10] We found that the basic response of the system was either symmetric or antisymmetric in the "wave-number" (ζ, ν, ξ). [Parenthetically we point out that this is a property of the exact equations involving the $R^i(n, m, l)$ given in Sec. II].

For long wavelength disturbances both the symmetric and antisymmetric modes give rise to kinematic dynamo action with an e-folding time

$$\tau = O(L^2/N^2\pi^2\epsilon^2).$$

These limiting cases are sufficient to establish the basic point that centrally symmetric turbulent fluid motions confined to a 'box' give rise, on their own, to kinematic dynamo action.

This result, together with earlier work demonstrating kinematic dynamo action in infinite media with helical turbulence[12] and isotropic turbulence[3,4] suggests that it is unlikely that any particular system will possess a distribution of velocity turbulence which does *not* give rise to kinematic dynamo action, be it a finite or an infinite medium.

ACKNOWLEDGMENTS

The author is grateful to Professor E. N. Parker for much valuable criticism and encouragement. The work reported here was supported by the United States Air Force, Air Force Systems Command under contract F-19628-72-C-0056.

1 E. N. Parker, Astrophys. J. **122**, 293 (1955).
2 S. I. Braginskii, Zh. Eksp. Teor. Fiz. **20**, 726 (1964) [Sov. Phys. JETP **47**, 1084 (1964)].
3 I. Lerche, Astrophys. J. **166**, 627 (1971).
4 I. Lerche and B.-C. Low, Astrophys. J., **168**, 503 (1971).
5 I. Lerche and E. N. Parker, Astrophys. J., **168**, 231 (1971).
6 I. Lerche, J. Math. Phys. **12**, 1538 (1971).
7 I. Lerche, Astrophys. J. **166**, 639 (1971).
8 I. Lerche, J. Math. Phys. **12**, 1858 (1971).
9 G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. **36**, 823 (1930).
10 M. Kac, in *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover, New York, 1954).
11 M. Smoluchowski, Phys. Z. **17**, 557 (1916).
12 M. Steenbeck, F. Krause, and K. H. Radler, Z. Naturforsch. **21**, 369 (1966).

General Structure of the Distribution Functions for the Heisenberg Model and the Ising Model

Tohru Morita*
Department of Physics, Ohio University, Athens, Ohio 45701
(Received 9 June 1971)

The general structure for the distribution functions (reduced density matrices) for systems composed of a number of elements is given by taking the variation with respect to the distribution functions in the formalism of the cluster variation method. The parameters or the Lagrange multipliers occurring in the distribution functions must be determined by the reducibility condition of the distribution functions or by the stationariness condition of the free energy.

1. INTRODUCTION

The cluster variation method for the cooperative phenomena has been formulated on the basic variational principle for the free energy.[1-3] Some years ago, the present author and Tanaka[4] showed that the constant coupling approximation in the form given by Callen and Callen,[5] where a parameter is determined by a stationariness condition, is obtained on this more concrete basis. At the same time, they derived also the constant coupling approximation in the original form proposed by Kasteleijn and Kranendonk,[6] where the parameter is determined by the consistency relation between the distribution functions for one spin and a pair of spins, on the same basis.

In addition to Kasteleijn and Kranendonk, various authors[7-9] have introduced approximation methods where the form of distribution functions is assumed and the parameters appearing in them are determined by some kinds of consistency conditions. The purpose of the present paper is to give the general structures for the distribution functions where the parameters are connected by the consistency relations. This is done on the basis of the variational principle for the free energy.

The methods using series expansions and their extrapolations are developed to an extent that one can obtain the exact knowledge near the critical point.[10] That became possible by calculating higher expansion coefficients which are related to very large clusters. The cluster variation method is also expected to give detailed knowledge about the phase transition when such large clusters are considered and the limiting behaviors can be guessed as the larger clusters are considered. The general knowledge about the structure of the distribution function is considered to provide a concrete method of extending the constant coupling approximation to larger clusters.

modes here. The mathematical analysis is similar to that already given in this section and the dispersion relations are encompassed in one or more of Eqs. (81).

VI. CONCLUSION AND DISCUSSION

In this paper we have demonstrated for the first time that turbulent velocity fields confined to a 'box' give rise to kinematic dynamo action using an *exact* statistical set of kinematic dynamo equations. The analysis carried through to the end of Sec. II is *exact* and we also outline there how to obtain the *exact* infinite determinant which describes the normal modes (and their eigenvalues σ) pertaining to the ensemble average magnetic field.

In view of the complexity of the general equations, we analyzed them only partially using a limit-theorem approach previously employed by Kac.[10] We found that the basic response of the system was either symmetric or antisymmetric in the "wave-number" (ζ, ν, ξ). [Parenthetically we point out that this is a property of the exact equations involving the $R^i(n, m, l)$ given in Sec. II].

For long wavelength disturbances both the symmetric and antisymmetric modes give rise to kinematic dynamo action with an e-folding time

$$\tau = O(L^2/N^2\pi^2\epsilon^2).$$

These limiting cases are sufficient to establish the basic point that centrally symmetric turbulent fluid motions confined to a 'box' give rise, on their own, to kinematic dynamo action.

This result, together with earlier work demonstrating kinematic dynamo action in infinite media with helical turbulence[12] and isotropic turbulence[3,4] suggests that it is unlikely that any particular system will possess a distribution of velocity turbulence which does *not* give rise to kinematic dynamo action, be it a finite or an infinite medium.

ACKNOWLEDGMENTS

The author is grateful to Professor E. N. Parker for much valuable criticism and encouragement. The work reported here was supported by the United States Air Force, Air Force Systems Command under contract F-19628-72-C-0056.

1 E. N. Parker, Astrophys. J. **122**, 293 (1955).
2 S. I. Braginskii, Zh. Eksp. Teor. Fiz. **20**, 726 (1964) [Sov. Phys. JETP **47**, 1084 (1964)].
3 I. Lerche, Astrophys. J. **166**, 627 (1971).
4 I. Lerche and B.-C. Low, Astrophys. J., **168**, 503 (1971).
5 I. Lerche and E. N. Parker, Astrophys. J., **168**, 231 (1971).
6 I. Lerche, J. Math. Phys. **12**, 1538 (1971).
7 I. Lerche, Astrophys. J. **166**, 639 (1971).
8 I. Lerche, J. Math. Phys. **12**, 1858 (1971).
9 G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. **36**, 823 (1930).
10 M. Kac, in *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover, New York, 1954).
11 M. Smoluchowski, Phys. Z. **17**, 557 (1916).
12 M. Steenbeck, F. Krause, and K. H. Radler, Z. Naturforsch. **21**, 369 (1966).

General Structure of the Distribution Functions for the Heisenberg Model and the Ising Model

Tohru Morita*
Department of Physics, Ohio University, Athens, Ohio 45701
(Received 9 June 1971)

The general structure for the distribution functions (reduced density matrices) for systems composed of a number of elements is given by taking the variation with respect to the distribution functions in the formalism of the cluster variation method. The parameters or the Lagrange multipliers occurring in the distribution functions must be determined by the reducibility condition of the distribution functions or by the stationariness condition of the free energy.

1. INTRODUCTION

The cluster variation method for the cooperative phenomena has been formulated on the basic variational principle for the free energy.[1-3] Some years ago, the present author and Tanaka[4] showed that the constant coupling approximation in the form given by Callen and Callen,[5] where a parameter is determined by a stationariness condition, is obtained on this more concrete basis. At the same time, they derived also the constant coupling approximation in the original form proposed by Kasteleijn and Kranendonk,[6] where the parameter is determined by the consistency relation between the distribution functions for one spin and a pair of spins, on the same basis.

In addition to Kasteleijn and Kranendonk, various authors[7-9] have introduced approximation methods where the form of distribution functions is assumed and the parameters appearing in them are determined by some kinds of consistency conditions. The purpose of the present paper is to give the general structures for the distribution functions where the parameters are connected by the consistency relations. This is done on the basis of the variational principle for the free energy.

The methods using series expansions and their extrapolations are developed to an extent that one can obtain the exact knowledge near the critical point.[10] That became possible by calculating higher expansion coefficients which are related to very large clusters. The cluster variation method is also expected to give detailed knowledge about the phase transition when such large clusters are considered and the limiting behaviors can be guessed as the larger clusters are considered. The general knowledge about the structure of the distribution function is considered to provide a concrete method of extending the constant coupling approximation to larger clusters.

In the following two sections, the variation with respect to the distribution functions is taken in the formalism of the cluster variation method, taking account of the subsidiary conditions. The main conclusion of the present paper is summarized at the last paragraph of Sec. 3. In Secs. 4–6, simple examples of approximations based on the present formalism are given for the purpose of illustration.

The presentation is given for the Heisenberg model. It is noticed that the results are applicable for the Ising model or lattice gases by suitable reinterpretations of the notations.

2. GENERAL STRUCTURE OF THE REDUCED DENSITY MATRICES: GENERAL THEORY

In Sec. 2 of the preceding paper,[4] a short account is given of the formulation of the cluster variation method due to the present author.[3] That is summarized as follows, by using more abstract notations for the convenience of the discussion in the following section.

The free energy F of the system is calculated by the following variational principle:

$$F = \min \mathfrak{F}, \tag{2.1}$$

$$\mathfrak{F} = \mathrm{Tr}\rho_t[H + kT \ln\rho_t]. \tag{2.2}$$

The minimum is taken with respect to the trial density matrix ρ_t under the normalization condition that

$$\mathrm{Tr}\rho_t = 1. \tag{2.3}$$

The system is assumed to consist of L localized parts, which are called spins for convenience.

In this paper we shall use a symbol i to denote a cluster of spins $(i_1, \ldots, i_n)$ ($n = 1, 2, \ldots, L$ and $i_1, i_2, \ldots, i_n = 1, 2, \ldots, L$). The fact that spin $i_j (j = 1, 2, \ldots,$ or $n)$ is involved in the cluster i, is denoted by $i_j \in$ i. When all the spins $j_1, j_2, \ldots,$ and j_m involved in a cluster j is also involved in a cluster i, we call that j is a subcluster of i. If the number of spins involved in j is less than that of i, this fact is denoted by $\mathrm{j} \subset \mathrm{i}$. Notation $\mathrm{j} \subseteq \mathrm{i}$ is used when j is either a subcluster of i or equal to i. We treat many summations of the type

$$\begin{aligned} &A(i_1, i_2, \ldots, i_n) \\ &\quad = \sum_{j=1}^{n} a(i_j) + \sum_{n \geq j > k \geq 1} a(i_j, i_k) + \ldots + a(i_1, i_2, \ldots, i_n) \\ &\quad = \sum_{m=1}^{n} \sum_{n \geq j_1 > j_2 > \cdots > j_m \geq 1} a(i_{j_1}, i_{j_2}, \ldots, i_{j_m}). \end{aligned} \tag{2.4}$$

In terms of the cluster notation, we write this equation as follows:

$$A(\mathrm{i}) = \sum_{\substack{\mathrm{j} \\ (\mathrm{j} \subseteq \mathrm{i})}} a(\mathrm{j}). \tag{2.4'}$$

When the cluster i is the cluster of all the spins in the system, (2.4) becomes

$$\begin{aligned} A &= A(1, 2, \ldots, L) \\ &= \sum_{m=1}^{L} \sum_{L \geq i_1 > i_2 > \cdots > i_m \geq 1} a(i_1, i_2, \ldots, i_m), \end{aligned} \tag{2.5}$$

which is simply written as

$$A + \sum_{\mathrm{i}} a(\mathrm{i}), \tag{2.6}$$

where the summation is taken over all clusters of spins in the system.

In the present notation, the Hamiltonian of the system is expressed as

$$H = \sum_{\mathrm{i}} h(\mathrm{i}) \tag{2.6'}$$

For the Heisenberg model where the formulation developed in this and the following sections will be applied, the Hamiltonian is

$$H = -\sum_{j=1}^{L} hS_{jz} - \sum_{L \geq j > k \geq 1}^{L} J_{jk} \mathbf{S}_j \cdot \mathbf{S}_k, \tag{2.7}$$

so that

$$h(\mathrm{i}) = \begin{cases} -hS_{jz}, & \text{if } \mathrm{i} = (j), \\ -J_{jk}\mathbf{S}_j \cdot \mathbf{S}_k & \text{if } \mathrm{i} = (j, k), \\ 0, & \text{if i is a cluster of more than two spins.} \end{cases} \tag{2.7'}$$

The trial-reduced density matrix $\rho_t(\mathrm{i})$ for the cluster i is introduced by

$$\rho_t(\mathrm{i}) = \mathrm{tr}_{\{L\}-\mathrm{i}}\rho_t, \tag{2.8}$$

where the subscript $\{L\} - \mathrm{i}$ of tr denotes that the trace is taken over the states for all the L spins excluding the spins involved in the cluster i. From (2.8), one obtains the recurrence relation

$$\rho_t(\mathrm{i}) = \mathrm{tr}_k \rho_t(\mathrm{i}, k), \tag{2.9}$$

where the spin at site k is assumed not to be included in the cluster i. Then the function $\Gamma(\mathrm{i})$ and $\gamma(\mathrm{j})$ are introduced by

$$\Gamma(\mathrm{i}) = kT \ \mathrm{tr}_{\mathrm{i}} \ \rho_t(\mathrm{i}) \ln\rho_t(\mathrm{i}) \tag{2.10}$$

and

$$\Gamma(\mathrm{i}) = \sum_{\mathrm{j}(\mathrm{j} \subseteq \mathrm{i})} \gamma(\mathrm{j}). \tag{2.11}$$

Using this formula for the case when i is the cluster of all L spins, one obtains

$$\mathfrak{F} = \sum_{\mathrm{i}} \mathrm{tr}_{\mathrm{i}} \ h(\mathrm{i})\rho_t(\mathrm{i}) + \sum_{\mathrm{i}} \gamma(\mathrm{i}), \tag{2.12}$$

from (2.2), (2.6), and (2.8).

The original variational principle (2.1)–(2.3) is equivalent to the following one: (2.1) where $\mathfrak{F}$ is given by (2.12) and the variation is taken with respect to $\rho_t(\mathrm{i})$ with the subsidiary conditions (2.9).

In practice, we introduce an approximation where $\gamma(\mathrm{i})$ for some clusters are considered correctly in (2.12) and the others are ignored. If one introduces $\ln g_t(\mathrm{i})$ by

$$\ln\rho_t(\mathrm{i}) = \sum_{\mathrm{i}(\mathrm{i} \subseteq \mathrm{i})} \ln g_t(\mathrm{i}), \tag{2.13}$$

substitute (2.13) into (2.10) and compare with (2.11), one finds that

$$\gamma(\mathrm{i}) = \mathrm{tr}_{\mathrm{i}}\rho_t(\mathrm{i}) \ln g_t(\mathrm{i}). \tag{2.14}$$

Hence this approximation may be said as the approximation where $\rho_t(\mathfrak{j})$ for the ignored $\mathfrak{j}$ is chosen such that the corresponding $\ln g_t(\mathfrak{j})$ is zero, instead of executing the stationariness condition of making the $\mathfrak{F}$ stationary.

When $\gamma(\mathfrak{i})$ is ignored but $h(\mathfrak{i})$ is not zero, an approximation must be introduced to express the $\rho_t(\mathfrak{i})$ in terms of $\rho_t(\mathfrak{j})$ for $\mathfrak{j}$ for which $\gamma(\mathfrak{j})$ are retained correctly. A natural choice will be to use

$$\rho_t(\mathfrak{i}) = \exp[\ln\rho_t(\mathfrak{i}) \text{ given by (2.13)}] \tag{2.15}$$

and neglect $\ln g_t(\mathfrak{j})$ in the exponent for the ignored clusters. A method of replacing the effect $h(\mathfrak{i})$ for the ignored clusters by the effective potentials for $\mathfrak{j}$ which are considered correctly, is presented in the Appendix. In the text, we shall assume that $h(\mathfrak{i})$ is zero for the ignored clusters.

Following the terminology given in Refs. 2 and 3, the cluster $\mathfrak{j}$ for which $\gamma(\mathfrak{j})$ is considered correctly will be called a "preserved cluster". Now the approximation is characterized by the set of the preserved clusters. For simplicity we shall assume that a subcluster of a preserved cluster is also a preserved cluster. The set of preserved clusters will be written as U. In the following, $\mathfrak{j} \in U$ means that $\mathfrak{j}$ is a preserved cluster and $\mathfrak{j} \notin U$ that $\mathfrak{j}$ is not a preserved cluster or an ignored cluster.

Neglecting $\gamma(\mathfrak{j})$ for the ignored clusters, one obtains

$$\mathfrak{F} = \sum_{(\mathfrak{j}\in U)} \mathrm{tr}_{\mathfrak{j}}\rho_t(\mathfrak{j})h(\mathfrak{j}) + \sum_{(\mathfrak{j}\in U)} \gamma(\mathfrak{j}). \tag{2.16}$$

The set of equilibrium distribution functions $\rho(\mathfrak{i})$ must be determined as $\rho_t(\mathfrak{i})$, which make this expression (2.16) stationary under the subsidiary conditions (2.9). This is the stage of the formulation given in the preceding papers.

In the next place, we shall consider the problem of taking the variations of (2.16) with respect to $\rho_t(\mathfrak{i})$, introducing the Lagrange multipliers to secure the subsidiary conditions (2.9). Before doing that, the variational function (2.16) is modified a little bit. $\Gamma_1(\mathfrak{i})$ and $\gamma_1(\mathfrak{j})$ are introduced by

$$\Gamma_1(\mathfrak{i}) = kT\mathrm{tr}_{\mathfrak{j}}\rho_t(\mathfrak{i})\ln\rho(\mathfrak{i}), \tag{2.17}$$

$$\Gamma_1(\mathfrak{i}) = \sum_{\mathfrak{j}(\mathfrak{j}\subseteq\mathfrak{i})} \gamma_1(\mathfrak{j}) \tag{2.18}$$

On the other hand, $\ln g(\mathfrak{i})$ is introduced by

$$\ln\rho(\mathfrak{i}) = \sum_{\mathfrak{j}(\mathfrak{j}\subseteq\mathfrak{i})} \ln g(\mathfrak{j}) \tag{2.19}$$

and then one notices that

$$\gamma_1(\mathfrak{j}) = kT\mathrm{tr}_{\mathfrak{j}}\rho_t(\mathfrak{j})\ln g(\mathfrak{j}). \tag{2.20}$$

Adding a sum of $\mathrm{tr}_{\mathfrak{j}}\rho_t(\mathfrak{j})\ln g(\mathfrak{j})$ over $\mathfrak{j}$ to the right-hand side of (2.16) and then subtracting the corresponding sum of $\gamma_1(\mathfrak{j})$ defined by (2.17) and (2.18) from it, one obtains

$$= \sum_{(\mathfrak{j}\in U)} \mathrm{tr}_{\mathfrak{j}}\rho_t(\mathfrak{j})(h(\mathfrak{j}) + kT\ln g(\mathfrak{j})) + \sum_{(\mathfrak{j}\in U)} \gamma_2(\mathfrak{j}), \tag{2.21}$$

where $\gamma_2(\mathfrak{j}) \equiv \gamma(\mathfrak{j}) - \gamma_1(\mathfrak{j})$ is considered to be defined by

$$\Gamma_2(\mathfrak{i}) = \sum_{\mathfrak{j}(\mathfrak{j}\subseteq\mathfrak{i})} \gamma_2(\mathfrak{j}), \tag{2.22}$$

where

$$\Gamma_2(\mathfrak{i}) = kT\ \mathrm{tr}\ \rho_t(\mathfrak{i})(\ln\rho_t(\mathfrak{i}) - \ln\rho(\mathfrak{i}) - 1) + kT; \tag{2.23}$$

$-kT\ \mathrm{tr}_{\mathfrak{i}}\rho_t(\mathfrak{i}) + kT$ has been added for the convenience of later calculation. Equation (2.21) with (2.22) is equal to (2.16) if the subsidiary conditions (2.9) are satisfied, and hence is considered as the function to be minimized under the subsidiary conditions (2.9).

Let us introduce a complete set of matrices 1 and $M_\nu(\mathfrak{i})$, such that any matrix A in the space of the cluster $\mathfrak{i}$ is expanded as follows:

$$A = a_0 + \sum_\nu a_\nu M_\nu(\mathfrak{i}). \tag{2.24}$$

For example, in the case of the Heisenberg model, the set of $Y_l^m(\mathrm{S}_k)$ with $0 \leqslant l \leqslant 2S$ and $-l \leqslant m \leqslant l$ or its equivalent represents the set of 1 and $M_\nu(k)$, when $\mathfrak{i} = (k)$ and the set of

$\Pi_{j=1}^n Y_{l_j}^{m_j}(\mathrm{S}_{i_j})$ with $0 \leqslant l_j \leqslant 2S_{i_j}$ and $-l_j \leqslant m_j \leqslant l_j$

or its equivalent represents the set of 1 and $M_\nu(\mathfrak{i})$ when $\mathfrak{i} = (i_1, i_2, \ldots, i_n)$.

With the aid of this complete set of matrices, the subsidiary conditions (2.9) for the variation is replaced by

$$\mathrm{tr}_{\mathfrak{i}}M_\nu(\mathfrak{i})[\mathrm{tr}_k\rho_t(\mathfrak{i}, k) - \rho_t(\mathfrak{i})] = 0, \tag{2.25}$$

where $k \notin \mathfrak{i}$ and

$$\mathrm{tr}_{\mathfrak{i}}\rho_t(\mathfrak{i}) - 1 = 0. \tag{2.26}$$

Denoting the Lagrange multipliers for the subsidiary conditions (2.25) and (2.26) as

$$\lambda_\nu(k; \mathfrak{i}) \text{ and } f(\mathfrak{i}), \tag{2.27}$$

the variational function is written as

$$\begin{aligned}\mathfrak{F} = \{&\mathfrak{F} \text{ given by (2.21)}\} \\ &- \sum_{(\mathfrak{i}\in U, k\notin\mathfrak{i}, (\mathfrak{i},k)\in U)} \sum_{\mathfrak{i}} \sum_{k} \mathrm{tr}_{\mathfrak{i}}\lambda(k; \mathfrak{i})[\mathrm{tr}_k\rho_t(\mathfrak{i}, k) - \rho_t(\mathfrak{i})] \\ &- \sum_{(\mathfrak{i}\in U)} f(\mathfrak{i})[\mathrm{tr}_{\mathfrak{i}}\rho_t(\mathfrak{i}) - 1],\end{aligned} \tag{2.28}$$

where

$$\lambda(k; \mathfrak{i}) \equiv \sum_\nu \lambda_\nu(k; \mathfrak{i})M_\nu(\mathfrak{i}). \tag{2.29}$$

The Lagrange multipliers must be determined such that $\rho_t(\mathfrak{i})$, which minimize this expression, satisfy the subsidiary conditions (2.25) and (2.26). It is noticed that the differentiations of the expression (2.28) with respect to $\lambda_\nu(k; \mathfrak{i})$ and $f(\mathfrak{i})$ give the subsidiary conditions (2.25) and (2.26). This implies that one can determine these Lagrange multipliers by the condition that the $\mathfrak{F}$ is stationary with respect to the variations of them, instead of the subsidary conditions.

The $\mathfrak{F}$ given by (2.28) is written explicitly as

$$\begin{aligned}\mathfrak{F} = &\sum_{(\mathfrak{j}\in U)} \mathrm{tr}_{\mathfrak{j}}\rho_t(\mathfrak{j})[h(\mathfrak{j}) + \lambda(\mathfrak{j}) - f(\mathfrak{j}) + kT\ln g(\mathfrak{j})] \\ &+ \sum_{(\mathfrak{j}\in U)} \gamma_3(\mathfrak{j}),\end{aligned} \tag{2.30}$$

where

$$\lambda(\mathrm{j}) = \sum_{\substack{k \\ (k\in \mathrm{j})}} \gamma(k;\mathrm{j}) - \sum_{\substack{k \\ (k\in \mathrm{j})}} \lambda(k;\mathrm{j}-k) \tag{2.31}$$

and $\gamma_3(\mathrm{j}) = [\gamma_2(\mathrm{j})$ given by (2.22)$] + f(\mathrm{j})$ is considered to be defined by

$$\Gamma_3(\mathrm{i}) = \sum_{\substack{\mathrm{j} \\ (\mathrm{j}\subseteq \mathrm{i})}} \gamma_3(\mathrm{j}) \tag{2.32}$$

and

$$\Gamma_3(\mathrm{i}) = kT\,\mathrm{tr}_{\mathrm{i}}\rho_t(\mathrm{i})[\ln\rho_t(\mathrm{i}) - \ln\rho(\mathrm{i}) - 1] + kT + F(\mathrm{i}), \tag{2.33}$$

where

$$F(\mathrm{i}) = \sum_{\substack{\mathrm{j} \\ (\mathrm{j}\subseteq \mathrm{i})}} f(\mathrm{j}). \tag{2.34}$$

The variation

$$[\delta\mathfrak{F}/\delta\rho_t(\mathrm{j})]_{\rho_t(\mathrm{i})=\rho(\mathrm{i})} = 0 \tag{2.35}$$

gives us

$$\ln g(\mathrm{j}) = \beta[f(\mathrm{j}) - h(\mathrm{j}) - \lambda(\mathrm{j})]. \tag{2.36}$$

Substituting this into (2.19), one obtains

$$\rho(\mathrm{i}) = \exp\beta[F(\mathrm{i}) - H(\mathrm{i}) - \Lambda(\mathrm{i})], \tag{2.37}$$

where

$$H(\mathrm{i}) = \sum_{\substack{\mathrm{j} \\ (\mathrm{j}\subseteq \mathrm{i})}} h(\mathrm{j}), \tag{2.38}$$

$$\Lambda(\mathrm{i}) = \sum_{\substack{\mathrm{j} \\ (\mathrm{j}\subseteq \mathrm{i})}} \lambda(\mathrm{j}) = \sum_{\substack{k \\ (k\notin \mathrm{i},\, \mathrm{j}\subseteq \mathrm{i},\, (\mathrm{i},k)\in\bar{U})}} \sum_{\mathrm{i}} \lambda(k;\mathrm{j}). \tag{2.39}$$

Note that, when (2.31) is substituted in (2.39), $\lambda(k;\mathrm{j})$ are cancelled if $k \in \mathrm{i}$ and $\mathrm{j} \subseteq \mathrm{i}$. With the aid of the normalization conditions for $\rho(\mathrm{i})$ corresponding to (2.26), one obtains

$$\exp[-\beta F(\mathrm{i})] = \mathrm{tr}_{\mathrm{i}}\,\exp\{-\beta[H(\mathrm{i}) + \Lambda(\mathrm{i})]\} \tag{2.40}$$

and

$$F = \sum_{\mathrm{i}} f(\mathrm{j}), \tag{2.41}$$

where $f(\mathrm{j})$ are calculated from the values of $F(\mathrm{i})$ with the aid of (2.34).

The other Lagrange multipliers $\lambda_\nu(k;\mathrm{j})$ are determined by the subsidiary conditions corresponding to (2.25);

$$\mathrm{tr}_{\mathrm{i}}M_\nu(\mathrm{i})[\mathrm{tr}_k\rho(\mathrm{i},k) - \rho(\mathrm{i})] = 0 \tag{2.42}$$

or the variational principle

$$\delta[F \text{ given by (2.41), (2.40), and (2.34)}]/\delta\lambda_\nu(k;\mathrm{i}) = 0. \tag{2.43}$$

As a result, one obtains general expressions for the distribution functions and the free energy, where the parameters or the Lagrange multipliers appearing in the expressions must be determined either by the self-consistency relation or by the stationariness of the free energy.

3. GENERAL STRUCTURE OF THE REDUCED DENSITY MATRICES: APPROXIMATION METHODS

We have proceeded our calculation in the last section quite formally without considering on which $\rho_t(\mathrm{i})$ the variational function $\mathfrak{F}$ given by (2.16) does depend actually. The results obtained are quite general but it contains much more Lagrange multipliers than is actually necessary.

As examples, two rather trivial cases will be considered. First, the case when all the clusters are considered correctly. Then the entropy term in the variational function (2.16) should be the original one (2.2), which depends only on ρ_t. All the reduced density matrices $\rho_t(\mathrm{i})$ must cancel with each other in the variational function $\mathfrak{F}$; and the variational calculation gives simply

$$\rho = \exp[\beta(F - H)], \tag{3.1}$$

where F is a constant determined by the normalization condition (2.3) or $\mathrm{Tr}\rho = 1$. It is obvious that if $\rho(\mathrm{i})$ is calculated by reducing this, this set of ρ and $\rho(\mathrm{i})$ is the set which makes the $\mathfrak{F}$ minimum under the subsidiary conditions (2.9).

Next example is the approximation where two clusters M and N and all their subclusters are retained correctly. According to (2.11), sum of $\gamma(\mathrm{j})$ for the subset j of i is $\Gamma(\mathrm{i})$ defined by (2.10). Then the second term of $\mathfrak{F}$ given by (2.16) is

$$\sum_{\substack{\mathrm{j} \\ (\mathrm{j}\subseteq M \text{ or } N)}} \gamma(\mathrm{j}) = \Gamma(M) + \Gamma(N) - \Gamma(M\cap N). \tag{3.2}$$

In the sum of $\Gamma(M) + \Gamma(N)$, we have summed $\gamma(\mathrm{i})$ for i which is included in the common part $M \cap N$ and N twice, and so that contribution is substracted. Then the variational function includes only the distribution functions for the clusters, M, N, and $M\cap N$ if we assume that the first term is also expressed in terms of $\rho(M), \rho(N)$, and $\rho(M\cap N)$. When they are determined to make the $\mathfrak{F}$ stationary with the subsidiary conditions that those for M and N reduce to that for $M \cap N$ correctly, then the set of distribution functions obtained by reducing the distribution functions for M and N and $M\cap N$ is obviously the set which is determined so as to make the variational function $\mathfrak{F}$ given by (2.16) stationary under the subsidiary conditions (2.9). The result obtained in this way has less Lagrange multipliers than the set (2.37).

We have considered two trivial examples, where the result contains far less Lagrange multipliers compared with the general result in the last section. The simplicity is obtained by noticing which $\rho_t(\mathrm{i})$ actually appear in the expression of the entropy in the variational function, $\mathfrak{F}$ given by (2.16), in the approximation considered.

In general, when the approximation is specified by the set of preserved clusters M_i and their subclusters, the expression for the entropy is expressed as

$$\sum_{\substack{\mathrm{j} \\ (\mathrm{j}\subseteq M_i,\, i=1,2,\ldots)}} \gamma(\mathrm{j}) = \sum_i (M_i) - \sum_{i>j} \Gamma(M_i\cap M_j)$$

$$+ \sum_{i>j>k} \Gamma(M_i\cap M_j\cap M_k) - + \cdots \pm\Gamma(\cap M_i), \tag{3.3}$$

which is expressed in terms of the distribution functions for the M_i and their common parts; $\cap M_i$ denotes the common part of all the clusters M_i. Keeping this fact in mind, we proceed in a similar way as in the preceding section. Let us call the common parts of M_i, including themselves as m, n, etc., and define $\gamma^\dagger(n)$ for n as

$$\Gamma(m) = \sum_{(n\subseteq m)}^{\dagger} \gamma^\dagger(n), \tag{3.4}$$

where the dagger (†) on the summation sign means that the summation is taken over the clusters M_i and their common parts. Then the entropy term in F given by (2.16) is expressed as

$$\sum_{(j\subseteq M_i; i=1,2,\dots)} \gamma(j) = \sum_n{}^{\dagger}\gamma^\dagger(n), \tag{3.5}$$

where $\gamma^\dagger(n)$ is defined in terms of $\Gamma(m)$ by (3.4), and so expressed in terms of $\rho_t(n)$ for M_i and their common parts.

In the next place, $h^\dagger(n)$ is introduced by

$$H(m) = \sum_{(n\subseteq m)}^{\dagger} h^\dagger(n), \tag{3.6}$$

and then the first term in (2.16) is expressed as

$$\sum_{(j\subseteq M_i; i=1,2,\dots)} \mathrm{tr}_j \rho_t(j) h(j) = \sum_n{}^{\dagger}\rho_t(n) h^\dagger(n), \tag{3.7}$$

so that the variational function $\mathfrak{F}$ in our approximation is given by

$$\mathfrak{F} = \sum_m^{\dagger} [\mathrm{tr}_m \rho_t(m) h^\dagger(m) + \gamma^\dagger(m)]. \tag{3.8}$$

Introducing $\Gamma_1(m)$ and $\gamma_1^\dagger(n)$ by

$$\Gamma_1(m) = kT\ \mathrm{tr}_m \rho_t(m)\ \ln\rho(m), \tag{3.9}$$

$$\Gamma_1(m) = \sum_{(n\subseteq m)}^{\dagger} \gamma_1^\dagger(n), \tag{3.10}$$

and $\ln g^\dagger(m)$ by

$$\ln\rho(m) = \sum_{(n\subseteq m)} \ln g^\dagger(n), \tag{3.11}$$

so that

$$\gamma_1^\dagger(n) = kT\ \mathrm{tr}_n \rho_t(n)\ \ln g^\dagger(n). \tag{3.12}$$

We add $\sum_n^{\dagger}\gamma_1^\dagger(n)$ given by (3.12) to the right-hand side of (3.8) and then subtract the same quantity $\sum_n^{\dagger}\gamma_1^\dagger(n)$ which we express in terms of $\gamma_1(m)$ defined by (3.9). Then the variational function $\mathfrak{F}$ is expressed as

$$\mathfrak{F} = \sum_m{}^{\dagger}\mathrm{tr}_m \rho_t(m)[h(m) + kT\ \ln g^\dagger(m)] + \sum_m{}^{\dagger}\gamma_1^\dagger(m), \tag{3.13}$$

where $\gamma_2^\dagger(m)$ is defined by

$$\Gamma_2^\dagger(m) = kT\ \mathrm{tr}_m \rho_t(m)[\ln g_t^\dagger(m) - \ln g^\dagger(m) - 1] + kT \tag{3.14}$$

and

$$\Gamma_2^\dagger(m) = \sum_{(n\subseteq m)} \gamma_2{}^\dagger(n). \tag{3.15}$$

The Lagrange multipliers $\lambda_\nu(m - m'; m')$ are introduced to secure the reducibility of $\rho(m)$ to $\rho(m')$ when m' is one of such common parts of m and M_i that no common part of m and M_i includes m', that means no m'' included in U exits for which $m' \subset m'' \subset m$. The Lagrange multiplier to secure the normalization condition

$$\mathrm{tr}_m \rho(m) = 1, \tag{3.16}$$

is called $f^\dagger(m)$. Then the variational function $\mathfrak{F}$ is given by

$$\mathfrak{F} = \sum_m{}^{\dagger}\mathrm{tr}_m \rho_t(m)(h^\dagger(m) + \lambda^\dagger(m) + kT\ \ln g^\dagger(m) - f^\dagger(m)) + \sum_m{}^{\dagger}\gamma_3^\dagger(m). \tag{3.17}$$

Here $\gamma_3^\dagger(m)$ is given by

$$\Gamma_3^\dagger(m) = \{\Gamma_2^\dagger(m) \text{ given by (3.14)}\} + F(m), \tag{3.18}$$

$$\Gamma_3^\dagger(m) = \sum_{(n\subseteq m)}^{\dagger} \gamma_3^\dagger(n), \tag{3.19}$$

where $F(m)$ is coupled with $f^\dagger(m)$ by

$$F(m) = \sum_{(n\subseteq m)}^{\dagger} f^\dagger(n) \tag{3.20}$$

and $\lambda^\dagger(m)$ for a cluster m is

$$\lambda^\dagger(m) = -\sum_{\substack{(m'\subset m)\\(\text{no } m''\in U: m\subset m''\subset m)}}^{\dagger} \lambda^\dagger(m - m', m') + \sum_{\substack{(m\subset n)\\(\text{no } n'\in U: m\subset n'\subset n)}}^{\dagger} \lambda^\dagger(n - m, m). \tag{3.21}$$

It is noticed here that the differentiations of (3.17) with respect to the Lagrange multipliers give the subsidiary conditions which are to be secured by them. Hence they may be determined by the stationariness conditions, instead of the subsidiary conditions.

The variation with respect to $\rho_t(m)$ gives for $\ln g^\dagger(m)$

$$\ln g^\dagger(m) = \beta[f^\dagger(m) - h^\dagger(m) - \lambda^\dagger(m)], \tag{3.22}$$

so that

$$\rho(m) = \exp\{\beta[F(m) - H(m) - \Lambda^\dagger(m)]\}, \tag{3.23}$$

where

$$\Lambda^\dagger(m) = \sum_{\substack{(n\not\subset m\not\subset n)\\(\text{no } n'\in U: n\cap m\subset n'\subset n)}}^{\dagger} \lambda^\dagger(n - n\cap m; n\cap m). \tag{3.24}$$

Here the restriction for the summation is such that it is taken over all those n that there exists no common part of M_i, which is a subcluster of n and has $n\cap m$ as its subcluster. Note that in the effective Hamiltonian $H(m) + \Lambda^\dagger(m)$ for the cluster m, $\lambda^\dagger(n - n\cap m; n\cap m)$ may be regarded as an effect to the subcluster $n\cap m$ due to the cluster $n - n\cap m$ which is out-

side of m. $F(\mathsf{m})$ is determined by the normalization of $\rho(\mathsf{m})$, so that

$$\exp[-\beta F(\mathsf{m})] = \mathrm{tr}_{\mathsf{m}} \cdot \exp\{-\beta[H(\mathsf{m}) + \Lambda^{\dagger}(\mathsf{m})]\}. \tag{3.25}$$

Substituting these into (3.17), one obtains

$$F = \sum_{\mathsf{m}}^{\dagger} f^{\dagger}(\mathsf{m}), \tag{3.26}$$

where $f^{\dagger}(\mathsf{m})$ is calculated from the values of $F(\mathsf{m})$ calculated by (3.25), with the aid of (3.20).

As the result, in this formulation, the general expressions for the distribution functions for the cluster M_i and their common parts and that for the free energy have been obtained. The parameters or the Lagrange multipliers must be determined either by the self-consistency conditions or by the variational principle. If one determines the distribution function for those subclusters of M_i that are not common parts of M_i by reducing those for M_i or their common parts, then the set of the distribution functions makes the $\mathfrak{F}$ stationary under the subsidiary conditions (2.42). The same result is obtained by introducing less Lagrange multipliers in the present formalism, and this formalism is more convenient when some subclusters of M_i are not common parts of M_i.

The conclusion of this section is that the general expression for the distribution function for the clusters M_i and their common parts are given by (3.23) with (3.24) in terms of the parameters or the Lagrange multipliers $\lambda^{\dagger}$ and $F(\mathsf{m})$. $F(\mathsf{m})$ is determined by the normalization of $\rho(\mathsf{m})$; that gives (3.25). $\lambda^{\dagger}$ are determined by either of the following two procedures: (i) $\lambda^{\dagger}(\mathsf{n}-\mathsf{m},\mathsf{m})$ must be determined by the reducibility of $\rho(\mathsf{n})$ to $\rho(\mathsf{m})$. Once $\lambda^{\dagger}$ are determined, $F(\mathsf{m})$ is calculated by (3.25) and the free energy F is obtained via (3.26) and (3.20). (ii) we first calculate $F(\mathsf{m})$ in terms of $\lambda^{\dagger}$ by (3.25) and then F via (3.20) and (3.26). Now F is a function of $\lambda^{\dagger}$. $\lambda^{\dagger}$ are so determined as to make the obtained expression for F stationary. The obtained value of F is the free energy of our system.

4. TRIANGLE APPROXIMATION

The purpose of the present and the following sections is to illustrate the method of applying the formalism presented in the preceding sections to simple examples. The discussions in those sections will be restricted to the Heisenberg model where the exchange integral is nonzero only between nearest neighbor pairs of spins. In this section, we consider the approximation where clusters of three spins are preserved. The general results in Sec. 2 read as follows:

$$\rho(j) = \exp\beta[F(j) + hS_{jz} + \sum_{k} \lambda(k;j)], \tag{4.1}$$

$$\rho(j,k) = \exp\beta\{F(j,k) + h(S_{jz} + S_{kz}) + \sum_{\substack{l \\ (l=j,k)}} [\lambda(l;j) + \lambda(l;k) + \lambda(l;j,k)] + J_{jk}S_j\cdot S_k\}, \tag{4.2}$$

$$\rho(j,k,l) = \exp\beta\{F(j,k,l) + h(S_{jz} + S_{kz} + S_{lz}) + \sum_{\substack{m \\ (m\neq j,k,l)}} [\lambda(m;j) + \lambda(m;k) + \lambda(m;l) + \lambda(m;j,k) + \lambda(m;k,l) + \lambda(m;j,l)] + J_{jk}S_j\cdot S_k + J_{jl}S_j\cdot S_l + J_{kl}S_k\cdot S_l\} \tag{4.3}$$

For the case of spin $\frac{1}{2}$, the Lagrange multipliers take the form

$$\lambda(k;j) = \lambda_1(k;j)S_{jz}, \tag{4.4}$$

$$\lambda(l;j,k) = \lambda_1(l;j,k)S_{jz} + \lambda_2(l;j,k)S_{kz} + \lambda_3(l;j,k)S_{jz}S_{kz} + \lambda_4(l;j,k)S_j\cdot S_k \quad . \tag{4.5}$$

The coefficients are determined either by the self-consistency conditions that $\langle S_{jz}\rangle$, $\langle S_{jz}S_{kz}\rangle$, and $\langle S_j\cdot S_k\rangle$ are calculated consistently by these reduced density matrices or by the stationariness conditions that the free energy must be stationary with respect to the variations with respect to λ.

For the pair approximation, $\lambda(k;j)$ have been found to be zero for a pair of spins j and k which are farther apart than nearest neighbors,[11] and so that approximation has been found to be equivalent to the bond approximation where $f(j,k)$ are retained only for the nearest neighbor pairs of j and k. For our present case it is not likely to occur that some $\lambda(l,k)$ or $\lambda(l;j,k)$ are automatically zero. Hence the problem of taking account of all triplet correctly becomes a formidable one. However, we believe that the contribution $f(j,k,l)$ becomes less important when j,k and l are far apart with each other. In this situation, it will be sensible to consider either of the following two approximations:

(1) Triangle approximation where $f(j,k,l)$ are retained only when $h(\mathrm{i})$ is nonzero for all three pairs from $j,k,$ and l.

(2) An approximation where $f(j,k,l)$ are retained only when $h(\mathrm{i})$ is nonzero between two or three pairs from $j,k,$ and l.

The triangle approximation will be considered in this section. For the case of spin $\frac{1}{2}$, the reduced density matrices are given by

$$\rho(j) = \exp\beta[F^{(1)} + (h + z\lambda)S_{jz}], \tag{4.6}$$

$$\rho(j,k) = \exp\beta\{F^{(2)} + [h + (z-1)\lambda + \nu_3\lambda'](S_{jz} + S_{kz}) + \nu_3\lambda''S_{jz}S_{kz} + (J + \nu_3\lambda'')S_j\cdot S_k\}, \tag{4.7}$$

$$\rho(j,k,l) = \exp\beta\{F^{(3)} + [h + (z-2)\lambda + 2(\nu_3 - 1)\lambda'](S_{jz} + S_{kz} + S_{lz}) + (\nu_3 - 1)\lambda''[S_{jz}S_{kz} + S_{jz}S_{lz} + S_{kz}S_{lz}] + [J + (\nu_3 - 1)\lambda''](S_j\cdot S_k + S_j\cdot S_l + S_k\cdot S_l)\}, \tag{4.8}$$

where ν_3 is the number of triangles which consist of one fixed bond and two others in the lattice. $F^{(1)}$, $F^{(2)}$, and $F^{(3)}$ are calculated with the aid of the normalization conditions for the $\rho(i)$ from these expressions. The free energy is given by

$$\frac{F}{L} = F^{(1)} + \tfrac{1}{2}z(F^{(2)} - 2F^{(1)}) + \tfrac{1}{6}z\nu_3(F^{(3)} - 3F^{(2)} + 3F^{(1)}) \tag{4.9}$$

in terms of them. The Lagrange multipliers λ must be determined by the reducibility conditions or the variational principle to make this F stationary.

For the case when spin is more than $\frac{1}{2}$, it is practical to introduce an approximation to use (4.6)–(4.9), the self consistencies for the expectation values other than $S_{jz}, S_{jz}S_{kz}$, and $S_j \cdot S_k$ are neglected. This is similar to the constant coupling approximation compared to the pair or bond approximation.

Oguchi et al.[8] introduced an approximation where $\rho(j,k)$ for a pair and $\rho(j,k,l)$ for a triangle cluster are given with a set of parameters and determined the parameters by the reducibility conditions. Comparing with our formula, their expressions for the reduced density matrices have been given on a more intuitive standpoint.

5. SQUARE APPROXIMATION

For an Ising system of a square, simple cubic, and body-centered cubic lattices, Kikuchi[1] introduced the square approximation next to the bond approximation. In the square approximation, the reduced density matrices for four spins constitute a square where each edge connects a pair of nearest neighbors. We shall consider to apply this approximations to the Heisenberg model of ferromagnets.

The preserved clusters in the approximations are the square and their subclusters. The common parts of the squares are clusters of two spins at nearest neighbor sites and clusters of one spin. The reduced density matrices for them are read from the result of Sec. 3 as follows, when spin is $\frac{1}{2}$,

$$\rho(j) = \exp\beta[F^{(1)} + (h + z\lambda)S_{jz}], \tag{5.1}$$

$$\rho(j,k) = \exp\beta\{F^{(2)} + [h + (z-1)\lambda + \nu_4\lambda'](S_{jz} + S_{kz}) + \nu_4\lambda'' S_{jz}S_{kz} + (J + \nu_4\lambda''')S_j \cdot S_k\}, \tag{5.2}$$

$$\begin{aligned}\rho(i,j,k,l) = \exp\beta\{ & F^{(4)} + [h + (z-2)\lambda + 2(\nu_4 - 1)\lambda'] \\ & \times (S_{iz} + S_{jz} + S_{kz} + S_{lz}) + (\nu_4 - 1)\lambda'' \\ & \times (S_{iz}S_{jz} + S_{jz}S_{kz} + S_{kz}S_{lz} + S_{lz}S_{iz}) \\ & + [J + (\nu_4 - 1)\lambda'''] \\ & \times (S_i \cdot S_j + S_j \cdot S_k + S_k \cdot S_l + S_l \cdot S_i)\}.\end{aligned} \tag{5.3}$$

$F^{(1)}, F^{(2)}$, and $F^{(4)}$ are calculated by the normalization of these. The expression for the free energy is given by

$$\frac{F}{L} = F^{(1)} + \tfrac{1}{2} z[F^{(2)} - 2F^{(1)}] + \tfrac{1}{8} z\nu_4[F^{(4)} - 4F^{(2)} + 4F^{(1)}]. \tag{5.4}$$

Here ν_4 is the number of squares which have a given bond as an edge; $\nu_4 = 2, 4$, and 9 for the square, simple cubic, and body-centered cubic lattices, respectively. The Lagrange multipliers λ are determined either by the self consistency conditions that the averages of $S_{jz}, S_{jz}S_{kz}$, or $S_j \cdot S_k$ calculated by these are the same or by the stationary conditions that the F/L is stationary with respect to the variations λ.

We may introduce an approximation where the above formulae are used when spin is more than $\frac{1}{2}$.

6. ON THE P. R. WEISS APPROXIMATION

Here we shall consider an approximation, for the Heisenberg model, which resembles the P. R. Weiss approximation. The P. R. Weiss approximation was introduced for a lattice where the lattice can be divided into two sublattices, A and B, such that all the nearest neighbors to a lattice site, belonging to A, belong to the other sublattice B, and vice versa. The examples are the square, simple cubic, body-centered cubic lattices. In this approximation, the distribution function of $z + 1$ spins, which consists of a spin and its z nearest neighbors. The reduced density matrix is assumed to be

$$\rho^{(z+1)}(j,k_1,k_2,\ldots,k_z) = \exp\beta\Big(F^{(z+1)} + hS_{jz} + [h + (z-1)\lambda]\sum_{l=1}^{z} S_{k_l z} + \sum_{l=1}^{z} J S_j \cdot S_{kl}\Big). \tag{6.1}$$

λ is determined such that $\langle S_{jz}\rangle$ and $\langle S_{kz}\rangle$ are equal with each other.

Let us consider in our scheme an approximation where the preserved clusters are the clusters of $z + 1$ spins, which consists of a spin on the sublattice A and its z nearest neighbors and their subclusters. For this case, the common parts of the clusters of $z + 1$ spins are clusters of a pair of next neighbor spins on sublattice B and clusters of one spin on B. The reduced density matrices for them are given, according to the results of Sec. 3 as

$$\rho^{(1)}(j) = \exp\beta[F^{(1)} + (h + z\lambda + z'\lambda')S_{jz}], \tag{6.2}$$

$$\rho_2^{(2)}(j,k) = \exp\beta\{F_2{}^{(2)} + [h + z\lambda + (z' - 1)\lambda' + 2\lambda''](S_{jz} + S_{kz}) + 2\lambda'''S_{jz}S_{kz} + 2\lambda^{IV}S_j \cdot S_k\}, \tag{6.3}$$

$$\rho^{(z+1)}(j,k_1,k_2,\ldots,k_z) = \exp\beta\Big(F^{(z+1)} + hS_{jz} + [h + (z-1)\lambda + (z' - z'')\lambda' + \lambda'']\sum_{l=1}^{z} S_{k_l z} + J\sum_{l=1}^{z} S_j \cdot S_{k_l}\Big). \tag{6.4}$$

for the case of spin $\frac{1}{2}$.

For the case when the spin is larger than $\frac{1}{2}$, we have to introduce more Lagrange multipliers and self-consistency conditions. However, we consider the approximation to neglect them. Then we have the same expressions and the same self-consistency conditions as for the case of spin $\frac{1}{2}$. If one neglects further $f_2^{(2)}, \ln g_2^{(2)}$ so that λ'', λ''', and λ^{IV}, then one gets

$$\rho^{(1)}(j) = \exp\beta[F^{(1)} + (h + z\lambda)S_{jz}], \tag{6.5}$$

$$\rho^{(z+1)}(j,k_1,k_2,\ldots,k_z) = \exp\beta\Big(F^{(z+1)} + hS_{jz} + [h + (z-1)\lambda]\times\sum_{l=1}^{z} S_{k_l z} + J\sum_{l=1}^{z} S_j \cdot S_{k_l}\Big). \tag{6.6}$$

The structure of this $\rho^{(z+1)}$ is completely the same as the one in the P. R. Weiss approximation. However, the condition to determine λ is that $\langle S_{jz}\rangle$ calculated by $\rho^{(z+1)}$ and $\rho^{(1)}$ are equal to each other. Hence $\langle S_{kz}\rangle$ will be different from $\langle S_{jz}\rangle$. It is not easy to argue which gives the better result. In this approximation all the nearest neighbor pairs of spins are considered correctly, and so this is considered to be better than the constant coupling approximation.

7. VALIDITY OF THE PRESENT SCHEME

In the preceding sections, we discussed how the cluster variation method is applied when the distribution functions or the reduced density matrices are correctly considered for a certain type of clusters of lattice sites. In this section, a discussion is given about when the method is useful.

When the exact macroscopic properties cannot be calculated, we consider the problem of how properties of a small system can represent the properties of the macroscopic systems. The answer will be that a reasonable fit is expected if the length of correlation R_c— e.g., the distance at which the spin-pair correlation function $\sigma(R_{if}) = \langle S_i^z S_f^z \rangle - \langle S_i^z \rangle \langle S_f^z \rangle$, decays to zero—is fairly small compared with the dimension L of the small system; e.g., $R_c < \frac{1}{5} L$. In the cluster variation method, if we consider the clusters of dimension up to L, we include all the effects of the correlations involved within this length and hence we expect very good results if $R_c < L$, because the contribution from the larger clusters can, in this case, be ignored when the free energy is minimized. If we consider a system at high temperatures, the correlation length will be short and we can easily calculate the properties of a macroscopic system by considering the clusters of reasonable dimensions.

At low temperatures, the spin-wave excitations are good excitation modes for the Heisenberg magnet,[12] and the correlation length is expected to be very large. But if the temperature is not so low, the wavelength of the spin waves which plays an important role will not be very large. In that case, the spin-pair correlation function $\sigma(R_{if})$ will decay to a small value when the distance R_{if} becomes of the order of the wavelength. Thus we may expect reasonable results by including the clusters up to the dimension of that wavelength. The other case when the correlation length becomes large is the critical region.

For the high temperature expansions, the extrapolation methods have been shown to be very powerful up to the critical point.[10] For the present scheme of approximations, we hope that extrapolation of the results obtained with the aid of successively larger clusters will again lead to reasonable results for the low temperature as well as the critical region. Application of the present method to the square Ising model is under consideration. It is hoped that the exact results of Onsager[13] and Yang[14] will be reproduced in a good approximation.

ACKNOWLEDGMENT

This work was mostly developed when the author stayed at the Catholic University of America (1962–1966). The author is grateful to Professor Tomoyasu Tanaka for valuable discussions which stimulated this work.

APPENDIX: EFFECT OF POTENTIALS FOR THE IGNORED CLUSTERS AND THE WEISS MOLECULAR FIELD APPROXIMATION

When the cluster i is not a preserved cluster and $h(\mathrm{i})$ is not zero, $\rho_t(\mathrm{i})$ must be approximately expressed in terms of $\rho_t(\mathrm{j})$ for the preserved clusters j. A method of doing this has been suggested in Sec. 2. Then

$$\mathrm{tr}_{\mathrm{i}} h(\mathrm{i})\rho_t(\mathrm{i}) \tag{A1}$$

in Eq. (2.9) is replaced by

$$\mathrm{Tr}\, h(\mathrm{i})\rho_t(\mathrm{i};\{\rho_t(\mathrm{j})\}), \tag{A2}$$

where Tr means the trace over i and also over spins in j which are not equal to any of i, if any. In the variational calculations, we need the first derivative in addition to the value itself for $\rho_t(\mathrm{j}) = \rho(\mathrm{j})$. Hence instead of (A2), one may use

$$\mathrm{Tr}\, h(\mathrm{i})\rho(\mathrm{i};\{\rho(\mathrm{j})\}) + \sum_{\mathrm{j}} \mathrm{Tr}\, h(\mathrm{i})[\delta\rho(\mathrm{i};\{\rho(\mathrm{j})\})/\delta\rho(\mathrm{j})][\rho_t(\mathrm{j}) - \rho(\mathrm{j})]. \tag{A3}$$

It is noticed that (A2) and (A3) have the same value and the same first derivatives when $\rho_t(\mathrm{j}) = \rho(\mathrm{j})$.

If one uses (A3) instead of (A1) in Eq. (2.9), the coefficient of $\rho_t(\mathrm{j})$ plays the role of the effective potential for the cluster j and hence added to $h(\mathrm{j})$. Then the argument in the later part of Sec. 2 is valid, where $h(\mathrm{j})$ must be replaced by the one which includes the effective potentials due to the potentials for the ignored clusters. The terms which do not depend on $\rho_t(\mathrm{j})$ in (A3) should be added to the final expression for the free energy.

Let us consider the well-known example of the Weiss molecular field approximation. It is well known that it is obtained by retaining only $\Gamma(j)$ for clusters composed of one spin in the second line in Eq. (2.9), approximating $\rho_t(j,k)$ by $\rho_t(j)\rho_t(k)$ in the first line and taking $\rho_t(j)$ to be diagonal in the representation in which S_{jz} is diagonal. For this case, the expression corresponding to (A3) is

$$\mathrm{tr}_{j,k} h(j,k)\rho(j)\rho(k) + \mathrm{tr}_j[\mathrm{tr}_k h(j,k)\rho(k)][\rho_t(j) - \rho(j)] + \mathrm{tr}_k[\mathrm{tr}_j h(j,k)\rho(j)][\rho_t(k) - \rho(k)], \tag{A4}$$

where $h(j,k) = -J_{jk}\mathbf{S}_j \cdot \mathbf{S}_k$. Using the assumption that $\rho_t(j)$ and hence $\rho(j)$ is diagonal with S_{jz}, this reduces to

$$J_{jk}\langle S_{jz}\rangle\langle S_{kz}\rangle - \mathrm{tr}_j J_{jk}\langle S_{kz}\rangle S_{jz}\rho_t(j) - \mathrm{tr}_k J_{jk}\langle S_{jz}\rangle S_{kz}\rho_t(k). \tag{A5}$$

In the approximation where only $f(j)$ are retained in Eq. (2.37), including the second and third terms in (A5) in the effective potential in $h(j)$ and $H(k)$, and (2.34) and (2.37) are read as

$$\rho(j) = \exp\beta\left[F(j) + \left(h + \sum_k J_{jk}\langle S_{kz}\rangle\right)S_{jz}\right], \tag{A6}$$

$$F = \sum_{j>k} J_{jk}\langle S_{jz}\rangle\langle S_{kz}\rangle + \sum_j F(j). \tag{A7}$$

Here $F(j)$ is determined by the normalization of (A6) and $\langle S_{kz}\rangle$ is determined by calculating the average of S_{jz} using the distribution function (A6) and solving the resulting equation for $\langle S_{jz}\rangle = \langle S_{kz}\rangle$ for the ferromagnet. These are exactly the expressions in the Weiss molecular field approximation.

* On leave of absence from Department of Applied Science, Tohoku University, Sendai, Japan.

1 R. Kikuchi, Phys. Rev. **81**, 988 (1951).
2 T. Morita, J. Phys. Soc. Japan **12**, 753 (1957).
3 T. Morita, J. Phys. Soc. Japan **12**, 1060 (1957).
4 T. Morita and T. Tanaka, Phys. Rev. **145**, 288 (1966).
5 H. B. Callen and E. R. Callen, Phys. Rev. **136**, A1675 (1964).
6 P. W. Kasteleijn and J. Van Kranendonk, Physica **22**, 317 (1956).
7 T. Oguchi, Progr. Theoret. Phys. (Kyoto) **13**, 148 (1955).
8 T. Fujishiro, F. Takano, and T. Oguchi, J. Phys. Soc. Japan **19**, 1666 (1964).
9 P. R. Weiss, Phys. Rev. **74**, 105 (1948).
10 E.g., C. Domb, Advan. Phys. **9**, 245, 362 (1960); G. A. Baker, Jr., Phys. Rev. **136**, A1376 (1964); G. A. Baker, Jr., H. E. Gilbert, J. Eve, and G. S. Rushbrooke, Phys. Rev. **164**, 800 (1967).
11 *Note added in proof:* In the pair approximation, the reduced density matrix is given by (2.37) for each pair of sites, $i = (j, l)$. If $h(j, l) = 0$, the exponent of (2.37) for the pair (j, l) is a sum of operators involving only site j or only site l. For this case, it follows from reducibility conditions (2.9) and (2.26) that $\rho(j, l)$ must be equal to $\rho(j)\rho(l)$.
12 F. J. Dyson, Phys. Rev. **102**, 1217, 1230 (1956).
13 L. Onsager, Phys. Rev. **65**, 117 (1944).
14 C. N. Yang, Phys. Rev. **85**, 808 (1952).