Regular perturbations, Brillouin–Wigner expansion, and continued fractions

S. Graffi

Istituto Nazionale di Fisica Nucleare, Sezione di Bologna, Bologna, Italy
(Received 24 September 1973)

The least eigenvalue of an operator $H = H_0 + gV$ is considered, where H_0 is a semibounded self-adjoint operator in a Hilbert space and V is symmetric. It is shown that the J-type continued fraction (i.e., the sequence of the $[N-1, N]$ Padé approximants) to the Brillouin–Wigner perturbation expansion converges to the eigenvalue, provided V is a regular perturbation of H_0. An application of this result to some quantum mechanical systems, such as the helium atom, is briefly discussed.

1. INTRODUCTION

In quantum theory one is very often faced with the problem of computing the eigenvalues of a self-adjoint operator H in a Hilbert space X written under the form $H = H_0 + gV$, where g is a scalar parameter. Usually an exact solution for H_0 is known, and V is treated as a perturbation.

It is well known from the Rellich-Kato theory that under some appropriate conditions on H_0 and V the perturbation theory is regular, i.e., the usual (Rayleigh-Schrödinger) perturbation expansion in powers of g of the eigenvalues and eigenprojections has a nonzero radius of convergence.[1] It will be shown in this paper that the same conditions imply also the convergence of the J-type continued fraction (i.e., of the $[N-1, N]$ Padé approximants sequence) corresponding to the Brillouin-Wigner implicit perturbation expansion of the least eigenvalue $E(g)$ of H.

The utility of this result is twofold: First, the convergence takes place for any real g belonging to the maximal analyticity interval I of $E(g)$, i.e., g may lie outside the radius of convergence of the Rayleigh-Schrödinger expansion,[2] secondly, whereas the rate of convergence of the Rayleigh-Schrödinger expansion may be very slow, the Padé approximants yield a good approximation method, due to their well-known monotonicity and bounding properties. The paper is organized as follows: In the next Sec. we state our hypotheses and establish the validity of the Brillouin-Wigner implicit formula for the case under discussion; in Sec. 3 we prove the convergence statement and point out briefly an application to the He atom and He-like ions.

For an explicit expression of the approximants to $E(g)$, as well as the relationship with the Rayleigh-Schrödinger expansion, the reader is referred to a recent paper of McClary,[3] where the same result is proven for a $\Phi^4(x)$ quantum field theory model in two-dimensional space-time.

2. BRILLOUIN-WIGNER PERTURBATION THEORY

Let X be a Hilbert space, and H_0 a positive self-adjoint operator in X. Let the eigenvalue $E_0 > 0$, isolated and nondegenerate, with normalized eigenvector Φ, be the least point of $\sum(H_0)$. Let V be symmetric and H_0-bounded in X, i.e., let $D(V) \supset D(H_0) \equiv D$, and let there exist positive constants a and b such that

$$\|Vu\| \leq a\|u\| + b\|H_0 u\|, \qquad u \in D \tag{2.1}$$

These are the well-known conditions under which

$$H \equiv H(g) = H_0 + gV \tag{2.2}$$

is a self-adjoint holomorphic family of operators of type (A) and also of type (B_0), defined for any complex g.[1]

If we indicate with $E(g)$ the least point of $\sum(H)$, the above conditions imply the existence of a maximal interval $I \subset R$, containing the origin, such that for any $g \in I$ $E(g)$ is an analytic function representing an isolated nondegenerate eigenvalue of $H = H(g)$.[4] If $P(g)$ is the corresponding eigenprojection, it will be analytic in the same interval, $P(g)X$ being one-dimensional there.

We proceed now to obtain the Brillouin-Wigner implicit formula for $E(g)$.

Let $X^\perp$ be the Hilbert space $X \ominus \Phi$. If T is an operator in X, and $P^\perp$ the orthogonal projection onto X, let $T^\perp$ the operator in $X^\perp$ defined by $T^\perp = P^\perp T P^\perp$.

If $\psi \equiv \psi(g)$ is the normalized eigenvector corresponding to $E(g)$,

$$(H_0 + gV)\Psi(g) = E(g)\Psi(g), \tag{2.3}$$

write

$$\Psi(g) = \Phi + \Psi^\perp(g),$$
$$\Psi^\perp(g) = \Psi^\perp \in X^\perp, \qquad (\Phi, \Psi^\perp) = 0$$

Applying $P^\perp$ to (2.3) and then taking in the same formula the scalar product with Φ, we get the two equations

$$[(H_0 + gV)^\perp - E(g)]\,\Psi^\perp + gP^\perp V\Phi = 0, \tag{2.4}$$

$$E(g) = E_0 + g(\Phi, V\Phi) + g(P^\perp V\Phi, \Psi^\perp), \tag{2.5}$$

where $H_0\Phi = E_0\Phi$ has been used. By elimination of $\Psi^\perp$ between (2.4) and (2.5), we get the Brillouin-Wigner implicit equation for $E(g)$:

$$E(g) - E_0 - g(\Phi, V\Phi) = f(E, g), \tag{2.6}$$

where

$$f(E, g) = -g^2(P^\perp V\Phi, [(H_0 + gV)^\perp - E]^{-1} P^\perp V\Phi). \tag{2.7}$$

By Lemma 2.1 below, for any $g \in I$, $f(E, g)$ exists in the interval $-\infty < E < B(g)$, with $B(g) > E(g)$, and since it is clearly a monotonically decreasing function of E, the eigenvalue $E(g)$ must be the unique solution of (2.6) in the above defined interval, for any $g \in I$.

By Theorem 2.1, we can write $f(E, g)$ under the form

$$f(E, g) = -g^2(K^{-1/2} P^\perp V\Phi, (1 + A)^{-1} K^{-1/2} P^\perp V\Phi), \tag{2.8}$$

where $K = P^\perp H_0 P^\perp$ is of course a self-adjoint and strictly positive operator in $X^\perp$ so that $K^{1/2}$ and $K^{-1/2}$ exist as positive self-adjoint operators in $X^\perp$, and $A =$

$K^{-1/2}(V-E)^{\perp}K^{-1/2}$ with $A+1$ positive by Lemma 2.2. The usual Brillouin-Wigner implicit perturbation expansion comes from the expansion of $(1+A)^{-1}$ in (2.8), and an approximate expression for the eigenvalue is obtained by truncating the expansion, substituting in (2.6), and solving the resulting algebraic equation.

Let us turn now to the above mentioned statements.

Lemma 2.1: For any real g the operator $H^{\perp} = (H_0+gV)^{\perp} = P^{\perp}HP^{\perp}$ is self-adjoint in $X^{\perp}$, and for any $g \in I$ its lower bound $B(g)$ is strictly greater than $E(g)$.

Proof: The first part of the statement is obvious. As for the second part, it will be shown that by the existence of a $\delta(g) > 0$ that the interval $(-\infty, E+\delta)$ belongs to the resolvent set of H. Since $P^{\perp}$ is a projection of nullity 1, we can use the well-known Weinstein-Aronszajn formula,[5] which gives the relationship between the isolated eigenvalues of H and PHP if P is a projection of finite nullity. According to the Weinstein-Aronszajn method, we have to compute

$$\omega(z,H) = (u,(H-z)^{-1}u), \tag{2.9}$$

u being any basis vector of $(1-P^{\perp})X$, i.e., any non-zero vector proportional to Φ. By our hypotheses, there exists a transformation function[6] $U(g)$, analytic and unitary for $g \in I$, such that $P(g) = U(g)P(0)U(g)^{-1} = U(g)P^{\perp}U(g)^{-1}$. Taking $u = U(g)\Phi$ the spectral theorem yields

$$\omega(z,H) = \frac{1}{E(g)-z} + \int_{E(g)+d}^{\infty} \frac{d\rho(\lambda)}{\lambda - z}, g \in I,$$

where $\overline{E}(\lambda)$ is the spectral family associated with H, $\rho(\lambda) = (u, \overline{E}(\lambda)u)$, and d is the isolation distance of the eigenvalue $E(g)$, positive for any $g \in I$.

It follows that $\omega(z,H)$ is meromorphic in the whole complex z plane cut along the real axis from $E+d$ to $+\infty$, with a simple pole at $z = E(g)$. In addition, since zeroes and poles of a meromorphic function are isolated points, there is a $\delta(g) > 0$ such that $\omega(z)$ is different from zero for $E(g) < z < E(g) + \delta(g)$. Hence by the second Weinstein-Aronszajn formula we can conclude that no point of $\sum(H^{\perp})$ lies in the interval $(-\infty, E(g)+\delta(g))$, and this proves the lemma.[7]

Lemma 2.2: Let A be the symmetric operator in $X^{\perp}$ defined as $A = K^{-1/2}(V-E)^{\perp}K^{-1/2}$. Then A is a bounded and strictly positive operator in $X^{\perp}, g \in I$.

Proof: Since V is H_0-bounded in X, it is easy to see that $(V-E)^{\perp}$ is K-bounded in $X^{\perp}$. Now K is self-adjoint and semibounded in $X^{\perp}$ and $(V-E)^{\perp}$ is symmetric. It is known that this implies that the quadratic form $((V-E)^{\perp}u, u)$ is relatively bounded with respect to the form (Ku,u) and the same is true for their closures.[8] It follows that the symmetric quadratic form $(K^{-1/2}(V-E)^{\perp}K^{-1/2}v, v)$ $v \in X^{\perp}$, is bounded. Then A is bounded and hence self-adjoint.

As for the positivity, we make use of an argument of McClary.[3] Since $B(g) > E(g)$ and $E(g)$ is continuous, there exists $\epsilon > 0$ such that

$$B\left(\frac{g}{1-\epsilon}\right) > \frac{1}{(1-\epsilon)} E(g)$$

and there is a $C(g) > E(g)$ such that

$$B\left(\frac{g}{1-\epsilon}\right) > \frac{1}{1-\epsilon} C(g) > \frac{1}{1-\epsilon} E(g).$$

Then if $E < C(g)$ we have as a quadratic form on $D^{\perp} \otimes D^{\perp}$

$$\left(H_0 + \frac{g}{1-\epsilon} V\right)^{\perp} > B\left(\frac{g}{1-\epsilon}\right) > \frac{1}{1-\epsilon} E \quad (D^{\perp} = X^{\perp} \cap D),$$

i.e., $(H_0+gV)^{\perp} - E > H_0^{\perp} = \epsilon K$, which is equivalent to the relation $1 + A > \epsilon$ on $K^{1/2}D^{\perp} \otimes K^{1/2}D^{\perp}$. Since $K^{1/2}$ is positive and essentially self-adjoint on[9] $D^{\perp}$, the set of vectors $K^{1/2}D^{\perp}$ is dense in $X^{\perp}$ and the assertion follows.

Theorem 2.1: As an operator identity in $X^{\perp}$ we have

$$[(H_0+gV)^{\perp} - E]^{-1} = K^{-1/2}(1+A)^{-1}K^{-1/2}. \tag{2.10}$$

Proof: For any $f \in D^{\perp}$ we can write

$$[(H_0+gV)^{\perp} - E]f = K^{1/2}(1+A)K^{1/2}f. \tag{2.11}$$

Now, as it has been seen before, $D' = K^{1/2}D^{\perp}$ is dense in $X^{\perp}$ and by Lemma 2.2 $(1+A)D'$ is also dense. Hence for $E < C$ we can invert (2.11) and conclude

$$[(H_0+gV)^{\perp} - E]^{-1} = K^{-1/2}(1+A)K^{-1/2}.$$

3. CONVERGENCE OF THE CONTINUED FRACTION

The validity of the Brillouin-Wigner implicit formula $E(g) - E_0 - (\Phi, V\Phi) = f(E,g)$

$$f(E,g) = -g^2(K^{-1/2}P^{\perp}V\Phi, (1+A)^{-1}K^{-1/2}P^{\perp}V\Phi) \tag{3.1}$$

has been justified in the former Sec. Let us turn now to the convergence questions.

Theorem 3.1: The Brillouin-Wigner expansion of $E(g)$ has a non-zero "radius of convergence".

Proof: Since A is bounded, we can expand $(1+A)^{-1}$ in geometrical series in (3.1):

$$f(E,g) = \sum_{n=0}^{\infty} (-1)^n C_n ; C_n = (K^{-1/2}P^{\perp}V\Phi, A^nK^{-1/2}V\Phi) \tag{3.2}$$

and there is a $D(E)$ such that

$$|C_n| < D(E)^n, \quad n = 0,1,2,\cdots. \tag{3.3}$$

This shows that the power series $\sum_{n=0}^{\infty} C_n(-z)^n$ has a nonzero radius of convergence, depending on E and g.

Lemma 3.1: Let M be the upper bound of $\sum(A)$. The Hausdorff moment problem

$$C_n = \int_{-1+\epsilon}^{M} X^n d\varphi(x), \quad n = 0,1,2,\cdots, \tag{3.4}$$

where the numbers c_n are defined by (3.2), $n = 0,1,2\cdots$, and $d\varphi(x)$ is a positive measure on $[-1+\epsilon, M]$, has the unique solution $\varphi(x) = (K^{-1/2}P^{\perp}V\Phi, E(x)K^{-1/2}P^{\perp}V\Phi)$ where $E(x)$ is the resolution of the identity of A.

Proof: The fact that $d\varphi(x)$ is a solution of (3.4) is a direct consequence of the spectral theorem. The uniqueness comes from the general statement that whenever a Hausdorff moment problem has a solution it is determined.[10]

Consider now the sequence of orthonormal polynomials $P_n(z), n = 0,1,2,\cdots$, of degree n, obtained from the powers $1, z, \ldots, z^n$ through the Gram-Schmidt orthogonalization procedure, the orthogonality being defined by

$$\int_{-1+\epsilon}^{M} P_m(X)P_n(X)d\varphi(X) = \delta_{m,n}.$$

Consider also the polynomials $Q_n(z)$, of degree $n-1$, defined as

$$Q_n(z) = -\int_{-1+\epsilon}^{M} \frac{P_k(z) - P_k(x)}{z-x} d\varphi(x).$$

As is well known,[11] the rational functions $Q_n(z)/P_n(z)$ are the successive approximants of the J-type continued fraction associated with the power series $\sum_{n=0}^{\infty} C_n(z)^n$. It is also well known[12] that $Q_n(z)/P_n(z)$ coincides with the $[n-1,n]$ Padé approximant to the power series $\sum_{n=0}^{\infty} C_n(z)^{-n-1}$, uniquely defined as that rational function whose Taylor expansion at the origin coincides with $\sum_{n=0}^{\infty} C_n(+z)^{-n-1}$ up to the order $2n-1$. Let us indicate with Q_n/P_n the sequence $Q_n(z)/P_n(z)|_{z=-1}$. The convergence statement for the eigenvalue is the following:

Theorem 3.2: Let $f_n(E,g) = -Q_n/P_n$, and $E_n(g)$ be the least solution of the equation $E(g) = f_n(E,g) + E_0 + (\Phi, V\Phi)$. Then for any $g \in I$, $E_n(g)$ decreases monotonically to $E(g)$ as $n \to \infty$.

Proof: Since the coefficients c_n have the representation (3.2), it is well known[11] that for $z < -1+\epsilon$ the rational functions $Q_n(z)/P_n(z)$ are positive and monotonically decreasing as $\to \infty$. Furthermore, since the representation (3.2) is unique, $\lim_{n\to\infty} \frac{Q_n(z)}{P_n(z)} = \int_{-1+\epsilon}^{M} \frac{d\varphi(x)}{x-z}$ $= -(K^{-1/2}P^{\perp}V\Phi, (z-A)^{-1}K^{-1/2}P^{\perp}V_{\Phi})$, uniformly in any compact subset of the complex z plane cut along the real axis from $-1+\epsilon$ to M[11]. We may then take $z = -1$, so that $f_n(E,g)$ is negative and monotonically decreasing to $f(E,g)$ as $n \to \infty$, and consequently $E_n(g)$ is monotonically decreasing as well, with $\lim_{n\to\infty} E_n(g) = E(g)$.

Let us briefly indicate now an application to the Helium atom and Helium-like ions. In this connection the present result justifies earlier computations of the ground state energy of the above mentioned systems by means of the $[n-1,n]$ Padé approximants applied to the Brillouin-Wigner expansion.[13]

In atomic units, the Schrödinger operator for a system consisting of a fixed nucleus of charge Z and two electrons is given by $H = H_0 + gV$, where

$$H_0 = T - U, \quad T = -\tfrac{1}{2}\Delta_1 - \tfrac{1}{2}\Delta_2, \quad U = (1/\tau_1) + (1/\tau_2)$$
$$V = \frac{1}{\tau_{12}}, \quad g = 1/Z, \quad Z > 0. \tag{3.5}$$

Here $\bar{r}_i = (x_i, y_i, z_i)$, $i = 1, 2$, is the position of the ith electron, and $r_{12} = [(x_1-x_2)^2 + (y_1-y_2)^2 + (z_1-z_2)^2]^{1/2}$ is the relative distance. Δ_i, $i = 1, 2$, is the three-dimensional Laplacian $\partial^2/\partial x_i^2 + \partial^2/\partial y_i^2 + \partial^2/\partial z_i^2$.

It is well known that $H = H(g)$ can be realized in $L^2(R^6)$ as a self-adjoint holomorphic family of type (A) as well as (B_0) defined for any complex g.[14]

It is also known[15] that for any $g > 0$ the least (ground state) eigenvalue is isolated and nondegenerate. It is then possible to apply the former statements and conclude that the $[n-1,n]$ Padé approximants to the Brillouin-Wigner expansion converge to the ground state eigenvalue for the He atom and any He-like ion.[16]

[1] T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, 1966), especially Chap. VII.

[2] The analytic continuation of $E(g)$ in a region outside the radius of convergence of the Rayleigh-Schrodinger expansion has been obtained through the Mittag-Leffler method. See M. Reeken, J. Math. Phys. **11**, 822 (1970).

[3] W. K. McClary, Commun. Math. Phys. **24**, 171 (1972).

[4] $I \subset R$ here is the maximal analyticity interval, containing the origin, of $E(g)$ as the least eigenvalue of H. The maximal analyticity interval of $E(g)$ as an eigenvalue of H may contain I, because there may be level crossing without singularities. Our definition excludes that such a level crossing can take place for $g \in I$. (See Ref. 1. Chap. 11-6.4 and VII-3.1,2.

[5] Reference 1, Chap. IV-6.1, 2.

[6] Reference 1, Chap. II-6.2 and VII-3.2.

[7] The second Weinstein-Aronszajn formula does not hold for $z = 0$. This makes no difficulty, because in the present case H_0 is semi-bounded as well as H.

[8] Reference 1, Chap. VI-1.7.

[9] Reference 1, Chap. V-3.11.

[10] J. Shohat and J. D. Tamarkin, *The Problem of Moments* (A. M. S. Colloquium Publications, Providence, R. I., 1943).

[11] N. I. Akhiezer, *The Classical Moment Problem* (Oliver and Boyd, London, 1965).

[12] See, e. g., *The Padé Approximant in Theoretical Physics*, edited by G. A. Baker and J. L. Gammel (Academic, New York, 1970).

[13] G. L. Bendazzoli, O. Goscinski, and G. Orlandi, Phys. Rev. **2A**, 2 (1970).

[14] Reference 1, p. 410

[15] T. Kato, Trans. Americ. Math. Soc. **70**, 212 (1951).

[16] A constant must be added to H_0 to make it a positive operator. Of course this does not change anything in the procedure.

Solving linear stochastic differential equations

A. Brissaud

Ecole Nationale Supérieure de l'Aéronautique et de l'Espace, B.P. 4032, 31055-Toulouse-Cedex, France

U. Frisch

Centre National de la Recherche Scientifique, Observatoire de Nice, 06300-Nice, France
(Received 18 September 1973)

The aim of this paper is to provide the user with tools for the solution of linear differential equations with random coefficients. Only analytic methods which lead to expressions in closed form for first and second order moments and probability distributions of the solution are considered. The paper deals both with approximate methods which require the existence of a small (or large) dimensionless parameter and with the method of model coefficients, where the true coefficients of the stochastic equation are replaced by random step functions with the same first and second order moments and probability distributions, chosen in such a way that the equation can be solved analytically. The second procedure does not rely on the existence of a small parameter.

1. INTRODUCTION

Consider a linear system subject to time dependent stochastic perturbations (both in the external forces and in the parameters). The evolution of such a system is governed by a set of linear differential equations with random coefficients (stochastic equations) of the form

$$\frac{d}{dt} X_i(\omega;t) = \sum_j M_{ij}(\omega;t) X_j(\omega;t) + F_i(\omega;t), \qquad i,j = 1,\ldots,n, \quad (1.1)$$

where ω is an element of a probability space Ω, the X_i describe the state of the system in an n-dimensional space and where the parameters (coefficients) $M_{ij}(\omega;t)$ and the forces $F_i(\omega;t)$ are prescribed stationary random functions of the time variable t. To simplify the notation, ω will usually be omitted. In addition to Eq. (1.1), a set of initial conditions is given (usually nonrandom)

$$X_i(\omega;0) = X_i^0. \quad (1.2)$$

Examples of physical applications of linear stochastic differential equations are mentioned in the concluding section. Broadly speaking, by "solving" a stochastic equation we mean finding the statistical properties of the solution. Notice that most of the material covered in this paper can be extended to linear stochastic operational differential equations involving time dependent stochastic operators in an abstract finite- or infinite-dimensional space. However, the more difficult problem of stochastic partial differential equations is not covered here (see, e.g., Refs. 1–3).

When dealing with the linear stochastic equation (1.1), it is convenient to introduce the *Green's function* G satisfying an equation which in matrix notations reads

$$\frac{d}{dt} G(t,t') = M(t)G(t,t'), \qquad G(t',t') = I, \quad (1.3)$$

where I is the identity matrix. In terms of G, the solution of Eq. (1.1) with the initial condition (1.2) may be written

$$X(t) = G(t,0)X(0) + \int_0^t G(t,t')\,F(t')dt'. \quad (1.4)$$

The aim of this paper is to present the reader with a variety of methods which have proved to be useful in dealing with physical applications. We shall concentrate on analytic methods leading to exact or approximate solutions in closed form. Questions of existence, uniqueness, measurability, stability, etc., will not be considered here.[4,5,6]

It is useful to distinguish between two approaches: one either tries to find an approximate solution of the stochastic equation using the true random coefficients, or to find an exact solution using a model (e.g., a Markov process) for the random coefficients.

In Sec. 2, various approximation methods will be reviewed and their validity discussed. This includes the Born approximation, the static approximation, the Bourret and related approximations (diffusion and Hashminskii limits, Kraichnan direct interaction approximation). The concept of Kubo number, a measure of the effect of the stochastic perturbation over one correlation time, is introduced.

In Sec. 3, it is shown that the mean Green's function of a linear stochastic differential equation can be obtained explicitly for a rather large class of random coefficients called kangaroo processes (KP) for which the single time probability distribution and the two-time second order moments can be chosen in a rather arbitrary way. Particular attention is given to the validity of the approximation procedure where the true coefficients of a stochastic equation are replaced by KP coefficients.

In Sec. 4, the calculations are extended to second order moments and probability distributions of the solution, and also to the inhomogeneous case. Nonlinear stochastic differential equations are also briefly considered in connection with the Liouville equation approach. It is also shown that for certain conservative systems, the asymptotic probability distribution of the $X_i(t)$ for $t \to \infty$ can be obtained explicitly from ergodic theory.

Sections 2, 3 (excepting part C), and 4 (excepting part A) can be read independently.

Finally, we mention that, as far as the result are concerned, there is quite a bit of overlap between this paper and other papers on linear stochastic differential equations,[7,8] especially in Secs. 1B and 3A. The distinctive features of this paper are that

(i) many results usually obtained by Fokker–Planck techniques are here derived simply by averaging the equations and using the semigroup property of the Green's function;

(ii) a large class of exactly soluble equations is obtained;

(iii) the ranges of validity of the various methods are carefully examined and a guide for the user is given in the last section.

2. APPROXIMATION PROCEDURE FOR LIMITING CASES.

A. Short time perturbation expansions: The Born and mean Born

We start from Eq. (1. 3) for the Green's function, written as

$$\frac{d}{dt}\, G(t,t') = [M_0 + M_1(t)]G(t,t'), \quad G(t',t') = I, \tag{2. 1}$$

where we have separated the stationary random matrix $M(t)$ into its mean value M_0 and its fluctuating part $M_1(t)$; Eq. (2. 1) is easily recast into the following integral form

$$G(t,t') = e^{M_0(t-t')} + \int_{t'}^{t} e^{M_0(t-t'')} M_1(t'') G(t'',t')\, dt'' \tag{2. 2}$$

which, when iterated, yields the well-known von Neumann series

$$\begin{aligned} G(t,t') = {} & e^{M_0(t-t')} + \int_{t'}^{t} dt_1\, e^{M_0(t-t_1)} M_1(t_1)\, e^{M_0(t_1-t')} \\ & + \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2\, e^{M_0(t-t_1)} M_1(t_1)\, e^{M_0(t_1-t_2)} \\ & \times M_1(t_2)\, e^{M_0(t_2-t')} + \cdots . \end{aligned} \tag{2. 3}$$

To study the convergence of this expansion, we assume that M_0 and $M_1(t)$ are operators acting in a normed space. The norm of the vector X is denoted $\| X \|$. Furthermore, we assume that

$$\| e^{M_0 t} \| \leqq 1. \tag{2. 4}$$

This condition is satisfied if, e.g., M_0 is anti-Hermitian or dissipative. In the rest of this paper we shall denote by σ the order of magnitude of the fluctuations of the coefficients of the stochastic equation (1. 1). This can be measured, e.g., by the largest dispersion of the coefficients of $M(t)$ assumed to be finite. To avoid unnecessary complications, we assume in this section, the much stronger condition

$$\| M_1(t) \| \leqq \sigma \tag{2. 5}$$

almost surely and for any t. It is then easily seen that the norm of the nth term in the perturbation expansion (2. 3) is less than

$$|t-t'|^n \ \frac{\sigma^n}{n!}.$$

We conclude that the perturbation expansion is always convergent and that

$$G(t,t') = e^{M_0(t-t')} + O(|t-t'|\,\sigma); \tag{2. 6}$$

for moderate values of $|t-t'|\,\sigma$ we can use the *Born approximation*

$$\begin{aligned} G(t,t') = {} & e^{M_0(t-t')} + \int_{t'}^{t} e^{M_0(t-t_1)} M_1(t_1)\, e^{M_0(t_1-t')}\, dt_1 \\ & + O[(|t-t'|\,\sigma)^2]. \end{aligned} \tag{2. 7}$$

Consider now the mean Green's function $\langle G(t,t') \rangle$. Since $\langle M_1(t) \rangle = 0$, the second term in Eq. (2. 3) vanishes upon averaging. Expanding to second order we obtain the "mean Born approximation"

$$\begin{aligned} \langle G(t,t') \rangle \approx {} & e^{M_0(t-t')} + \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2\, e^{M_0(t-t_1)} \\ & \times \langle M_1(t_1)\, e^{M_0(t_1-t_2)} M_1(t_2) \rangle\, e^{M_0(t_2-t')}. \end{aligned} \tag{2. 8}$$

At first sight, the validity of (2. 8) as an approximation still requires $|t-t'|\,\sigma \ll 1$. However, let us assume that $M_1(t)$ has a finite correlation time, i.e., that its autocorrelation is integrable; define the correlation time T_{corr} as the integral scale of the autocorrelation [roughly speaking, T_{corr} is the time over which $M_1(t_1)$ and $M_1(t_2)$ are appreciably correlated]. Now, we notice that the major contribution to the double integral in (2. 8) comes from $|t_1 - t_2| \lesssim T_{corr}$; as a consequence the order of magnitude of the second term on the rhs of (2. 8) is only $\sigma^2 T_{corr}\, |t-t'|$ and not $\sigma^2 |t-t'|^2$. Hence, the validity of the mean Born approximation requires

$$\sigma^2 T_{corr}\, |t-t'| \ll 1, \tag{2. 9}$$

which is weaker than $|t-t'|\,\sigma \ll 1$ provided that $|t-t'| \gg T_{corr}$.

B. Weak perturbations: Bourret approximation, the white noise, and Hashminskii limits

Clearly, when $|t-t'|\,\sigma \gg 1$, the perturbation expansion is of little use. We now seek an approximate expression for $\langle G(t,t') \rangle$ valid for arbitrarily large $|t-t'|$.

Iterating the integral equation (2. 2) once averaging, we obtain

$$\begin{aligned} \langle G(t,t') \rangle = {} & e^{M_0(t-t')} + \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2\, e^{M_0(t-t_1)} \\ & \times \langle M_1(t_1)\, e^{M_0(t_1-t_2)} M_1(t_2)\, G(t_2,t') \rangle. \end{aligned} \tag{2. 10}$$

We notice that

$$\langle G(t,t') \rangle = \langle G(t-t',0) \rangle \tag{2. 11}$$

which is a consequence of the stationarity of $M_1(t)$; we may therefore as well set $t' = 0$. Differentiating with respect to t, we obtain

$$\begin{aligned} \frac{d}{dt} \langle G(t,0) \rangle = {} & M_0 \langle G(t,0) \rangle + \int_0^t \langle M_1(t)\, e^{M_0(t-t')} \\ & \times M_1(t')\, G(t',0) \rangle\, dt'. \end{aligned} \tag{2. 12}$$

Bourret[9] has proposed the following closure approximation

$$\begin{aligned} & \langle M_1(t)\, e^{M_0(t-t')} M_1(t')\, G(t',0) \rangle \\ & \quad \approx \langle M_1(t)\, e^{M_0(t-t')} M_1(t') \rangle \langle G(t',0) \rangle, \end{aligned} \tag{2. 13}$$

originally obtained by him as a first order approximation on the basis of a diagrammatic expansion;[9] this approximation can also be obtained quite differently as will be shown below.

Equation (2. 12) reduces upon use of (2. 13) to a simple integrodifferential equation for $\langle G(t,0) \rangle$ which we shall call the *Bourret equation:*

$$\begin{aligned} \frac{d}{dt} \langle G(t,0) \rangle = {} & M_0 \langle G(t,0) \rangle + \int_0^t \langle M_1(t)\, e^{M_0(t-t')} \\ & \times M_1(t') \rangle \langle G(t',0) \rangle\, dt', \quad G(0,0) = I. \end{aligned} \tag{2. 14}$$

Equivalent equations have been proposed by Keller[10] and Frisch [2]; closed equations of this type for mean quantities are generally called *master equations*. Notice that the Bourret equation is easily solved by Laplace transformation. Indeed, defining

$$\langle \tilde{G}(z) \rangle = \int_0^\infty e^{izt} \langle G(t,0) \rangle\, dt \tag{2. 15}$$

and

$$\tilde{K}(z) = \int_0^\infty e^{izt} \langle M_1(t)\, e^{M_0 t} M_1(0) \rangle\, dt, \tag{2. 16}$$

we obtain

$$\langle \tilde{G}(z) \rangle = [-iz - M_0 - \tilde{K}(z)]^{-1}. \tag{2. 17}$$

Now we investigate the validity of the closure assumption (2. 13). Let us assume that

$$\| G(t, 0) \| \leqq 1 . \tag{2. 18}$$

This condition is usually satisfied, since in most applications $M_0 + M_1(t)$ is anti-Hermitian and, hence, $G(t, 0)$ is unitary. It is known that any Green's function satisfies a semigroup property

$$G(t', 0) = G(t', s)\, G(s, 0). \tag{2. 19}$$

From the preceding section, we have

$$G(t', s) = e^{M_0(t'-s)} + O(\sigma(t' - s)). \tag{2. 20}$$

From (2. 18), (2. 19), and (2. 20) we obtain

$$G(t', 0) = e^{M_0(t'-s)}\, G(s, 0) + O(\sigma(t' - s)). \tag{2. 21}$$

Assuming

$$\sigma \mid t' - s \mid \ll 1, \tag{2. 22}$$

we can write the lhs of (2. 13) in the following form:

$$\langle M_1(t)\, e^{M_0(t-t')}\, M_1(t')\, G(t', 0)\rangle \approx \langle M_1(t)\, e^{M_0(t-t')} M_1(t') \times e^{M_0(t'-s)}\, G(s, 0)\rangle . \tag{2. 23}$$

We now make the fundamental assumption that $M_1(t)$ has a finite correlation time T_{corr}. For

$$\mid t' - s \mid \gg T_{\mathrm{corr}} \tag{2. 24}$$

the stochastic Green's function $G(s, 0)$, which is a functional of $M_1(\tau)$ for $0 < \tau < s$, is only very weakly correlated to $M_1(t')$ and $M_1(t)(t > t')$. It is therefore legitimate to factorize the rhs of (2. 23) to obtain

$$\langle M_1(t)\, e^{M_0(t-t')}\, M_1(t')\, G(t', 0)\rangle \approx \langle M_1(t)\, e^{M_0(t-t')}\, M_1(t')\rangle \times \langle e^{M_0(t'-s)}\, G(s, 0)\rangle . \tag{2. 25}$$

Using again (2. 20), we obtain the desired closure approximation (2. 13).

The compatibility of (2. 22) and (2. 24) obviously requires

$$K = \sigma\, T_{\mathrm{corr}} \ll 1 . \tag{2. 26}$$

The dimensionless number K, sometimes called the generalized Reynolds number,[2] will be called the *Kubo number*, because it was first introduced by Kubo.[11]

In deriving the Bourret equation we implicitly assumed $t \gg T_{\mathrm{corr}}$; this is indeed a consequence of $t > t' > s > 0$ and of (2. 24). In fact, the Bourret equation is also valid for small times since it can be checked that the perturbation expansion solution of Eq. (2. 14) agrees with the mean Born approximation (2. 8) up to the order of σ^2.

It is interesting to notice that the closure approximation (2. 13) and the Bourret equation become exact, whatever the Kubo number, if $M_1(t)$ is of the form

$$M_1(t) = m(t)\, L_1, \tag{2. 27}$$

where $m(t)$ is a dichotomic Markov process (also called random telegraph process) and L_1 is a constant matrix.[12] Recall that the dichotomic Markov process is defined as a step function with values ± 1, the transitions occurring at Poisson distributed times; this process is a special case of the KAP introduced in Sec. 3A.

The Bourret equation (2. 14) is a "non-Markovian" master equation, i.e., the derivative of $\langle G(t, 0)\rangle$ involves an integral over past values of the mean Green's function. Yet, the Bourret equation can be used as starting point for the derivation of various Markovian approximations which we shall now consider.

For $t \gg T_{\mathrm{corr}}$ the Bourret equation can be reduced to the following Markovian form, first given by Kubo:[11]

$$\frac{d}{dt}\langle G(t, 0)\rangle = \Big(M_0 + \int_0^\infty \langle M_1(s)\, e^{M_0 s}\, M_1(0)\, e^{-M_0 s}\rangle\, ds\Big) \times \langle G(t, 0)\rangle, \quad \langle G(0, 0)\rangle = I. \tag{2. 28}$$

To derive Eq. (2. 28) from the Bourret equation (2. 4), we notice that, as a consequence of $K \ll 1$, we have for $\mid t - t' \mid < T_{\mathrm{corr}}$

$$G(t'0) \approx e^{-M_0(t-t')}\, G(t, 0). \tag{2. 29}$$

To obtain (2. 28) we then put $t - t' = s$ and integrate over s from zero to infinity, rather than from zero to t; this is legitimate provided that the covariance of $M_1(t)$ is integrable, since the integrand will be negligible for $t \gg T_{\mathrm{corr}}$.

The Kubo equation (2. 28) has two limiting cases which actually cover all situations as we shall find later. First, the *white noise limit:* write $M_1(t) = \sigma M_1'(t/T_{\mathrm{corr}})$ and let $T_{\mathrm{corr}} \to 0, \sigma \to \infty$ in such a way that $\sigma^2\, T_{\mathrm{corr}} \to D$. It is easily seen that in this limit the factors $e^{\pm M_0 s}$ in Eq. (2. 28) cancel out and that the Kubo equation goes over into

$$\frac{d}{dt}\langle G(t, 0)\rangle = M_0 \langle G(t, 0)\rangle + D \int_0^\infty \langle M_1'(s)\, M_1'(0)\rangle\, ds \langle G(t, 0)\rangle . \tag{2. 30}$$

Since, in the white noise limit, the Kubo number $K = \sigma T_{\mathrm{corr}}$ goes to zero, Eq. (2. 30) becomes exact. (Notice that, whereas the amplitude of white noise is infinite, its strength, measured by the Kubo number, is zero.) In Ref. 13 the reader will find another derivation of a master equation equivalent to (2. 30) which uses the fact that white noise can be defined as the limit of shot noise.

We turn now to the *Hashminskii limit.* If we let the strength σ of the stochastic perturbation go to zero, the variations of the Green's function over a finite time interval will be entirely due to M_0. We now factor out the variation to M_0 by introducing the "interaction representation"

$$\langle G(t, 0)\rangle = e^{+M_0 t}\, \langle G_I(t, 0)\rangle . \tag{2. 31}$$

Then, we let $t \to \infty$ in such a way that $\sigma^2 t$ remains finite; this results in a finite variation of $\langle G_I\rangle$. Indeed, writing $M_1 = \sigma M_1'$ and $t = \tau/\sigma^2$, we find that in the limit $\sigma \to 0$, the Kubo equation (3. 28) goes over into the Hashminskii equation

$$\frac{d}{d\tau}\, \langle G_I(\tau, 0)\rangle = H \langle G_I(\tau, 0)\rangle , \tag{2. 32}$$

wherein

$$H = \lim_{\sigma \to 0}\, e^{-M_0 \tau/\sigma^2} \int_0^\infty \langle M_1'(s)\, e^{M_0 s}\, M_1'(0)\, e^{-M_0 s}\rangle\, ds\, e^{+M_0 \tau/\sigma^2}. \tag{2. 33}$$

Limits of the form $\lim_{\sigma \to 0} e^{-M_0 \tau/\sigma^2} A e^{+M_0 \tau/\sigma^2}$ are frequently used in the quantum mechanical theory of S matrices.[14] The existence of the limit requires that M_0 be anti-Hermitian; it is then easily checked that H commutes with M_0 (hint: take a representation where M_0 is diagonal). This result greatly simplifies the resolution

of the Hashminskii equation.[15] Other derivations of the Hashminskii equation, based on Fokker–Planck techniques may be found in Refs. 16 and 17.

Let us now investigate more closely the validity of the white noise equation (2. 30) and the Kubo equation (2. 28); we see that the only difference is the drop out of the factors $e^{\pm M_0 s}$. Since the integral over s extends over roughly one correlation time $T_{\rm corr}$, we may safely neglect the exponentials if the following condition is fulfilled

$$\| M_0 \| \, T_{\rm corr} \ll 1. \tag{2. 34}$$

For the Hashminskii limit the problem is somewhat more difficult. Consider the Kubo equation (2. 28); the first operator M_0 on the rhs, which is usually anti-Hermitian, does not contribute to the relaxation of the mean Green's function as $t \to \infty$. This relaxation comes entirely from the second operator

$$R = \int_0^\infty \langle M_1(s) e^{M_0 s} M_1(0) e^{-M_0 s} \rangle \, ds. \tag{2. 35}$$

By dimensional analysis we find that this operator is of the order of $\sigma^2 T_{\rm corr}$. Hence, the relaxation time $t_{\rm rel}$ must be of the order of $(\sigma^2 T_{\rm corr})^{-1}$. If we now rewrite the Kubo equation in the interaction representation, we obtain

$$\frac{d}{dt} \langle G_I(t, 0) \rangle = e^{-M_0 t} R e^{+M_0 t} \langle G_I(t, 0) \rangle. \tag{2. 36}$$

If t is large enough we can replace $e^{-M_0 t} R e^{+M_0 t}$ by its limit for $t \to \infty$ which is precisely the Hashminskii operator H (within trivial changes of notations). Now the times t of interest are of the order of $t_{\rm rel} \approx (\sigma^2 T_{\rm corr})^{-1}$, hence the condition to be fulfilled is

$$\| M_0 \| \, (\sigma^2 T_{\rm corr})^{-1} \gg 1. \tag{2. 37}$$

If we recall that $K = \sigma T_{\rm corr} \ll 1$, we find that one of the two conditions (2. 34) and (2. 37) is automatically satisfied; there is even some overlap. We thus arrive at the important conclusion that if the Kubo number is small and if $t \gg T_{\rm corr}$ it is always possible to use one of the two white noise and Hashminskii limits.

Remark: Kraichnan[18] has proposed another master equation for the mean Green's function called the direct interaction equation. With

$$\frac{d}{dt} \langle G(t, 0) \rangle = M_0 \langle G(t, 0) \rangle + \int_0^t \langle M_1(t) \langle G(t - t', 0) \rangle \times M_1(t') \rangle \langle G(t', 0) \rangle \, dt'. \tag{2. 38}$$

This nonlinear equation is an exact consequence of Kraichnan's random coupling model. It can also be obtained from the theory of parastochastic operators.[3] The usefulness of the Kraichnan equation as an approximation is questionable. For small values of the Kubo number, the Kraichnan equation is equivalent to the much simpler Bourret equation, and for large values of the Kubo number, it can no longer be used as an approximation. For linear stochastic ordinary differential equation, a much more powerful method will be described in the subsequent sections. Nevertheless, the Kraichnan method of stochastic models, the essence of which is to introduce an additional stochastic element into the equation to make it tractable, remains very useful in dealing with nonlinear stochastic equations, particularly in the field of turbulence.[19-21]

C. Strong perturbations: The static approximation

The Bourret equation is limited to the case when the effect of the stochastic perturbation over one correlation time is weak, i.e., when $\sigma T_{\rm corr} \ll 1$. In the opposite case,

$$K = \sigma T_{\rm corr} \gg 1, \tag{2. 39}$$

the mean Green's function $\langle G(t, 0) \rangle$ can be substantially affected by the stochastic perturbation for times t satisfying

$$t \ll T_{\rm corr}. \tag{2. 40}$$

Since the stochastic operator $M(\tau)$ undergoes insignificant changes for $0 \leqq \tau \leqq t \ll T_{\rm corr}$, we may as well neglect its time dependence (but not its randomness) and integrate Eq. (1. 3) to obtain

$$\langle G(t, 0) \rangle \approx G_S(t) = \langle \exp\{tM\} \rangle. \tag{2. 41}$$

$G_S(t)$ will be called the static mean Green's function. It is often useful to deal with the static resolvent, the Laplace transform of the static Green's function

$$\tilde{G}_S(z) = \int_0^\infty e^{izt} G(t) \, dt = \langle \{ -iz - M \}^{-1} \rangle. \tag{2. 42}$$

The explicit calculation of $\tilde{G}_S(z)$ requires only a matrix inversion and an averaging over the *probability distribution* of the coefficients of M.

At first sight, the static approximation is restricted to $t \ll T_{\rm corr}$. However, in many problems, the mean Green's function is damped by phase mixing or dissipation for $t \to \infty$. If the damping time is small compared to $T_{\rm corr}$, the static Green's function can be used to describe the full relaxation. An example is provided by the pseudo-oscillator discussed in Sec. 3C.

3. THE METHOD OF MODEL COEFFICIENTS

From the preceding chapter, we know that the Born approximation is limited to short times and that the Bourret and static approximation are limited to respectively small and large Kubo numbers. Such approximations are of no use if the Kubo number is of the order of one and if it is necessary to follow the evolution of the mean Green's function over times long enough so that there is an appreciable damping by phase mixing. There is thus need for a method which puts no restriction on the Kubo number; of course, the results of Sec. 2 should be recovered in the corresponding limits.

If we recall that the Bourret equation involves in an essential way the two time second order moments (covariance) of the stochastic coefficients, whereas the static Green's function involves the single-time probability distribution of the coefficients, it is clear that an approximate master equation for the mean Green's function should involve both the probability distribution and the covariance. It turns out that it is possible to construct a class of stepwise constant Markovian random functions with arbitrary probability distributions and rather arbitrary covariances. Such functions, when used as coefficients in a linear stochastic equation, lead to a closed analytic expression for the mean Green's function.

A. The Kubo-Anderson process (KAP)

The Kubo–Anderson process (KAP) is a stepwise constant random function which jumps at randomly chosen times between random step-values. The times

$t_1, t_2, \ldots$ will be called jumping times. A more precise definition will be given below. The KAP has been introduced in connection with nuclear magnetic resonance[22,23] and introduced again, in a special case, by Bourret[24] as a tool for linear stochastic equations (see also Ref. 25).

Definition: The step-wise constant random function $m(t)$ is called a Kubo–Anderson process (KAP) if the jumping times $t_i\,(i = -\infty, \ldots, +\infty)$ are uniformly and independently distributed in $(-\infty, +\infty)$ with density ν (Poisson distribution) and $m(t)$ is a constant, $m(t) = m_i$, for $t_i \leqq t < t_{i+1}$; the m_i' are independent random variables with the same probability density $P(m)$.

We notice that $m(t)$ is a stationary Markov process with probability density $P(m)$. Assuming $\langle m \rangle = 0$ for simplicity, we obtain for the covariance of $m(t)$

$$\langle m(t)\, m(t') \rangle = \langle m^2 \rangle\, e^{-\nu |t-t'|}. \tag{3.1}$$

We see that for a KAP the probability density $P(m)$ and the correlation time

$$T_{\text{corr}} = \nu^{-1} \tag{3.2}$$

may be chosen arbitrary, but not the functional form of the covariance, which is always exponential.

We shall now show that the linear stochastic equation

$$\frac{d}{dt}\, G(t,0) = M(t)\, G(t,0), \qquad G(0,0) = I, \tag{3.3}$$

can be solved analytically for the mean Green's function provided that the coefficients of the stochastic matrix are KAP's with all the same jumping-times or, in short, when $M(t)$ is a KAP.

If there is no jumping-time between 0 and t, $M(\tau)$ remains constant for $0 \leqq \tau \leqq t$ and we are back to the static case (Sec. 2C). The probability of this event is $e^{-\nu t}$ and the corresponding contribution to $\langle G(t,0) \rangle$ is

$$\begin{aligned} \langle G(t,0) \rangle_{\text{no jump}} &= e^{-\nu t} \langle e^{tM} \rangle \\ &= e^{-\nu t}\, G_S(t). \end{aligned} \tag{3.4}$$

In the opposite case, let t' denote the last jumping-time before t. It is a well-known property of the Poisson process that the probability for this jump to occur between t' and $t' + dt'$ is $\nu e^{-\nu(t-t')}\, dt'$. Using the semigroup property of the Green's function, $G(t,0) = G(t,t')\, G(t',0)$, we can write the corresponding contribution to $\langle G(t,0) \rangle$ as

$$\langle G(t,0) \rangle_{\text{jumps}} = \int_0^t \nu e^{-\nu(t-t')} \langle G(t,t')\, G(t'0) \rangle_{t'} dt', \tag{3.5}$$

where $\langle \cdot \rangle_{t'}$, is a conditional average knowing that a jump occurred at t'. Using the fact that M remains constant between t' and t and that its value is independent of the values of $M(\tau)$, $0 \leqq \tau \leqq t'$, and therefore also independent of $G(t',0)$ which is a functional of $M(\tau)$, we obtain

$$\langle G(t,t')\, G(t',0) \rangle_{t'} = G_S(t-t') \langle G(t',0) \rangle_{t'}. \tag{3.6}$$

We claim that

$$\langle G(t',0) \rangle_{t'} = \langle G(t',0) \rangle. \tag{3.7}$$

Indeed, the knowledge that a jump occurred at t' imposes no constraints on previous jumping-times and previous values of $M(\tau)$. Adding the contribution of (3.4) and (3.5) and using (3.6) and (3.7), we obtain the Kubo–Anderson master equation

$$\begin{aligned} \langle G(t,0) \rangle_{\text{KAP}} = G_S(t)\, e^{-\nu t} + \nu \int_0^t e^{-\nu t(t-t')}\, G_S(t-t') \\ \times \langle G(t',0) \rangle_{\text{KAP}}\, dt'. \end{aligned} \tag{3.8}$$

This master equation can be solved for the mean resolvent, the Laplace transform of the mean Green's function,

$$\langle \tilde{G}(z) \rangle_{\text{KAP}} = \int_0^\infty e^{izt} \langle G(t,0) \rangle_{\text{KAP}}\, dt. \tag{3.9}$$

The solution reads

$$\langle \tilde{G}(z) \rangle_{\text{KAP}} = [I - \nu \tilde{G}_S(z + i\nu)]^{-1}\, \tilde{G}_S(z + i\nu), \tag{3.10}$$

where the static resolvent $\tilde{G}_S(z)$ is given by Eq. (3.3). Another derivation of the master equation may be found in Ref. 25 where the KAP is called the Poisson step process.

Equation (3.10) constitutes a strikingly simple result: the KAP resolvent is an algebraic function of the static resolvent; it turns out that in many applications, the quantity of interest is the resolvent and not the Green's function itself.[25,26]

Remark: The KAP can be slightly generalized to include the case of jumping-times selected according to a compound Poisson process, i.e., when the density $\nu(t)$ is a (deterministic) function of the time. A straightforward modification of Eq. (3.8) yields

$$\begin{aligned} \langle G(t,0) \rangle = G_S(t) \exp\Big(-\int_0^t \nu(\tau)\, d\tau\Big) + \int_0^t dt'\, \nu(t') \\ \times \exp\Big(-\int_{t'}^t \nu(\tau)\, d\tau\Big)\, G_S(t-t') \langle G(t',0) \rangle. \end{aligned} \tag{3.11}$$

This equation may be useful in the study of nonstationary processes.

B. The kangaroo process (KP)

We recall that a KAP has an exponential covariance. The study of the problem of stochastic Stark broadening,[25] where the covariance is proportional to $1/t$ and is not even integrable, has led us to modify the KAP by requiring that the frequency of jumping times is a function $\nu(m)$ of the value of the process itself. The new process is called a "kangaroo process" (KP).

Definition: A KP is a step-wise constant Markov process[27] with stationary transition probability given for infinitesimal time intervals by

$$\begin{aligned} P_{t\tau}(m, \Delta t \mid m', 0) = \{1 - \nu(m')\Delta t\}\, \delta(m' - m) \\ + \nu(m')\, \Delta t\, Q(m), \end{aligned} \tag{3.12}$$

where $Q(m)$ is a given probability density.

$P_{t\tau}\, dm$ is the probability that the kangaroo process at time Δt is between m and $m + dm$ knowing that it was equal to m' at time 0. The meaning of Eq. (3.12) is clear: $\{1 - \nu(m')\, \Delta t\}$ is the probability that no jump occurred in the time interval $(0, \Delta t)$ and $\nu(m')\, \Delta t$ the probability that at last one jump occurred. Immediately after such a jump, the probability density of m becomes $Q(m)$. We stress the fact that $Q(m)$ is not the stationary probability density of $m(t)$. Indeed, the Fokker–Planck equation[27] for the kangaroo process reads

$$\begin{aligned} \frac{\partial}{\partial t}\, P(m,t) &= \lim_{\substack{\Delta t \to 0 \\ \Delta t \geqq 0}} \Big\{ \int P_{t\tau}(m, \Delta t \mid m', 0)\, P(m',t)\, dm' - P(m,t) \Big\} / \Delta t \\ &= -\nu(m)\, P(m,t) + Q(m) \int \nu(m')\, P(m',t)\, dm'. \end{aligned} \tag{3.13}$$

Hence, the stationary probability density $P(m)$ of $m(t)$ is related to $Q(m)$ by

$$Q(m) = \nu(m)P(m)/\int \nu(m')P(m')\,dm' = \nu(m)P(m)/\langle\nu\rangle. \tag{3.14}$$

Next, we evaluate the covariance of a KP. The calculation of $\Gamma(t) = \langle m(t)\,m(0)\rangle$ requires the summation of a series to take into account the possible occurrence of an arbitrary number of jumps between 0 and t. After some algebra we obtain for the Laplace transform

$$\tilde{\Gamma}(z) = \int_0^\infty e^{izt}\,\Gamma(t)\,dt, \tag{3.15}$$

the following result

$$\tilde{\Gamma}(z) = \left\langle \frac{m^2}{\nu(m) - iz} \right\rangle_S - \frac{1}{iz\langle \nu(m)/[\nu(m) - iz]\rangle_S} \left\langle \frac{m}{\nu(m) - iz} \right\rangle_S^2. \tag{3.16}$$

In many cases, e.g., if $P(m)$ and $\nu(m)$ are *even*, we have

$$\langle m/(\nu(m) - iz)\rangle_S = 0 \quad \text{or, equivalently,} \quad \langle me^{-\nu(m)t}\rangle_S = 0;$$

then, the above result simplifies to

$$\tilde{\Gamma}(z) = \left\langle \frac{m^2}{\nu(m) - iz} \right\rangle_S \tag{3.17}$$

or

$$\Gamma(t) = \int_{-\infty}^{+\infty} m^2 e^{-\nu(m)|t|}\,P(m)\,dm, \tag{3.18}$$

which is just an ordinary variance conditioned by the probability $e^{-\nu(m)t}$ that no jump occurs between 0 and t. The interesting point, about formula (3.18) is that it can easily be inverted: Given $P(m)$ and the covariance $\Gamma(t)$, the jumping frequency $\nu(m)$ can be calculated as follows. Assume that $\nu(m)$ is a monotonic increasing function of $|m|$ such that $\nu(\infty) = \infty$; this is a reasonable assumption since in most applications very strong values of the stochastic perturbations last only for a very short time. Taking ν as new integration variable, we obtain

$$\Gamma(t) = 2\int_{\nu(0)}^\infty m^2 P(m)\,\frac{dm}{d\nu}\,e^{-\nu|t|}\,d\nu \tag{3.19}$$

which is essentially a Laplace integral. A calculation of $\nu(m)$ requires the inversion of the Laplace transformation and the solution of a simple differential equation. An example may be found in Ref. 25 (Sec. 5). In connection with this inversion, Table I gives some useful results. Notice that the inversion is not always feasible. Indeed, from (3.18) we see that the derivative of the covariance is necessarily discontinuous at the origin; we do not know whether this is a sufficient condition for inversion. We conclude that it is always possible to construct a KP with an arbitrary probability distribution and a (quite) arbitrary covariance.

A linear stochastic equation with KP coefficients can again be solved implicitly for the mean Green's function. Indeed, the KP resolvent can be put in the following form:

TABLE I. Kangaroo process: expression of the jumping frequency $\nu(m)$ in terms of the one-time probability distribution $P(m)$ for several types of covariance $\Gamma_{KP}(t)$.

To obtain $\Gamma_{KP}(t) = \langle m(t)m(0)\rangle_{KP} =$	Use $\nu(m) =$	Remarks
$e^{-\nu_0\|t\|}$	ν_0	
$\dfrac{\Gamma(1/n)}{\|t\|^{1/n}}$	$\left(2\int_0^{\|m\|} m^2P(m)\,dm\right)^n$	$\Gamma(\cdot)$ is the gamma function. $P(m)$ even
$\sigma^2/(1 + \|t\|)$	$-\log\left(2\int_{\|m\|}^\infty m^2P(m)\dfrac{dm}{\sigma^2}\right)$	$P(m)$ even
$\sigma^2\dfrac{1 - \|t\|e^{-\pi\|t\|/2}}{1 + t^2}$	$\arccos\left(2\int_{\|m\|}^\infty m^2P(m)\dfrac{dm}{\sigma^2}\right)$	$P(m)$ even
$\sigma^2\dfrac{\|t\| + e^{-\pi\|t\|/2}}{1 + t^2}$	$\arcsin\left[2\int_0^{\|m\|} m^2P(m)\dfrac{dm}{\sigma^2}\right]$	$P(m)$ even

$$\langle\tilde{G}(z)\rangle_{KP} = \langle\tilde{G}\rangle_S + \langle\nu\tilde{G}\rangle_S\{\langle\nu(I - \nu\tilde{G})\rangle_S\}^{-1}\langle\nu\tilde{G}\rangle_S, \tag{3.20}$$

where ν denotes $\nu(M)$,

$$\tilde{G} = \{[\nu(M) - iz]\,I - M\}^{-1}, \tag{3.21}$$

and $\langle\cdot\rangle_S$ denotes the averaging over the stationary distribution of M (static averaging). Again, the resolvent is expressed in terms of purely static quantities. A proof of Eq. (3.20) may be given which parallels the proof for the KAP given in Sec. 3A. Another more constructive proof will be found in Ref. 25, Sec. 4.

Remark 1: In some applications the stochastic operator $M(t)$ appears naturally as the sum of two (or more) processes with quite different correlation times. To deal with these situations, a *compound KP* has been defined and the corresponding mean resolvent has been calculated (cf., Ref. 25, Sec. 6).

Remark 2: The KP has a non-Markovian generalization which allows an arbitrary probability distribution for the step-length, still conditioned by the step-value of M. Let

$$B(M, t) = \text{prob}\{t_{i+1} - t_i \geqq t \mid M(\tau) = M \text{ for } t_i < \tau < t_{i+1}\} \tag{3.22}$$

be the conditional probability distribution of any step-length. Let $Q(M)$ and $P(M)$ denote respectively the probability density of step-values and of the stationary process $M(t)$; they are related by

$$Q(M) = \frac{P(M)}{\{\int_0^\infty B(M,\tau)\,d\tau \quad \langle 1/\int_0^\infty B(M,\tau)\,d\tau\rangle_S\}}, \tag{3.23}$$

where $\langle\cdots\rangle_S$ denotes averaging over $P(M)$. With these notations the resolvent is again expressible in closed form as

$$\langle\tilde{G}(z)\rangle = L_z\left(\int_t^\infty C(M,\tau)\,d\tau\right) + L_z[C(M,t)] \times \left\{L_z\left[\left(\delta(t) + \frac{d}{dt}\right)C(M,t)\right]\right\}^{-1} L_z[C(M,t)] \tag{3.24}$$

with

$$C(M,t) = B(M,t)/\int_0^\infty B(M,\tau)\,d\tau \tag{3.25}$$

and

$$L_z[f(M,t)] = \int_0^\infty dt e^{izt}\langle e^{Mt}f(M,t)\rangle_S. \tag{3.26}$$

C. The method of model coefficients used as approximation

We have seen in Sec. 2 that for weak (resp. strong) perturbations, the mean Green's function of a linear stochastic equation depends essentially on the covariance (resp. the probability density) of the coefficients. The question naturally arises how close the KP solution will fit the true solution when the true coefficients are replaced by KP's with the same probability distributions and covariances.

Let us first check that the KP solution (3.20) is in agreement with the true solution for short times satisfying $t_\sigma \ll 1$. To the lowest nontrivial order, the mean Green's function is then given by the mean Born approximation (2.8) which involves only the covariance of $M_1(t)$. Hence, for short times, the mean Green's function depends only on the covariance of $M_1(t)$.

We show now that the KP solution (3.20) reduces indeed to the previously obtained approximations of Sec. 2 for very weak and very strong perturbations. We recall that the strength of perturbations is measured by

$K = \sigma T_{\text{corr}} \approx \sigma/\langle\nu\rangle$, where σ is the order of magnitude of the fluctuating part of $M(t)$. We thus obtain the limit of very strong perturbations by letting $\nu(M) \to 0$. In this limit the KP resolvent (3. 20) reduces obviously to $G_S(z)$, which is the static resolvent; this is in agreement with the result of Sec. 2C.

To study the opposite limit, we write

$$M(t) = M_0 + \sigma M_1'(t) \quad \text{and} \quad \nu(M) = T_{\text{corr}}^{-1}\, \nu'(M_1'). \tag{3. 27}$$

As in Sec. 2B when deriving the white noise limit, we let $T_{\text{corr}} \to 0$, $\sigma \to \infty$, $\sigma^2 T_{\text{corr}} \to D$, and obtain from equation (3. 20)

$$\langle \tilde{G}(z)\rangle_{\text{KP}} = [-izI - M_0 - D\langle M_1'^2/\nu'(M_1')\rangle]^{-1} \quad \text{(diffusion limit).} \tag{3. 28}$$

This is equivalent to the following master equation

$$\frac{d}{dt}\langle G(t,0)\rangle_{\text{KP}} = (M_0 + D\int_0^\infty \langle M_1'(\tau)\, M_1'(0)\rangle_{\text{KP}}\, d\tau)\langle G(t,0)\rangle_{\text{KP}} \tag{3. 29}$$

which is identical with the white noise limit (2. 30).

Remark: The proof given here that the KP is correct in the limit of weak perturbations relies implicitly on the assumption that the covariance of $M(t)$ is integrable. In Refs. 25 and 28, we have checked, in a special case, that the KP can still be used as an approximation when the covariance is proportional to $1/t$.

So far we have only checked the agreement between the KP solution and the exact solution in limiting cases. In intermediate range of moderate perturbations, it is very difficult to draw any general conclusions. Special cases have been investigated which show indeed very good agreement; in particular, when the KP is applied to Stark broadening of spectral lines.[25-28] As a quantitative test for the validity of the intermediate range, we have compared the true and KP solutions for a randomly frequency modulated pseudo-oscillator satisfying the scalar equation

$$\frac{d}{dt}\, g(t,0) = im(t)g(t,0), \tag{3. 30}$$

where $m(t)$ is a real zero mean value stationary Gaussian process with covariance $\langle m(t)\, m(t')\rangle = \sigma^2 e^{-\nu|t-t'|}$.

Equation (3. 30) has an exact solution[2]

$$\langle g(t,0)\rangle = \exp\{-K^2[t_1 + e^{-t_1} - 1]\} \tag{3. 31}$$

with $K = \sigma/\nu$ and $t_1 = \nu t$. The corresponding exact resolvent can be written as

$$\langle \tilde{g}(z)\rangle_E = e^{K^2} \sum_{n=0}^{\infty} \frac{(-K^2)^n}{n!(n - iz + K^2)}. \tag{3. 32}$$

The KP resolvent for the same problem reduces to a KAP resolvent, since the covariance of $m(t)$ is an exponential.

The KP resolvent reads

$$\langle \tilde{g}(z)\rangle_{\text{KP}} = \left(\int \frac{-iz - iKm}{1 - iz - iKm}\, e^{-m^2/2}\, dm\right)^{-1} \cdot \int \frac{e^{-m^2/2}\, dm}{1 - iz - iKm}. \tag{3. 33}$$

The real parts of $\langle \tilde{g}(z)\rangle_E$ and $\langle \tilde{g}(z)\rangle_{\text{KP}}$ for z real, i.e., the Fourier transform of $\langle g(t,0)\rangle_E$ and $\langle g(t,0)\rangle_{\text{KP}}$, have been calculated numerically for different values of the Kubo number K and plotted as a function of z: results are shown on Fig. 1. For small and large values of K, the agreement is almost perfect. In the intermediate range $K \approx 1$ the discrepancy is at most 10%.

D. Stochastic equations with Markovian and shot noise coefficients.

The KAP and the KP processes constitute special cases of Markov processes. A general theory can be given for linear stochastic differential equations with Markovian coefficients,[2] based on the fact that the joint process $\{M(t), G(t,0)\}$ is also a Markov process; this leads to a Fokker–Planck or a Chapman–Kolmogorov equation for the joint probability density. This method should not be recommended since it leads to rather complicated partial differential or integral equations for which closed analytic solutions are generally not available.

Another case worth mentioning has been considered by Blume.[29] The coefficients are taken in the form of shot noise

$$M(t) = M_0 + \sum_{-\infty}^{+\infty} M_i\, \delta(t - t_i), \tag{3. 34}$$

where the t_i's are Poisson-distributed with density ν and the M_i's are independent identically distributed random matrices. This case, which is very similar to the KAP, leads to the following resolvent:

$$\langle G(z)\rangle = [-iz - M_0 + \nu(I - \langle \exp M\rangle)]^{-1}. \tag{3. 35}$$

4. CALCULATION OF VARIOUS STATISTICAL QUANTITIES

A. Simultaneous and time-displaced second-order moments

By second-order moments, we understand the quantities $\langle X_i(t)\, X_j(t)\rangle$ or, in short, $\langle X(t) \otimes X(t)\rangle$. We notice that, for zero, the right-hand side in Eq. (1. 1) we have

$$\langle X(t) \otimes X(t')\rangle = \langle G(t,0) \otimes G(t',0)\rangle\, (X^0 \otimes X^0), \tag{4. 1}$$

by definition of the tensor product of two matrices.

Let us first consider the case of simultaneous moments, i.e., $t = t'$. We introduce the *double Green's function*

$$\mathcal{G}(t,t') = G(t,t') \otimes G(t,t'). \tag{4. 2}$$

Differentiating (4. 2) with respect to t and using the fundamental stochastic equation (1. 3), we find that $\mathcal{G}(t,t')$, satisfies another linear stochastic equation, namely

$$\frac{d}{dt}\, \mathcal{G}(t,t') = \mathfrak{M}(t)\, \mathcal{G}(t,t'), \qquad \mathcal{G}(t,t') = I, \tag{4. 3}$$

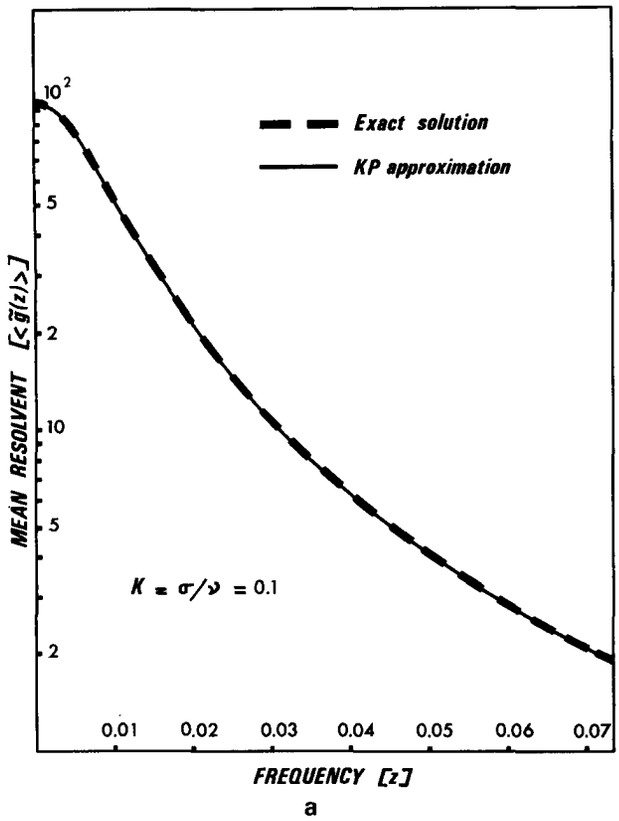
where

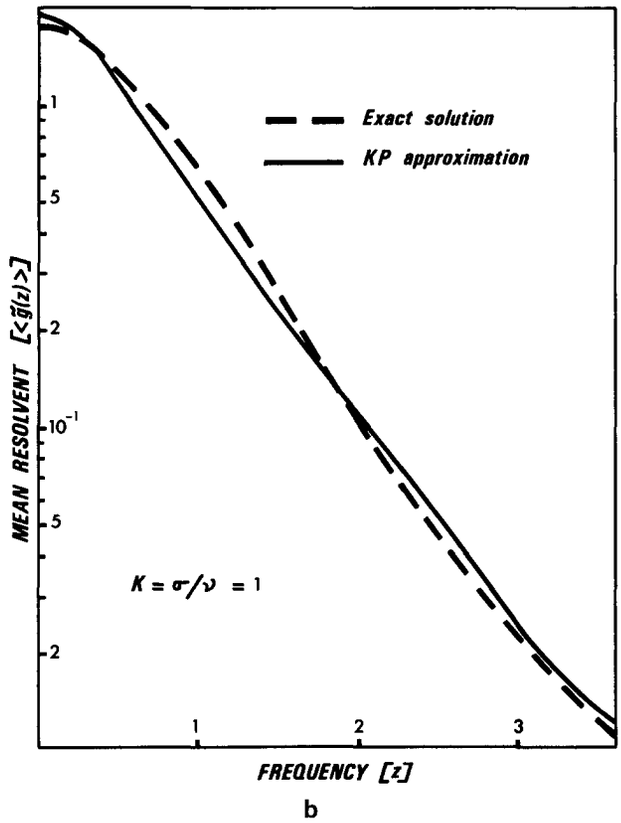
$$\mathfrak{M}(t) = M(t) \otimes I + I \otimes M(t). \tag{4. 4}$$

Clearly, Eq. (4. 3) is a linear stochastic equation of the standard form. Hence, the calculation of *simultaneous* second-order moments has been reduced to the calculation of first-order moments.[30,31]

The calculation of time-displaced moments of the form $\langle G(t,0) \otimes G(t',0)\rangle$ is somewhat more involved, except when the random coefficients are of white noise type (i.e., have zero correlation time). Indeed, assuming that $t \geqq t'$, which is no loss of generality, and using the semigroup property, we have

$$\langle G(t,0) \otimes G(t',0)\rangle = \langle G(t,t')\, G(t',0) \otimes G(t',0)\rangle. \tag{4. 5}$$





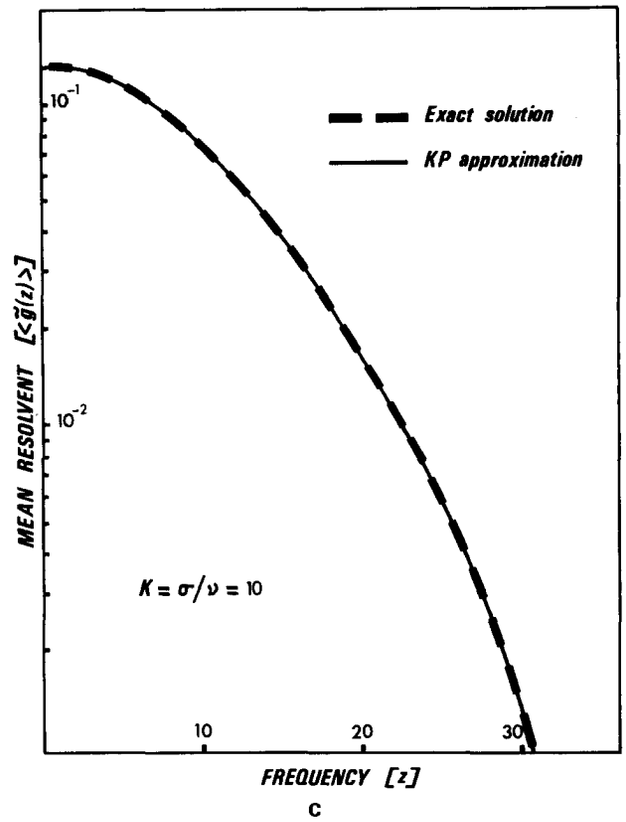


FIG. 1. *Comparison of the exact mean resolvent and of its KP approximation for the randomly modulated pseudo-oscillator defined in Section 7.*

a: $K = \sigma/\nu = 0.1$: the agreement is perfect.
b: $K = \sigma/\nu = 1$: the discrepancy is at most 10%.
c: $K = \sigma/\nu = 10$: the agreement is perfect.

Using the independence of $G(t,t')$ and $G(t',0)$ which are functionals of $M(\tau)$ for $t \geqq \tau \geqq t'$ and $t' \geqq \tau \geqq 0$, respectively, we finally obtain

$$\langle G(t,0) \otimes G(t',0)\rangle = \langle G(t-t',0)\rangle \langle \mathcal{G}(t',0)\rangle, \qquad (4.6)$$

which is a shorthand notation for $(\langle G\rangle \otimes I)\langle \mathcal{G}\rangle$.

The general case can also be dealt with if we assume that $M(t)$ is a KAP. The essential idea is to notice that the KAP is a Markov Process. It follows that if $M(t')$ is given, the past ($\tau < t'$) and the future ($\tau > t'$) become independent. If we define the conditional mean Green's functions $\langle G(t,t';M)\rangle_c$ and $\langle \mathcal{G}(t,t';M)\rangle_c$, conditioned by $M(t') = M$, we obtain, using (4.6),

$$\langle G(t,0) \otimes G(t',0)\rangle = \langle\langle G(t,t';M)\rangle_c \langle \mathcal{G}(t,t';M)\rangle_c\rangle_M, \qquad (4.7)$$

where $\langle\cdot\rangle_M$ denotes an averaging over the probability distribution of M (static averaging). It remains to calculate the conditional mean Green's function. First, we notice that from stationarity

$$\langle G(t,t';M)\rangle_c = \langle G(t-t',0;M)\rangle_c. \qquad (4.8)$$

Denoting by $\langle \tilde{G}(z;M)\rangle_c$ and $\langle \tilde{\mathcal{G}}(z;M)\rangle_c$ the Laplace transforms of $\langle G(t,0;M)\rangle_c$ and $\langle \mathcal{G}(t,0;M)\rangle_c$, we easily obtain, using the same method and notations as in Sec. 2C

$$\langle \tilde{G}(z;M)\rangle_c = [I - \nu \tilde{G}_S(z+i\nu)]^{-1}(-iz+\nu-\mathfrak{M})^{-1} \qquad (4.9)$$

and, similarly,

$$\langle \tilde{\mathcal{G}}(z;M)\rangle_c = [I - \nu \mathcal{G}_S(z+i\nu)]^{-1}(-iz+\nu-\mathfrak{M})^{-1}. \qquad (4.10)$$

This method, which is somewhat reminiscent of a method introduced by Morrison and McKenna[30,31] may lead to rather tedious calculations. A much simpler method is described in Ref. 11 for the special case when $M(t)$ is of the form

$$M(t) = M_0 + m(t)\, L_1,$$

where $m(t)$ is a dichotomic Markov process (random telegraph process).

B. Inhomogeneous equations

In general, because of dissipation, the solution of a linear stochastic equation with initial conditions and no rhs relaxes to zero as $t \to \infty$ (there is, however, a notable exception for conservative systems, see Sec. 4C). A nonzero stationary solution may be obtained in the presence of a random rhs (random driving forces). In this section we shall therefore be concerned with the inhomogeneous case

$$\frac{d}{dt}\, X_i(t) = \sum_j M_{ij}(t)\, X_j(t) + F_i(t); \qquad i,j = 1,\ldots,n. \tag{4.11}$$

$M_{ij}(t)$ is a stationary random matrix as before, and the $F_i(t)$ constitute a set of stationary random functions independent of the $M_{ij}(t)$'s. For the sake of simplicity we shall assume that $F_i(t)$ is a real zero mean value white noise, i.e.,

$$\langle F_i(t)\rangle = 0, \qquad \langle F_i(t)\, F_j(t')\rangle = S_{ij}\, \delta(t-t'), \tag{4.12}$$

where S_{ij} is a constant positive definite matrix.

In order to get stationary solutions we assume that $M_{ij}(t)$ has a dissipative part, so that

$$\lim_{t\to\infty} G_{ij}(t,t') = 0, \tag{4.13}$$

where $G_{ij}(t,t')$ is the random Green's function defined by (1.3). In terms of the Green's function we may write the solution of (4.11) as

$$X_i(t) = \sum_j G_{ij}(t,0)\, X_j^0 + \int_0^t \sum_j G_{ij}(t,t')\, F_j(t')\, dt'. \tag{4.14}$$

We shall be interested in the statistical properties of $X(t)$ for $t \to \infty$, and especially in the first- and second-order moments.

Taking the average of (4.14) and using (4.12) and (4.13), we obtain

$$\lim_{t\to\infty} \langle X_i(t)\rangle = 0. \tag{4.15}$$

Let us now evaluate the time displaced second-order moment,

$$\Gamma_{ij}(\tau) = \lim_{t\to\infty} \langle X_i(t)\, X_j(t+\tau)\rangle. \tag{4.16}$$

From (4.14), we obtain, using (4.13) and the independence of the Green's function and the driving forces,

$$\Gamma_{ij}(\tau) = \lim_{t\to\infty} \int_0^t \int_0^{t+\tau} \sum_{nm} \langle G_{in}(t,t')G_{jm}(t+\tau,t'')\rangle \times \langle F_n(t')\, F_m(t'')\rangle\, dt'\, dt''. \tag{4.17}$$

Finally, using (4.12) and the fact that

$$\langle G_{in}(t,t')\, G_{jm}(t+\tau,t')\rangle = \langle G_{in}(s,0)\, G_{jm}(s+\tau,0)\rangle, \qquad s = t-t', \tag{4.18}$$

we obtain

$$\Gamma_{ij}(\tau) = \int_0^\infty \sum_{n,m} \langle G_{in}(s,0)\, G_{jm}(s+\tau,0)\rangle\, ds\, S_{nm}, \tag{4.19}$$

which is the desired result. The calculation has been reduced to that of the second-order moment of the Green's function (see preceding section).

In the same context another (obvious) result is worth mentioning: The mean Green's function is the inter-correlation function of the solution and the right-hand side when $S_{ij} = \delta_{ij}$,

$$\lim_{t\to\infty} \langle X_i(t+\tau)\, F_j(t)\rangle = \langle G_{ij}(\tau)\rangle. \tag{4.20}$$

The main interest of the results (4.19) and (4.20) is that $\Gamma_{ij}(\tau)$ and $\lim_{t\to\infty}\langle X_i(t+\tau)\, F_j(t)\rangle$ are easily measurable quantities using a time average over a single realization, whereas the mean Green's function is not directly measurable.

C. Probability distributions: The Liouville equation

In this section, we shall consider the following problem: Let there be given the *nonlinear* stochastic differential equation

$$\frac{d}{dt}\, X_i(\omega;t) = A_i[m(\omega;t);X_j], \qquad j = 1,\ldots,n, \tag{4.21}$$

where A_i is a nonlinear real deterministic function of the real (scalar or vector-valued) random function $m(\omega;t)$ and of X_j $(j = 1,\ldots,n)$. In addition, we assume deterministic real initial conditions X_j^0.

Our purpose is to evaluate the joint probability density $P(t;X_1,\ldots,X_n)$ of $X_1, X_2, \ldots, X_n$ at time t. We shall use the so-called "Liouville equation" method which reduces the present problem to a linear stochastic equation of type studied in previous sections.

For each realization $m(\omega;t)$, let us denote by $X_i(\omega;t)$ the solution of Eq. (4.21). We shall assume existence and uniqueness. Let us think of $X_i(\omega,t)$ as a point which, starting from X_i^0, moves around in a stochastic fashion in an n-dimensional phase space. Introducing the "fine grained density,"

$$\rho(\omega;t;X_1,\ldots,X_n) = \delta\{X_1 - X_1(\omega;t)\}\cdots\delta\{X_n - X_n(\omega;t)\}, \tag{4.22}$$

we notice that the joint probability P is the average of the fine grained density, i.e.,

$$P(t;X_1,\ldots,X_n) = \langle \rho(\omega;t;X_1,\ldots,X_n)\rangle \tag{4.23}$$

which follows immediately from (4.22).

From the equation of motion (4.21), we may derive an equation of continuity, or Liouville equation, for the fine grained density, which reads

$$\frac{\partial}{\partial t}\,\rho + \sum_{i=1}^n \frac{\partial}{\partial X_i}\,[A_i\{m(\omega;t),X_i\}\,\rho\,] = 0. \tag{4.24}$$

To derive the Liouville equation (4.24), let us introduce an indefinitely differentiable test function $\varphi(X_1,\ldots,X_n)$. Integrating the lhs of (4.24) after multiplication by φ, we obtain, using (4.22),

$$\begin{aligned}\int\left(\frac{\partial\rho}{\partial t} + \sum_1^n \frac{\partial}{\partial X_i}\,[A_i p]\right)\varphi\, dX_1\ldots dX_n &= \frac{d}{dt}\int \rho\varphi\, dX_1\ldots dX_n - \int \sum_1^n \rho A_i\,\frac{\partial\varphi}{\partial X_i}\, dX_1\ldots dX_n \\ &= \sum_{i=1}^n \frac{dX_i}{dt}\,\frac{\partial\varphi[X_j(t)]}{\partial X_i} - \sum_{i=1}^n A_i\,\frac{\partial\varphi[X_j(t)]}{\partial X_i};\end{aligned} \tag{4.25}$$

this quantity vanishes identically because of (4.21). Since this property holds for an arbitrary test function, we have proved (4.24).

In spite of the fact that it contains partial derivatives, the Liouville equation can be treated as a stochastic ordinary differential equation since the random function

$m(\omega;t)$ does not depend on $X_1,\ldots,X_n$. The calculation of the joint probability distribution can now be carried out using the same methods as described in previous sections.

D. Asymptotic behavior of linear conservative systems from ergodic theory

Let us come back again to the linear stochastic equation

$$\frac{d}{dt}X = M(\omega;t)X, \qquad X(0) = X_0. \tag{4.26}$$

We assume now that X is a vector and M is a random matrix in a real n-dimensional space, and that equation (4.26) is *conservative* in the sense that

$$\sum_{i=1}^{n} X_i^2(t) = \text{const}, \tag{4.27}$$

This is obviously equivalent to the requirement that $M(\omega;t)$ be antisymmetric. For convenience, we shall assume that

$$\|X_0\|^2 = \sum_{i=1}^{n} |X_i(0)|^2 = 1. \tag{4.28}$$

The phase space of the system under consideration is then the n-dimensional unit sphere S.

Since $M(t)$ is antisymmetric, the mapping $X(t') \to X(t)$ is unitary. Therefore, the motion on the unit sphere S preserves the uniform measure dm. This situation is reminiscent of a problem in classical statistical mechanics: Given a system of interacting classical particles enclosed in a box, it is known that the point representing the system in the phase space remains on the energy surface, that the motion on the energy surface has an invariant measure, and that the point will eventually fill up the whole energy surface with a density proportional to the invariant measure, provided that a certain condition of metrical transitivity is satisfied.[32] In statistical mechanics, these results are proved by means of the Birkhoff ergodic theorem.[33]

Because of these similarities we expect that, under certain conditions to be specified later, the point $X(\omega;t)$ will eventually fill up the whole "energy surface" $\|X\| = 1$ with a uniform density. We do not want, here, to go into the mathematical details of the ergodic theory of the stochastic equations. We shall just state the main conditions to be satisfied by $M(\omega;t)$, and the results.

TABLE II. Range of validity of various approximations for the mean Green's function.

Approximation and equation number	Condition on Kubo number $K = \sigma T_{corr}$	Condition on t	Further condition	Remarks (See Sec. 4E)
Born (2.7)		$\sigma t \ll 1$		
Mean Born (2.8)	$K<1$ $K>1$	$K\sigma t \ll 1$ $\sigma t \ll 1$		
Bourret (2.14)	$K \ll 1$			(1)
White noise (2.30)		$t \gg T_{corr}$	$\|M_0\| T_{corr} \ll 1$	
Hashminskii (2.32)			$\|M_0\| T_{corr} \gg K^2$	
Static (2.41)		$t \ll T_{corr}$		(2)
KAP (3.10) and KP (3.20)				(3)

We require that

(i) $M(\omega;t)$ is a stationary ergodic (matrix-valued) random function of t; roughly, this means that ensemble averages of functionals of $M(\omega;t)$ are equal to time averages.

(ii) For any X_0 on the unit sphere S, and any set A of positive measure on the unit sphere, the probability that a solution of (4.26), starting from X_0, will never penetrate A is zero (stochastic metric transitivity).

It may then be shown that, for any function $f(X)$ which is measurable with respect to the uniform measure dm on the unit sphere,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t f[X(\omega;\tau)]d\tau \underset{\text{a.s.}}{=} \left(\int_S f[X]dm\right)/\left(\int_S dm\right) = \lim_{t\to\infty}\langle f[X(\omega;t)]\rangle. \tag{4.29}$$

As an illustration of this result, let us calculate the asymptotic values of the first and second moment of $X_i(\omega;t)$, assuming that dm is normalized.

For the first-order moment we obtain

$$\lim_{t\to\infty}\langle X_i(\omega;t)\rangle = \int_S X_i dm = 0 \tag{4.30}$$

from a symmetry argument.

For the second-order moments we obtain

$$\lim_{t\to\infty}\langle X_i^2(\omega;t)\rangle = \int_S X_i^2 dm = \frac{1}{n}\int_S (X_1^2+\cdots+X_n^2)dm = \frac{1}{n}\int_S dm = \frac{1}{n}, \tag{4.31}$$

and

$$\lim_{t\to\infty}\langle X_i(\omega;t)X_j(\omega;t)\rangle = 0 \quad \text{if} \quad j \neq i. \tag{4.32}$$

It is interesting to notice that the asymptotic distribution of the "energy" $\langle X_i^2\rangle$ is simply equipartition. This has interesting applications to the energy transfer between randomly coupled oscillators.[34]

Remark: The above ideas may sometimes be extended to nonlinear stochastic equations if there is an invariant measure. For an application to the stochastic Ricatti equation

$$\frac{dz}{dx} + z^2 + n^2(\omega;x) = 0. \tag{4.33}$$

which is encountered in the theory of wave propagation in a one-dimensional random medium, the reader is referred to Ref. 35.

E. A guide for the user

We give now a few practical indications for the user who wants to calculate the mean Green's function of a linear stochastic equation. Calculation of other statistical quantities are usually reducible to the former as we have seen in Sec. 4.

First check if the equation falls into one of the classes of exactly soluble equations: white noise, shot noise, KAP, KP, Markovian coefficients, etc. If it does, the equation is usually soluble in closed analytic form except for the case of Markovian coefficients where the solution of a Fokker–Planck equation is required. If not, some approximation procedure must be used. Then, separate the stochastic evolution operator into its mean part M_0 and its fluctuating part $M_1(t)$. Estimate the correlation time T_{corr} and the dispersion of $M_1(t)$ and the norm $\|M_0\|$. Then, evaluate the dimensionless Kubo number $K = \sigma T_{corr}$. Recall that K is a measure of the effect of the stochastic perturbation over one

correlation time. According to the values of K Table II indicates the optimal method(s) for each case. A number of remarks on this table are in order.

(1) If the stochastic perturbation is a dichotomic Markov process [c f. Eq. (2. 27)] the Bourret equation (2. 14) is exact. Its solution by Laplace transformation usually requires only a little algebra. Notice that the Bourret equation may also be considered as a first semiquantitative approach to any stochastic equation by suitably adjusting the dispersion and correlation time of a dichotomic Markov process.

(2) The static approximation describes the full relaxation of the mean Green's function only if $K \gg 1$.

(3) The KAP and KP approximations should be used when a wide range of Kubo numbers is involved (including $K \approx 1$). If a good accuracy is wanted (e.g., in line broadening problems) and if the covariance is a suitable candidate for the inversion problem, use the KP method. If not, in particular, if the covariance of the random coefficients is not too well known, use the KAP method with

$$\nu = T_{\mathrm{corr}}^{-1} .$$

To conclude, we stress that the KAP and KP methods are probably the most flexible tools presently available when it is required to solve a linear stochastic differential equation over a large range of values of the Kubo number. In contrast to most approximate methods in mathematical physics, they do not rely on the existence of a small expansion parameter.

Applications of the methods described in this paper to problems of physical interest are discussed in other papers. Among the possible fields of application, let us mention: Stark broadening,[25,28] line formation in turbulent stellar atmospheres,[26] stability and Brownian motion of linear and nonlinear dynamical systems with random parameters,[12] energy transfer between randomly coupled dynamical systems,[7,34] and propagation of waves in a one-dimensional random medium.[2,7,30,31,35-38]

ACKNOWLEDGMENTS

We are indebted to R. C. Bourret and P-L Sulem for very helpful comments. We wish to thank the "Société des amis de l'ENSAE" who were kind enough to cover publication charges.

[1]R. H. Kraichnan, Proc. Symp. Appl. Math. 13, 199 (1962).
[2]U. Frisch, "Wave Propagation in Random Media," in *Probabilistic Methods in Applied Mathematics,* edited by A. T. Bharucha Reid (Academic, New York, 1968).
[3]U. Frisch and R. Bourret, J. Math. Phys. **11**, 364 (1970).
[4]Vo-Khac Khoan, Comptes-Rendus **265**, 63 (1967).
[5]P. Krée, Comptes-Rendus **272**, 330 (1971).
[6]A. T. Bharucha-Reid, *Random Operator Equations* (Academic, New York, 1972).
[7]G. C. Papanicolaou, J. Math. Phys. **13**, 1912 (1972).
[8]J. A. Morrison and J. McKenna, "Analysis of some stochastic ordinary differential equations," Bell Labs., Murray Hill, N. J. (1972).
[9]R. Bourret, Nuovo Cimento **26**, 1 (1962).
[10]J. B. Keller, Proc. Symp. Appl. Math. **16**, 145 (1964).
[11]R. Kubo, J. Math. Phys. **4**, 174 (1963).
[12]R. Bourret, U. Frisch, and A. Pouquet, Physica **65**, 303 (1973).
[13]M. A. Leibowitz, J. Math. Phys. **4**, 852 (1963).
[14]K. O. Friedrichs, "Perturbation of spectra in Hilbert space," American Mathematical Society, Providence, R. I., 1965.
[15]G. C. Papanicolaou and R. Hersh, Indiana U. Math. J. **21**, 815 (1972).
[16]R. Z. Hashminskii, Theor. Prob. Appl. **11**, 390 (1966).
[17]G. C. Papanicolaou and S. R. S. Varadhan, "A limit theorem with strong mixing in Banach space and two applications to stochastic differential equations," Courant Institute, N. Y. U. (1973).
[18]R. H. Kraichnan, J. Math. Phys. **2**, 124 (1961).
[19]J. R, Herring and R. H. Kraichnan, in *Statistical models and turbulence,* Lecture Notes in Physics **12**, 148 (Springer, New York, 1972).
[20]M. Lesieur, U. Frisch, and A. Brissaud, Ann. Geophys. (Paris) **27**, 151 (1971).
[21]U. Frisch, M. Lesieur, and A. Brissaud, "A Markovian Random Coupling Model for Turbulence," to appear in J. Fluid Mech. (1974).
[22]P. W. Anderson, J. Phys. Soc. Japan **9**, 316 (1954)
[23]R. Kubo, J. Phys. Soc. Japan **9**, 935 (1954).
[24]R. Bourret, Can. J. Phys. **44**, 2519 (1966).
[25]A. Brissaud and U. Frisch, J. Quant. Spectrosc. Radiat. Transfer **11**, 1767 (1971).
[26]M. Auvergne, H. Frisch, U. Frisch, Ch. Froeschlé, and A. Pouquet, Astron. and Astrophys. **29**, 93 (1973).
[27]A. T. Bharucha-Reid, *Elements of the Theory of Markov Processes and their Applications* (Mc Graw-Hill, New York, 1960).
[28]U. Frisch and A. Brissaud, J. Quant. Spectrosc. Radiat. Transfer **11**, 1753 (1971).
[29]M. Blume, Phys. Rev. **174**, 351 (1968).
[30]J. McKenna and J. A. Morrison, J. Math. Phys. **11**, 2348 (1970).
[31]J. A. Morrison and J. McKenna, J. Math. Phys. **11**, 2361 (1970).
[32]G. E. Uhlenbeck and G. W. Ford, *Lectures in Statistical Mechanics* (American Mathematical Society, Providence, R. I, 1963).
[33]P. R. Halmos, *Lectures on ergodic theory* (Chelsea, New York, 1956). A. Tortrat, *Calcul des Probabilités* (Masson, Paris 1971).
[34]M. Auvergne and A. Pouquet, Physica **66**, 403 (1973).
[35]P. L. Sulem and U. Frisch, J. Plasma Phys. **8**, 217 (1972).
[36]G. Papanicolaou and J. B. Keller, SIAM J. Appl. Math. **21**, 287 (1971).
[37]G. Papanicolaou, SIAM J. Appl. Math. **21**, 13 (1971).
[38]J. A. Morrison, G. C. Papanicolaou and J. B. Keller, Commun. Pure Appl. Math. **24**, 473 (1971).

Nonradiative algebraically special space-times

D. W. Trim

Department of Applied Mathematics, The University of Manitoba, Winnipeg, Manitoba, Canada

J. Wainwright

Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada

(Received 24 September 1973)

A recently characterized class of (in general nonvacuum) algebraically special space–times with twisting rays is studied. The Weyl tensor satisfies the peeling-off property along the repeated principal null congruence, and the Ricci tensor exhibits an equally simple asymptotic behavior, which is in fact compatible (via the Einstein field equations) with the presence of a suitably restricted electromagnetic or neutrino field. If the gravitational and source fields are nonradiative, the above asymptotic behavior is restricted. In this case we explicitly solve the Einstein vacuum field equations, the Einstein–Maxwell equations and the Einstein–Weyl (combined gravitational–neutrino) equations. The solutions obtained are related to the known algebraically special solutions of these equations.

1. INTRODUCTION

Motivated by the 1962 paper of Robinson and Trautman,[1] the problem of finding exact solutions of the vacuum gravitational field equations in general relativity which admit a geodesic, shearfree and expanding null congruence[2] (and hence are algebraically special[2]) has been extensively studied in the past decade,[3-16] with emphasis on the case in which the congruence is twisting. The general procedure that has been used is as follows.[4,8,9,12] One chooses a coordinate system (u, r, x, y) such that r is an affine parameter along the geodesics of the given congruence, while u, x, y are constant along these curves. It then follows, by virtue of the assumed properties of the null congruence, that the subset of the vacuum field equations

$$R_{ab} = 0 \tag{1.1}$$

contained in the algebraic restriction

$$k_{[a}R_{b][c}k_{d]} = 0, \tag{1.2}$$

where k^a is tangent to the given congruence, completely determines[17] the dependence of the metric components on the affine parameter r. Because of this one can obtain a reasonably simple canonical form for the line element using (some of) the remaining freedom in the choice of the coordinate system. One of the vacuum field equations (1.1) not included in (1.2), that corresponding to

$$R = 0, \tag{1.3}$$

is identically satisfied[12] on account of (1.2). The remaining three field equations yield a set of partial differential equations to be satisfied by the remaining arbitrary functions of u, x, y which appear in the line element. In this way one obtains a considerable simplification of the problem.

The remaining field equations have not been solved in general, even in the case in which the given null congruence is in addition twist-free, considered by Robinson and Trautman.[1] However, by imposing simplifying assumptions on the line element one can obtain explicit solutions. In particular, this approach led to the discovery of a physically important generalization of the Schwarzschild solution, namely the Kerr solution,[4] in which the given null congruence has nonzero twist.

Both the Schwarzschild and Kerr space-times are asymptotically flat (see Sachs[18] or Penrose[19]), although this is not true of the general class of space-times under consideration. The general space-time does, however, display one feature of asymptotically flat vacuum space-times in that its Riemann tensor exhibits the "peeling off" property[20] along a generic geodesic of the given null congruence:

$$R_{abcd} = [4]_{abcd}r^{-1} + [31]_{abcd}r^{-2} + [211]_{abcd}r^{-3} + O(r^{-4}), \tag{1.4}$$

where r is an affine parameter along the given null congruence.[21] The coefficient tensors have the symmetries of the vacuum Riemann tensor, and admit the tangent field k^a as a repeated principal null direction, of multiplicity indicated by the kernel.[2] These tensors are also covariantly constant along the null congruence.

One of the aims of the present paper is to explicitly solve the remaining vacuum field equations referred to earlier, subject to the assumption that the first two terms in the expansion (1.4) are identically zero. This assumption corresponds physically to the absence of gravitational radiation. In order to justify this statement, however, it is necessary to relate the resulting class of exact vacuum solutions to the Newman-Unti[22] asymptotically flat approximate vacuum solutions, using the correspondence developed by Aronson and Newman.[23] We will not elaborate on this here.

The solution of the preceding problem is obtained as a special case of a more general situation. In a recent paper,[24] the present authors showed that the integration procedure based on (1.2) (leading to a canonical form for the line-element) could also be carried out subject to an algebraic assumption on the Ricci tensor weaker than (1.2) (see Condition II below). This weaker restriction is in fact compatible, via the gravitational field equations

$$G_{ab} = -T_{ab} \tag{1.5}$$

with the presence of a fairly wide class of electromagnetic or neutrino fields. [On the other hand, condition (1.2) is only compatible with a *null* electromagnetic field, or a *null* neutrino field.] In this paper then we study in some detail the significance and consequences of this weaker restriction.

More specifically, we begin by imposing (locally) the following conditions on space-time.

Condition I: There exists a null vector field k^a which satisfies

$$k_{a;\,b}k^b = 0, \qquad 2k_{(a;\,b)}k^{a;\,b} - (k^a{}_{;\,a})^2 = 0, \tag{1.6a}$$

$$k^a{}_{;\,a} \neq 0. \tag{1.6b}$$

Equations (1.6) assert that the integral curves of this vector field form a *geodesic, shearfree, but expanding null congruence.*[2]

Condition II: The Ricci tensor satisfies

$$k_{[a}R_{b][c}k_{d]} = f k_{[a}g_{b][c}k_{d]}, \tag{1.7a}$$

$$R = 0, \tag{1.7b}$$

for some function f, where k^a is the given vector field.

Note that Conditions I and II imply that *k^a is a repeated principal null direction of the Weyl tensor, which is thus algebraically special.* This is a straight-forward application of a result due to Kundt and Trumper.[25]

In Sec. 2, we describe the canonical form for the line-element of a space-time satisfying Conditions I and II, which reduces to the previously given vacuum canonical form when we impose $R_{ab} = 0$. It is shown that the asymptotic behavior (1.4) of the vacuum Riemann tensor (subject to Condition I) is preserved provided we consider the Weyl tensor:

$$C_{abcd} = [4]_{abcd}r^{-1} + [31]_{abcd}r^{-2} + [211]_{abcd}r^{-3} + O(r^{-4}), \tag{1.8}$$

where the symbols have the same meaning as in Eq. (1.4). In addition, the Ricci tensor has a simple expansion in inverse powers of r:

$$R_{ab} = R^{(1)}_{ab}r^{-2} + R^{(2)}_{ab}r^{-3} + R^{(3)}_{ab}r^{-4} + O(r^{-5}). \tag{1.9}$$

The coefficients $R^{(\alpha)}_{ab}$, $\alpha = 1, 2, 3$ are tracefree symmetric tensors which are covariantly constant along the given null congruence, and are related algebraically to k^a by the following conditions[26]:

$$R^{(1)}{}_{a[b}k_{c]} = 0, \tag{1.10a}$$

$$k_{[a}R^{(2)}{}_{b][c}k_{d]} = 0, \tag{1.10b}$$

$$k_{[a}R^{(3)}{}_{b][c}k_{d]} = f k_{[a}g_{b][c}k_{d]}. \tag{1.10c}$$

In Sec. 3 we impose the following restrictions on the asymptotic behavior (1.8), (1.9) of the Weyl and Ricci tensors.

Condition III:

$$C_{abcd} = [211]_{abcd}r^{-3} + O(r^{-4}), \tag{1.11}$$

with $[211]_{abcd} \neq 0$.

Condition IV:

$$R_{ab} = R^{(3)}_{ab}r^{-4} + O(r^{-5}). \tag{1.12}$$

We regard Condition III as excluding the possibility of gravitational radiation, by analogy with the vacuum case. Furthermore, if the Ricci tensor is determined, via (1.5), by an electromagnetic field, then Condition IV excludes the possibility of electromagnetic radiation, as will be seen in Sec. 5.

It is shown (in Sec. 3) that subject to Conditions III and IV, the canonical line element of Sec. 2 can be further simplified so that it is determined explicitly up to a number of completely arbitrary functions of at most two variables. In addition, *it is found that the non-vanishing of the Ricci tensor affects the line element only via the appearance, in a purely algebraic fashion,*[27] *of a single real function of two real variables.* The general solution of the vacuum gravitational field equations subject to Conditions I and III is thus obtained from the line element of Sec. 3 by simply setting this function to zero, thereby solving the problem which was mentioned in the paragraph following Eq. (1.4). We in fact find that Condition III, together with the various additional restrictions to be introduced in Sec. 4, enable us to characterize almost all the explicitly known twisting solutions of the vacuum field equations subject to Condition I. Sections 5 and 6 deal respectively with the specialization of the results of Sec. 2, 3, and 4, to obtain and characterize a class of exact solutions of the Einstein-Maxwell and Einstein-Weyl[24] (combined gravitational-neutrino) field equations.

As regards techniques for dealing with the type of problem under consideration here, the spin coefficient formalism of Newman and Penrose[31] is particularly convenient. In the remainder of this paper familiarity with this formalism on the part of the reader is assumed. Our conventions for the Riemann, Ricci and Weyl tensors are those of Newman and Penrose.

2. SPACE-TIMES SATISFYING CONDITIONS I AND II: LINE ELEMENT AND ASYMPTOTIC BEHAVIOR

In this section we give the canonical form for the line element of a space-time satisfying Conditions I and II, and describe the asymptotic behavior of the associated Weyl and Ricci tensors. We use coordinates $(u, r, x, y) \equiv (x^1, x^2, x^3, x^4)$, where u, x, y are constant along the geodesics of the given null congruence, and r is an affine parameter for the congruence, chosen so that the tangent field k^a in (1.6) has the form

$$k^a = \delta^a_2. \tag{2.1}$$

Theorem 2.1: A space-time satisfies Conditions I and II of Sec. 1 if and only if the line element can be written locally in the form

$$ds^2 = -(2G\overline{G})^{-1}dzd\bar{z} + 2(k_a dx^a)[dr - \mathrm{Re}(\overline{W}dz) - U(k_b dx^b)], \tag{2.2a}$$

where

$$k_a dx^a = du - \mathrm{Re}(\overline{Q}dz), \qquad dz \equiv dx + i\,dy, \tag{2.2b}$$

$$G = -P(r - i\Sigma)^{-1}, \tag{2.2c}$$

$$W = -rD_1Q - i(D_3\Sigma + \Sigma D_1Q), \tag{2.2d}$$

$$U = U^0 + rD_1 \ln P + (mr + M\Sigma - \phi^0_{11})(r^2 + \Sigma^2)^{-1}. \tag{2.2e}$$

The complex function Q, and the real functions m, P, ϕ^0_{11} are arbitrary functions of $u, z, \bar{z}$, while the real functions Σ, U^0, and M are defined in terms of P, Q by

$$2i\Sigma = P^2(D_3\overline{Q} - D_4Q), \tag{2.3a}$$

$$U^0 = -\mathrm{Re}[P^2D_3(D_1\overline{Q} + D_4 \ln P)], \tag{2.3b}$$

$$M = -2\Sigma U^0 + P^2\,\mathrm{Re}[D_3(D_4\Sigma + \Sigma D_1\overline{Q}) + D_1Q(D_4\Sigma + \Sigma D_1\overline{Q})]. \tag{2.3c}$$

The differential operators D_1, D_3, D_4 are defined by

$$D_1 \equiv \partial_u, \quad D_3 \equiv 2\partial_{\bar{z}} + Q\partial_u \equiv \partial_x + i\partial_y + Q\partial_u, \quad D_4 \equiv \bar{D}_3. \tag{2.4}$$

An outline of the proof of this theorem, using the Newman-Penrose formalism, was given in Trim and Wainwright.[24] A more detailed treatment can be found in the thesis of Trim.[32]

A remarkable feature of the line element (2.2) is the fact that its dependence on the affine parameter r is in terms of rational functions of r. This permits r to assume arbitrarily large values, so that *a generic null geodesic of the given congruence can be extended to infinite values of its affine parameter*. This is borne out by the dependence on r of the expansion θ and twist ω of the congruence[2]:

$$\theta = r(r^2 + \Sigma^2)^{-1}, \quad \omega = -\Sigma(r^2 + \Sigma^2)^{-1}, \tag{2.5}$$

[see Eq. (A 4a), noting that θ, ω are related to the spin coefficient ρ according to $\rho = -(\theta + i\omega)$].

In the derivation of the line element (2.2), the vector fields $k^a, n^a, m^a, \bar{m}^a$ of the associated null tetrad[31] were simplified to the following form:

$$k^a = \delta^a_2, \tag{2.6a}$$

$$n^a = \delta^a_1 + U\delta^a_2, \tag{2.6b}$$

$$m^a = G(Q\delta^a_1 + W\delta^a_2 + \delta^a_3 + i\delta^a_4). \tag{2.6c}$$

This tetrad is adapted to the given null congruence in the sense that the tetrad vector field k^a is tangent to the congruence [see Eq. (2.1)]. In addition the tetrad vectors are parallel transferred along the congruence [see Eq. (A 2)]. One can thus study the variation of the Weyl and Ricci tensors along the given congruence by considering the dependence of their components, taken with respect to this null tetrad, on the affine parameter r. The detailed expressions for these components, and for the spin coefficients are given in Appendix A for completeness. In this paper we are mainly interested in the *asymptotic behavior* (i.e., as the affine parameter $r \to \infty$) of these quantities, and the form of their leading coefficients.

Firstly, for the nonzero components of the Weyl tensor we obtain from (A10), on expanding $\rho = -(r + i\Sigma)^{-1}$ in inverse powers of r:

$$\Psi_2 = \Psi_2^0 r^{-3} + O(r^{-4}), \tag{2.7a}$$

$$\Psi_3 = \Psi_3^0 r^{-2} + O(r^{-3}), \tag{2.7b}$$

$$\Psi_1 = \Psi_4^0 r^{-1} + [(D_4 + 4D_1\bar{Q})(P\Psi_3^0) - i\Sigma\Psi_4^0]r^{-2} + O(r^{-3}). \tag{2.7c}$$

For our purposes it is convenient to write the leading coefficients in the form

$$\Psi_2^0 = -(m + iM), \tag{2.8a}$$

$$\Psi_3^0 = P^3 D_3 J, \tag{2.8b}$$

$$\Psi_4^0 = -P^2 D_1 J, \tag{2.8c}$$

where[33]

$$J = D_4 L + L^2, \tag{2.9a}$$

$$L = D_1\bar{Q} + D_4 \ln P. \tag{2.9b}$$

The expressions (2.8b, c), which are the same as in the vacuum case,[34] are obtained from Eqs. (A5a), (A11b, c) using the commutator relations for the operators D_1, D_3, D_4:

$$D_{13} - D_{31} = (D_1 Q)D_1, \tag{2.10a}$$

$$D_{14} - D_{41} = (D_1\bar{Q})D_1, \tag{2.10b}$$

$$D_{34} - D_{43} = (D_3\bar{Q} - D_4 Q)D_1, \tag{2.10c}$$

where $D_{13} \equiv D_1 D_3$, etc. Equations (2.7), together with the fact that $\Psi_0 = \Psi_1 = 0$, imply that the Weyl tensor exhibits the "peeling off" property (1.8) along the given null congruence.

Secondly, for the nonzero tetrad components of the Ricci tensor, we obtain, from Eqs. (A7) and (A8)

$$\Phi_{11} = \Phi_{11}^0 r^{-4} + O(r^{-5}), \tag{2.11a}$$

$$\Phi_{12} = \Phi_{12}^0 r^{-3} + O(r^{-4}) \equiv \bar{\Phi}_{21}, \tag{2.11b}$$

$$\Phi_{22} = \Phi_{22}^0 r^{-2} + \{2P^2 \mathrm{Re}[(D_4 + 3D_1\bar{Q})(P^{-1}\Phi_{12}^0)] + P^4 D_1(P^{-4}\Phi_{11}^0)\} r^{-3} + O(r^{-4}), \tag{2.11c}$$

where Φ_{11}^0 is an arbitrary real function of $u, z, \bar{z}$, and[35]

$$\Phi_{12}^0 = \tfrac{1}{2} P(D_3 + 3D_1 Q)\Psi_2^0, \tag{2.12a}$$

$$\Phi_{22}^0 = P^2(D_3 + 2D_1 Q)(P^{-1}\Psi_3^0) + P^3 D_1(P^{-3}\Psi_2^0). \tag{2.12b}$$

The expansion (1.9) and the relations (1.10) follow from (2.11) on making use of the expansion (A13) for R_{ab}.

In connection with the Ricci tensor it is important to note that

$$R_{ab} = 0 \iff \Phi_{11}^0 = \Phi_{12}^0 = \Phi_{22}^0 = 0. \tag{2.13}$$

We thus regard the functions $\Phi_{11}^0, \Phi_{12}^0, \Phi_{22}^0$ as representing the freedom available in specifying the energy tensor T_{ab} in the gravitational field equations (1.5). Thus when T_{ab} has been specified (in a way consistent with Condition II), Eqs. (2.12a, b) represent restrictions on the functions P, Q and $\Psi_2^0 \equiv -(m + iM)$. However, when we recall that M is expressed in terms of P and Q by means of Eqs. (2.3), we see that *(2. 12a, b) together with (2. 3c) form a system of partial differential equations in $u, z, \bar{z}$ to be solved for P, Q and m.* We will refer to this set of equations as *the reduced gravitational field equations*. For convenience we rewrite them in a slightly different form below:

$$P(D_3 + 3D_1 Q)(m + iM) = -2\Phi_{12}^0, \tag{2.14a}$$

$$P^2(D_3 + 2D_1 Q)(P^{-1}\Psi_3^0) - P^3 D_1[P^{-3}(m + iM)] = \Phi_{22}^0, \tag{2.14b}$$

$$M = \tfrac{1}{2} iP^3(D_{4433}V - D_{3344}V), \tag{2.14c}$$

where the real function V is defined by[36]

$$D_1 V = P, \tag{2.15}$$

up to a real additive function of z and $\bar{z}$.

In the case of *vacuum* gravitational fields [$\Phi_{11}^0 = \Phi_{12}^0 = \Phi_{22}^0 = 0$] the canonical form (2.2) of the line element and the reduced field equations (2.14) were first derived by Kerr[4] using a null tetrad approach. They have been subsequently derived in more detail by Debney, Kerr, and Schild[10] (using differential forms), by Robinson, Robinson, and Zund[8], and, using the Newman-Penrose formalism, by Talbot.[9] [Our Eqs. (2.14a, b), with $\Phi_{12}^0 = \Phi_{22}^0 = 0$ are identical with Talbot's equations (5.15) and (5.17), and our Eq. (2.14c), in the equivalent form (2.3c)

is identical with his equation (5.16), on noting that our Ψ_2^0 is the negative of Talbot's.]

To conclude this section, we state the general coordinate transformation which preserves the form of the line element (2.2):

$$u' = h(u, z, \bar{z}), \quad r' = R^{-1}r, \quad z' = f(z) \tag{2.16a}$$

with

$$R(u, z, \bar{z}) = \frac{\partial h}{\partial u}. \tag{2.16b}$$

This is in fact the same as in the vacuum case.[9] In order that the form (2.6) of the null tetrad be preserved under (2.16), it is necessary to perform simultaneously with (2.16) the following tetrad transformation:

$$\tilde{k}^a = Rk^a, \quad \tilde{n}^a = R^{-1}n^a, \quad \tilde{m}^a = e^{iS}m^a, \tag{2.17a}$$

where R is given by (2.16b) and the real function S is related to $f(z)$ in (2.16a) according to

$$e^{2iS} = \frac{f'(z)}{\bar{f}'(\bar{z})}. \tag{2.17b}$$

In order to use this freedom to simplify the solutions of the reduced field equations, it is necessary to know the transformation laws of various expressions under (2.16)-(2.17). These are listed in Appendix B.

3. SPACE-TIMES SATISFYING CONDITIONS I, II, III AND IV AND ASSOCIATED VACUUM SOLUTIONS

In this section, we impose Conditions III and IV on the line element of Sec. 2. The reduced gravitational field equations (2.14) can then be solved to yield an explicit form for the line element, containing three (complex) analytic functions of z (and their complex conjugates) and one real function of z and $\bar{z}$. Condition IV in fact makes it unnecessary to specify the exact nature of the sources corresponding to T_{ab} in Eq. (1.5) (though it does of course further restrict the nature of the sources). Thus *the general line element to be derived in this section is characterized by the purely geometric restrictions contained in Conditions I-IV.*

Firstly, the expansions (2.7) imply that Condition III [see equations (1.11) and (A12)] is equivalent to

$$\Psi_3^0 = \Psi_4^0 = 0, \quad \Psi_2^0 \neq 0. \tag{3.1}$$

On account of Eqs. (2.8b, c) and (2.4), this restricts the function J, defined by (2.9a), to be of the form $J = J(z)$. Thus the coordinate freedom (2.16a) [See Eq. (B8)] may be used to set

$$J \equiv D_4 L + L^2 = 0. \tag{3.2}$$

This condition is preserved if and only if the transformation (2.16a) is restricted by

$$f(z) = (Az + B)/(Cz + D), \quad AD - BC \neq 0, \tag{3.3}$$

where A, B, C, D are complex constants.

The next step is to transform

$$D_1 P = 0, \tag{3.4}$$

using the freedom in the function $R(u, z, \bar{z})$ in Eqs. (2.16) [see Eq. (B1)]. This condition is preserved if and only if R is restricted by

$$R = R(z, \bar{z}) \Rightarrow h(u, z, \bar{z}) = Ru + H(z, \bar{z}). \tag{3.5}$$

It is now necessary to make use of Condition IV. On account of the expansions (2.11) and Eq. (3.4), this condition is equivalent to

$$\Phi_{12}^0 = \Phi_{22}^0 = 0, \tag{3.6a}$$

$$\Phi_{11}^0 = \Phi_{11}^0(z, \bar{z}). \tag{3.6b}$$

Equations (3.1), (3.4), and (3.6) simplify the reduced gravitational field equations (2.14a, b) to the form

$$(D_3 + 3D_1 Q)(m + iM) = 0, \tag{3.7a}$$

$$D_1(m + iM) = 0. \tag{3.7b}$$

Using the commutators (2.10), we find that these equations entail

$$D_{11} Q = 0, \tag{3.8}$$

since $m + iM \neq 0$, by assumption.

We are now in a position where we can follow the vacuum integration procedure of Robinson and Robinson.[12] In fact, Eq. (3.8) implies, on account of (2.9) and (3.4), that $D_1 L = 0$, so that Eq. (3.2) can be integrated to give

$$L \overset{\vee}{=} 2k\bar{z}[kz\bar{z} + \bar{l}(\bar{z})]^{-1}, \tag{3.9}$$

where $\bar{l}(\bar{z})$ is the constant of integration.[37] The real constant k, if nonzero, can be set equal to 1 by redefining $\bar{l}$. Without loss of generality we may assume $\bar{l} \neq 0$. [If $\bar{l} = 0$, we simply perform a transformation (2.16), with f given by (3.3), which reintroduces $\bar{l}$.] Equation (2.9b) gives an expression for $D_1 Q$, which permits us to write Q in the form

$$Q(u, z, \bar{z}) = (\bar{L} - D_3 \ln P)u + P^{-1}q(z, \bar{z}), \tag{3.10}$$

where the factor P^{-1} in the second term has been introduced for later convenience.

Equations (3.7) can now be integrated, once $D_1 Q$ is eliminated using Eq. (3.10). By suitably choosing the "constant" of integration $n(z)$, one obtains

$$m + iM = 2^{3/2}n(z)P^3 l^3(kz\bar{z} + l)^{-3}. \tag{3.11}$$

There remains to be integrated Eq. (2.14c), which is a partial differential equation for the function $q(z, \bar{z})$ in (3.10), although this is not immediately obvious. On account of (3.4), the function V, as defined by (2.15), has the form

$$V = Pu + v(z, \bar{z}). \tag{3.12}$$

A straightforward calculation using Eqs. (2.4), (2.9b), (2.10), (3.2) and (3.10) implies that

$$D_{33} V = 2\bar{L}\partial_{\bar{z}}(q\bar{L}^{-1}) + 4\partial_{\bar{z}}\partial_{\bar{z}} v.$$

Thus using (3.11) the reduced field equation (2.14c) can be written, when $k \neq 0$, in the form

$$\bar{n}\bar{l}^3(kz\bar{z} + \bar{l})^{-3} - nl^3(kz\bar{z} + l)^{-3} = 2^{3/2}\{\partial_z\partial_z[\bar{L}\partial_{\bar{z}}(q\bar{L}^{-1})] - \partial_{\bar{z}}\partial_{\bar{z}}[L\partial_z(\bar{q}L^{-1})]\}. \tag{3.13}$$

This is a *linear* nonhomogeneous partial differential equation for $q(z, \bar{z})$. Its general solution is thus of the form

$$q(z, \bar{z}) = q_0 + q_1, \tag{3.14a}$$

where q_0 is the general solution of the corresponding homogeneous equation and q_1 is a particular solution of the nonhomogeneous equation. It can be directly verified using (3.9) that a possible choice for q_1 is

$$q_1 = 2^{-1/2}\bar{L}\int(2k\bar{z})^{-3}\bar{n}\bar{l}^3 L\bar{L}^{-1}d\bar{z}. \tag{3.14b}$$

To obtain the form of q_0, note that (3.13), with $n = 0$, implies the existence, at least locally, of a real function $K(z, \bar{z})$ such that

$$\bar{L}\partial_{\bar{z}}(q_0\bar{L}^{-1}) = 2\partial_{\bar{z}}\partial_z K.$$

The solution of this equation can be written in the form

$$q_0(z,\bar{z}) = 2^{1/2}q(z)l(kz\bar{z} + l)^{-1} + 2\partial_z K - K\bar{L}, \tag{3.14c}$$

provided that the "constant" of integration $q(z)$ is suitably chosen. Note that the second and third terms are obtained by integration by parts using the fact that $\bar{L}$ satisfies $2\partial_{\bar{z}}(\bar{L}^{-1}) = 1$ [see (3.2)]. The function Q, in the case $k \neq 0$, is thus given by Eqs. (3.10) and (3.14). Finally, by performing a coordinate transformation of the form (2.16a), with $h(u, z, \bar{z}) = u - KP^{-1}$, $f(z) = z$, it follows, using the transformation laws (B1), (B2), that we can set $K = 0$ in (3.14c).

If $k = 0$, so that $L = 0$ [see (3.9)], the only change that is in fact necessary is in the form of the particular solution (3.14b). A suitable choice is given in the summary below.

We have thus shown that the function $Q(u, z, \bar{z})$ can be written in the form

$$Q = (\bar{L} - D_3 \ln P)u + P^{-1}[2^{1/2}q(z)l(kz\bar{z} + l)^{-1} + q_1(z,\bar{z})], \tag{3.15a}$$

where

$$q_1 = \begin{cases} 2^{-1/2}\bar{L}\int(2k\bar{z})^{-3}\bar{n}\bar{l}^3 L\bar{L}^{-1}d\bar{z}, & k \neq 0 \\ 2^{-5/2}z^2\int\bar{n}\,d\bar{z}, & k = 0, \end{cases} \tag{3.15b}$$

and L is given by Eq. (3.9).

This completes the integration procedure. The function $P(z, \bar{z})$ is arbitrary, but no in an essential way, as we could for example transform $P = 1$ locally, using the function R in the coordinate freedom (2.16) [see (3.5) and (B1)]. On the other hand, *the functions* $l(z)$, $n(z)$, $q(z)$ *and* $\Phi^0_{11}(z, \bar{z})$ *are essentially arbitrary.*

We summarize the results in the following:

Theorem 3.1: A space-time satisfies Conditions I, II, III, and IV of Sec. 1 if and only if its line element can be written locally in the form (2.2), (2.3) with Q, m, M given by Eqs. (3.9), (3.11), and (3.15).

Note that the nonvanishing of the Ricci tensor affects the line element only *algebraically*, in the following sense. The metric tensor g_{ab} of Theorem 3.1 can be written in the form

$$g_{ab} = g^{(v)}_{ab} + 2\Phi^0_{11}\rho\bar{\rho}k_a k_b, \tag{3.16}$$

where $g^{(v)}_{ab}$ is the associated vacuum metric (obtained by setting $\Phi^0_{11} = 0$), k_a is given by (2.2b) and ρ by (A4a). We can thus state

Corollary 1: A *vacuum* space-time satisfies Conditions I and III of Sec. 1 if and only if its line element can be written locally in the form (2.2), (2.3) with $\Phi^0_{11} = 0$, and Q, m, M given by Eqs. (3.9), (3.11), and (3.15).

This class of explicit vacuum solutions is in fact precisely the class found by Robinson and Robinson.[12] To verify this we note that in vacuum our Condition III is equivalent to the assumptions (in addition to Condition I) imposed by Robinson and Robinson, which in our notation read

$$D_1 L = D_3 J = 0. \tag{3.17}$$

[In vacuum (3.1) => (3.8) => (3.17), using the definitions (2.8b, c). The converse is obvious.]

To end this section we note that the presence of the arbitrary functions $l(z), n(z)$, and $q(z)$ in the line element is directly related to the presence of twist in the given null congruence. In fact, we have

Theorem 3.2: A space-time satisfies Conditions I, II, III, and IV of Sec. 1, and the given null congruence has zero twist, if an only if its line element can be written locally in the form

$$ds^2 = -r^2(1 - \tfrac{1}{2}U^0 z\bar{z})^{-2}dzd\bar{z} + 2dudr - 2(U^0 + mr^{-1} - \Phi^0_{11}r^{-2})du^2,$$

where U^0 assumes the values $\pm\frac{1}{2}$, or 0, m is a constant and Φ^0_{11} is an arbitrary function of z and $\bar{z}$.

Outline of the Proof: The vanishing of the twist [=> $\Sigma = 0$, see (2.5)] implies that one is in Case 1 of the following section, and that in addition one can transform Q to be zero. The result then follows from Eqs. (2.2), (2.3), and (4.12).

The vacuum subclass consists of the *DS*-spaces of Robinson and Trautman,[1] and includes the Schwarzschild solution when $U^0 = -\frac{1}{2}$. The usual angular coordinates are defined by Eq. (4.19a).

4. SPECIALIZATIONS

The line element of Sec. 3 can be specialized, firstly by eliminating the u dependence, and secondly by requiring that some of the higher order coefficients in the expressions (A7) and (A10) for the Ricci and Weyl tensors vanish. We choose the additional assumptions in such a way that the arbitrary functions $l(z), n(z), \Phi^0_{11}(z, \bar{z})$, and $q(z)$ are successively determined, and in this way, we arrive at the well-known Demianski-Newman,[7] Kerr,[4] and NUT[3] solutions.

Case 1: $D_1\Sigma = 0.$ (4.1)

This condition is invariant under the coordinate and tetrad freedom, by virtue of (3.5) [see Eq. (B3)]. On account of (2.3a) and the form (B10) of the transformation law for $D_1 Q$, we may set

$$D_1 Q = 0. \tag{4.2}$$

This condition is preserved provided that

$$R = \text{const.} \tag{4.3}$$

The metric is now completely independent of u and thus admits $\xi^a = \delta^a_1$ as a Killing vector.

Additional simplifications follow from (4.2). Equations (2.9b) and (3.2) imply that $\partial_z\partial_z P = 0$, so that P, being real, is of the form

$$P = az\bar{z} + bz + \bar{b}\bar{z} + c,$$

where a, c are real constants and b is a complex constant. This implies via (2.3b) that $U^0 = -4(ac - b\bar{b}) =$ const. Using the freedom (2.16) in the choice of the z coordinate, with $f(z)$ given by (3.3), we can write P in the form

$$P = 2^{-1/2}(1 - \tfrac{1}{2}U^0 z\bar{z}). \tag{4.4a}$$

By virtue of (B4) we can choose R [now restricted to be constant by (4.3)] so that if $U^0 \neq 0$, it has the value $-\frac{1}{2}$ or $+\frac{1}{2}$:

$$U^0 = -\tfrac{1}{2}, +\tfrac{1}{2}, 0. \tag{4.4b}$$

Using Eqs. (2.9b), (4.2), and (4.4), one finds that

$$L = -2^{-1/2}U^0\bar{z}P^{-1}, \tag{4.5}$$

which, on comparison with (3.9), yields

$$2k + U^0\bar{l}(\bar{z}) = 0. \tag{4.6}$$

Thus if $U^0 \neq 0$, $l(z)$ is a *real* constant, while $U^0 = 0$ implies $k = 0$, so that $L = 0$ and $l(z)$ does not appear in the solution [see (3.11) and (3.15)]. In both cases (4.5) is valid. The expression (3.11) then simplifies to

$$m + iM = n(z). \tag{4.7}$$

The expressions (3.15) for $Q(z, \bar{z})$ likewise simplify, and using Eqs. (4.4), (4.5), and (4.6), we obtain

$$Q(z, \bar{z}) = P^{-2}q(z) + Q_1, \tag{4.8a}$$

where

$$Q_1 = \begin{cases} \tfrac{1}{2}(U^0P)^{-2}\int \bar{z}^{-2}\bar{n}d\bar{z}, & U^0 \neq 0, \\ \tfrac{1}{4}z^2\int \bar{n}d\bar{z}, & U^0 = 0. \end{cases} \tag{4.8b}$$

The function $\Phi^0_{11}(z, \bar{z})$ is unaffected by the specialization (4.1). *The line element is thus determined by Eqs. (2.2), (2.3), (4.4), (4.7), and (4.8), with $\Phi^0_{11}(z, \bar{z})$ arbitrary.* It depends on three arbitrary functions $n(z), q(z), \Phi^0_{11}(z, \bar{z})$, together with the indicator U^0 [see Eq. (4.4)].

The associated class of vacuum solutions, (obtained by setting $\Phi^0_{11} = 0$) is in fact that obtained by Robinson, Robinson, and Zund[8] who, prior to the paper of Robinson and Robinson,[12] solved the reduced vacuum field equations subject to additional conditions, which in our notation read

$$D_{11}Q = 0, \quad D_1\Sigma = 0, \quad U^0 = \text{const}.$$

(These conditions can be shown to be equivalent in vacuo to our conditions $\Psi^0_3 = \Psi^0_4 = D_1\Sigma = 0$.)

Case 2: $Y^0_1 = 0$. (4.9)

Note that Y^0_1, as given by (A5b), is the coefficient of ρ^3 in the expression (A10b) for Ψ_3. By virtue of Eqs. (A10b, c) and (A11e), the assumption (4.9) implies that $\Psi_3 = O(r^{-4})$, $\Psi_4 = O(r^{-4})$, while as before $\Psi_2 = O(r^{-3})$. Thus the asymptotic behavior (1.11) assumed in Condition III specializes to

$$C_{abcd} = [22]_{abcd}r^{-3} + O(r^{-4}), \tag{4.10}$$

with n^a being the second repeated principal null direction of $[22]_{abcd}$.

Condition (4.9) also implies (4.1) [evaluate the commutator (2.10c) applied to $\Psi^0_2 \equiv -(m + iM)$ using (3.7a, b), (4.9), (A5b), (A11a), and (2.3a)]. Additional simplifications arise since (4.9), with (4.2), (A11a), and A5b) implies that the function $n(z)$ in (4.7) must be a constant:

$$m + iM = n = \text{const}. \tag{4.11}$$

The function Q_1, as given by (4.8b), thus simplifies to

$$Q_1 = \begin{cases} -\tfrac{1}{2}(m - iM)(U^0P)^{-2}\bar{z}^{-1}, & U^0 \neq 0, \\ \tfrac{1}{4}(m - iM)z^2\bar{z}, & U^0 = 0. \end{cases}$$

In each case the part of Q_1 containing m is of the form $\partial H/\partial\bar{z}$, for a real function H, and can thus be eliminated by redefining the u-coordinate according to $u' = u - \frac{1}{2}H$. [We have this freedom again since now the metric is independent of u.] *In case 2 then, the line-element is given by (2.2), (2.3) with*

$$P = 2^{-1/2}(1 - \tfrac{1}{2}U^0 z\bar{z}), \quad U^0 = -\tfrac{1}{2}, \tfrac{1}{2}, 0, \tag{4.12a}$$

$$m + iM = \text{const}, \tag{4.12b}$$

$$Q = q(z)P^{-2} + Q_1, \tag{4.12c}$$

$$Q_1 = \begin{cases} \tfrac{1}{2}iM(U^0P)^{-2}\bar{z}^{-1}, & U^0 \neq 0, \\ -(i/4)Mz^2\bar{z}, & U^0 = 0. \end{cases} \tag{4.12d}$$

$$\Phi^0_{11} = \Phi^0_{11}(z, \bar{z}), \quad \text{arbitrary}. \tag{4.12e}$$

If we further set $M = 0$, so that the leading coefficient of Ψ_2 is real

$$\Psi^0_2 = \bar{\Psi}^0_2, \tag{4.13}$$

then the corresponding vacuum solutions comprise the class of solutions first discovered by Kerr and Schild,[5,6] characterized by having a line element of the form

$$g_{ab} = \eta_{ab} + l_a l_b,$$

where η_{ab} is the flat space metric and l_a is an expanding null vector field (necessarily geodesic and shearfree). This class of vacuum solutions was also later derived by Debney, Kerr, and Schild,[10] by solving the reduced field equations subject to an additional restriction, which in our notation reads

$$D_{33}V = 0.$$

Case 3: $Y^0_1 = Z^0_3 = B^0_1 = 0$. (4.14)

Using (A10), these conditions imply that $2\Psi^2_3 - 3\Psi_2\Psi_4 = 0$, so that *the Weyl tensor is of type* $\{22\}$ (although n^a is in general not the second repeated principal null direction).

As in Case 2, *the line element can be written in the form (2.2), (2.3), and (4.12).* The additional restriction $B^0_1 = 0$ implies directly that

$$\Phi^0_{11} = \text{const} \tag{4.15}$$

[see (A8b)], while a straightforward calculation using Eqs. (A11f), (A5c, d), (4.12b, c, d), and (2.3a), shows that $Z^0_3 = 0$ restricts the function $q(z)$ in (4.12c) according to $q'''(z) = 0$, so that

$$q(z) = b_0 + b_1 z + b_2 z^2, \tag{4.16}$$

where b_0, b_1, b_2 are complex constants.

It can be shown that *the vacuum solutions contained in this case* [obtained by setting $\Phi_{11}^0 = 0$] *are the only vacuum solutions of Petrov-Penrose type* $\{22\}$ *which satisfy Condition III* (type $\{22\} <=> 2\Psi_3^2 - 3\Psi_2\Psi_4 = 0 <=> Y_1^0 = Z_3^0 = 0$, in vacuo). These solutions thus comprise classes I and II in Kinnersley's[11] classification. [His class III satisfies (1.8) but not (1.11), while his class IV solutions do not satisfy (1.8) since the repeated principal null congruences are nonexpanding.] Note that of the six real parameters contained in b_0, b_1, b_2 in (4.16) only one is essential. It is not feasible to give complete details here. The reader is referred to the paper of Kinnersley,[11] which contains a detailed classification of these solutions in the vacuum case, although using a different coordinate system.

In the remainder of this section *we restrict our attention to the class of solutions in case 3 for which* $U^0 < 0$. By suitably choosing the coordinates we are able to show how the line element specializes to give the well-known algebraically special vacuum solutions referred to at the beginning of this section. Firstly, it is shown in Appendix C that by using the remaining coordinate freedom one can specialize $q(z)$, as given by (4.16), to the form

$$q(z) = \tfrac{1}{2} ibz, \tag{4.17}$$

where b is a real constant. In order to obtain agreement with standard notation we write

$$M \equiv l, \quad b \equiv a + l.$$

Then using (4.17), the functions P, Q, Σ, W, and U which appear in the line element (2.2) assume the form

$$P = 2^{-1/2}(1 + \tfrac{1}{4} z\bar{z}), \tag{4.18a}$$

$$Q = \tfrac{1}{2} iazP^{-2} + 2^{3/2} il(\bar{z}P)^{-1}, \tag{4.18b}$$

$$\Sigma = -a(1 - \tfrac{1}{4} z\bar{z})(1 + \tfrac{1}{4} z\bar{z})^{-1} + l, \tag{4.18c}$$

$$W = -\tfrac{1}{2} iazP^{-2}, \tag{4.18d}$$

$$U = -\tfrac{1}{2} + (mr + \Sigma l - \Phi_{11}^0)(r^2 + \Sigma^2)^{-1}. \tag{4.18e}$$

[P, Q are given by (4.12) with (4.17), noting that $U^0 = -\frac{1}{2}$, while Σ, W, U are obtained from Eqs. (2.2d, e) and (2.3a)]. We now introduce angular coordinates θ, ϕ by

$$z = 2e^{i\varphi} \tan\theta/2, \tag{4.19a}$$

and a new u coordinate by

$$u = u' + 2l\phi. \tag{4.19b}$$

The line element (2.2) then assumes the form (dropping the prime on u)

$$\begin{aligned} ds^2 = &-(r^2 + \Sigma^2)d\theta^2 - (r^2 + \Sigma^2)^{-1}[(R^2 \sin^2\theta \\ &- H^2\Delta)d\phi^2 - 2(aR \sin^2\theta - H\Delta)du d\phi \\ &- (\Delta - a^2 \sin^2\theta)du^2] - 2H dr d\phi + 2dr du, \end{aligned} \tag{4.20a}$$

with

$$R = r^2 + a^2 + l^2, \tag{4.20b}$$

$$H = a \sin^2\theta + 2l \cos\theta, \tag{4.20c}$$

$$\Delta = r^2 - 2mr + a^2 - l^2 + 2\Phi_{11}^0, \tag{4.20d}$$

$$\Sigma = -a \cos\theta + l. \tag{4.20e}$$

The vacuum solution contained in this line element, obtained by setting $\Phi_{11}^0 = 0$, *is in fact the Demianski-Newman*[7] *solution,* in the form given by Kinnersley.[11] The constants, m, a, and l are the mass parameter, the Kerr angular momentum parameter, and the NUT parameter, respectively. The Kerr solution,[4] NUT solution,[3] and Schwarzschild solution (in retarded Eddington-Finkelstein coordinates) are obtained by setting $l = 0$, $a = 0$, and $l = a = 0$, respectively.

In Sec. 5 we show that *the line element (4.20) is compatible with the combined Einstein-Maxwell field equations,* and hence, by suitably specifying Φ_{11}^0 in terms of the electromagnetic field, we arrive at the charged Demianski-Newman solution. On the other hand, it is shown in Sec. 6 that *the above line element is not compatible with the combined Einstein-Weyl equations for the case of a (nontrivial) neutrino field whose principal null congruence is geodesic and shearfree.*

5. EINSTEIN-MAXWELL FIELDS

A. The reduced Einstein-Maxwell equations

The Einstein-Maxwell equations for a source-free electromagnetic field read:[38]

$$R_{ab} = F_a{}^c F_{bc} - \tfrac{1}{4} g_{ab} F_{cd} F^{cd}, \tag{5.1a}$$

$$F_{[ab,c]} = 0, \quad F^{ab}{}_{;b} = 0. \tag{5.1b}$$

We assume that *one of the principal null directions*[2] *of the electromagnetic field, say* k^a, *is tangent to a geodesic, shearfree, and expanding null congruence.* This assumption clearly implies that Condition I of Sec. 1 holds. In addition, on account of (5.1a), it also ensures that Condition II is valid. This is most easily verified using the Newman-Penrose formalism. In fact, Eq. (5.1a) assumes the form[31]

$$\Phi_{AB} = \Phi_A \bar{\Phi}_B, \quad A, B = 0, 1, 2, \tag{5.2}$$

where the Φ_A are the tetrad components[31] of F_{ab}. Since we are using the principal null direction k^a as one of the tetrad vectors, it follows that[31]

$$\Phi_0 = 0. \tag{5.3}$$

Thus (5.2) implies that $\Phi_{00} = \Phi_{01} = \Phi_{02} = 0$, which is equivalent to (1.7a). The remainder of Condition II (i.e., $R = 0$) follows directly from (5.1a).

In order to apply the results of Sec. 2 it is necessary to relate the functions $\Phi_{11}^0, \Phi_{12}^0, \Phi_{22}^0$ to the tetrad components Φ_A of F_{ab}, using (5.2). This requires a knowledge of the r dependence of the Φ_A, which is determined by the first set of Maxwell's equations (5.1b). In the Newman-Penrose formalism these equations read[31]

$$D\Phi_1 = 2\rho\Phi_1, \tag{5.4a}$$

$$D\Phi_2 - \delta\Phi_1 = \rho\Phi_2, \tag{5.4b}$$

because of (5.3) and the restrictions (A1), (A2), (A3) on the spin coefficients. These can be integrated with respect to r to yield

$$\Phi_1 = \Phi_1^0 \rho^2, \tag{5.5a}$$

$$\Phi_2 = \rho(-\Phi_2^0 + F_1^0 \rho + F_2^0 \rho^2), \tag{5.5b}$$

where Φ_1^0, Φ_2^0 are the complex "constants" of integration (functions of $u, z, \bar{z}$) and

$$F_1^0 = P(D_4 + 2D_1\overline{Q})\Phi_1^0, \tag{5.6a}$$

$$F_2^0 = 2\Phi_1^0 N, \tag{5.6b}$$

with

$$N = iP(D_4 + D_1\overline{Q})\Sigma.$$

For future reference we expand the expressions (5.5) in inverse powers of r:

$$\Phi_1 = \Phi_1^0 r^{-2} + O(r^{-3}), \tag{5.7a}$$

$$\Phi_2 = \Phi_2^0 r^{-1} + O(r^{-2}), \tag{5.7b}$$

or in tensor form:

$$F_{ab} = [2]_{ab}\, r^{-1} + [11]_{ab}\, r^{-2} + O(r^{-3}). \tag{5.8}$$

The coefficient tensors in (5.8) are antisymmetric, covariantly constant along the congruence k^a, and admit k^a as a principal null direction of multiplicity two and one, respectively. The r^{-1} term is the radiation part of the electromagnetic field.

The second set of Maxwell's equations (5.1b) provides a set of partial differential equations for Φ_1^0 and Φ_2^0. In the Newman-Penrose formalism, they are equivalent, by virtue of (5.3), (A1)-(A3), to[31]

$$\delta\Phi_1 = 0, \tag{5.9a}$$

$$\delta\Phi_2 - \Delta\Phi_1 = 2\mu\Phi_1 - 2\beta\Phi_2. \tag{5.9b}$$

Using Eqs. (5.5), (A4c, e), (2.2c, d, e), and (2.6) these can be simplified to

$$(D_3 + 2D_1Q)\Phi_1^0 = 0, \tag{5.10a}$$

$$(D_3 + D_1Q)(P^{-1}\Phi_2^0) + D_1(P^{-2}\Phi_1^0) = 0. \tag{5.10b}$$

Finally, by comparing the leading terms in the expressions for $\Phi_{11}, \Phi_{12}, \Phi_{22}$ in the form (A7) with the expressions for these quantities calculated from (5.2) using (5.5), one obtains

$$\Phi_{11}^0 = \Phi_1^0\bar{\Phi}_1^0, \quad \Phi_{12}^0 = \Phi_1^0\bar{\Phi}_2^0, \quad \Phi_{22}^0 = \Phi_2^0\bar{\Phi}_2^0. \tag{5.11}$$

Using these identifications and the reduced Maxwell equations (5.10), one finds that the remaining terms in the two expressions for $\Phi_{11}, \Phi_{12}, \Phi_{22}$ agree identically.

We use Eqs. (5.11) to eliminate $\Phi_{12}^0 (\equiv \bar{\Phi}_{21}^0)$ and Φ_{22}^0 from Eqs. (2.14a, b). Equations (2.14), together with (5.10) then form a system of partial differentiation equations for the functions $P, Q, m, \Phi_1^0, \Phi_2^0$. We will refer to these equations as *the reduced Einstein-Maxwell equations*. These equations and associated canonical line element have also been obtained by Robinson, Schild, and Strauss,[30] and, using the Newman-Penrose formalism, by Lind.[39] These authors, however, did not make use of the fact that the r dependence of the line element is completely determined (via the Bianchi identities) by the *algebraic* restrictions (i.e., Condition II) which the presence of the electromagnetic field imposes on the Ricci tensor.

B. Exact solutions

We now wish to apply the results of Sec. 3. We thus assume that the Weyl tensor satisfies Condition III. By virtue of Eqs. (3.4), (3.6), (5.10b), and (5.11), Condition IV is valid if and only if the electromagnetic field satisfies

$$\Phi_2^0 = 0. \tag{5.12}$$

This is equivalent to the vanishing of the first term in the tensor expansion (5.8) and means physically that the field is nonradiative.

The reduced gravitational field equations (2.14) can now be integrated as in Sec. 3. The line element is then completely determined, apart from the term $\Phi_{11}^0 = \Phi_1^0\bar{\Phi}_1^0$ [see (5.11)] in U [see (2.2e)]. To obtain Φ_1^0, we have to solve the reduced Maxwell equations (5.10), subject to (3.4) and (5.12). Using Eqs. (2.9) and (3.9) one readily obtains

$$\Phi_1^0(z, \bar{z}) = 2f(z)P^2l^2(kz\bar{z} + l)^{-2}, \tag{5.13}$$

provided that the "constant" of integration $f(z)$ is suitably chosen. We summarize the results in the following

Theorem 5.1: A space-time satisfies the following conditions:

(1) the Einstein-Maxwell field equations hold,

(2) one of the principal null directions of the electromagnetic field is tangent to a geodesic, shearfree, and expanding null congruence,

(3) the electromagnetic field tensor and the Weyl tensor exhibit the following asymptotic behavior along the null geodesics:

$$F_{ab} = [11]_{ab}\, r^{-2} + O(r^{-3}), \tag{5.14a}$$

$$C_{abcd} = [211]_{abcd}\, r^{-3} + O(r^{-4}), \tag{5.14b}$$

if and only if the line element can be written locally in the form (2.2), (2.3), (3.9), (3.11), (3.15), with Φ_{11}^0 given by (5.11) and the electromagnetic field in the form (5.5), (5.6), (5.12), (5.13).

This class of exact solutions of the Einstein-Maxwell equations, although not its characterization as contained in Theorem 5.1, has been given by Robinson, Schild, and Strauss.[30]

C. Specializations

The assumption (4.1) of Case 1 in Sec. 4 implies (4.4) and (4.6), and thus reduces Φ_1^0 as given by (5.13), to the form

$$\Phi_1^0 = f(z). \tag{5.15}$$

From Eq. (5.11), it follows that

$$\Phi_{11}^0 = f(z)\bar{f}(\bar{z}), \tag{5.16}$$

and from Eqs. (5.5) we find

$$\Phi_1 = f(z)\rho^2, \tag{5.17a}$$

$$\Phi_2 = 2P\rho^2[f'(z) + 2if(z)\partial_z\Sigma\rho]. \tag{5.17b}$$

The line element of this class of solutions of the Einstein-Maxwell equations is given by Eqs. (2.2), (2.3), (4.4), (4.7), (4.8), (5.16) and the electromagnetic field tensor by (A14) and (5.17). This class of solutions was given by Lind,[39] independently of the paper of Robinson, Schild, and Strauss.[30] The subclass of these solutions satisfying (4.13) [so that Q can be specialized to the form (4.12c)] with $Q_1 = 0$ was given earlier by Debney, Kerr, and Schild.[10]

The assumption (4.9) of Case 2 does not affect Φ_1^0, while the assumption $B_1^0 = 0$ of Case 3 [see (4.14)] implies that

$$\Phi_1^0 = f = \text{const}. \tag{5.18}$$

The charged Demianski-Newman[7] solution is the unique member of the Case 3 solutions subject to

$$U^0 < 0, \quad \Phi_1^0 = \bar{\Phi}_1^0. \tag{5.19}$$

In order to obtain agreement with the usual notation we write

$$\Phi_1^0 = 2^{-1/2}e. \tag{5.20}$$

Since $U^0 < 0$, and we are in Case 3 of Sec. 4, the line element can be written in the form (4.20), with the arbitrary constant Φ_{11}^0 given, on account of (5.11) and (5.20) by

$$\Phi_{11}^0 = \tfrac{1}{2}e^2. \tag{5.21}$$

Using Eqs. (5.15), (5.17a, b), (5.20), and (4.18a, c), the nonzero tetrad components of the electromagnetic field assume the form

$$\begin{aligned}\Phi_1 &= 2^{-1/2}e\rho^2,\\ \Phi_2 &= iae\bar{z}(1 + \tfrac{1}{4}z\bar{z})^{-1}\rho^3.\end{aligned} \tag{5.22}$$

The electromagnetic field tensor F_{ab} is expressed in terms of Φ_1, Φ_2 according to (A14). After performing the coordinate transformation (4.19), one obtains

$$F_{ab}\,dx^a \wedge dx^b = 2^{3/2}e(r^2 + \Sigma^2)^{-2}[(r^2 - \Sigma^2)(du - Hd\phi) \wedge dr + 2r\Sigma \sin\theta d\theta \wedge (Rd\phi - adu)], \tag{5.23}$$

where Σ, H, and R are given by (4.20b, c, e), and $\wedge$ denotes the operation of taking the antisymmetrized tensor product. The *charged Demianski-Newman[7] solution of the Einstein-Maxwell equations is thus given by Eqs. (4.20), (5.21), and (5.23).* Note that if we set $l = 0$ and make the changes $a \to -a$, $u \to -u$, the above expressions agree with Carter's[40] form of the charged Kerr solution (first discovered by Newman et al.[28]). The Reissner-Nordstrom solution (in retarded Eddington-Finkelstein coordinates) is obtained when we set $l = a = 0$.

6. EINSTEIN-WEYL FIELDS

A. The reduced Einstein-Weyl equations

The Einstein-Weyl equations governing the interaction of a neutrino field and a gravitational field read[41]

$$R_{ab} = -i[\sigma_{aAX'}(\phi^A\phi^{X'}{}_{;b} - \phi^A{}_{;b}\,\phi^{X'}) + \sigma_{bAX'}(\phi^A\phi^{X'}{}_{;a} - \phi^A{}_{;a}\,\phi^{X'})], \tag{6.1a}$$

$$\sigma^{aAX'}\phi_{A;a} = 0, \tag{6.1b}$$

where $\phi^A(x^a)$ is the spinor field which describes the neutrino field.

We assume that *the principal null direction of the neutrino field,*[41] defined by

$$k^a = \sigma^{aAX'}\phi_A\phi_{X'},$$

is tangent to a geodesic, shearfree, and expanding null congruence. This assumption clearly implies that Condition I of Sec. 1 holds, and in fact also ensures the validity of Condition II, as was shown in a previous paper.[24]

With respect to a null tetrad and associated spinor dyad (based on the principal null direction) the neutrino spinor is described[41] by a complex function ϕ. The Weyl equations (6.1b) determine the r dependence of this function according to

$$\phi = -\phi^0\rho \equiv \phi^0 r^{-1} + O(r^{-2}) \tag{6.2}$$

[see (A4a)] and the "constant" of integration $\phi^0(u, z, \bar{z})$ is required to satisfy

$$(D_3 + D_1Q)(P^{-1/2}\phi^0) = 0, \tag{6.3}$$

in the notation of Sec. 2. In addition, the functions Φ_{11}^0, Φ_{12}^0, and Φ_{22}^0 are expressed in terms of ϕ^0 by the equations

$$\Phi_{11}^0 = \Sigma\phi^0\bar{\phi}^0, \tag{6.4a}$$

$$\Phi_{21}^0 = \tfrac{1}{2}iP^{1/2}\bar{\phi}^0(D_4 + 2D_1\bar{Q})(P^{1/2}\phi^0), \tag{6.4b}$$

$$\Phi_{22}^0 = i(\phi^0 D_1\bar{\phi}^0 - \bar{\phi}^0 D_1\phi^0). \tag{6.4c}$$

For the results contained in Eqs. (6.2)-(6.4) we refer to Trim and Wainwright.[24]

As in Sec. 4 we use Eqs. (6.4) to eliminate Φ_{12}^0 and Φ_{22}^0 from Eqs. (2.14a, b). Equations (2.14), together with (6.3) then form a system of partial differential equations to be solved for the functions P, Q, m, ϕ^0. We will refer to these equations as *the reduced Einstein-Weyl equations.*

B. Exact solutions

We now wish to apply the results of Sec. 3. We thus assume that the Weyl tensor satisfies Condition III. By virtue of (3.6) and (6.4) Condition IV is valid if and only if

$$(D_4 + 2D_1\bar{Q})(P^{1/2}\phi^0) = 0, \tag{6.5a}$$

$$\phi^0 D_1\bar{\phi}^0 - \bar{\phi}^0 D_1\phi^0 = 0, \tag{6.5b}$$

$$D_1(\Sigma\phi^0\bar{\phi}^0) = 0. \tag{6.5c}$$

On evaluating the commutator $(D_{34} - D_{43})\ln\phi^0$ using (6.3) and (6.5a), we obtain

$$2i\Sigma D_1\phi^0 = P^2(D_4\bar{L} - 2D_3L)\phi^0. \tag{6.6}$$

In conjunction with (6.5b) this implies

$$D_4\bar{L} + D_3L = 0, \tag{6.7}$$

and hence, on account of (2.3b) and (2.9b),

$$U^0 = 0. \tag{6.8}$$

Using the explicit form (3.9) for L, Eq. (6.7) in fact entails $l(z) = Kz$, K a constant, so that we may use (B7) with $f(z)$ given by (3.3) to set

$$L = 0. \tag{6.9}$$

Thus by (6.6)

$$D_1\phi^0 = 0, \tag{6.10}$$

provided that $\Sigma \neq 0$. (The case $\Sigma = 0$, corresponding to zero twist of the neutrino principal null congruence, has been treated previously.[24]) Equation (6.5c) now implies, for a nonvanishing neutrino field, that

$$D_1\Sigma = 0, \tag{6.11}$$

and we are in Case 1 of Sec. 4. By virtue of Eqs. (6.8), (4.2), and (4.4a), Eqs. (6.3) and (6.5a) can be integrated to yield

$$\phi^0 = k = \text{const}. \tag{6.12}$$

From Eq. (6.4a) we obtain

$$\Phi_{11}^0 = \Sigma k\bar{k}. \tag{6.13}$$

The line element is thus determined by Eqs. (2.2), (2.3), (4.7), (4.8), and (6.13), with $P = 2^{-1/2}$, and the neutrino field by

$$\phi = k(r + i\Sigma)^{-1}. \tag{6.14}$$

It is clear that the class of solutions of the Einstein-Weyl equations compatible with Conditions I-IV is much more restricted [see (6.8), (6.11), and (6.12)] than the corresponding class of solutions of the Einstein-Maxwell equations. In particular, on account of (6.8), *there exists no solution of the Einstein-Weyl equations analogous to the charged Demianski-Newman[7] solution of the Einstein-Maxwell equations*, for which $U^0 < 0$.

To conclude this section, we note that the reduced Einstein-Weyl equations (2.14), (6.3), can be solved to obtain a more general class of solutions if we replace Condition IV [which entails (6.5)] by the weaker restriction $D_1\phi^0 = 0$ on the neutrino field. Details are given in the thesis of Trim.[32]

7. CONCLUSIONS

We have given a unified approach to the derivation of a simple form for the line element of a wide class of algebraically special solutions of the Einstein, Einstein-Maxwell, and Einstein-Weyl equations. In addition, we have derived the well known twisting solutions of the Einstein and Einstein-Maxwell equations (i.e., the charged Demianski-Newman[7] solution and its specializations) by systematically restricting the asymptotic behavior of the Weyl and Ricci tensors, instead of specifying an arbitrary function in an ad hoc fashion, as is usually done.[8,9,10]

Not all explicitly known exact solutions of the (vacuum) Einstein equations satisfying Condition I also satisfy Condition III. As mentioned earlier, the Case III solutions of Kinnersley[11] satisfy $\Psi_4^0 \neq 0$ in our notation, so that Condition III does not hold. The second twisting example known to us is a particular vacuum solution found by Kerr and Debney[13] which in our notation satisfies $\Psi_3^0 \neq 0$, $\Psi_4^0 = 0$, so that again Condition III does not hold. In addition, in the twistfree case, several classes of vacuum solutions which satisfy Condition I but not III are known,[1,43] including the general solution for the type {4} case. However, to the best of the writers' knowledge all explicitly known exact vacuum solution satisfying Condition I, with *nonzero twist, and containing* (essentially) *arbitrary functions* also satisfy our Condition III. This justifies the remarks at the end of Sec. 1.

ACKNOWLEDGMENTS

One of the authors (J.W.) would like to thank E. T. Newman for helpful discussions in connection with Condition III of Sec. 1. This work was supported in part by grants from the National Research Council of Canada.

APPENDIX A

This appendix contains the detailed expressions for the spin coefficients and components of the Weyl and Ricci tensors associated with the null tetrad (2.6).

The fact that k^a is tangent to a geodesic and shearfree null congruence (Condition I) implies that[31]

$$\kappa = \sigma = 0. \tag{A1}$$

The further simplification of the null tetrad arising in the proof of Theorem 2.1 results in[24]

$$\epsilon = \pi = 0, \tag{A2}$$

$$\tau = \lambda = 0. \tag{A3}$$

The remaining nonzero spin coefficients are given by[42]

$$\rho = -(r + i\Sigma)^{-1}, \tag{A4a}$$

$$\alpha = (PD_1\overline{Q} + \tfrac{1}{2}D_4P)\rho, \tag{A4b}$$

$$\beta = -\tfrac{1}{2}D_3P\bar{\rho}, \tag{A4c}$$

$$\gamma = -\tfrac{1}{2}D_1\ln P - \tfrac{1}{2}\Psi_2^0\rho^2 + \Phi_{11}^0\rho^2\bar{\rho}, \tag{A4d}$$

$$\mu = (-U^0 + iD_1\Sigma - i\Sigma D_1\ln P)\bar{\rho} - \tfrac{1}{2}\Psi_2^0(\rho^2 + \rho\bar{\rho}) + \Phi_{11}^0\rho^2\bar{\rho}, \tag{A4e}$$

$$\nu = \nu^0 + \Psi_3^0\rho + \tfrac{1}{2}Y_1^0\rho^2 + \tfrac{1}{3}Y_2^0\rho^3 + \bar{\rho}(-\Phi_{21}^0\rho + B_1^0\rho^2 + B_2^0\rho^3), \tag{A4f}$$

where

$$\nu^0 = PD_1(D_1\overline{Q} + D_4\ln P), \tag{A5a}$$

$$Y_1^0 = -P(D_4 + 3D_1\overline{Q})\Psi_2^0, \tag{A5b}$$

$$Y_2^0 = -3N\Psi_2^0, \tag{A5c}$$

with

$$N = iP(D_4 + D_1\overline{Q})\Sigma, \tag{A5d}$$

and $\Psi_2^0, \Phi_{21}^0, B_1^0, B_2^0$ are given below.

Condition II restricts the components of the Ricci tensor with respect to the tetrad (2.6) according to

$$\Phi_{00} = \Phi_{01} = \Phi_{02} = \Lambda = 0. \tag{A6}$$

The remaining nonzero components are given by[42]

$$\Phi_{11} = \Phi_{11}^0\rho^2\bar{\rho}^2, \tag{A7a}$$

$$\Phi_{21} = \rho\bar{\rho}^2(-\Phi_{21}^0 + B_1^0\rho + B_2^0\rho^2), \tag{A7b}$$

$$\Phi_{22} = \rho\bar{\rho}\{\Phi_{22}^0 + C_1^0\rho + \overline{C}_1^0\bar{\rho} - 2\Phi_{12}^0N\rho^2 + C_2^0\rho\bar{\rho} - 2\Phi_{21}^0\overline{N}\bar{\rho}^2 + 2\overline{B}_1^0N\rho^2\bar{\rho} + 2B_1^0\overline{N}\rho\bar{\rho}^2 + 4N\overline{N}\Phi_{11}^0\rho^2\bar{\rho}^2\}, \tag{A7c}$$

where

$$\Phi_{21}^0 = \tfrac{1}{2}P(D_4 + 3D_1\overline{Q})\overline{\Psi}_2^0 \equiv \overline{\Phi}_{12}^0, \tag{A8a}$$

$$B_1^0 = P(D_4 + 4D_1\overline{Q})\Phi_{11}^0, \tag{A8b}$$

$$B_2^0 = 2\Phi_{11}^0N, \tag{A8c}$$

$$\Phi_{22}^0 = P^2(D_3 + 2D_1Q)(P^{-1}\Psi_3^0) + P^3D_1(P^{-3}\Psi_2^0), \tag{A8d}$$

$$C_1^0 = -P^2(D_4 + 3D_1\overline{Q})(P^{-1}\Phi_{12}^0) - \tfrac{1}{2}P^4D_1(P^{-4}\Phi_{11}^0), \tag{A8e}$$

$$C_2^0 = P^2\,Re[(D_3 + 4D_1Q)(P^{-1}B_1^0)]. \tag{A8f}$$

As mentioned in the introduction, Conditions I and II imply that the tangent field k^a is at each point a repeated principal null direction of the Weyl tensor, so that[31]

$$\Psi_0 = \Psi_1 = 0 \tag{A9}$$

with respect to the null tetrad (2.6). The nonzero components of the Weyl tensor are given below[42]:

$$\Psi_2 = -\Psi_2^0\rho^3 + 2\Phi_{11}^0\rho^3\bar{\rho}, \tag{A10a}$$

$$\Psi_3 = \rho^2\{\Psi_3^0 + Y_1^0\rho + Y_2^0\rho^2 + \bar{\rho}(-\Phi_{21}^0 + 2B_1^0\rho + 3B_2^0\rho^2)\}, \tag{A10b}$$

$$\Psi_4 = \rho\{-\Psi_4^0 + Z_1^0\rho + \tfrac{1}{2}Z_2^0\rho^2 + \tfrac{1}{3}Z_3^0\rho^3 + \tfrac{1}{4}Z_4^0\rho^4 + \bar{\rho}(R_1^0\rho + R_2^0\rho^2 + R_3^0\rho^3 + R_4^0\rho^4)\}, \quad \text{(A10c)}$$

where

$$\Psi_2^0 = -(m + iM), \quad \text{(A11a)}$$

$$\Psi_3^0 = P(D_4 + 2D_1\bar{Q})(-U^0 + iD_1\Sigma - i\Sigma D_1 \ln P) + 2i\Sigma\nu^0, \quad \text{(A11b)}$$

$$\Psi_4^0 = -(D_4 + 3D_1\bar{Q})(P\nu^0), \quad \text{(A11c)}$$

$$Z_1^0 = (D_4 + 4D_1\bar{Q})(P\Psi_3^0), \quad \text{(A11d)}$$

$$Z_2^0 = (D_4 + 5D_1\bar{Q})(PY_1^0) + 4N\Psi_3^0, \quad \text{(A11e)}$$

$$Z_3^0 = (D_4 + 6D_1\bar{Q})(PY_2^0) + 6NY_1^0, \quad \text{(A11f)}$$

$$Z_4^0 = 8NY_2^0, \quad \text{(A11g)}$$

$$R_1^0 = -(D_4 + 5D_1\bar{Q})(P\Phi_{21}^0), \quad \text{(A11h)}$$

$$R_2^0 = (D_4 + 6D_1\bar{Q})(PB_1^0) - 2N\Phi_{21}^0, \quad \text{(A11i)}$$

$$R_3^0 = (D_4 + 7D_1\bar{Q})(PB_2^0) + 4NB_1^0, \quad \text{(A11j)}$$

$$R_4^0 = 6NB_2^0. \quad \text{(A11k)}$$

For convenience we list the general expansions for the Weyl tensor, Ricci tensor, and electromagnetic field tensor in terms of their tetrad components and the null tetrad vectors:

$$\tfrac{1}{2}(C_{abcd} - iC^*_{abcd}) = -\Psi_0 U_{ab}U_{cd} + 2\Psi_1(U_{ab}M_{cd} + M_{ab}U_{cd}) - \Psi_2(U_{ab}V_{cd} + 4M_{ab}M_{cd} + V_{ab}U_{cd}) + 2\Psi_3(V_{ab}M_{cd} + M_{ab}V_{cd}) - \Psi_4 V_{ab}V_{cd}, \quad \text{(A12)}$$

where

$$U_{ab} = 2\bar{m}_{[a}n_{b]}, \qquad V_{ab} = 2k_{[a}m_{b]},$$

$$M_{ab} = k_{[a}n_{b]} - m_{[a}\bar{m}_{b]},$$

and

$$C^*_{abcd} \equiv \tfrac{1}{2}\epsilon_{cdef}C_{ab}{}^{ef},$$

$$R_{ab} = -2\Phi_{00}n_a n_b + 4\Phi_{01}n_{(a}\bar{m}_{b)} + 4\Phi_{10}n_{(a}m_{b)} - 2\Phi_{02}\bar{m}_a\bar{m}_b - 2\Phi_{20}m_a m_b - 4\Phi_{11}\{k_{(a}n_{b)} + m_{(a}\bar{m}_{b)}\} + 4\Phi_{12}k_{(a}\bar{m}_{b)} + 4\Phi_{21}k_{(a}m_{b)} - 2\Phi_{22}k_a k_b + 6\Lambda g_{ab}, \quad \text{(A13)}$$

$$\tfrac{1}{2}(F_{ab} - iF^*_{ab}) = \Phi_0 U_{ab} - 2\Phi_1 M_{ab} + \Phi_2 V_{ab}, \quad \text{(A14)}$$

where

$$F^*_{ab} = \tfrac{1}{2}\epsilon_{abcd}F^{cd}.$$

APPENDIX B

The combined coordinate and tetrad transformation (2.16), (2.17) preserves the form of the canonical line element (2.2), (2.3), and of the associated null tetrad (2.6) We list below the induced transformation laws of some of the functions of Sec. 2 and Appendix A:

$$\tilde{P}' = R^{-1}|f'(z)|P, \quad \text{(B1)}$$

$$\tilde{Q}' = (\bar{f}')^{-1}[RQ + 2\partial_{\bar{z}}h], \quad \text{(B2)}$$

$$\tilde{\Sigma}' = R^{-1}\Sigma, \quad \text{(B3)}$$

$$\tilde{U}^{0\prime} = R^{-2}U^0, \quad \text{(B4)}$$

$$\tilde{\Psi}_2^{0\prime} = R^{-3}\Psi_2^0, \quad \text{(B5)}$$

$$\tilde{\Phi}_{11}^{0\prime} = R^{-4}\Phi_{11}^0, \quad \text{(B6)}$$

$$\tilde{L}' = (f')^{-1}[L + f''/f'], \quad \text{(B7)}$$

$$\tilde{J}' = (f')^{-2}[J + 2f'''/f' - 3(f''/f')^2], \quad \text{(B8)}$$

$$\tilde{N}' = R^{-2}(\bar{f}'/f')^{1/2}N, \quad \text{(B9)}$$

$$(\tilde{D}_1\tilde{Q})' = (\bar{f}')^{-1}[D_1Q + D_3 \ln R]. \quad \text{(B10)}$$

In addition, the transformation laws of the coefficients in the expressions (A7) and (A10) for the Ricci and Weyl tensors are all of the form

$$\tilde{C}^{0\prime} = (f'/\bar{f}')^p R^q C^0,$$

for suitable values of p and q.

Finally, note that the basic differential operators D_1, D_3, and D_4, as defined by Eq. (2.4), transform according to

$$\tilde{D}_1' = R^{-1}D_1, \qquad \tilde{D}_3' = (\bar{f}')^{-1}D_3, \qquad \tilde{D}_4' = (f')^{-1}D_4, \quad \text{(B11)}$$

provided we regard them as operating only on functions of $u, z, \bar{z}$, as is the case here.

APPENDIX C

In Case 3 of Sec. 4, with $U^0 < 0$, the functions P, Q are given by

$$P = 2^{-1/2}(1 + \tfrac{1}{4}z\bar{z}), \quad \text{(C1)}$$

$$Q = P^{-2}[b_0 + b_1 z + b_2 z^2 + 2iM\bar{z}^{-1}], \quad \text{(C2)}$$

where b_0, b_1, b_2 are complex constants, and M is a real constant [see Eqs. (4.12a, c, d) and (4.16)]. The form of P is preserved under the following coordinate transformation:

$$z' = 4(Az + B)(\bar{B}z - 4\bar{A})^{-1}, \quad \text{(C3a)}$$

$$u' = u + h(z, \bar{z}), \quad \text{(C3b)}$$

where A, B are complex constants satisfying $4A\bar{A} + B\bar{B} = 4$, and h is a real function [c.f. (2.16), (3.3), and (3.5)].

The transformation (C3) does not in general preserve the form of Q. Our aim is to show that by making a suitable choice for A, B, and h, we can transform Q to the following form

$$Q = P^{-2}[ibz + 2iM\bar{z}^{-1}], \quad \text{(C4)}$$

where b is a *real* constant and M is as above, thereby justifying Eq. (4.17).

Under the transformation (C3), it follows from (B1) and (B2) (with $R = 1$) that P^2Q transforms according to

$$(P^2Q)' = \left[\frac{dz'}{dz}\right][P^2Q + 2P^2\partial_{\bar{z}}h].$$

In order to achieve our aim we require

$$ibz' + 2iM(\bar{z}')^{-1} = \left[\frac{dz'}{dz}\right][b_0 + b_1 z + b_2 z^2 + 2iM\bar{z}^{-1} + (1 + \tfrac{1}{4}z\bar{z})^2\partial_{\bar{z}}h],$$

where z' is given by (C3a). This equation will have a solution for $h(z, \bar{z})$ provided that $\partial_z\partial_{\bar{z}}h = \partial_{\bar{z}}\partial_z h$. A straight-

forward calculation shows that this condition holds provided that the ratio A/B is chosen to satisfy

$$(\overline{A/B}) - \tfrac{1}{4}(B/A) + (b_1 - \bar{b}_1 - iM)(\bar{b}_0 + 4b_2)^{-1} = 0,$$

in terms of the given arbitrary constants appearing in Q. (If $\bar{b}_0 + 4b_2 = 0$, choose $B = 0$.) This is clearly permissible, and the transformation to the form (C4) is thus justified.

[1]I. Robinson and A. Trautman, Proc. Roy. Soc. London, **A265**, 463 (1962).
[2]F. A. E. Pirani, in *Lectures on General Relativity*, Brandeis Summer Institute in Theoretical Physics, (Prentice-Hall, Englewood Cliffs, N. J., 1964), Vol. 1. See Chaps. 3 and 4.
[3]E. T. Newman, L. Tamburino, and T. Unti, J. Math. Phys. **4**, 915 (1963).
[4]R. P. Kerr, Phys. Rev. Lett. **11**, 237 (1963).
[5]R. P. Kerr and A. Schild, Proc. Symp. Appl. Math. Amer. Math. Soc. **XVII**, 199 (1964).
[6]R. P. Kerr and A. Schild, "Atti del convegno sulla relativita Generale: Problemi dell'energia e onde gravitazionale," Firenze (1964), p. 173.
[7]M. Demianski and E. T. Newman, Bull. Acad. Polon. Sci. Ser. Math. Astron. Phys. **14**, 653 (1966).
[8]I. Robinson, J. Robinson, and J. D. Zund, J. Math. Mech. **18**, 881 (1969).
[9]C. J. Talbot, Commun. Math. Phys. **13**, 45 (1969).
[10]G. C. Debney, R. P. Kerr, and A. Schild, J. Math. Phys. **10**, 1842 (1969).
[11]W. Kinnersley, J. Math. Phys. **10**, 1195 (1969).
[12]I. Robinson and J. R. Robinson, Int. J. Theoret. Phys. **2**, 231 (1969).
[13]R. P. Kerr and G. C. Debney, J. Math. Phys. **11**, 2807 (1970).
[14]C. D. Collinson, J. Phys. A **2**, 621 (1969).
[15]J. D. Zund, Lett. al Nuovo Cimento **4**, 879 (1972); **7**, 233 (1973).
[16]R. Debever, Bull. Soc. Math. Belg. **23**, 360 (1971).
[17]This is brought out most clearly in the papers of I. Robinson and his coworkers. See refs. 8 and 12.
[18]R. K. Sachs, Proc. Roy. Soc. London A **270**, 103 (1962).
[19]R. Penrose, in *Battelle Rencontres: 1967 Lectures in Mathematics and Physics*, edited by B. De Witt and J. Wheeler (Benjamin, New York, 1968). See p. 184.
[20]P. Jordan, J. Ehlers, and R. K Sachs, Akad. Wiss. Mainz. No. 1 (1961). See p. 54. In this reference the expansion is given in an exact form applicable to this class of vacuum solutions, using a complex quantity z analogous to our expression (A 4a) for the spin coefficient ρ. For a more general statement see Refs. 2 (p.323), 18 (p.116), and 19 (p.163).
[21]In this and future similar equations a symbol such as $O(r^{-4})$ indicates the higher order terms in an infinite series expansion in powers of r^{-1}.
[22]E. T. Newman and T. W. J. Unti, J. Math. Phys. **3**, 891 (1962). For a review of these solutions see R. W. Lind, J. Messmer, and E. T. Newman, J. Math. Phys. **13**, 1884 (1972).
[23]B. Aronson and E. T. Newman, J. Math. Phys. **13**, 1847 (1972).
[24]D. W. Trim and J. Wainwright, J. Math. Phys. **12**, 2494 (1972). See the Appendix.
[25]W. Kundt and M. Trümper, Akad. Wiss. Mainz No. 12 (1962). The required result is an obvious combination of Theorem 3.4 and Lemma 3.5 on pp. 984-985.
[26]Using the fact that $R_{ab}{}^{(1)}$ is symmetric and tracefree, and (A13), condition (1.10a) implies that $R_{ab}{}^{(1)} = h\, k_a\, k_b$, for some function h.
[27]The simplest example of this phenomenon occurs when one generalizes from the Schwarzschild vacuum solution to the Reisnner-Nordstrom solution of the Einstein-Maxwell equations. The line element is changed only via the appearance of the constant charge parameter e. For more general examples, see Refs. 28, 29, and 30.
[28]E. T. Newman, E. Couch, K Chinnapared, A. Exton, A. Prakash, and R. Torrence, J. Math. Phys. **6**, 918 (1965).
[29]W. M. Kinnersley, Ph. D. Thesis, California Institute of Technology, 1969. See Chap. 5.
[30]I. Robinson, A. Schild, and H. Strauss, Int. J. Theoretical Physics **2**, 243 (1969). H. R. Strauss, Ph. D. Dissertation, The University of Texas at Austin, 1969.
[31]E. T. Newman and R. Penrose, J. Math Phys. **3**, 566 (1962).
[32]D. W. Trim, Ph. D. Thesis, University of Waterloo, 1972.
[33]The function J was first introduced by Robinson and his co-workers in the vacuum case. See Refs. 8 and 12.
[34]See Ref. 9, p. 54. Our Ψ_2^0 and Ψ_4^0 differ in sign from Talbot's. This change has the effect of making the leading terms of the expansions (2.7) formally identical with those of Newman and Unti (Ref. 22).
[35]The imaginary part of the right-hand side of (2.12b) is identically zero as a consequence of (2.3c) and (2.8a), so that (2.12b) is in fact consistent with the fact that Φ_{22}^0 is real.
[36]This function was first introduced by Robinson and Robinson, who also derived (2.14c). See Ref. 12, eq. (4.5), noting the correspondences $V \leftrightarrow U$, $P \leftrightarrow \exp(-u)$.
[37]This expression for L differs from that given by Robinson and Robinson. We find this form more convenient when we further specialize the line element in Sec. 4.
[38]See for example R. Adler, M. Bazin and M. Schiffer, *Introduction to General Relativity* (McGraw Hill, New York, 1965).
[39]R. W. Lind, Ph. D. Thesis, University of Pittsburgh, 1970.
[40]B. Carter, Phys. Rev. **174**, 1559 (1968).
[41]J. Wainwright, J. Math. Phys. **12**, 828 (1971).
[42]As is customary a degree sign as superscript indicates that the corresponding quantity is independent of r.
[43]J. Foster and E. T. Newman, J. Math. Phys. **8**, 189 (1967).

An action principle for superconductivity

J. Chela-Flores

Centro de Física, IVIC, Apartado 1827, Caracas, Venezuela

International Centre for Theoretical Physics, Trieste, Italy
(Received 28 August 1973)

We study a Lagrangian theory of superconductivity which, for slowly varying field variables, is shown to be equivalent to the BCS theory; we obtain the conservation of pairs, clearly showing its relation with the second kind of gauge invariance of the Euler–Lagrange equations. Finally, by treating the gauge function as a further field variable, we generalize the Maxwell-type Ginzburg–Landau equation and use it to discuss flux quantization.

1. INTRODUCTION

There are many physical systems whose dynamic equations can be derived from an action principle, namely a functional of the dynamic variables takes a stationary value with respect to the small variations of the field variables. Such a formalism has an outstanding feature: It establishes a direct connection between symmetry principles and conservation laws; furthermore, it is a very convenient and economical way of formulating dynamical laws. In the present paper, an action principle is set up which, when the order parameter Φ and the density matrix Ω_1 defined below are slowly varying in space, is shown to be equivalent to the Bardeen–Cooper–Schrieffer (BCS) theory.[1] The Lagrangian theory, however, clearly displays the connection between gauge invariance and pair conservation. Let us first recall that in superconductivity theory an important role is played by:

(a) The first reduced density matrix

$$\Omega_1(x_1|x_2) = \mathrm{Tr}\rho(t)\psi^\dagger(x_2)\psi(x_1), \tag{1.1}$$

where $\rho(t)$ is the statistical density matrix, $\psi(x)$ is the fermion field operator at the space point $\mathbf{x}$; the spin label σ is also included with the convention $x = (\mathbf{x}, \sigma)$.

(b) The order parameter or macroscopic wave function defined as

$$\Phi(x_1|x_2) = \mathrm{Tr}\rho(t)\psi(x_2)\psi(x_1). \tag{1.2}$$

In the BCS theory a good description of the superconducting sample is obtained by letting the electrons interact by means of a two-body potential $V(\mathbf{x}_1 - \mathbf{x}_2)$ in such a way that the interactions are effective between electrons of opposite spin and momentum; then, by postulating an ansatz for the ground state, an expression for the energy is written and minimized subject to two constraints. This procedure yields the Valatin commutator equations,[2] which we have shown how to write in configuration space as a pair of partial differential equations[3]:

$$\begin{aligned} i\partial_t\Omega_1(x_1|x_2) = {} & \frac{1}{2m}(\partial^2_{\mathbf{x}_1} - \partial^2_{\mathbf{x}_2})\Omega_1(x_1|x_2) \\ & + \int dx'\Phi(x_1|x')\mu^*(x'|x_2) \\ & - \int dx'\mu(x_1|x')\Phi^*(x'|x_2), \end{aligned} \tag{1.3}$$

together with

$$\begin{aligned} i\partial_t\Phi(x_1|x_2) = {} & \left(\frac{1}{2m}(\partial^2_{\mathbf{x}_1} + \partial^2_{\mathbf{x}_2}) - V(\mathbf{x}_1 - \mathbf{x}_2)\right)\Phi(x_1|x_2) \\ & - \int dx'\mu(x_1|x')\Omega_1(x'|x_2) \\ & - \int dx'\Omega_1(x_1|x')\mu(x'|x_2), \end{aligned} \tag{1.4}$$

where the energy scale has been defined such that the Fermi momentum is null. Here, however, we have written the BCS equations in a more general situation than in Ref. 3, since we have not limited our discussion to a delta function potential, neither have we restricted ourselves to a time-independent theory. The definition of the pair potential μ is given in Sec. 2.

2. THE ACTION PRINCIPLE

Let us introduce a superconducting vector field θ_j where

$$\theta_j = (\Phi, \Phi^*, \Omega_1, \Omega_1^*)^T, \tag{2.1}$$

T denoting the transpose. The action principle requires that if I is the action, then a first-order variation yields

$$\delta I = 0 = \delta\int dt dx_1 dx_2\, \mathcal{L}',$$

provided that the surface Σ of the region of integration is kept constant and that variations on Σ are null. This leads to

$$\delta I = 0 = \int\bar{\delta}\theta_j\left(\frac{\partial\mathcal{L}'}{\partial\theta_j} - \partial_K\frac{\partial\mathcal{L}'}{\partial\partial_K\theta_j}\right)dt dx_1 dx_2, \tag{2.2}$$

where $\bar{\delta}$ represents a change of form.[4] Now for paired electrons the fields Φ and Φ^* ought to satisfy the constraints

$$\Phi(x_1|x_2) = -\Phi(x_2|x_1), \tag{2.3a}$$

$$\Phi^*(x_1|x_2) = -\Phi^*(x_2|x_1). \tag{2.3b}$$

Further, from definition (1.1) we find that the fields Ω_1, Ω_1^* should satisfy

$$\Omega(x_1|x_2) = \Omega^*(x_2|x_1), \tag{2.4a}$$

$$\Omega^*(x_1|x_2) = \Omega(x_2|x_1) \tag{2.4b}$$

(henceforth the suffix "1" is dropped from Ω_1). In order to include the constraints (2.3) and (2.4), we shall appeal to the method of Lagrange multipliers; therefore, consider the new Lagrangian density,

$$\begin{aligned} \mathcal{L} = {} & \mathcal{L}' + \xi(\Phi^*_{12} + \Phi^*_{21}) + \xi^*(\Phi_{12} + \Phi_{21}) \\ & + \lambda(\Omega^*_{12} - \Omega_{21}) + \lambda^*(\Omega_{12} - \Omega^*_{21}), \end{aligned} \tag{2.5}$$

where we have adopted the following convention:

$$f(x_1|x_2) \equiv f_{12}, \quad f = \Phi, \Omega.$$

From Eq. (2.2) we infer, with the new Lagrangian density,

$$\frac{\partial\mathcal{L}}{\partial\theta_j} - \partial_K\frac{\partial\mathcal{L}'}{\partial\partial_K\theta_j} = 0, \tag{2.6}$$

since in (2.5) there are no derivatives in $\mathcal{L} - \mathcal{L}'$.

We find it convenient to separate $\mathcal{L}$,

$$\mathcal{L} = \mathcal{L}_{LT} + \mathcal{L}_{NLT},$$

where $\mathcal{L}_{LT}$ will lead us, through the action principle, to linear terms in the field equations, while $\mathcal{L}_{NLT}$ will lead us to nonlinear terms. We submit that

$$\begin{aligned}\mathcal{L}_{LT} = {} & i\Omega_{12}^{*}\partial_t\Omega_{12} \\ & + \frac{1}{2m}\partial_{\mathbf{x}_1}\Omega_{12}^{*}\cdot\partial_{\mathbf{x}_1}\Omega_{12} - \frac{1}{2m}\partial_{\mathbf{x}_2}\Omega_{12}^{*}\cdot\partial_{\mathbf{x}_2}\Omega_{12} \\ & + \lambda(\Omega_{12}^{*} - \Omega_{21}) + \lambda^{*}(\Omega_{12} - \Omega_{21}^{*}) + i\Phi_{12}^{*}\partial_t\Phi_{12} \\ & + \frac{1}{2m}\partial_{\mathbf{x}_1}\Phi_{12}^{*}\cdot\partial_{\mathbf{x}_1}\Phi_{12} + \frac{1}{2m}\partial_{\mathbf{x}_2}\Phi_{12}^{*}\cdot\partial_{\mathbf{x}_2}\Phi_{12} \\ & + V(\mathbf{x}_1 - \mathbf{x}_2)\Phi_{12}^{*}\Phi_{12} + \xi(\Phi_{12} + \Phi_{21}) + \xi^{*}(\Phi_{12}^{*} + \Phi_{21}^{*}).\end{aligned}$$

The pair potential μ is defined by

$$\mu_{12} = V(\mathbf{x}_1 - \mathbf{x}_2)\Phi_{12}.$$

We find that $\mathcal{L}_{NLT}$ is of the following form:

$$\mathcal{L}_{NLT} = \Phi_{12}^{*}P_{12} + \Omega_{12}^{*}S_{12} + T_{12} + U_{12} + \text{c.c.},$$

where c.c. means the complex conjugate expression; the P and S functions are given by

$$P_{12} = \int dx'\mu(x_1|x')\Omega^{*}(x'|x_2) + \int dx'\Omega(x_1|x')\mu(x'|x_2),$$

$$S_{12} = \int dx'\Phi(x_1|x')\mu^{*}(x'|x_2) - \int dx'\mu(x_1|x')\Phi^{*}(x'|x_2).$$

Finally, the "counter terms" T and U are given by

$$T_{12} = -\tfrac{1}{2}\int dx'\Omega(x_1|x')\mu^{*}(x'|x_2)\Phi(x'|x_2),$$

$$U_{12} = -\tfrac{1}{2}\int dx'\Omega(x'|x_2)\mu^{*}(x_1|x')\Phi(x_1|x').$$

Following the standard procedure, we consider ξ, ξ^{*}, λ, and λ^{*} as additional variables. This leads us to the appropriate field equations.

(a) *The Ω equation.* Making variations with respect to Ω_{12}, Ω_{12}^{*}, λ, and λ^{*}, respectively, we find,

$$-\lambda^{*} + \frac{\partial\mathcal{L}_{NLT}}{\partial\Omega_{12}} = i\partial_t\Omega_{12}^{*} + \frac{1}{2m}\partial_{\mathbf{x}_1}^{2}\Omega_{12}^{*} - \frac{1}{2m}\partial_{\mathbf{x}_2}^{2}\Omega_{12}^{*}, \tag{2.7}$$

$$\lambda + i\partial_t\Omega_{12} + \frac{\partial\mathcal{L}_{NLT}}{\partial\Omega_{12}^{*}} = \frac{1}{2m}\partial_{\mathbf{x}_1}^{2}\Omega_{12} - \frac{1}{2m}\partial_{\mathbf{x}_2}^{2}\Omega_{12}, \tag{2.8}$$

$$\Omega_{12}^{*} = \Omega_{21}, \tag{2.9}$$

$$\Omega_{21}^{*} = \Omega_{12}. \tag{2.10}$$

Equations (2.7) and (2.8) imply

$$-2i\partial_t\Omega_{12}^{*} = \frac{1}{m}\left(\partial_{\mathbf{x}_1}^{2} - \partial_{\mathbf{x}_2}^{2}\right)\Omega_{12}^{*} - \frac{\partial\mathcal{L}_{NLT}}{\partial\Omega_{12}} - \left(\frac{\partial\mathcal{L}_{NLT}}{\partial\Omega_{12}^{*}}\right)^{*},$$

but the last equations can be simplified to

$$-i\partial_t\Omega_{12}^{*} = \frac{1}{2m}\left(\partial_{\mathbf{x}_1}^{2} - \partial_{\mathbf{x}_2}^{2}\right)\Omega_{12}^{*} - \frac{\partial\mathcal{L}_{NLT}}{\partial\Omega_{12}}. \tag{2.11}$$

We further notice that the explicit form of the equation complex conjugate to (2.11) is identical to the BCS Eq. (1.3).

(b) *The Φ equation.* In a similar manner we find that the equation of motion for the macrowave function is given by

$$\begin{aligned}i\partial_t\Phi_{12} = {} & \left(\frac{1}{2m}(\partial_{\mathbf{x}_1}^{2} + \partial_{\mathbf{x}_2}^{2}) - V(\mathbf{x}_1 - \mathbf{x}_2)\right)\Phi_{12} \\ & - \int dx'\mu(x_1|x')\Omega^{*}(x'|x_2) - \int dx'\Omega(x_1|x')\mu(x'|x_2) + \alpha_{12}.\end{aligned} \tag{2.12}$$

With the given Lagrangian density we find that the α_{12} function is given by

$$\alpha_{12} = (\Omega_{12}^{*}\Phi_{11} - \Omega_{12}\Phi_{22})V(\mathbf{x}_1 - \mathbf{x}_2). \tag{2.13}$$

We remark that in theories where α is negligible Eq. (2.12) coincides with the corresponding equation of the BCS theory, i.e., with Eq. (1.4).

We now restrict our discussion to bulk superconductors in the presence of a weak magnetic field H, such that $0 < H \lesssim H_{c1}$, where H_{c1} is the first critical value for penetration of magnetic flux. Under these conditions it is reasonable to assume that the field variables θ_j vary slowly as functions of the spatial coordinates $\mathbf{r}$ and $\mathbf{R}$, where

$$\mathbf{R} = \tfrac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2), \qquad \mathbf{x}_1 - \mathbf{x}_2 = \mathbf{r}.$$

In the Appendix we have shown that α is proportional to the small quantity $\partial_{\mathbf{R}}\Phi(\mathbf{R}, o)$ and is therefore negligible.

3. CONSEQUENCES OF THE ACTION PRINCIPLE

In this section we proceed to discuss some simple consequences of the theory when the magnetic field is coupled. As discussed in Sec. 4 below, this is simply achieved by modifying the expression for the energy of the Bloch electrons: The velocity $\mathbf{k}/m$ is replaced by $[\mathbf{k} - e\mathbf{A}(\mathbf{x})/c]/m$, where $\mathbf{A}$ is the vector potential; the total potential in (2.12) becomes

$$V(\mathbf{x}_1 - \mathbf{x}_2) + \sum_{i=1}^{2} e\phi(\mathbf{x}_i),$$

and in Eq. (1.3) the term $(1/2m)\left(\partial_{\mathbf{x}_1}^{2} - \partial_{\mathbf{x}_2}^{2}\right)\Omega_{12}$ becomes

$$\frac{1}{2m}\left[\left(\partial_{\mathbf{x}_1} - (e/c)\mathbf{A}(\mathbf{x}_1)\right)^{2} - \left(\partial_{\mathbf{x}_2} - (e/c)\mathbf{A}(\mathbf{x}_2)\right)^{2}\right] + \sum_{i=1}^{2}(-1)^{i}e\phi(\mathbf{x}_i),$$

where ϕ is the scalar potential. Then, the field equations are invariant under the following infinitesimal transformations:

$$\phi(\mathbf{x}) \to \phi(\mathbf{x}) - (1/c)\,\partial_t\Lambda(\mathbf{x}, t), \tag{3.1a}$$

$$\mathbf{A}(\mathbf{x}) \to \mathbf{A}(\mathbf{x}) + \partial_{\mathbf{x}}\Lambda(\mathbf{x}, t), \tag{3.1b}$$

$$\psi(\mathbf{x}) \to \psi(\mathbf{x})\ \exp\Gamma(ie/c)\Lambda(\mathbf{x}, t)], \tag{3.1c}$$

$$\psi^{\dagger}(\mathbf{x}) \to \psi^{\dagger}(\mathbf{x})\ \exp[-(ie/c)\Lambda(\mathbf{x}, t)]. \tag{3.1d}$$

For *null* electric field one can demonstrate the invariance of the field equations under the transformation (3.1) in the case we discuss here, namely $\alpha_{12} = 0$ (cf. Ref. 2, pp. 74–77). Such gauge invariance has the following consequences.

A. Conservation of Cooper pairs

We first notice that when we make the gauge transformation the field variables transform as

$$\theta_j \to \theta_j' = \theta_j + \Lambda(\mathbf{x}_1, t)G_j + \Lambda(\mathbf{x}_2, t)H_j, \tag{3.2}$$

where $G_j = G_j[\{\theta\}]$ and $H_j = H_j[\{\theta\}]$. For example, for $j = 1$, $G_1 = i\Phi_{12}$, $H_1 = i\Phi_{12}$.

Then, as shown in the Appendix (Sec. B), making first-order variations in the gauge function Λ we are able to define a bilocal current

$$j^{\alpha}(x,t) = \int dx_1 \frac{\partial \mathcal{L}(x_1|x)}{\partial \partial_\alpha \Lambda(\mathbf{x},t)} + \int dx_2 \frac{\partial \mathcal{L}(x|x_2)}{\partial \partial_\alpha \Lambda(\mathbf{x},t)}, \tag{3.3}$$

whose divergence may be found as usual (with the Euler–Lagrange equations), we find a null divergence and hence the theory will contain a conserved charge, given by

$$\int dx\, j^0(x,t) = \text{const} = -2Q. \tag{3.4}$$

Using the definition of j^0 in Eq. (3.3), we find

$$-2Q = \int dx_1 dx_2 \frac{\partial \mathcal{L}(x_1|x_2)}{\partial \partial_t \theta_j} G_j + \int dx_1 dx_2 \frac{\partial \mathcal{L}(x_1|x_2)}{\partial \partial_t \theta_j} H_j,$$

having changed from the gauge functions Λ to the field variables (using the chain rule), the appropriate relation being given by Eq. (3.2). Finally, making the relevant differentiations, we find

$$Q = \int dx_1 dx_2 \, |\Phi(x_1|x_2)|^2.$$

In other words, we have shown that the number of pairs is the conserved quantity which corresponds to the gauge symmetry of the second kind of the field equations. Noether's theorem[4] assured us that such a conserved quantity exists in our formalism, once we knew that the field equations were invariant under the transformation (3.1).

B. Generalization of a Ginzburg-Landau equation

In order to extend the theory to allow a description of type II superconductors, we modify the Lagrangian density $\mathcal{L}$ by the coupling of the electromagnetic field, as explained at the beginning of Sec. 3. We therefore add a further piece to the density $\mathcal{L}$, to wit, $\mathcal{L}_{\text{EM}} = -H^2/2\pi$. This does not alter any of the Euler–Lagrange equations (because when we make variations with respect to the fields θ_j, $\mathcal{L}_{\text{EM}}$ does not contribute). In the Appendix we have made variations with respect to the new field variable $\mathbf{A}$ (cf. Sec. C of the Appendix) and find

$$\frac{1}{2\pi}\partial_{\mathbf{x}_\Lambda}\partial_{\mathbf{x}_\Lambda}\mathbf{A}(\mathbf{x}) + \frac{ie\hbar}{mc}\int dy\{\Omega_1(x|y)\partial_{\mathbf{x}}\Omega_1^*(x|y) - \Omega_1^*(x|y)\partial_{\mathbf{x}}\Omega_1(x|y)\} - \frac{2e^2}{mc^2}\mathbf{A}(\mathbf{x})\int dy\,|\Phi(x|y)|^2 = 0. \tag{3.5}$$

This equation is a natural generalization of the Maxwell-type Ginzburg–Landau equation and, in particular, will define a supercurrent (generated by the magnetic field $\mathbf{H}$) which will be typical of the Lagrangian theory.[5]

If in analogy with the corresponding case of the order parameter of the Ginzburg–Landau equations, we write the density matrix into its modulus and phase

$$\Omega_1(x|y) = \Omega_0(x|y)\,\exp[-i\omega(x|y)], \tag{3.6}$$

we find that the current defined by (3.5) becomes

$$\mathbf{J}_s(\mathbf{x}) = -\frac{4\pi e\hbar}{mc}\int dy\,\Omega_0^2(x|y)\partial_{\mathbf{x}}\omega(x|y) + \frac{4\pi e^2}{mc^2}\mathbf{A}(\mathbf{x})\int dy\,|\Phi(x|y)|^2. \tag{3.7}$$

C. Quantization of flux in type II superconductors

Let us consider a large sample of pure type II superconductor at very low temperatures (essentially at absolute zero, where our equations are valid). We examine the problem of magnetic flux penetration for a field close to the first critical value H_{c_1}; for this purpose we need to consider the local limit of $\Omega_1(x'|x'')$, namely, $x' = x''$, $\Omega = \tilde{\sigma}$, the macroscopic density of the whole system. Our first simplifying assumption is *spatial homogeneity*, $\tilde{\sigma}(x) = \tilde{\sigma}$ where $\tilde{\sigma}$ is a constant. For our discussion we also need the second reduced density matrix

$$\Omega_2(x'y'|x''y'') = \text{Tr}\rho(t)\psi^\dagger(x'')\psi^\dagger(y'')\psi(y')\psi(x'),$$

which may be shown to satisfy the following properties:

$$\int \Omega_2(x'y|x''y)\,dy = (N-1)\Omega_1(x'|x''), \tag{3.8}$$

$$\Omega_2(x'y|x''y) = -\Omega_1(x'|y)\Omega_1(y|x'') + \tilde{\sigma}(y)\Omega_1(x'|x'') + \Phi^*(x''|y)\Phi(x'|y), \tag{3.9}$$

where N is the total number of particles.

We integrate both sides of Eq. (3.9) in the local case $x' = x''$,

$$\int \Omega_2(xy|xy)\,dy = -\int |\Omega_1(x|y)|^2 dy + \tilde{\sigma}(x)\int \tilde{\sigma}(y)\,dy + \int |\Phi(x|y)|^2 dy. \tag{3.10}$$

We may deduce, from Eq. (3.10), the following relation:

$$\int |\Phi(x|y)|^2 dy = \int |\Omega_1(x|y)|^2 dy, \tag{3.11}$$

since $N - 1 \cong N$, N being a very large number. Another property which the density matrices obey, in the case of superconductivity, is ODLRO,[6]

$$\Omega_1(x'|x'') \to 0, \qquad |x' - x''| \to \infty$$
$$\Omega_2(x'x_1'|x''x_1'') \to \Phi(x'x_1')\Phi^*(x''x_1''),$$

for x', x_1' sufficiently far from x'', x_1''. We only need the first part of the above statement, which we use to make the following approximation:

$$\Omega_0^2(x|y) = \gamma\delta_{\sigma_1,\sigma_2}\Omega_0^2(y|y)\delta(\mathbf{x}-\mathbf{y}). \tag{3.12}$$

This expresses the *strong localization* of the square of the real part of Ω_1 [the notation was introduced in Eq. (3.7)]. In writing (3.12) we have also included the assumption of antiparallel spin pairing and, finally, the constant factor γ has been inserted, whose units are those of volume, since by definition $\delta(\mathbf{x}-\mathbf{y})$ has units of (volume)$^{-1}$. We are now in a position to take a second look at our new current (3.7). Using our expression for the square of the real part of the density matrix (3.12), we find

$$\mathbf{J}_s(\mathbf{x}) = -\frac{4\pi e\hbar}{mc}\gamma\tilde{\sigma}^2\lim_{\mathbf{x}\to\mathbf{y}}\partial_{\mathbf{x}}\omega(\mathbf{x}|\mathbf{y}) + \frac{4\pi e^2}{mc^2}\gamma\tilde{\sigma}^2\mathbf{A}(\mathbf{x}), \tag{3.13}$$

where the second term on the right-hand side of (3.13) follows from (3.11). Since $\omega(\mathbf{x}|\mathbf{y})$ is the phase of Ω_1, we suppose it is a smooth enough function such that, after differentiating, the limit will exist and be written as

$$\lim_{\mathbf{x}\to\mathbf{y}}\partial_{\mathbf{x}}\omega(\mathbf{x}|\mathbf{y}) = \mathbf{W}(\mathbf{x}) \equiv (W_1, W_2).$$

The pair of functions is best studied in the complex plane for the following reason: From the nature of the problem we find that by choosing a cylindrical triad of coordinates with z axis along the single flux line, the problem has z symmetry. Let us take $\mathbf{x} = (x,y)$, where x and y are Cartesian coordinates. We find it is useful to introduce the new function of a single complex variable $\zeta = x + iy$,

$$W(\zeta) = W_2(x,y) + iW_1(x,y).$$

Then, if L is some arbitrary closed contour in the ζ plane,

$$\int_L W(\zeta)\,d\zeta = \int_L (W_2 dx - W_1 dy) + i\int_L (W_1 dx + W_2 dy).$$

Hence, we find that,

$$\mathrm{Im}\int_L W(\zeta)\,d\zeta = \int_L \mathbf{W}(\mathbf{x})\cdot d\mathbf{l}, \qquad d\mathbf{l} = (dx, dy). \tag{3.14}$$

Integrating Eq. (3.13) round some contour L inside the superconducting sample, we may neglect the line integral of $\mathbf{J}_s$ for, by appropriate choice of contour, this quantity is negligible, since $\mathbf{J}_s$ will be strongly damped (in London's theory, it would decay exponentially). By Stokes' theorem, we have

$$\int_L \mathbf{A}\cdot d\mathbf{l} = \int_S \mathbf{B}\cdot d\mathbf{S} = F. \tag{3.15}$$

Hence, from Eq. (3.13) we infer

$$(e/\hbar c)F = Im\int_L d\zeta W(\zeta), \tag{3.16}$$

having used Eqs. (3.14) and (3.15), together with the strong damping of $\mathbf{J}_s$. For nonvanishing magnetic flux penetration, we find from Eq. (3.16) that there will be a singularity in the complex ζ plane inside the contour L. We shall assume the simplest type, namely a simple pole at the origin of the ζ plane. Therefore, we can make the following Laurent expansion for W,

$$W(\zeta) = f(\zeta) + b_1/\zeta,$$

where f is its Taylor part and b_1 its residue at the origin. By Cauchy's residue theorem, the integral round the contour L does not contribute and all reduces to the contribution of a small circle round the origin, which we call D: $\{\zeta = \delta\exp(i\theta);\ |\zeta| = \delta\}$. Hence, we have

$$\int_D W(\zeta)\,d\zeta = \int_0^{2\pi} [b_1\delta i\exp(i\theta)/\delta\ \exp i\theta]\,d\theta = 2\pi i b_1.$$

Therefore, we find, from Eq. (3.16) that $F = 2\pi b_1 c\hbar/e$. Since we know independently that flux is quantized in units of hc/e where e is twice the electronic charge, we may determine the residue $b_1 = 1$, and hence the Laurent expansion is also determined. Thus although the formal expression for the superconducting current is somewhat complicated, Eq. (3.7), in this section we have found that by simple physical assumptions it may be brought into a form which is equivalent to the phenomenological expression of Ginzburg and Landau.

4. DISCUSSION AND CONCLUSION

In studying the coupling between pairing and magnetic field, as we did in the previous section, several reasonable assumptions were taken for granted; we mention the following[7]:

(i) The two-body pairing interaction is taken to be independent of the magnetic field.

(ii) The effect of the field manifests itself only through the single particle energy operator

$$\epsilon(x) = \frac{1}{2m}(-i\hbar\partial_{\mathbf{x}})^2 - E_F, \tag{4.1}$$

which is assumed to become

$$\epsilon(x)_A = \frac{1}{2m}(-i\hbar\partial_{\mathbf{x}} - e\mathbf{A}/c)^2 - E_F. \tag{4.2}$$

(iii) The effect of the magnetic field on the Fermi energy E_F is neglected.

(iv) Small contributions from the pairing interaction, which one would have to include in ϵ_A are neglected.

With the limits of validity discussed in this work, we summarize by saying that the basic result of the paper is the action principle of Sec. 2 with the corresponding Lagrangian density given in Eq. (2.5). The principle gives back some of the physical consequences of the BCS theory under appropriate slowly varying conditions. Looking at superconductivity from the Lagrangian viewpoint clarifies the connection between gauge symmetry and conservation of pairs.

ACKNOWLEDGMENTS

The author is grateful to Professor Abdus Salam, the International Atomic Energy Agency and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste. He is also grateful for an Associateship at the ICTP which allowed him to conclude this work.

APPENDIX

A. Proportionality between α_{12} and $\partial_R\Phi$

For simplicity we omit the spin label σ. In the $\mathbf{R}, \mathbf{r}$ coordinates, we have

$$\Omega_{12} = \bar{\Omega}(\mathbf{R},\mathbf{r}) = \bar{\Omega}(\mathbf{R},o) + \mathbf{r}\cdot\partial_{\mathbf{R}}\bar{\Omega}(\mathbf{R},o) + \cdots,$$

$$\Omega_{12}^* = \bar{\Omega}^*(\mathbf{R},\mathbf{r}) = \bar{\Omega}^*(\mathbf{R},o) + \mathbf{r}\cdot\partial_{\mathbf{R}}\bar{\Omega}^*(\mathbf{R},o) + \cdots.$$

Then, retaining terms up to first derivatives, $\bar{\Omega}(\mathbf{R},\mathbf{r}) = \bar{\Omega}^*(\mathbf{R},\mathbf{r})$, since $\bar{\Omega}(\mathbf{R},o)$ is real. We remark that in the new coordinates

$$\Phi(\mathbf{x}_1|\mathbf{x}_1) = \bar{\Phi}\left(\frac{\mathbf{x}_1+\mathbf{x}_2}{2}, \mathbf{x}_1-\mathbf{x}_2\right)\Big|_{\mathbf{x}_1=\mathbf{x}_2} = \bar{\Phi}(\mathbf{x}_1,o) = \bar{\Phi}(\mathbf{R}+\mathbf{r}/2,o),$$

$$\Phi(\mathbf{x}_2|\mathbf{x}_2) = \bar{\Phi}\left(\frac{\mathbf{x}_1+\mathbf{x}_2}{2}, \mathbf{x}_1-\mathbf{x}_2\right)\Big|_{\mathbf{x}_1=\mathbf{x}_2} = \bar{\Phi}(\mathbf{x}_2,o) = \bar{\Phi}(\mathbf{R}-\mathbf{r}/2,o).$$

Now, since the field variable varies slowly in $\mathbf{R}$,

$$\bar{\Phi}(\mathbf{R}+\mathbf{r}/2,o) \cong \bar{\Phi}(\mathbf{R},o) + \mathbf{r}/2\cdot\partial_{\mathbf{R}}\bar{\Phi}(\mathbf{R},o),$$

$$\bar{\Phi}(\mathbf{R}-\mathbf{r}/2,o) \cong \bar{\Phi}(\mathbf{R},o) - \mathbf{r}/2\cdot\partial_{\mathbf{R}}\bar{\Phi}(\mathbf{R},o).$$

Then, we find from Eq. (4.8) that

$$\alpha_{12} = [\bar{\Omega}(\mathbf{R},o) + \mathbf{r}\cdot\partial_{\mathbf{R}}\bar{\Omega}(\mathbf{R},o)][\mathbf{r}\cdot\partial_{\mathbf{R}}\bar{\Phi}(\mathbf{R},o)]V(\mathbf{r}),$$

and conclude that α_{12} is proportional to the small quantity $\partial_{\mathbf{R}}\Phi(\mathbf{R},o)$.

B. The bilocal current and its divergence

Making variations up to first order in $\delta\Lambda$, we obtain

$$O = \int dt dx_1 dx_2 \delta\mathcal{L} = \int dt dx_1 dx_2 \left(\frac{\partial\mathcal{L}_{12}}{\partial\Lambda(x_1,t)}\delta\Lambda(\mathbf{x}_1,t) + \frac{\partial\mathcal{L}_{12}}{\partial\Lambda(\mathbf{x}_2,t)}\delta\Lambda(\mathbf{x}_2,t) + \frac{\partial\mathcal{L}_{12}}{\partial\partial_t\Lambda(\mathbf{x}_1,t)}\delta\partial_t\Lambda(\mathbf{x}_1,t) + \frac{\partial\mathcal{L}_{12}}{\partial\partial_t\Lambda(\mathbf{x}_2,t)}\delta\partial_t\Lambda(\mathbf{x}_2,t) + \frac{\partial\mathcal{L}_{12}}{\partial\partial_{\mathbf{x}_1}\Lambda(\mathbf{x}_1,t)}\delta\partial_{\mathbf{x}_1}\Lambda(\mathbf{x}_1,t) + \frac{\partial\mathcal{L}_{12}}{\partial\partial_{\mathbf{x}_2}\Lambda(\mathbf{x}_2,t)}\delta\partial_{\mathbf{x}_2}\Lambda(\mathbf{x}_2,t)\right).$$

Equivalently, since x_1 and x_2 are dummy variables,

$$O = \int dt \int dx \left[\delta\Lambda(\mathbf{x}, t)\left(\int \frac{\partial \mathcal{L}(x \mid x_2)}{\partial \Lambda(\mathbf{x}, t)}\, dx_2 + \int \frac{\partial \mathcal{L}(x_1 \mid x)}{\partial \Lambda(\mathbf{x}, t)}\, dx_1\right)\right.$$
$$\left. + \delta\partial_\alpha \Lambda(\mathbf{x}, t)\left(\int \frac{\partial \mathcal{L}(x \mid x_2)}{\partial \partial_\alpha \Lambda(\mathbf{x}, t)}\, dx_2 + \int \frac{\partial \mathcal{L}(x_1 \mid x)}{\partial \partial_\alpha \Lambda(\mathbf{x}, t)}\, dx_1\right)\right] \quad \text{(A1)}$$

Then we define the current (3.3) and evaluate its divergence applying the chain rule,

$$\partial_\alpha j^\alpha(x, t) = \int dx_2 \partial_\alpha \frac{\partial \mathcal{L}(x \mid x_2)}{\partial \partial_\alpha \theta_j} H_j + \int dx_1 \partial_\alpha \frac{\partial \mathcal{L}(x_1 \mid x)}{\partial \partial_\alpha \theta_j} G_j \quad \text{(A2)}$$

If we now invoke the Euler–Lagrange equations, we obtain

$$\partial_\alpha j^\alpha(x, t) = \int dx_2 \frac{\partial \mathcal{L}(x \mid x_2)}{\partial \Lambda(\mathbf{x}, t)} + \int dx_1 \frac{\partial \mathcal{L}(x_1 \mid x)}{\partial \Lambda(\mathbf{x}, t)}. \quad \text{(A3)}$$

Then, if the condition

$$\int dx_2 \frac{\partial \mathcal{L}(x \mid x_2)}{\partial \Lambda(\mathbf{x}, t)} + \int dx_1 \frac{\partial \mathcal{L}(x_1 \mid x)}{\partial \Lambda(\mathbf{x}, t)} = 0$$

is fulfilled or, equivalently, if

$$\frac{\partial}{\partial \Lambda(\mathbf{x}, t)} \{\mathcal{L}(x_1 \mid x) + \mathcal{L}(x \mid x_1)\} = 0 \quad \text{(A4)}$$

is fulfilled, then the theory will contain a zero divergence current and, hence, a charge. It is not difficult to verify that our Lagrangian density does satisfy condition (A4).

C. The superconducting current

The third Euler–Lagrange equation is obtained making variations with respect to the gauge field **A**:

$$\int dt dx_1 dx_2 \left\{\frac{\partial \mathcal{L}(x_1, x_2 \mid t)}{\partial \mathbf{A}(\mathbf{x}_1)} \cdot \delta\mathbf{A}(\mathbf{x}_1) + \frac{\partial \mathcal{L}(x_1, x_2 \mid t)}{\partial \mathbf{A}(\mathbf{x}_2)} \cdot \delta\mathbf{A}(\mathbf{x}_2)\right\}$$
$$+ \int dx \frac{\partial \mathcal{L}_{\mathrm{EM}}(\mathbf{x})}{\partial \mathbf{A}(\mathbf{x})} \cdot \delta\mathbf{A}(\mathbf{x}) = 0.$$

We remark that under the integral sign, x_1 and x_2 may be treated as dummy variables; hence

$$\int dt \int d\mathbf{x}\, \delta\mathbf{A}(\mathbf{x}) \cdot \left\{\int dx_2 \frac{\partial \mathcal{L}(x, x_2 \mid t)}{\partial \mathbf{A}(\mathbf{x})} + \int dx_1 \frac{\partial \mathcal{L}(x_1, x \mid t)}{\partial \mathbf{A}(\mathbf{x})}\right.$$
$$\left. + \frac{\partial \mathcal{L}_{\mathrm{EM}}(\mathbf{x})}{\partial \mathbf{A}(\mathbf{x})}\right\} = 0.$$

For arbitrary variations of $\delta\mathbf{A}(\mathbf{x})$ we have

$$-\frac{\partial \mathcal{L}_{\mathrm{EM}}}{\partial \mathbf{A}(\mathbf{x})} = \int dx_1 \frac{\partial \mathcal{L}(x_1, x \mid t)}{\partial \mathbf{A}(\mathbf{x})} + \int dx_2 \frac{\partial \mathcal{L}(x, x_2 \mid t)}{\partial \mathbf{A}(\mathbf{x})}.$$

After some manipulation, we find that the above equation becomes Eq. (3.5).

[1] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. 108, 1175 (1957).
[2] J. G. Valatin, "Superconducting electron and nucleon systems," in *Lectures in Theoretical Physics*, edited by W. S. Brittin, B. W. Downs, and J. Downs (Interscience, New York, 1962), Vol. IV, pp. 1-120.
[3] J. Chela-Flores, Collective Phenomena 1, 5 (1972).
[4] E. L. Hill, Rev. Mod. Phys. 23, 253 (1951).
[5] G. V. Ginzburg and L. Landau, Zh. Eksp. Teor. Fiz. 20, 1064 (1950).
[6] C. N. Yang, Rev. Mod. Phys. 34, 694 (1962).
[7] W. Young, C. di Castro, and J. G. Valatin, Nuovo Cimento 48B, 19 (1967).

Some consequences of the strengthened interpretative rules of quantum mechanics*

Paul Benioff

Argonne National Laboratory, Argonne, Illinois 60439
(Received 26 July 1973)

In this paper some consequences of the strengthened interpretative rules of quantum mechanics, which were proposed in an earlier paper, are obtained. It is also seen that, in general, the usual interpretative rules are too weak to obtain these results. For example, it is proved from the strengthened rules that if $f: R \to R$ is a Borel function which is also τ-definable, then for each observable A the procedure "measure A and compute $f(A$ outcome)" is an $f(A)$ measurement procedure. It is also shown that there exist Borel f and observables A such that the above procedure is not an $f(A)$ measurement procedure. Two methods of measuring the sum $A + B$ of two observables are then considered: The measurement of A and B on different systems followed by addition of the results and (if A and B commute) the simultaneous measurement of A and B on the same system followed by addition. It is proved from the strengthened rules that the first method is not a valid measurement procedure and the second is valid. Besides these, other processes such as procedures for preparing mixtures of different states, and the empirical generation of probability measures from outcome sequences are considered.

I. INTRODUCTION

In a previous paper[1] denoted here by I, a strengthening of the usual interpretative rules of quantum mechanics was proposed and discussed. In this paper some further consequences of the strengthened rules will be investigated. In Sec. II some definitions are given along with a review of the material given in I. In Sec. III Borel functions of observables are considered. It is first shown constructively that for any observable A with a nonempty continuous spectra, there exist Borel functions f such that the procedure "measure A and compute f on the outcome" is not an $f(A)$ measuring procedure.

It is then proved that the strengthened rule $3)_\tau$ implies that for each τ-definable Borel function the above procedure is an $f(A)$ measuring procedure for *all* observables. This includes the τ-definable polynomials. However, for those polynomial f which are not τ-definable it is open whether or not the above procedure measures $f(A)$ for all observables. By "open" is meant that we are not able to *prove* for these f whether or not the above procedure measures $f(A)$. This use of "open" in different contexts will appear several times later on. On the other hand, the usual interpretative rules are seen to be too weak in that one can not prove useful results from them even for polynomial f.

In the next section the linear operations are considered. After a brief discussion about multiplication by real numbers as a special case of polynomial f, two addition operations are considered.

It is first proved from a slight extension of rule $3)_\tau$ that the procedure "measure A on a system prepared in some state, measure B on a different system prepared in the same state and add the results" is not an $A + B$ measuring procedure. This result does not follow from the usual interpretative rules.

Then it is proved from another extension of rule $3)_\tau$ that if A and B commute, the simultaneous measurement (if such a procedure exists) of A and B on one system followed by addition of the results is an $A + B$ measuring procedure. It is satisfying that both these results are in accord with one's intuition.

In the previous sections much of the full power of the strengthened rules was not used. In Sec. IV this is remedied by considering the generation of empirical probability measures on the Borel subsets of R^ω (the set of all infinite sequences of real numbers) from outcome sequences. It is shown that rule $3)_\tau$ implies that for each preparation procedure s and each observation procedure α, the empirical probability measure $U_T(\psi_{s\alpha})$, generated in a standard way from the outcome sequence $\psi_{s\alpha}$, obtained from an infinite repetition of doing α and s, equals the theoretical measure generated from the state and spectral measure assigned to s and α. It is noted that one has here an even stronger incompleteness result than for the Borel functions in that for no outcome sequence $\psi_{s\alpha}$ is the limit relative frequency function $\bar{I}^T_{(-)}(\psi_{s\alpha})$ a probability measure on $\mathfrak{B}(R^\omega)$.

In Sec. V state preparation procedures are considered with particular attention paid to the mixing of different procedures. Let s_φ and $s_{s_2\beta}$ be two different procedures each of which mixes the inequivalent preparation procedures s_0 and s_1 in the ratio λ_0/λ_1 with $\lambda_0 + \lambda_1 = 1$. s_φ is the procedure "for the jth preparation compute φ, if $\varphi(j) = 0(1)$ carry out $s_0(s_1)$" (φ is a 0-1 sequence). $s_{s_2\beta}$ is the procedure "carry out s_2 and β where β is a projection operator measuring procedure, if outcome is $0(1)$ carry out $s_0(s_1)$." Both φ and the outcome sequence $\psi_{s_2\beta}$ of an infinite repetition of doing s_2 and β are such that $\bar{M}\varphi = \bar{M}\psi_{s_2\beta} = \lambda_1$.

It is proved from rule $3)_\tau$ that if φ is τ-definable then s_φ is not a state preparation procedure. If φ is τ-random then the above proof fails and it is an open question whether or not s_φ is a state preparation procedure. Similarly, it is open whether or not $s_{s_2\beta}$ is a state preparation procedure.

In conclusion it is noted that these problematic aspects disappear for those s which do not include mixing processes, such as preparation procedures for the pure states.

II. DEFINITIONS AND REVIEW

Let $\mathbb{S}$ and $\mathbb{O}$ denote the respective collections of nonmathematical[2] state preparation procedures and observation procedures. Let $B(\mathcal{H})$ be the set of all bounded linear operators over some Hilbert space $\mathcal{H}$ and let $\mathfrak{A}$ be a von Neumann subalgebra of $B(\mathcal{H})$ and let $S(\mathfrak{A})$ be the set of states on $\mathfrak{A}$. Let Φ and Ψ be maps from $\mathbb{O}$ *into* $\mathfrak{A}$ and from $\mathbb{S}$ *into* $S(\mathfrak{A})$, respectively, where for each α in $\mathbb{O}$, $\Phi(\alpha)$ is self adjoint.

Since the maps Φ and Ψ are in general many-one, they decompose $\mathbb{O}$ and $\mathbb{S}$ into equivalence classes of pro-

cedures in an obvious fashion. Thus for each operator A in the range of Φ, let $\Phi^{-1}(A)$ be the set of all α such that $\Phi(\alpha) = A$ and let $[\alpha]_\Phi$ be the set of all observation procedures which are equivalent under Φ to α. For states ρ and preparation procedures s, one defines $\Psi^{-1}(\rho)$ and $[s]_\Psi$ analogously.

It should be noted that, unlike other work[2] we do not assume that (the set of equivalence classes of) $\mathcal{O}$ and $\mathcal{S}$ are necessarily an algebra and a convex set, respectively. This is taken care of by allowing the maps Ψ and Φ to be into and not necessarily onto. In this way allowance is made for such problems as the possible nonexistence of a measurement procedure for the sum of any pair of noncommuting observables, the possibility that $\mathcal{O}$ and $\mathcal{S}$ may contain at most a denumerable infinity of procedures, etc.

One reason that this work is presented in terms of a von Neumann algebra and set of positive linear functionals over the algebra is that the existence of superselection rules is thereby allowed for. Of course at the cost of excluding superselection, everthing said in this paper can also be given in the usual formalism by considering Φ as a map into $B(\mathcal{H})$ (which it is anyway) and Ψ as a map into the set of density operators in $B(\mathcal{H})$. In this approach every occurrence of $\Psi(s)(\cdot)$ in this paper is replaced by $\mathrm{Tr}(\Psi(s) -)$.

In I it was proposed to replace the expectation rule and the spectrum rule of the usual interpretative rules by the following stronger rule:

$3)_\tau$: For each s in $\mathcal{S}$ and α in $\mathcal{O}$, the outcome sequence $\psi_{s\alpha}$ obtained from an infinite repetition of carrying out s and α must satisfy: each property of elements of R^ω which is

(a) τ-definable from $\tilde{P}_{s\alpha}$

(b) true $\tilde{P}_{s\alpha}$ almost everywhere

must be true for $\psi_{s\alpha}$.

$\tilde{P}_{s\alpha}$ is the product probability measure on $\mathfrak{B}(R^\omega)$, the set of all Borel subsets of R^ω (the set of all infinite sequences of real numbers) which is generated from $\Psi(s)$ and $\Phi(\alpha)$ according to

$$\tilde{P}_{s\alpha}E_{Fj} = \Psi(s)(\mathcal{E}^{\Phi(\alpha)}(F)) = P_{s\alpha}F \tag{1}$$

for each $j = 0, 1\cdots$ and each F in $\mathfrak{B}(R)$. $E_{Fj} = [\psi|\psi(j)\epsilon F]$ and $\mathcal{E}^{\Phi(\alpha)}$ is the spectral measure for $\Phi(\alpha)$. The right-hand part of Eq. (1) just defines the probability measure $P_{s\alpha}$ on $\mathfrak{B}(R)$ from s and α and will be used later on.

The various concepts appearing above have been discussed in I and thus will be given briefly here. A property p of elements of R^ω is τ-definable from $\tilde{P}_{s\alpha}$ with τ a mathematical theory if there is some formula $Q(\psi, \tilde{\mathbf{P}}_{s\alpha})$ in the language of τ to which a name, $\tilde{\mathbf{P}}_{s\alpha}$, of $\tilde{P}_{s\alpha}$ has been added such that for all sequences ψ in R^ω, ψ has p if and only if $Q(\psi, \tilde{P}_{s\alpha})$ is true. A property is true $\tilde{P}_{s\alpha}$ almost everywhere if the set of all elements of R^ω which have the property is a set of $\tilde{P}_{s\alpha}$ measure 1.

The strength of τ was discussed in I. There it was seen that τ must be sufficiently strong to include probability theory over $\mathfrak{B}(R^\omega)$ and possibly should be as strong as set theory. Here τ will not be specified although it will always be assumed to be sufficiently strong for the various proofs and discussions. If necessary, one can always choose τ to be Zermelo-Frankel set theory which is certainly strong enough for our purposes here.

It was also seen that rule $3)_\tau$ implies the usual expectation value rule and that $3)_\tau$ implies the spectrum rule to the extent that one can prove from $3)_\tau$ that $\sigma_d(\Phi(\alpha)) \subseteq S^\alpha$ and that S^α is dense in $\sigma_c(\Phi(\alpha))$. S^α is the outcome set of α and $\sigma_c(A)$ and $\sigma_d(A)$ are the respective continuous and discrete spectra of A.

The definition of randomness on which rule $3)_\tau$ is based is the following: A sequence ψ in R^ω is *τ-random* if there exists a nontrivial product measure $\tilde{P}$ on $\mathfrak{B}(R^\omega)$ [which is a product of the same measure P on $\mathfrak{B}(R)$] such that every property of elements of R^ω which is τ definable from $\tilde{P}$ and true $\tilde{P}$ almost everywhere must hold for ψ. A product measure $\tilde{P}$ is nontrivial if the measure P on $\mathfrak{B}(R)$, which generates $\tilde{P}$, is not concentrated on a single point of R.

The adequacy of this definition as a proper definition of randomness was briefly discussed in I in light of the fact that it can be proved[3] that if ψ is τ-random, then ψ is not τ-definable from any product measure. Also it is clear that rule $3)_\tau$ implies that if s and α are such that $0 < \Psi(s)(\mathcal{E}^{\Phi(\alpha)}(\{r\})) < 1$ for each real number r, then $\psi_{s\alpha}$ is τ-random. If $\Psi(s)(\mathcal{E}^{\Phi(\alpha)}(\{r\})) = 1$, then rule $3)_\tau$ implies that $\psi_{s\alpha}$ is a constant r sequence.

It should be noted that the essential concept of rule $3)_\tau$ and of the definition of randomness is that of τ-correctness of a probability measure for an outcome. This concept, which is easily extended to general probability measures and not just product measures is given as follows:[3,4] Let Ω be a set, $\mathfrak{B}(\Omega)$ a σ field of subsets of Ω, P a probability measure on $\mathfrak{B}(\Omega)$, and w an element of Ω. *P is τ-correct* for w if all properties of elements of Ω which define sets in $\mathfrak{B}(\Omega)$ and are τ-definable from P and are true P almost everywhere are true for w. Thus rule $3)_\tau$ says for all s and α, $\tilde{P}_{s\alpha}$ defined by Eq. (1) is τ-correct for $\psi_{s\alpha}$. Also a sequence ψ in R^ω is τ-random if there exists a nontrivial product measure on $\mathfrak{B}(R^\omega)$ which is τ-correct for ψ.

Finally, much of the following discussion will use the following precise statement of the meaning of s and α being respective state preparation and observation procedures. s is a state preparation procedure for $\Psi(s)$ if and only if s is in the domain of Ψ and for every observation procedure α, $\tilde{P}_{s\alpha}$, defined by Eq. (1), is τ-correct for $\psi_{s\alpha}$. α is an observation procedure for $\Phi(\alpha)$ if and only if α is in the domain of Φ and for every state preparation procedures s, $\tilde{P}_{s\alpha}$ is τ-correct for $\psi_{s\alpha}$.

The reasonableness of these conditions can be seen as follows: If the strengthened interpretative rules are valid, then these conditions are clearly necessary. The sufficiency of these conditions stems from the following intuitive requirement: If s (or α) lie in the domain of Ψ (or Φ) and if for every α (or s) the infinite repetition of s and α gives an outcome sequence $\psi_{s\alpha}$ which is random [or is a constant r sequence in case $\Psi(s)(\mathcal{E}^{\Phi(\alpha)}(\{r\})) = 1$ for some r] and is such that $\bar{M}\psi_{s\alpha} = \Psi(s)(\Phi(\alpha))$, then s (or α) is a state preparation (or observation) procedure for $\Psi(s)$ [or $\Phi(\alpha)$]. It is also assumed that τ-randomness is a valid precise definition for the intuitive concept of randomness.[1]

III. BOREL FUNCTIONS ON R

As is well known in quantum mechanics each bounded Borel function $f: R \to R$ induces a map $A \to f(A)$ on the set of all self-adjoint operators in $B(\mathcal{H})$ where $f(A)$ is defined for each A by

$$f(A) = \int_R f(r)d\mathcal{E}^A((-\infty, r]) \tag{2}$$

and $\mathcal{E}^A$ is the spectral measure on A. Since no confusion

will result the same symbol is used for the function f and the operator map induced by it.

It is usually assumed in quantum mechanics that if A can be measured then for any bounded Borel f, $f(A)$ can be measured by first measuring A and then computing f(outcome). For example, if $f(A) = A^2$, then one can always measure A^2 by measuring A and squaring the outcome. This assumption can be expressed more precisely as follows: For each $s \in \mathcal{S}$ and $\alpha \in \Theta$ and bounded Borel f,

$$\overline{M} F_f \psi_{s\alpha} = \Psi(s)(f(\Phi(\alpha))) \tag{3}$$

is assumed to hold. $\overline{M}$ denotes the limit mean operation and $F_f \psi_{s\alpha}$ is the outcome sequence of an infinite repetition of s and α followed by an f computation on each outcome. [$F_f : R^\omega \to R^\omega$ is defined by $(F_f \psi)(j) = f(\psi(j))$ for each $j = 0, 1 \cdots$ and each ψ in R^ω.]

Equation (3) relates a measurement of $\Phi(\alpha)$ and subsequent computations to an expectation value computed from $f(\Phi(\alpha))$. Equivalently, one can relate the measurement of $\Phi(\alpha)$ and subsequent f computation directly to a measurement of $f(\Phi(\alpha))$ as follows: For each α and for each f such that $\Phi^{-1}(f(\Phi(\alpha)))$ is not empty,

$$\overline{M} F_f \psi_{s\alpha} = \overline{M} \psi_{s\alpha_f} \tag{4}$$

must hold for each s and each α_f in $\Phi^{-1}(f(\Phi(\alpha)))$.

Equation (4) is the statement that if a carrying out of α followed by an f computation on the outcome is an $f(\Phi(\alpha))$ measuring procedure, then it is empirically equivalent to any other $f(\Phi(\alpha))$ measuring procedure. Of course, if $\Phi^{-1}(f(\Phi(\alpha)))$ includes only the f computation on the result of doing α then Eq. (4) becomes an identity.

The first point to be made is that there exist operators A and Borel f such that measuring f and computing f on the outcome is not an $f(A)$ measuring procedure. More precisely, the validity of Eq. (3) for all s is clearly a necessary condition for the procedure "do α and compute f(outcome)" to be a $f(\Phi(\alpha))$ measuring procedure. We show that if the range of Φ includes operators with nonempty continuous spectra (this condition is not exactly the correct one[5] but is good enough for our purposes) then there exist α and Borel f for which Eq. (3) fails for some s.

To see this let α be such that $\Phi(\alpha)$ has a nonempty continuous spectrum $\sigma_c(\Phi(\alpha))$. Let s be such that $\Psi(s)(\mathcal{E}^{\Phi(\alpha)}[\sigma_c(\Phi(\alpha))]) = 1$, i.e., the probability that carrying out α and s gives an outcome in $\sigma_c(\Phi(\alpha))$ is equal to 1. It follows from this and rule 3)$_\tau$ that $Rng\psi_{s\alpha}$ $\{Rng\psi_{s\alpha} = [r \mid \psi_{s\alpha}(j) = r \text{ for some } j]\}$ lies entirely in $\sigma_c(\Phi(\alpha))$. Define f to be the characteristic function for $Rng\psi_{s\alpha}$ [$f(r) = 1(0)$ if $r \in (\notin) Rng\psi_{s\alpha}$]. f is Borel as $Rng\psi_{s\alpha}$ is a Borel subset of R.

Now $\overline{M} F_f \psi_{s\alpha} = 1$ as $F_f \psi_{s\alpha}$ is a constant sequence of 1's. However, $f(\Phi(\alpha)) = 0$ as by Eq. (1)

$$\int f(r) d\mathcal{E}^{\Phi(\alpha)}((-\infty, r]) = \sum_{r \in Rng\psi_{s\alpha}} \mathcal{E}^{\Phi(\alpha)}(\{r\}) = 0$$

as $Rng\psi_{s\alpha}$ is a countable set and by the definition[6] of continuous spectrum. $\mathcal{E}^{\Phi(\alpha)}\{r\} = 0$ for each r in $Rng\psi_{s\alpha}$. Thus $\Psi(s)(f(\Phi(\alpha))) = 0$ and Eq. (3) is false. Also by rule 3)$_\tau$ $\overline{M}\psi_{s\alpha_f} = \Psi(s)(f(\Phi(\alpha)))$ for any α_f in $\Phi^{-1}(f(\Phi(\alpha)))$ and thus Eq. (4) is false. Note that in this case $\Phi^{-1}(f(\Phi(\alpha)))$ is the equivalence class of procedures for measuring the zero operator.

The following example will help to understand this result better. Let α be such that $\Phi(\alpha)$ is the truncated position operator which gives the position in the interval [2, 3] and 0 elsewhere. Let s be such that $\Psi(s)$ corresponds to a wave packet concentrated within [2, 3]. Suppose too that the experimenter has available a collection of black boxes such that for each bounded Borel function f there is a black box B_f for computing $f(r)$ given r for each r in R.

After each measurement of $\Phi(\alpha)$ on $\Psi(s)$ the experimenter feeds the result of the measurement into each of the black boxes and records the output of B_f for each f. Now the usual assumption is that for each f this output of B_f is a measurement of $f(\Phi(\alpha))$ on $\Psi(s)$. In particular, if this assumption is true, then an infinite repetition of the above will yield for each f a sequence $F_f \psi_{s\alpha}$ of B_f outputs which must satisfy Eqs. (3) and (4).

The point of the above proof is that there are f such that for the $\Phi(\alpha)$ and $\Psi(s)$ in the example, applying B_f to the outcome of a measurement of $\Phi(\alpha)$ on $\Psi(s)$ is *not* a measurement of $f(\Phi(\alpha))$ on $\Psi(s)$. In particular, if B_f is the box for f = characteristic function for $Rng\psi_{s\alpha}$ [$Rng\psi_{s\alpha}$ is the set of all outcomes of an infinite repetition of measuring $\Phi(\alpha)$ on $\psi(s)$ and for each real number r, B_f outputs 1 or 0 if r is or is not in $Rng\psi_{s\alpha}$, respectively] then B_f outputs 1 for every measurement of $\Phi(\alpha)$ on $\Psi(s)$ in the infinite repetition and thus $F_f \psi_{s\alpha}$ is a constant sequence of 1's. The failure of Eq. (3) results then from the fact that $f(\Phi(\alpha))$ is the zero operator.

Besides the particular f given above, there are many other Borel f for which Eqs. (3) and (4) fail. For example, for any real number between 0 and 1 there are Borel f for which $\overline{M} F_f \psi_{s\alpha} = r$. One can also find a Borel f for which $\overline{M} F_f \psi_{s\alpha}$ does not exist. All these Borel f have in common the fact that they are not τ-definable although they are τ-definable from $\psi_{s\alpha}$. Recall that from rule 3)$_\tau$, $\psi_{s\alpha}$ is not τ-definable.

It should be noted that the condition on s used in the above proof can be weakened to $\Psi(s)(\mathcal{E}^{\Phi(\alpha)}[\sigma_c(\Phi(\alpha))]) > 0$. Thus for each α with $\sigma_c(\Phi(\alpha))$ nonempty there are Borel f for which Eqs. (3) and (4) fail for any s which satisfies this weaker condition.

On the other hand, if $\Phi(\alpha)$ is pure discrete then the above proof fails as by rule 3)$_\tau$ for no s does one have $\mathcal{E}^{\Phi(\alpha)}\{r\} = 0$ for any r in $Rng\psi_{s\alpha}$. In fact, for this case it is an open question whether or not Eqs. (3) and (4) are valid for all Borel f. Also it is open whether or not for such $\Phi(\alpha)$ computing f on an outcome of doing f is an $f(\Phi(\alpha))$ measuring procedure for all Borel f.

It will now be seen that there are some Borel f for which one can *prove* that for each α f(outcome α) is an $f(\Phi(\alpha))$ measuring procedure. More precisely, it will be shown that rule 3)$_\tau$ implies that for each τ-definable Borel f, the probability measure $\tilde{P}_{sf(\alpha)}$ defined by Eq. (1) with $f(\Phi(\alpha))$ replacing $\Phi(\alpha)$ is τ-correct for $F_f \psi_{s\alpha}$ for each s and α. Clearly this is a sufficient condition for f (α outcome) to be a $f(\Phi(\alpha))$ measuring procedure.

To prove this first define the measure $\tilde{P}^f_{s\alpha}$ according to

$$\tilde{P}^f_{s\alpha} E = \tilde{P}_{s\alpha}(F_f^{-1} E) \tag{5}$$

for each E in $\mathfrak{B}(R^\omega)$ where

$$F_f^{-1} E = [\psi \mid F_f \psi \epsilon E], \tag{6}$$

F_f is defined as before and $\tilde{P}_{s\alpha}$ is given by Eq. (1).

Suppose it has been shown that for each s and α, $\tilde{P}^f_{s\alpha} = \tilde{P}_{sf(\alpha)}$. Then it is sufficient to show that the τ-correctness of $\tilde{P}_{s\alpha}$ for $\psi_{s\alpha}$ implies the τ-correctness

of $\tilde{P}^f_{s\alpha}$ for $F_f\psi_{s\alpha}$ as in this case rule $3)_\tau$ gives the desired result immediately. First assume $\tilde{P}_{s\alpha}$ is τ-correct for $\psi_{s\alpha}$, and let E be any Borel subset of R^ω which is τ-definable from $\tilde{P}^f_{s\alpha}$ and is such that $\tilde{P}^f_{s\alpha}E = 1$. We must show that $F_f\psi_{s\alpha}$ is in E.

Since f is τ-definable, any set E which is τ-definable from $\tilde{P}^f_{s\alpha}$ is, by Eq. (5) τ-definable from $\tilde{P}_{s\alpha}$ and thus $F_f^{-1}E$ is τ-definable from $\tilde{P}_{s\alpha}$. If $\tilde{P}^f_{s\alpha}(E) = 1$, one has by Eq. (5) $\tilde{P}_{s\alpha}(F_f^{-1}E) = 1$. Thus from the τ-correctness of $\tilde{P}_{s\alpha}$ for $\psi_{s\alpha}$ and Eq. (6), $\psi_{s\alpha}\epsilon F_f^{-1}E$ or $F_f\psi_{s\alpha}\epsilon E$.

To show that $\tilde{P}^f_{s\alpha} = \tilde{P}_{sf(\alpha)}$ it is sufficient to show that $\tilde{P}^f_{s\alpha}$ is a product measure and that it is generated from $\Psi(s)$ and $f(\Phi(\alpha))$ according to Eq. (1). The first follows from Eq. (5) and the facts that $\tilde{P}_{sf(\alpha)}$ is a product measure and that for any cylinder set $E = E_1 \times \cdots \times E_n \times R \times R \times \cdots$, $F_f^{-1}E = f^{-1}E_1 \times \cdots \times f^{-1}E_n \times R \times R \cdots$. The second follows from $\tilde{P}_{s\alpha}F_f^{-1}E_{Gj} = \Psi(s)(\mathcal{E}^{\Phi(\alpha)}(f^{-1}G)) = \Psi(s)(\mathcal{E}^{f(\Phi(\alpha))}(G))$ independent of j. Here $E_{Gj} = [\varphi\,|\,\varphi(j)\epsilon G]$ and the last equality follows from the fact that[7] $\mathcal{E}^{f(A)}(G) = \mathcal{E}^A(f^{-1}G)$ where $f^{-1}G = [r\,|\,f(r)\epsilon G]$ for each $G\epsilon\mathfrak{B}(R)$ and $\mathcal{E}^A$ is the spectral measure of A. This completes the proof.

This theorem has several consequences. First one sees immediately that Eqs. (3) and (4) hold for all s and α for each τ-definable Borel f. Equation (4) follows from Eq. (3) as rule $3)_\tau$ implies that[8]

$$\bar{M}\psi_{s\alpha_f} = \Psi(s)(f(\Phi(\alpha_f))) \tag{7}$$

holds for each α_f in $\Phi^{-1}(f(\Phi(\alpha)))$.

The class of Borel f for which Eqs. (3) and (4) can be proved to hold is larger than the class of τ-definable f. To see this one notes that any polynomial f with rational coefficients is τ-definable. Since both $\bar{M}$ (over its domain of definition) and $\Psi(s)$ are linear it follows that Eqs. (3) and (4) hold for all polynomial f, τ-definable or not.

At present it is an open question[9] whether or not for each α, Eqs. (3) and (4) hold for each s for all functions continuous on the spectrum of $\Phi(\alpha)$. However, it is suspected that such an extension does hold.

As a further consequence of the above theorem, one has the result that if f is a τ-definable polynomial then f (α outcome) is a $f(\Phi(\alpha))$ measuring procedure. Since for each α the polynomials with rational coefficients are[10] uniformly dense in the set of continuous functions on $\sigma(\Phi(\alpha))$, one has the result that the set of f for which $f(\alpha$ outcome) is an $f(\Phi(\alpha))$ measuring procedure is uniformly dense in the set of continuous functions on $\sigma(\Phi(\alpha))$. It is also an open question whether this result extends to all continuous functions on $\sigma(\Phi(\alpha))$.

Finally, for any reasonable τ the set of τ-definable Borel functions includes a countable generating set for the Borel functions. Thus one has the result that for each α, f (α outcome) is a $f(\Phi(\alpha))$ measuring procedure for every f in a generating set[11] for the Borel functions. However, as was seen earlier, this result cannot be extended to all Borel functions no matter how strong τ is.

In contrast to these results the usual interpretative rules (i.e., the spectrum rule and the expectation value rule) are too weak to give useful results. In particular, the usual rules are not sufficient to prove even the validity of Eqs. (3) and (4) for all s and α for all f in a countable generating set.[11]

One can prove this by constructing a sequence[12] ψ such that $Rng\psi$ is a finite set of rationals and such that $\bar{M}\psi$ exists but for no r in $Rng\psi$ does the limit relative frequency of finding r in ψ exist. Thus it is possible that for some s, α, $\psi_{s\alpha} = \psi$ as the usual interpretative rules require only that $\bar{M}\psi_{s\alpha} = \Psi(s)(\Phi(\alpha))$ hold. They say nothing about the limit relative frequencies of finding r in $\psi_{s\alpha}$.

One can also show that the expectation value condition of the usual rule implies that Eqs. (3) and (4) are valid for all s, α and all f in a set which is closed under the linear operations and which contains the identity function $f(r) = r$ and the unit function $f(r) = 1$. This set is clearly not a generating set for the Borel functions.

However, if one adds to the usual interpretative rules the requirement that Eq. (3) or Eq. (4) holds for the functions $f(r) = r^n$ for each $n = 0, 1, 2 \cdots$, then it follows that Eqs. (3) and (4) hold for all polynomial f. Note that it is not sufficient for this result to add to the usual rules that Eq. (3) or (4) holds for $f(r) = r^2$ only. In this case closure under composition is also needed.

In conclusion it must be stressed that in this and the following sections, in those arguments which use interpretative rules (and their extensions) we do *not* make the added assumption that an outcome sequence is random. That is, we assume that an outcome sequence satisfies only those conditions given in the interpretative rule and nothing more. Of course, it is a consequence of satisfying rule $3)_\tau$ that an outcome sequence is τ-random. Whether or not such a sequence is, in fact, random is a deeper problem which was discussed in I.

IV. LINEAR OPERATIONS

Since multiplication by a real number is an especially simple Borel function one can give the results for this operation directly from those obtained in the last section. Thus the usual interpretative (and, of course, rule $3)_\tau$] rules are sufficient to prove the validity of

$$\bar{M}c\psi_{s\alpha} = \Psi(s)(c\Phi(\alpha)) \tag{8}$$

and

$$\bar{M}c\psi_{s\alpha} = \bar{M}\psi_{sc\alpha} \tag{9}$$

for each s, α, and real number c. Here $c\psi$ denotes the sequence obtained by multiplying each element of ψ by c and $c\alpha$ is any observation procedure in $\Phi^{-1}(c\Phi(\alpha))$.

Also by the results of the last section, one proves from rule $3)_\tau$ that the procedure $c\cdot(\alpha$ outcome) is a $c\Phi(\alpha)$ measuring procedure for the τ-definable real numbers. For the real numbers, c, which are not τ-definable it is again open whether or not the above is a $c\Phi(\alpha)$ measuring procedure as for such c the proof given in the last section fails. However, since the rational numbers are all τ-definable one has the result that the class of numbers c for which $c\cdot(\alpha$ outcome) is a $c\Phi(\alpha)$ measuring procedure is dense in R.

The situation for the + operation is more interesting. Corresponding to Eqs. (3) and (4), one has

$$\bar{M}(\psi_{s\alpha} + \psi_{s\alpha'}) = \Psi(s)(\Phi(\alpha) + \Phi(\alpha')) \tag{10}$$

and

$$\bar{M}(\psi_{s\alpha} + \psi_{s\alpha'}) = \bar{M}\psi_{s\alpha+\alpha'}. \tag{11}$$

These equations are usually assumed to hold for all s, α, and α' with $\alpha + \alpha'$ any procedure in $\Phi^{-1}(\Phi(\alpha) + \Phi(\alpha'))$. Of course, Eq. (11) makes sense for only those α and α' for which $\Phi^{-1}(\Phi(\alpha) + \Phi(\alpha'))$ is not empty. $\psi_{s\alpha} + \psi_{s\alpha'}$ denotes the termwise addition of $\psi_{s\alpha} + \psi_{s\alpha'}$. As with Eqs. (8) and (9) the usual interpretative rules [and of

course rule $3)_\tau$] are sufficient to prove the vaidity of Eqs. (10) and (11).

A difficult problem in quantum mechanics is the question of whether or not for each α and α' in $\mathfrak{O}$, there exist procedures for measuring $\Phi(\alpha) + \Phi(\alpha')$. Two types of procedures will next be examined for their suitability for measuring $\Phi(\alpha) + \Phi(\alpha')$ for all α, α' in $\mathfrak{O}$. Both procedures involve measuring $\Phi(\alpha)$ and $\Phi(\alpha')$ and adding the results.

First consider the observation procedure $[\alpha + \alpha']_t$ defined as "carry out α on a system prepared in some state, then t time units later carry out α' on a *different* system prepared in the *same* state and add the α and α' outcomes." Now intuitively one feels that $[\alpha + \alpha']_t$ is not a valid $\Phi(\alpha) + \Phi(\alpha')$ measurement procedure. We shall prove from a slight extension of rule $3)_\tau$ that for almost all α and α' $[\alpha + \alpha']_t$ is not a valid $\Phi(\alpha) + \Phi(\alpha')$ measurement procedure.

Consider an infinite sequence of measurements in which the even (odd) number measurements consist in carrying out s and α (s and α') and such that for each $j = 0, 1, \cdots$ the jth α' measurement occurs t time units after the jth α measurement. Let $\psi_{s\alpha}$ and $\psi_{s\alpha'}$ be the respective outcome sequences of the even and odd numbered measurements and let $\psi_{s\alpha} * \psi_{s\alpha'}$ be the interleaving of $\psi_{s\alpha}$ and $\psi_{s\alpha'}$ such that $(\psi_{s\alpha} * \psi_{s\alpha'})(2j) = \psi_{s\alpha}(j)$ and $(\psi_{s\alpha} * \psi_{s\alpha'})(2j + 1)\ \psi_{s\alpha}(j)$ for each j. Thus $\psi_{s\alpha} * \psi_{s\alpha'}$ is an outcome sequence for the infinite sequence of measurements. Also, by construction, $\psi_{s\alpha} + \psi_{s\alpha'}$ is an outcome sequence $\psi_{s[\alpha+\alpha]_t}$ for an infinite repetition of doing $[\alpha + \alpha']_t$ and s.

Let $\tilde{P}_{s\alpha\alpha'}$ be the product measure on $\mathfrak{B}(R^\omega)$ generated from $\Psi(s)$, $\Phi(\alpha)$, and $\Phi(\alpha')$ according to

$$\tilde{P}_{s\alpha\alpha'} E_{Fj} = \begin{cases} \Psi(s)(\mathcal{E}^{\Phi(\alpha)}(F)) & \text{if } j \text{ even} \\ \Psi(s)(\mathcal{E}^{\Phi(\alpha')}(F)) & \text{if } j \text{ odd} \end{cases} \tag{12}$$

for each j and F in $\mathfrak{B}(R)$. Thus $\tilde{P}_{s\alpha\alpha'}$ is a product measure but it is not the product of the same measure on $\mathfrak{B}(R)$.

One next assumes that rule $3)_\tau$ extends to the above sequence of measurements also. That is one assumes that for each s, α, and α' $\tilde{P}_{s\alpha\alpha'}$ is τ-correct for $\psi_{s\alpha} * \psi_{s\alpha'}$. Now if $[\alpha + \alpha']_t$ is a valid $\Phi(\alpha) + \Phi(\alpha')$ measurement procedure then rule $3)_\tau$ applies and for all s $\tilde{P}_{s\alpha+\alpha'}$, defined by Eq. (1) with $\Phi(\alpha) + \Phi(\alpha')$ replacing $\Phi(\alpha)$, must be τ-correct for $\psi_{s[\alpha+\alpha']_t}$.

However, it is easy to see that if α and α' are such that neither $\Phi(\alpha)$ nor $\Phi(\alpha')$ is a multiple of the identity, then $\tilde{P}_{s\alpha+\alpha'}$ cannot be τ-correct for $\psi_{s[\alpha+\alpha']_t}$ for all s. Proof: Assume $\tilde{P}_{s\alpha+\alpha'}$ is τ-correct for $\psi_{s[\alpha+\alpha']_t}$ for all s. Then since $f(r) = r^2$ is τ-definable,

$$\bar{M}[(\psi_{s\alpha} + \psi_{s\alpha'})^2] = \Psi(s)[(\Phi(\alpha) + \Phi(\alpha'))^2] \tag{13}$$

must hold for all s. By the τ-correctness of $\tilde{P}_{s\alpha\alpha'}$ for $\psi_{s\alpha} * \psi_{s\alpha'}$ this is equivalent to requiring that

$$[\Psi(s)(\Phi(\alpha))] \cdot [\Psi(s)(\Phi(\alpha'))] = \Psi(s)([(\Phi(\alpha)\Phi(\alpha') + \Phi(\alpha')\Phi(\alpha)]/2) \tag{14}$$

hold for all s.

But this is impossible. To see this let s be such that $\Psi(s) = \lambda_0\rho_0 + \lambda_1\rho_1$ with ρ_0 and ρ_1 states over $\mathfrak{A}$ such that $\rho_i(\Phi(\alpha))$ and $\rho_i(\Phi(\alpha'))$ $\neq 0$ for $i = 0, 1$. Then Eq. (14) gives a condition λ_0 and λ_1 must satisfy besides $\lambda_0 + \lambda_1 = 1$. But clearly this condition is not satisifed in general for s for which $\Psi(s)$ has the above form. Thus a contradiction is reached and $\tilde{P}_{s,\alpha+\alpha'}$ cannot be τ-correct for $\psi_{s[\alpha+\alpha']_t}$ for all s and thus $[\alpha + \alpha']_t$ is not a valid $\Phi(\alpha) + \Phi(\alpha')$ measurement procedure.[13]

For the case in which either $\Phi(\alpha)$ or $\Phi(\alpha')$ is a multiple of the identity one can show that $[\alpha + \alpha']_t$ is a $\Phi(\alpha) + \Phi(\alpha')$ measurement procedure. In this case Eqs. (13) and (14) clearly are valid for all s.

Some points about this result are worth noting. Even though $[\alpha + \alpha']_t$ is not in general a valid $\Phi(\alpha) + \Phi(\alpha')$ measuring procedure, one can prove from the τ-correctness of $P_{s\alpha\alpha'}$ for $\psi_{s\alpha} * \psi_{s\alpha'}$ that $\psi_{s[\alpha+\alpha']_t}$ is τ-random. Second, the usual interpretative rules of quantum mechanics, extended to cover the alternating measurement sequence are clearly too weak to prove the above. This holds even if one adds to them the requirement that Eq. (13) hold for all s, α, and α'. The reason is that there does not seem to be a way to prove from them that $\bar{M}(\psi_{s\alpha} \cdot \psi_{s\alpha'}) = (\bar{M}\psi_{s\alpha})(\bar{M}\psi_{s\alpha'})$ which is necessary to obtain Eq. (14). Finally, one notes that, although the assumption that $\tilde{P}_{s\alpha\alpha'}$ is τ-correct for $\psi_{s\alpha} * \psi_{s\alpha'}$ is an extension of rule $3)_\tau$ to nonrepetitive measurement sequences, it is clearly a reasonable assumption. In fact it is just as reasonable as is the use of rule $3)_\tau$ for all s in $\mathfrak{S}$ and α in $\mathfrak{O}$.

It has been seen that, in general, carrying out α and α' on different systems prepared in the same state and adding the results is not a valid $\Phi(\alpha) + \Phi(\alpha')$ measurement procedure. On the other hand, one can prove from another extension of rule $3)_\tau$ that if $\Phi(\alpha)$ and $\Phi(\alpha')$ commute, then a simultaneous measurement of $\Phi(\alpha)$ and $\Phi(\alpha')$ (provided such a procedure exists) on one system followed by addition of the results is a valid $\Phi(\alpha) + \Phi(\alpha')$ measurement procedure.

Here we shall only outline how one proves this. First let $\{\alpha, \alpha'\}$ denote a procedure (provided one exists) for a simultaneous measurement of $\Phi(\alpha)$ and $\Phi(\alpha')$ on *one* system. The outcomes of $\{\alpha, \alpha'\}$ are pairs of real numbers. Let $\{\alpha + \alpha'\}$ denote the procedure $\{\alpha, \alpha'\}$ followed by addition of the pair of $\{\alpha, \alpha'\}$ outcomes. Let $\psi_{s,\{\alpha,\alpha'\}}$ and $\psi_{s,\{\alpha+\alpha'\}}$ denote respective outcome sequences of doing an infinite repetition of $\{\alpha, \alpha'\}$ and s or $\{\alpha + \alpha'\}$ and s. Note that $\psi_{s\{\alpha,\alpha'\}}$ is an infinite sequence of pairs of real numbers.

Let $\mathcal{E}^{\Phi(\alpha),\Phi(\alpha')}$ denote the unique spectral measure on $\mathfrak{B}(R \times R)$, the set of Borel subsets of $R \times R$, which satisfies

$$\mathcal{E}^{\Phi(\alpha),\Phi(\alpha')}(F \times G) = \mathcal{E}^{\Phi(\alpha)}(F) \cdot \mathcal{E}^{\Phi(\alpha')}(G) \tag{15}$$

for each F, G in $\mathfrak{B}(R)$. For each s, let $\tilde{P}_{s\{\alpha,\alpha'\}}$ be the product probability measure on $\mathfrak{B}((R \times R)^\omega)$ which satisfies

$$\tilde{P}_{s\{\alpha,\alpha'\}} E_{F\times G,j} = \Psi(s)(\mathcal{E}^{\Phi(\alpha),\Phi(\alpha')}(F \times G)) \tag{16}$$

for each j and F, G in (R). $E_{F\times G,j} = [\varphi \mid \varphi(j) \epsilon F \times G]$.

One now extends rule $3)_\tau$ to cover this case by assuming that for all α and α' for which a simultaneous measurement procedure for $\Phi(\alpha)$ and $\Phi(\alpha')$ exists, $\tilde{P}_{s,\{\alpha,\alpha'\}}$ is τ-correct for $\psi_{s\{\alpha,\alpha'\}}$ for all s. From this one can prove[14] that the measure $\tilde{P}_{s,\{\alpha+\alpha'\}}$, defined by Eq. (1) with $\Phi(\alpha) + \Phi(\alpha')$ replacing $\Phi(\alpha)$, is τ-correct for $\psi_{s\{\alpha+\alpha'\}} \cdot \psi_{s\{\alpha+\alpha'\}}$ is defined by $\psi_{s\{\alpha+\alpha'\}}(j) =$ sum of pair of numbers in $\psi_{s\{\alpha,\alpha'\}}(j)$. From this it follows

that $\{\alpha + \alpha'\}$ is a valid $\Phi(\alpha) + \Phi(\alpha')$ measurement procedure. As was the case for the $[\alpha + \alpha']_t$, this result is also in accord with intuition.

Finally, one notes that if $\Phi(\alpha)$ and $\Phi(\alpha')$ do not commute then any procedure for measuring $\Phi(\alpha) + \Phi(\alpha')$ is essentially different from α and α'. In particular, it does not involve a simultaneous measurement of $\Phi(\alpha)$ and $\Phi(\alpha')$ followed by addition of the results. Examples of procedures $\alpha + \alpha'$ for which $\Phi(\alpha)$ and $\Phi(\alpha')$ do not commute which are usually assumed to exist[15] are $J_x \cos\theta + J_y \sin\theta$ and $\mathbf{q} + \mathbf{p}t/2m$. $J_x \cos\theta + J_y \sin\theta$ corresponds to a measurement of the angular momentum projection along a line at angle θ from the x axis and $J_x \cos\theta$ corresponds to measuring the angular momentum projection along the x axis and multiplying the result by $\cos\theta$. $J_y \sin\theta$ is measured similarly. $\mathbf{q} + \mathbf{p}t/m$ is the Heisenberg operator corresponding to a position measurement of a free nonrelativistic particle of mass m, t time units after state preparation.

One can generalize the angular momentum example as follows: Let a and b be any pair of real numbers. Then

$$aJ_x + bJ_y = (a^2 + b^2)^{1/2} \times \left(J_x \frac{a}{(a^2 + b^2)^{1/2}} + J_y \frac{b}{(a^2 + b^2)^{1/2}}\right) \quad (17)$$

corresponds to multiplying by $(a^2 + b^2)^{1/2}$ an angular momentum projection measurement along a line at angle $\cos^{-1}(a/(a^2 + b^2)^{1/2})$, $\sin^{-1}(b/(a^2 + b^2)^{1/2})$ to the x axis.

V. GENERATION OF PROBABILITY MEASURES

In the previous sections only a small part of the properties included in rule $3)_\tau$ were actually used to carry out the proofs. Almost none of the properties on R^ω were used which express the requirement that the various single measurement outcomes be statistically independent of one another. In this section this will be partially remedied by consideration of the empirical generation of probability measures from outcome sequences.

It will be seen here that for each s and α, rule $3)_\tau$ is sufficient to prove that the empirical measure, generated from the outcome sequence in a standard way, is equal to $\tilde{P}_{s\alpha}$ (Eq. (1)).

Let $T: R^\omega \to R^\omega$ be the one-sided shift operator [i.e., $(T\psi)(j) = \psi(j + 1)$ each j and ψ]. For each Borel subset E of R^ω let I_E be the characteristic function for E [$I_E(\psi) = 1(0)$ if $\psi \epsilon E(\psi \notin E)$]. Define $\bar{I}_E^T(\psi)$ by

$$\bar{I}_E^T(\psi) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} I_{T^j E}(\psi)$$

if the limit exists where $T'E = [\psi\,|T\psi \epsilon E]$. Clearly, $\bar{I}_E^T(\psi)$ is the limit relative frequency that $\psi \epsilon E$, $T\psi \epsilon E$, $T^2\psi \epsilon E, \cdots$. In particular, if $E = E_{r,0} = [\psi\,|\psi(0) = r]$. then $\bar{I}_{E_{r,0}}^T(\psi)$ is the limit relative frequency that $\psi(0) = r$, $\psi(1) = r, \cdots$.

Now consider $\bar{I}_E^T(\psi)$ as a function on $\mathfrak{B}(R^\omega)$ by fixing ψ and letting E range over the elements of $\mathfrak{B}(R^\omega)$. Since $\bar{I}_E^T(\psi)$ is a limit relative frequency in E one might consider $\bar{I}_{(\cdot)}{}^T(\psi)$ to be a probability measure on $\mathfrak{B}(R^\omega)$. In particular, one might consider that for each s and α

$$\bar{I}_E^T(\psi_{s\alpha}) = \tilde{P}_{s\alpha}(E) \quad (18)$$

holds for each E in $\mathfrak{B}(R^\omega)$ with $\tilde{P}_{s\alpha}$ given by Eq. (1).

We first show that Eq. (18) cannot hold for all E [except for the trivial case where $\Psi(s)$ lies entirely in an eigenspace of $\Phi(\alpha)$]. Let $Or^T(\psi_{s\alpha}) = [\psi\,|\psi = T^j\psi_{s\alpha}$ for some j] denote the orbit of $\psi_{s\alpha}$ and define E by

$$E = R^\omega - Or^T(\psi_{s\alpha}) = [\psi\,|\psi \neq T^j\psi_{s\alpha} \text{ for all } j].$$

Since $Or^T(\psi_{s\alpha})$ is countable, E is a Borel set with $\tilde{P}_{s\alpha}E = 1$. However, $\bar{I}_E^T(\psi_{s\alpha}) = 0$ and thus Eq. (18) does not hold for this E. [One can, of course, find many other Borel E for which Eq. (18) fails.] Also unlike the case of Eq. (3), the failure of Eq. (18) holds for *all* α and is *not* restricted to just those α with $\Phi(\alpha)$ having a continuous spectrum. Furthermore, this failure is quite general and shows that for *no* ψ is $\bar{I}_{(\cdot)}{}^T(\psi)$ a probability measure on $\mathfrak{B}(R^\omega)$.

On the other hand let E be any τ-definable Borel subset of R^ω. Then $\bar{I}_E^T(\psi) = \tilde{P}_{s\alpha}(E)$ is clearly τ-definable from $\tilde{P}_{s\alpha}$ and by the ergodic and indecomposability theorems[8] this expression is true $\tilde{P}_{s\alpha}$ almost everywhere. Thus rule $3)_\tau$ gives the result that Eq. (18) holds for each s and α for all τ-definable Borel E.

Now the class $\mathfrak{F}^\tau$ of all τ-definable Borel subsets of R^ω is a field and thus by Rule $3)_\tau$, $\bar{I}_-^T(\psi_{s\alpha})$ is a probability measure on $\mathfrak{F}^\tau$. For sufficiently strong τ, $\mathfrak{F}^\tau$ includes the field $\mathfrak{F}^\omega$ which is a countable[11] generating set for $\mathfrak{B}(R^\omega)$. That is, $\mathfrak{B}(R^\omega)$ is the smallest σ field containing $\mathfrak{F}^\omega$.

Define $U_T(\psi_{s\alpha})$ to be the Hahn extension of $\bar{I}_{(\cdot)}{}^T(\psi_{s\alpha})$ from $\mathfrak{F}^\omega$ to $\mathfrak{B}(R^\omega)$. Since by rule $3)_\tau$ and the Hahn extension theorem,[16] $U_T(\psi_{s\alpha})$ exists and is unique with

$$U_T(\psi_{s\alpha}) = \bar{I}_{(\cdot)}{}^T(\psi_{s\alpha}) = \tilde{P}_{s\alpha}$$

on $\mathfrak{F}^\omega$, one has the result that $U_T(\psi_{s\alpha}) = \tilde{P}_{s\alpha}$ on $\mathfrak{B}(R^\omega)$. Thus one sees that by rule $3)_\tau$ for each s and α, the empirical measure $U_T(\psi_{s\alpha})$ generated from the outcome sequence $\psi_{s\alpha}$ as the Hahn extension of $\bar{I}_{(\cdot)}{}^T(\psi_{s\alpha})$ on $\mathfrak{F}^\omega$ equals the measure $\tilde{P}_{s\alpha}$ generated from s and α and Ψ and Φ. Clearly since $\tilde{P}_{s\alpha}$ is a product measure, so is $U_T(\psi_{s\alpha})$.

There are several points to be made about the above construction. First the method of generating $U_T(\psi)$ from an outcome sequence ψ was applied above to product measures. However, it is considerably more general than that. It can be shown[17] that the class of all probability measures generated as the Hahn extension of $\bar{I}_{(\cdot)}{}^T(\psi)$ for all ψ for which $\bar{I}_-^T(\psi)$ is a probability measure on $\mathfrak{F}^\omega$, is exactly equal to the class of all T-ergodic T-invariant probability measures on $\mathfrak{B}(R^\omega)$. Furthermore, this result extends to sets other than R^ω and to T other than the one-sided shift.

Another point is that the measures $\tilde{P}_{s\alpha}$ and $U_T(\psi_{s\alpha})$ differ in the following way. $\tilde{P}_{s\alpha}$ is a theoretical measure in that it depends on the maps Ψ and Φ, and s and α, but is independent of $\psi_{s\alpha}$. On the other hand, $U_T(\psi_{s\alpha})$ depends on the outcome sequence $\psi_{s\alpha}$ but is independent of Ψ and Φ.

Now by means of $U_T(\psi_{s\alpha})$ one can express rule $3)_\tau$ in a different but equivalent way as follows:

For each s in $\mathcal{S}$ and α in $\mathcal{O}$
(a) the outcome sequence $\psi_{s\alpha}$ of an infinite repetition of doing s and α must be τ-random and
(b) Ψ and Φ must be such that $U_T(\psi_{s\alpha}) = \tilde{P}_{s\alpha}$ with $\tilde{P}_{s\alpha}$ given by Eq. (1).

In this form rule $3)_\tau$ becomes a condition on the maps Ψ and Φ and is in many ways more satisfactory.[3,4] It also suggests a more constructive approach in that one

can prove *without using rule* $3)_\tau$ that if $\psi_{s\alpha}$ is τ-random, then (1) $U_T(\psi_{s\alpha})$ exists, and (2) $U_T(\psi_{s\alpha})$ is τ-correct for $\psi_{s\alpha}$. This whole approach in which one generates an outcome sequence $\psi_{s\alpha}$ (which may or may not be random) and then uses the sequence so generated to construct a measure with respect to which the sequence is required to be τ-correct and which does not use Ψ and Φ will be investigated more in future work. Note in this connection that one can redefine Eqs. (4), (9), and (11) in terms of maps from Θ and $\Theta \times \Theta$ to the set of subsets of Θ which also do not use Ψ and Φ.

VI. STATE PREPARATIONS

If s_0 and s_1 are state preparation procedures, it is often assumed[2,18,19] that a mixture $\lambda_0\Psi(s_0) + \lambda_1\Psi(s_1)$ with λ_0, λ_1 nonnegative real numbers and $\lambda_0 + \lambda_1 = 1$ can be prepared by carrying out s_0 and s_1 in the ratio λ_0/λ_1. In such preparation procedures, the description usually ignores the process by which s_0 and s_1 are chosen.

Here we wish to show that the strengthened interpretative rules of quantum mechanics do say quite a bit about such procedures and the choice process involved.

To discuss this in a fairly general setting, let φ be a function from N, the natural numbers into $\{0, 1\}$ such that $\overline{M}\varphi = \lambda_1$ where $\overline{M}$ is the limit relative frequency of finding 1 in φ. Let s_0 and s_1 be such that $\Psi(s_0) \neq \Psi(s_1)$ and let s_φ be the procedure "for the jth preparation compute $\varphi(j)$, if $\varphi(j) = 1(0)$ do $s_1(s_0)$." Let $\psi_{s_{\varphi,\alpha}}$ be an outcome sequence obtained from an infinite sequence of measurements in which the jth measurement consists in computing $\varphi(j)$, carrying out s_0 or s_1 depending on whether $\varphi(j)$ is 0 or 1 and then carrying out on the state so prepared.

The first point to note is that the probability measure $\tilde{P}_{s_{\varphi\alpha}}$ on $\mathfrak{B}(R^\omega)$ generated by the above infinite measurement sequence is a product measure, that is

$$\tilde{P}_{s_{\varphi\alpha}} = \bigotimes_{j=0}^{\infty} P_{s_{\varphi(j)}\alpha},$$

where

$$\tilde{P}_{s\varphi\alpha}E_{Fj} = \Psi(s_{\varphi(j)})(\mathcal{E}^{\Phi(\alpha)}(F)) \tag{19}$$

for each j and F in $\mathfrak{B}(R)$. However, $\tilde{P}_{s_{\varphi\alpha}}$ differs from $\tilde{P}_{s\alpha}$ defined by Eq. (1) in that $\tilde{P}_{s_{\varphi\alpha}}$ is *not*, in general, the product of the *same* measure. This is evident from the righthand side of Eq. (19) which defines a probability measure on $\mathfrak{B}(R)$ and which is, in general, j dependent.

On the other hand the product measure $\tilde{Q}_{s_{\varphi\alpha}}$ defined by

$$\tilde{Q}_{s_{\varphi\alpha}}E_{Fj} = (\lambda_0\Psi(s_0) + \lambda_1\Psi(s_1))(\mathcal{E}^{\Phi(\alpha)}(F)) \tag{20}$$

with $\lambda_1 = \overline{M}\varphi$ and which corresponds to Eq. (1) is a product measure which is the product of the same measure on $\mathfrak{B}(R)$. However, it is *not* the measure one would assign to the process. This can be seen, for example, by letting α and s_1 be such that

$$\Psi(s_1)(\mathcal{E}^{\Phi(\alpha)}(\{r\})) = 1$$

for some r. Then intuitively, one requires that for each j for which $\varphi(j) = 1, \psi_{s_{\varphi\alpha}}(j) = r$. This result follows from the assignment of $\tilde{P}_{s_{\varphi\alpha}}$ to the process. It does not follow from the assignment of $\tilde{Q}_{s_{\varphi\alpha}}$ to the process. Note that this requirement is the same as the requirement that if $\Psi(s)(\mathcal{E}^{\Phi(\alpha)}(\{r\})) = 1$ for some r, then *each* carrying out of s and α gives the outcome r.

More precisely, this intuitive result can be derived from the assumption that $\tilde{P}_{s_{\varphi\alpha}}$ is τ-correct for $\psi_{s_{\varphi\alpha}}$. [As in earlier sections this assumption requires the extension of rule $3)_\tau$ to cover measurement sequences of the above type.] It cannot be derived from the assumption that $\tilde{Q}_{s_{\varphi\alpha}}$ is τ-correct for $\psi_{s_{\varphi\alpha}}$.

Now suppose s_0 and s_1 are state preparation procedures. The question then arises, given that s_0 and s_1 are proper preparation procedures under what conditions is s_φ a proper preparation procedure for some state? First, one notes that it can be proved that if φ is such that $\overline{M}\varphi$ exists and if for each α, $\tilde{P}_{s_{\varphi\alpha}}$, defined by Eq. (19), is τ-correct for $\psi_{s_{\varphi\alpha}}$, then[20]

$$\overline{M}\psi_{s_{\varphi\alpha}} = [\overline{M}\varphi\,\Psi(s_1) + (1 - \overline{M}\varphi)\Psi(s_0)][\Phi(\alpha)] \tag{21}$$

hold for each α.

Now it may be argued that the existence of $\overline{M}\varphi$ and the τ-correctness of $\tilde{P}_{s_{\varphi\alpha}}$ for each α, with the resultant validity of Eq. (21) are sufficient conditions for s_φ to be a proper preparation procedure for the state $\overline{M}\varphi\,\Psi(s_1) + (1 - \overline{M}\varphi)\Psi(s_2)$. However, one should also note the following further consequence of these conditions.

If φ is τ-definable, then for each α for which $P_{s_0\alpha} \neq P_{s_1\alpha}$ [Eq. (1) right-hand equality], $\psi_{s_{\varphi\alpha}}$ is not τ-random.

Proof: Let $F_\varphi : R^\omega \to R^\omega$ be the subsequence selection procedure which picks out of $\psi_{s_{\varphi\alpha}}$, in the natural order, the results corresponding to all and only those measurements, in an infinite sequence of carrying out s_φ and α which correspond to doing s_1 and α. Now since φ is τ-definable so is F_φ. Thus, by the τ-correctness of $\tilde{P}_{s_{\varphi\alpha}}$ for $\psi_{s_{\varphi\alpha}}$, $\overline{M}F_\varphi\psi_{s_{\varphi\alpha}} = \Psi(s_1)(\Phi(\alpha))$ which when combined with Eq. (21) gives $\overline{M}F_\varphi\psi_{s_{\varphi\alpha}} \neq \overline{M}\psi_{s_{\varphi\alpha}}$. Since τ-random sequences have the property that their limit mean is invariant under all τ-definable subsequence selection procedures[3,27] $\psi_{s_{\varphi\alpha}}$ is clearly not τ-random.

A simple example of the above is $\varphi(j) = 1(0)$ if j even (odd). Then an infinite sequence of doing s_φ and α prepares s_1 at the even numbered trials and prepares s_0 at the odd numbered trials. Clearly $\psi_{s_{\varphi\alpha}}$ is not τ-random as the limit mean of the odd numbered elements of $\psi_{s_{\varphi\alpha}} \neq \overline{M}\psi_{s_{\varphi\alpha}}$.

Furthermore, if the strengthened interpretative rules are to be taken seriously, the validity of Eq. (21) for all α is not sufficient for s_φ to be a proper state preparation procedure. For it follows from rule $3)_\tau$ that a necessary condition for s_φ to be a proper state preparation procedure is that for each α, $\tilde{Q}_{s_{\varphi\alpha}}$ defined by Eq. (20) with $\overline{M}\varphi = \lambda_1$ must be τ-correct for $\psi_{s_{\varphi\alpha}}$. Thus rule $3)_\tau$ gives the result that for no τ-definable φ is s_φ a proper state preparation procedure.

At present it remains an open question whether or not there exist φ for which s_φ is a proper preparation procedure. The difficulty lies in proving that there exist φ with $\overline{M}\varphi = \lambda_1$ and such that for *all* α if $\tilde{P}_{s_{\varphi\alpha}}$ is τ-correct for $\psi_{s_{\varphi\alpha}}$ then $\tilde{Q}_{s_{\varphi\alpha}}$ is also τ-correct for $\psi_{s_{\varphi\alpha}}$. We have been able to prove[22] that if φ is τ-random with $M\varphi = \lambda_1$ then this holds for *some* α, namely those for which the measures $P_{s_0\alpha}$ and $P_{s_1\alpha}$ [Eq. (1) right-hand side] are τ-definable.

These results extend to the more intuitive procedures which include physical methods for generating φ. Such procedures are particularly useful for φ's which are *not* τ-definable as it is not clear how such φ's are to be computed. Let β be a measurement procedure such that

$\Phi(\beta)$ is a projection operator and let s_2 be a state preparation procedure. Define $s_{s_2\beta}$ as follows: "Carry out β and s_2; if outcome is 1(0) carry out preparation procedure $s_1(s_0)$."

Again the question arises whether or not $s_{s_2\beta}$ can be regarded as a preparation procedure for the state $\lambda_0\Psi(s_0) + \lambda_1\Psi(s_1)$ with $\lambda_1 = (\Psi(s_2))(\Phi(\beta))$ and $\lambda_0 = 1 - \lambda_1$. Let $\psi_{s_2\beta}$ denote the outcome sequence obtained from an infinite repetition of doing β and s_2 and let $\psi_{s(s_2\beta)\alpha}$ denote the outcome sequence obtained from an infinite repetition of doing $s_{s_2\beta}$ and α. Let $\tilde{P}_{s(s_2\beta)\alpha}$ be the product measure on $\mathfrak{B}(R^\omega)$ defined by Eq. (19) with $\psi_{s_2\beta}$ replacing φ. As for s_φ, $\tilde{P}_{s(s_2\beta)\alpha}$ is not the product of the same measure.

Using the above results from s_φ, one has by rule $3)_\tau$ that $\bar{M}\psi_{s_2\beta} = \Psi(s_2)(\Phi(\beta))$. The extension of rule $3)_\tau$ to an infinite repetition of $s_{s_2\beta}$ and α, i.e., the assumption that $\tilde{P}_{s(s_2\beta)\alpha}$ is τ-correct for $\psi_{s(s_2\beta)\alpha}$, gives the equivalent of Eq. (21) or

$$\bar{M}\psi_{s(s_2\beta)\alpha} = [\bar{M}\psi_{s_2\beta}\Psi(s_1) + (1 - \bar{M}\psi_{s_2\beta})\Psi(s_0)][\Phi(\alpha)]$$

hold for each α.

However, as in the case for s_φ, it is an open question whether or not $s_{s_2\beta}$ is a proper state preparation procedure. Again the reason lies in the difficulty of proving that if $\tilde{P}_{s(s_2\beta)\alpha}$ is τ-correct for $\psi_{s(s_2\beta)\alpha}$ so is $\tilde{Q}_{s(s_2\beta)\alpha}$ [given by Eq. (20) with $\lambda_1 = \bar{M}_{s_2\beta}$].

The above results depend on the assumption that $\tilde{P}_{s_\varphi\alpha}$ and $\tilde{P}_{s(s_2\beta)\alpha}$ are τ-correct for $\psi_{s_\varphi\alpha}$ and $\psi_{s(s_2\beta)\alpha}$, respectively. If one replaces this assumption by the assumption that for *all* α, rule $3)_\tau$ holds directly for s_φ and α (with τ-random) and for $s_{s_2\beta}$ and α (that is, that $\tilde{Q}_{s_\varphi\alpha}$ and $\tilde{Q}_{s(s_2\beta)\alpha}$ are τ-correct for $\psi_{s_\varphi\alpha}$ and $\psi_{s(s_2\beta)\alpha}$), then it follows that s_φ and $s_{s_2\beta}$ are proper state preparation procedures and that for all $\alpha, \psi_{s_\varphi\alpha}$ and $\psi_{s(s_2\beta)\alpha}$ are τ-random. However, this replacement seems counterintuitive for the reasons mentioned before.

Finally, it should be noted that the above problematic aspects disappear for those procedures s which do not include a decision procedure for choosing among more than one inequivalent preparation procedures. In particular, this applies to those s which prepare pure states. For these s, the measure one would assign intuitively to an infinite repetition of s on any α is just that given by Eq. (1). Then rule $3)_\tau$ applies directly and it follows that $\psi_{s\alpha}$ is τ-random for all α.

*Work performed under the auspices of the U. S. Atomic Energy Commission.

[1]P. A. Benioff, Phys. Rev. D, 7, 3603 (1973).
[2]H. Ekstein, Phys. Rev. 184, 1315 (1969).
[3]P. A. Benioff, J. Math. Phys. 11, 2553 (1970); 12, 361 (1971); J. Symbol. Logic, 36, 377 (1971).
[4]P. A. Benioff, Found. Phys. 3, 359 (1973).

[5]The correct condition is that $\Phi(\ )$ include operators with nonzero continuous parts. To see this let $\sigma_d(A)$ be the discrete spectrum of A and define the projection operator P_D by $P_D = \Sigma r \in \sigma_d(A)\mathcal{E}^A(\{r\})$. $\mathcal{E}^A(\{r\})$ is the eigenprojector of A for eigenvalue r. If $P_D < 1$, then A has a nonzero continuous part and $P_C = 1 - P_D > 0$. The reason this condition is correct rather than the one in the main text is that there exist self-adjoint operators with nonempty continuous spectra but for which $P_D = 1$. For an example see Ref. 1, footnote 4.

[6]M. H. Stone, *Linear Transformations in Hilbert Space*, American Mathematical Society Colloquium Publ. 1932, Vol. XV, Chap. V, Sec. 5.
[7]N. Dunford and J. Schwartz, *Linear Operators* (Wiley-Interscience, New York, 1963), Vol. II, Chap. X, p. 902.
[8]M. Loeve, *Probability Theory* (Van Nostrand, Princeton, New Jersey. 1963), 3rd ed., Secs. 30-32.

[9]By the Stone-Weierstrass theorem any continuous $f = \lim_n f_n$ where f_n is a polynomial. The difficulty is to prove $\lim_m \lim_n \bar{M}_m F_{f_n} = \lim_n \lim_m \bar{M}_m F_{f_n}$ where $\bar{M}_m\psi$ is the mean of the first m elements of ψ.

[10]Let f_n be a sequence of polynomials for which uniform $\lim_n f_n = f$ on $\sigma(\Phi(a))$. For each n let g_{nm} be a sequence of polynomials with rational coefficients such that for each $m |g_{nm}(r) - f_n(r)| < 1/m$ for all r on $\sigma(\Phi(a))$. Then the diagonal sequence g_{nn} is such that uniform $\lim_n g_{nn} = f$ on $\sigma(\Phi(a))$.

[11]The usual definition (Ref. 8, pp. 102-110) defines the generating set of all functions f of the form $f(r) = c_k$ if $r \epsilon E_k$, $k = 0, 1 \ldots n$, $f(r) = 0$ if $r \notin \cup_{k=0}^{n} E_k$ with c_k a real number and E_k, $k = 0, 1, \ldots, n$ any finite set of pairwise disjoint Borel sets. Now $\mathfrak{B}(R)$ [and $\mathfrak{B}(R\omega)$] are countably generated [K. R. Parthasarathy, *Probability Measures on Metric Spaces* (Academic, New York, 1967), Chap. I, Sec. 2 and Chap. V, Sec. 2] from the set of intervals $[a, b)$ with rational endpoints, and any real number is the limit of a sequence of rationals. Thus the set of $f(r)$ defined as above except that c_k is rational and $E_k = [a_k, b_k)$ with a_k and b_k rational is a set of τ-definable functions which is also a countable generating set for the Borel functions.

[12]Let ϕ be a $0, r_1, r_2$, sequence such that $\phi(k) = 0$ if k odd, $\phi(k) = r_1[r_2]$ if $k = 2j$ and $2^n < j < 2^{n+1}$ for some even [odd] n. Let θ be a $0, r_3, r_4$ sequence such that $\theta(k+1) = 0$ if k odd, $\theta(k+1) = r_3[r_4]$ if $k = 2j$ and $2^n < j \leq 2^{n+1}$ for some even [odd] n. Define ψ by $\psi = \phi + \theta$. Then if r_1, r_2, r_3, and r_4 are four rational numbers such that $r_1 \neq r_2$, $r_3 \neq r_4$ but $r_1 + r_3 = r_2 + r_4 = r$, then $\bar{M}\psi$ exists and equals $r/2$ but for none of the four rational numbers does the limit relative frequency of finding it in ψ exist.

[13]This proof depends on there being a "sufficient" supply of preparation procedures available. Possible definitions of sufficiency are that $\Psi(\mathfrak{S})$ = set of normal states on $\mathfrak{A}$ or that $\Psi(\mathfrak{S})$ is dense in some topology on $\mathfrak{A}$.

[14]One first shows that for any commuting self-adjoint operators A and B, the spectral measure for $A + B$ is the convolution of the spectral measures of A and B, or $\mathcal{E}^{A+B}(E) = \int \mathcal{E}^A(E - r)\mathcal{E}^B(dr)$ for each E in $\mathfrak{B}(R)$ with $E - r = \{s | s + r \epsilon E\}$. From this one shows easily that $\tilde{P}_S\{a+a'\}(E) = \tilde{P}_S\{a,a'\}(+_\omega^{-1}E)$ for each E in $\mathfrak{B}(R^\omega)$. $+_\omega^{-1}E = [(\psi, \pi)|\psi + \pi \epsilon E] \epsilon \mathfrak{B}((R \times R)^\omega)$. From this it follows that if any set E in $\mathfrak{B}(R^\omega)$ is τ-definable from $\tilde{P}_S\{a+a'\}$ it is τ-definable from $\tilde{P}_S\{a, a'\}$ as $+_\omega^{-1}$ is τ-definable. The τ-correctness of $\tilde{P}_S\{a+a'\}$ for $\psi_S\{a+a'\}$ follows from this and the τ-correctness of $\tilde{P}_S\{aa'\}$ for $\psi_S\{aa'\}$.

[15]The author is indebted to Hans Ekstein for suggesting these examples.

[16]N. Dunford and J. Schwartz, *Linear Operators*, (Wiley-Interscience, New York, 1958), Chap. III, Sec. 5.

[17]P. A. Benioff, unpublished work.

[18]W. Mackey, *Mathematical Foundations of Quantum Mechanics* (Benjamin, New York, 1963), pp. 63, 64.
[19]R. Giles, J. Math. Phys. 11, 2139 (1970).

[20]If ψ is such that $\bar{M}\psi$ exists and equals λ_1, then the limit mean measure $\bar{P}^T_{s_\phi a}$ defined by $\bar{P}^T_{s_\phi a} = \lim_n n^{-1} \Sigma_{j=0}^{n-1} \tilde{P}_{s_\phi a} T'^j$ equals $\tilde{Q}_{s_\phi a}$, Eq. (20). T is the one sided shift and T' is the set transformation generated by T. The ergodic theorem (Ref. 8) and the fact that $\tilde{P}_{s_\varphi a}$ is a product measure give (Ref. 8, p. 230) the result that $\bar{M}\psi = [\lambda_0\Phi(s_0) + \lambda_1\Phi(s_1)\phi(a)]$ for $\tilde{P}_{s\phi a}$ almost all ψ. Equation (21) then follows from the τ-correctness of $\tilde{P}_{s\phi a}$.

[21]A. Wald, *Mathematischen Kolloquiums* (Vienna, 1973), Vol. 8 [*Selected Papers in Statistics and Probability of A. Wald* (McGraw-Hill, 1955), pp. 79-99.]

[22]One must show that if E in $\mathfrak{B}(R\omega)$ is τ-definable from $\tilde{Q}_{s\phi a}$ and $\tilde{Q}_{s\phi a}E = 1$ then $\psi_{s\phi a} \epsilon E$. Much of the proof consists in showing that if ϕ is τ-random and E is τ-definable from $\tilde{Q}_{s\phi a}$ and $\tilde{Q}_{s\phi a}E = 1$, then $\tilde{P}_{s\phi a}E = 1$. The proof of this is done by constructing a measure μ from λ_0 and λ_1 in a standard way on $\mathfrak{B}(\{0,1\}^\omega)$ and defining for each E in $\mathfrak{B}(R^\omega)$ a Borel function $f: \{0, 1\}^\omega \to R$ by $f_E(g) = \tilde{P}_{sga}E$. From $1 = \tilde{Q}_{s\phi a}E = \int f_E d\mu$ one has that $f_E(g) = 1$ is true μ almost everywhere. Thus if f_E is τ-definable from μ, which occurs only if $P_{s_0 a}$ and $P_{s_1 a}$ are τ-definable, the τ-randomness of ϕ together with $\bar{M}\phi = \lambda_1$ gives $f_E(\phi) = 1$.

The group chain $SO_{n,1} \supset SO_{1,1} \otimes SO_{n-1}$, a complete solution to the "missing label" problem

Charles P. Boyer and Kurt Bernardo Wolf

Centro de Investigación en Matemáticas Aplicadas y en Sistemas (CIMAS), Universidad Nacional Autónoma de México, México 20, D. F., México

(Received 2 August 1973; revised manuscript received 18 October 1973)

We discuss the decomposition $SO_{n,1} \supset SO_{1,1} \otimes SO_{n-1}$ by constructing multiplier representations over the group manifold of SO_n. Explicit orthogonal and complete bases in terms of functions diagonal with respect to the canonical ($SO_{n,1} \supset SO_n$) and noncanonical ($SO_{n,1} \supset SO_{1,1} \otimes SO_{n-1}$) chains are provided which give a complete solution to the "missing label" and multiplicity problems occuring in the latter decomposition. Moreover, an integral representation for the overlap functions between the two chains is given, for which the singularity structure can be immediately ascertained. Expressions for the cases $n = 3$ and 4 are given.

1. INTRODUCTION

The orthogonal and unitary groups SO_n and U_n exhibit canonical chains, i.e., chains of subgroups whose unitary irreducible representation (UIR) labels can be used to characterize uniquely their basis vectors. These are well known to be $SO_n \supset SO_{n-1} \supset \cdots \supset SO_2$ and $U_n \supset U_{n-1} \supset \cdots \supset U_1$. Many physical systems, however, require different chains. The foremost example of the latter is the Elliott chain $U_3 \supset SO_3$ in nuclear shell theory as studied by Beidenharn and Moshinsky[1] which has been the subject of extensive research. More in line with the problem we shall treat here are the reductions of the type $SO_{m+n} \supset SO_m \otimes SO_n$ and $U_{m+n} \supset U_m \otimes U_n$. The latter ones are of relevance in some elementary particle classification schemes[2] while the former, in its noncompact version $SO_{3,1} \supset SO_{1,1} \otimes SO_2$ has been used to work with helicity bases for the Lorentz group[3] and applied to the construction of solutions to the Dirac and Proca free field equations.[4]

The Elliott chain as well as the other examples mentioned (except the last one) exhibit the multiplicity and what is called the "missing label" problems,[2,5,6] that is, the UIR labels obtained from the subgroup do not specify completely the UIR bases of the group. The problem has been solved by constructive methods[2] whereby one starts with the highest or lowest weight state of a multiplet and, applying convenient lowering or raising operators, generates all the states of that multiplet labelling thus the states in the process. This is a valid procedure for finite multiplets or for "half-infinite" multiplets belonging to some discrete series of UIRs which have an extremum state.

Most of these procedures, however, are *algebraic* in the sense of making use of the enveloping algebra of the group. It is our contention that the labelling problem can be solved through the methods of harmonic analysis on the *group manifold*.[7] The procedure is essentially that of building complete and orthogonal sets of functions on a subgroup manifold and defining a multiplier representation of the group in both canonical and noncanonical chains. In the process of constructing such a set of functions we provide an explicit basis classified by labels, some of which are eigenvalues of operators which are *not* elements of the enveloping algebra.

In Sec. 2 we review the mathematical framework needed: Gel'fand states, Euler angles and their labels, the Wigner D-functions and generalizations of the spherical harmonics which we call E functions, and the relevant orthogonality and completeness relations. The multiplier representation of $SO_{n,1}$ in the canonical and noncanonical chains are set up in Secs. 3 and 4 and the overlap functions are found in Sec. 5 together with its pole structure and asymptotic behavior. In two appendixes we treat the cases $SO_{3,1}$, known from the work of Kuznetzov et al.[3] and Kalnins,[4] and $SO_{4,1}$.

2. THE BASIS FUNCTIONS

The Gel'fand kets and Euler angles for SO_n are very well known (see, e.g., Ref. 8). Only some points on notation will therefore be repeated. The SO_n single-valued UIRs are labeled by a set of $[\frac{1}{2}n]$ integer numbers ($[\frac{1}{2}n]$ stands for the largest integer $\leq \frac{1}{2}n$), $\{J_{n,1}, J_{n,2}, \cdots, J_{n,[n/2]}\} \equiv J_n$ and its bases for UIRs are completely classified by the canonical chain $SO_n \supset SO_{n-1} \supset \cdots \supset SO_2$, whereby the basis vectors have their rows labeled by $\{J_{n-1}, J_{n-2}, \cdots, J_2\} \equiv \underline{J}_{n-1}$. The Gel'fand ket

$$\left| \begin{matrix} J_{n,1} J_{n,2} \cdots J_{n,[n/2]} \\ J_{n-1,1} J_{n-1,2} \cdots J_{n-1,[1/2(n-1)]} \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ J_{41} J_{42} \\ J_{31} \\ J_{21} \end{matrix} \right\rangle \equiv \left| \begin{matrix} J_n \\ \underline{J}_{n-1} \end{matrix} \right\rangle \tag{2.1}$$

containing $\frac{1}{2}[\frac{1}{2}n^2]$ labels thus transforms as the J_p UIR of SO_p ($p = n, n-1, \ldots, 2$) and the SO_p content of SO_n is found through the orthogonal group branching relations.[8] In order to reduce the indexing to a bare minimum, we shall always denote the UIR labels of SO_n (resp. SO_{n-1} and SO_{n-2}) by the letters J (resp. L and M) and their row labels by $\underline{L} \equiv L, \underline{M}$ (resp. $\underline{M} \equiv M, \underline{N}$ and $\underline{N}$) and write the ket (2.1) horizontally as $|J\underline{L}\rangle$.

The elements r of SO_n can be parametrized by the set of $\frac{1}{2}n(n-1)$ Euler angles $\theta_{jk}^{(p)}$ ($1 \leq j < k \leq p \leq n$) which represent rotations in the j-k plane, written as

$$R_p(\{\theta\}^{(p)}) = R_{p-1}(\{\theta\}^{(p-1)}) S_{p-1}(\{\theta^{(p-1)}\}), \tag{2.2a}$$

$$S_{p-1}(\{\theta^{(p)}\}) = r(\theta_{p-1,p}^{(p)}) S_{p-2}(\{\theta^{(p)}\}), n \geq p > 2, \tag{2.2b}$$

$$R_2(\{\theta\}^{(2)}) = S_1(\{\theta^{(2)}\}) = r(\theta_{12}^{(2)}), \tag{2.2c}$$

where R_p is thus an element of SO_p and S_{p-1} a representative of the coset space SO_{p-1}/SO_p isomorphic to the $(p-1)$-dimensional sphere S_{p-1}. The Haar measure can be similarly split as

$$dR_n = \omega(Rn) d\{\theta\}^{(n)} = dR_{n-1} dS_{n-1}, \tag{2.3a}$$

$$dS_{n-1} = \sin^{n-2}\theta_{n-1,n} d\theta_{n-1,n} dS_{n-2}, \tag{2.3b}$$

$$dR_2 = dS_1 = d\theta_{12}, \tag{2.3c}$$

where $\omega(R_n)$ is the Haar weight of SO_n. The ranges of $\theta_{jk}^{(p)}$ are $[0, \pi]$ for $j, k > 1$, while $\theta_{12}^{(p)}$ runs over $[0, 2\pi)$. The volume of the group is thus $\mathrm{vol}SO_n = \mathrm{vol}SO_{n-1}\,|S_{n-1}|$, where $|S_{n-1}| = 2\pi^{n/2}/\Gamma(n/2)$ is the surface of the $(n-1)$-sphere, and $\mathrm{vol}SO_2 = |S_1| = 2\pi$.

The Wigner D-functions (UIR matrix elements) for SO_n are then constructed and decomposed as

$$D^J_{\overline{L}'\overline{L}}(R_n) \equiv \langle J\overline{L}'|R_n|J\overline{L}\rangle = \sum_{\overline{M}''} D^{L'}_{\overline{M}'\overline{M}''}(R_{n-1})\, E^J_{L'\overline{M}'',\overline{L}}(S_{n-1}), \tag{2.4}$$

where we have defined the E functions

$$E^J_{\overline{L}'\overline{L}}(S_{n-1}) \equiv \langle J\overline{L}'|S_{n-1}|J\overline{L}\rangle = d^J_{L'M'L}(\theta_{n-1,n})E^L_{\overline{M}'\overline{M}}(S_{n-2}), \tag{2.5}$$

using the Wigner d-functions

$$d^J_{L'ML}(\theta_{n-1,n}) \equiv \langle JL'M\overline{N}|\,r(\theta_{n-1,n})\,|JLM\overline{N}\rangle, \tag{2.6}$$

which are diagonal in M (UIR label of SO_{n-2}) and independent of $\overline{N}$ (row label of SO_{n-2}). The E functions are generalizations of the spherical harmonics: for SO_3, $E^l_{m'm}(\theta,\varphi) = d^l_{m'm}(\theta)e^{im\varphi}$ and for SO_4, $E^{J_1J_2}_{l'm',lm}(\xi\theta\varphi) = d^{J_1J_2}_{l'm'l}(\xi)E^l_{m'm}(\theta,\varphi)$, etc. Orthogonality and completeness relations for these functions will be analyzed below.

Consider the *right* action from the group on functions *on* the coset manifold S_{n-1} as

$$\mathcal{T}(R'_n)f(S_{n-1}) = f(S_{n-1}R'_n); \tag{2.7}$$

the $E^J_{\overline{L}'\overline{L}}(S_{n-1})$ functions then transform as the ket $|J\overline{L}\rangle$. The indices $\overline{L}'$ do not enter into the transformation properties of the ket, and will be called *redundant* labels. They are only asked to respect the branching relations. They *do* distinguish, however, between different E functions transforming in the same way but are *not* eigenlabels of any operator in the enveloping algebra of SO_n from the *right*. (They are the eigenvalues, however, of operators acting on the group from the left.)

We know from the Peter-Weyl theorem[7] that the D functions are orthogonal and complete over the space $\mathcal{L}^2(SO_n)$ of square integrable functions over the SO_n group manifold. The orthogonality relation is

$$\int dR_n D^{J_1}_{\overline{L}'_1\overline{L}_1}(R_n)^* D^{J_2}_{\overline{L}'_2\overline{L}_2}(R_n) = \frac{\mathrm{vol}SO_n}{\dim J_1}\,\delta_{J_1,J_2}\delta_{\overline{L}_1,\overline{L}_2}\delta_{\overline{L}'_1,\overline{L}'_2}, \tag{2.8}$$

where $\dim J$ is the dimension of the UIR labelled by J, and where the δ's in the collective indices $J, \overline{L}$, etc. are products of δ's in the individual indices. From (2.4) and (2.8) we can write the generalized orthogonality relations for the E functions as

$$\int dS_{n-1} \sum_{\overline{M}} E^{J_1}_{L\overline{M},\overline{L}_1}(S_{n-1})^* E^{J_2}_{L\overline{M},\overline{L}_2}(S_{n-1}) = |S_{n-1}|\,\frac{\dim L}{\dim J_2}\,\delta_{J_1,J_2}\delta_{\overline{L}_1,\overline{L}_2}, \tag{2.9}$$

where $\int dS_{n-1}$ stands for the integration, with the correct measure, over the coset space S_{n-1}. The completeness relation of the D's is given by

$$\sum_J \frac{\dim J}{\mathrm{vol}SO_n}\,\mathrm{Tr}[D^J(R_n)^\dagger D^J(R'_n)] = \frac{1}{\omega(R_n)}\,\delta(R_n, R'_n) \tag{2.10}$$

where the trace sums over all row and column indices and the δ in R_n and R'_n is the product of δ's in the individual angles. The right-hand side of (2.10) is the reproducing kernel under the Haar integral in (2.8) and conversely, the right-hand side of (2.8) is the reproducing kernel in the Plancherel sum in (2.10). We can find a generalized completeness relation for the E's if we write the right-hand side of (2.10) as

$$\frac{1}{\omega(R_n)}\,\delta(R_n, R'_n) = \frac{1}{\omega(S_{n-1})}\,\delta(S_{n-1}, S'_{n-1})\,\frac{1}{\omega(R_{n-1})}\,\delta(R_{n-1}, R'_{n-1}), \tag{2.11}$$

with the weight on S_{n-1} given by $\omega(S_{n-1}) = \omega(R_n)/\omega(R_{n-1})$. Writing the last two factors of (2.11) as (2.10) with $n \to n-1$ multiplying by $D^{L_0}_{\overline{M}_0\overline{M}_0}(R'_{n-1})^*$ and integrating over $R'_{n-1} \in SO_{n-1}$ with the appropriate Haar measure, we obtain

$$\sum_J \frac{1}{|S_{n-1}|}\,\frac{\dim J}{\dim L_0}\sum_{\overline{L}} E^J_{L_0\overline{M}'_0,\overline{L}}(S_{n-1})^* E^J_{L_0\overline{M}_0,\overline{L}}(S'_{n-1}) = \frac{1}{\omega(S_{n-1})}\,\delta(S_{n-1}, S'_{n-1})\,\delta_{\overline{M}'_0,\overline{M}_0}. \tag{2.12}$$

3. MULTIPLIER REPRESENTATIONS AND THE CANONICAL CHAIN

It has been shown in Refs. 9-12 that one can deform the SO_n algebra of elements M_{ij} $(i,j = 1, \ldots, n)$ into the $SO_{n,1}$ algebra through the addition of the "noncompact" generators

$$M^{(\sigma)}_{i,n+1} = \tfrac{1}{2}[x_i, \Phi] + \sigma x_i \tag{3.1}$$

where Φ is the second-order SO_n Casimir operator and x_i is a point on S_{n-1}. Since one can decompose an arbitrary element g of $SO_{n,1}$ into double cosets as $g = hbh'$ where $h, h' \in SO_n$ and $b \in SO_n\backslash SO_{n,1}/SO_n$, it suffices to consider[11] the "last boost" generated by

$$M^{(\sigma)}_{n,n+1} = \sin\theta\,\frac{\partial}{\partial\theta} + \lambda\cos\theta = \frac{\partial}{\partial\omega} + \lambda\tanh\omega, \qquad \lambda = \tfrac{1}{2}(n-1) - \sigma, \tag{3.2}$$

where $\theta \equiv \theta^{(n)}_{n-1,n}$ and $\omega = \ln\tan\frac{1}{2}\theta$. The range $[0, \pi]$ for θ implies $-\infty < \omega < \infty$. We construct a multiplier representation of $SO_{n,1}$ on $R_n = (R_{n-1}, \theta, S_{n-2})$ such that, for the "last boost" generated by (3.2), we have

$$\mathcal{T}(\exp[\xi M^{(\sigma)}_{n,n+1}])f(R_{n-1}, \theta, S_{n-2}) = \left(\frac{\sin\theta}{\sin\theta'}\right)^\lambda f(R_{n-1}, \theta', S_{n-2}), \tag{3.3a}$$

where

$$\tan\tfrac{1}{2}\theta' = e^\xi\tan\tfrac{1}{2}\theta, \tag{3.3b}$$

and the action (2.7) holds for the elements in $SO_n \subset SO_{n,1}$. When $\lambda = -\frac{1}{2}(n-1) - i\tau$ (τ real), the multiplier just offsets the change in SO_n measure and provides the principal series of unitary representations of $SO_{n,1}$ on the space $\mathcal{L}^2(SO_n)$. This is, essentially, Mackey's method of induced representations.

Now, we *do* have a complete and orthogonal set of functions over SO_n, namely, the $D^J_{\overline{L}'\overline{L}}(R_n)$. These were shown[11] to transform under $SO_{n,1}$ in this realization as the (λ, L') bases for UIRs with *row* $\overline{J} = J, L, \overline{M}$, i.e.,

$$\left\langle R_n \left| \begin{matrix} \lambda L' \\ J \end{matrix} \right\rangle_{(\overline{M}')} = \left[\frac{\dim J}{\mathrm{vol} SO_n} \right]^{1/2} D^{J}_{L',\overline{M}',\overline{L}}(R_n), \tag{3.4}$$

where the factor has been added to make the kets orthonormal over SO_n. The important points to notice are (i) the *column* labels L' enter as UIR labels for $SO_{n,1}$ satisfying the correct branching relations with respect to J, (ii) the principal series of UIRs is obtained for $\lambda = -\frac{1}{2}(n-1) - i\tau$ (τ real), and (iii) the column labels $\overline{M}'$ do not affect the transformation properties of (3.4) and are thus *redundant* labels in the same sense as these seen in the former section. There are $\frac{1}{2}[\frac{1}{2}(n-2)^2]$ of these labels.[8] For fixed $SO_{n,1}$ and redundant labels, the functions (3.4) form an irreducible basis. The functions (3.4) were used in Ref. 11 in order to find the Bargmann d-functions for $SO_{n,1}$ in a recursive fashion. Here, we only wish to stress that the orthogonality relation (2.8) is written for the kets (3.4) as

$$\underset{(\overline{M}'_1)}{\left\langle \begin{matrix} \lambda L'_1 \\ J_1 \end{matrix} \right|} \left. \begin{matrix} \lambda L'_2 \\ J_2 \end{matrix} \right\rangle_{(\overline{M}'_2)} = \delta_{J_1 J_2} \delta_{L'_1 L'_2} \delta_{\overline{M}'_1 \overline{M}'_2}, \tag{3.5}$$

and completeness (2.10) becomes

$$\sum_{JL'\overline{M}'} \left| \begin{matrix} \lambda L' \\ J \end{matrix} \right\rangle_{(\overline{M}')(\overline{M}')} \left\langle \begin{matrix} \lambda L' \\ J \end{matrix} \right| = \mathbf{1}, \tag{3.6}$$

i.e., the redundant labels $\overline{M}'$ *do* enter into the orthogonality and completeness relations, only λ is never invoked and stems from the realization of the $SO_{n,1}$ algebra we are working with. Notice that the total number of labels in the ket (3.4), excepting λ which is fixed, is $\frac{1}{2}n(n-1)$, equal to the number of parameters of SO_n. It should be emphasized that although the representation (3.3) is reducible over $\mathcal{L}^2(SO_n)$, the irreducible components are given explicitly by the redundant labels. These appear in block-diagonal form in Eq. (3.5).

4. THE NONCANONICAL CHAIN

We will now *construct* orthogonal and complete sets of basis functions with definite transformation properties under the subgroup $SO_{1,1} \otimes SO_{n-1}$ of $SO_{n,1}$. The number of ket labels specifying the $SO_{n,1}$ rows provided by the canonical chain is $\frac{1}{2}[\frac{1}{2}n^2]$ (see Ref. 8), while that provided by the noncanonical chain $SO_{1,1} \otimes SO_{n-1} \supset SO_{n-2} \supset \cdots \supset SO_2$ is $1 + \frac{1}{2}[\frac{1}{2}(n-1)^2]$. The number of *missing* labels in the noncanonical chain is thus $[\frac{1}{2}n] - 1$. The functions we wish to construct must first contain an E function on S_{n-2} in order to have the necessary SO_{n-1} labels. Secondly, they must be eigenfunctions of the generator of $SO_{1,1}$ so that they are classified by its label. This generator can be taken to be (3.2) in the variable $\omega = \ln \tan\frac{1}{2}\theta$ so that the eigenfunctions are obtained from the differential equation

$$M^{(\sigma)}_{n,n+1} \tilde{f}^{\lambda}_{\nu}(\omega) = \nu \tilde{f}^{\lambda}_{\nu}(\omega) \tag{4.1a}$$

which are

$$\tilde{f}^{\lambda}_{\nu}(\omega) = (2\pi)^{-1/2} \cosh^{-\lambda}\omega e^{i\nu\omega}, \tag{4.1b}$$

complete and orthogonal over the ν and ω real lines as

$$\int_{-\infty}^{\infty} d\nu \tilde{f}^{\lambda}_{\nu}(\omega)^* \tilde{f}^{\lambda}_{\nu}(\omega') = \cosh^{n-1}\omega \delta(\omega - \omega'), \tag{4.2a}$$

$$\int_{-\infty}^{\infty} \cosh^{n-1}\omega d\omega \tilde{f}^{\lambda}_{\nu}(\omega)^* \tilde{f}^{\lambda}_{\nu'}(\omega) = \delta(\nu - \nu'), \tag{4.2b}$$

when we recall that $\lambda = -\frac{1}{2}(n-1) - i\tau$ (τ real) for the principal series. Furthermore, the redefinition of the argument brings in the functions

$$f^{\lambda}_{\nu}(\theta) = \tilde{f}^{\lambda}_{\nu}(\ln \tan\tfrac{1}{2}\theta) = (2\pi)^{-1/2} 2^{\lambda} \sin^{\lambda + i\nu}\tfrac{1}{2}\theta \cos^{\lambda - i\nu}\tfrac{1}{2}\theta, \tag{4.3}$$

which are complete as in (4.2a) and orthogonal in θ under the measure $\sin^{n-2}\theta d\theta$.

Lastly, our noncanonical basis functions should be functions of the whole of the SO_n manifold, so that the whole of $SO_{n,1}$ can be applied to it as a multiplier representation. This can be achieved multiplying the functions by a D function on the remaining SO_{n-1} manifold as

$$\left\langle R_n \left| \begin{matrix} \lambda L' \\ \nu(M'') \\ \overline{L} \end{matrix} \right\rangle_{(\overline{M}')} = \left(\frac{\mathrm{vol} SO_{n-2}}{\mathrm{vol}^2 SO_{n-1}} \frac{\dim L \, \dim L'}{\dim M''} \right)^{1/2} \times \sum_{\overline{N}''} D^{L'}_{\overline{M}', M''\overline{N}''}(R_{n-1}) f^{\lambda}_{\nu}(\theta) E^{L}_{M''\overline{N}'',\overline{M}}(S_{n-2}), \tag{4.4}$$

where M'' are the "missing" labels found by our scheme and $\overline{M}'$ the "redundant" ones. In all there are again $\frac{1}{2}n(n-1)$ labels (excluding λ).

Using (2.8), (2.9), and (4.2b) we can verify that these functions are indeed orthogonal over SO_n, i.e.,

$$\underset{(\overline{M}'_1)}{\left\langle \begin{matrix} \lambda L'_1 \\ \nu_1(M''_1) \\ \overline{L}_1 \end{matrix} \right|} \left. \begin{matrix} \lambda L'_2 \\ \nu_2(M''_2) \\ \overline{L}_2 \end{matrix} \right\rangle_{(\overline{M}'_2)} = \delta(\nu_1 - \nu_2) \delta_{L'_1 L'_2} \delta_{M''_1 M''_2} \delta_{\overline{L}_1 \overline{L}_2} \delta_{\overline{M}'_1 \overline{M}'_2}, \tag{4.5}$$

and using (2.10), (2.12), and (4.2a), completeness over $\mathcal{L}^2(SO_n)$ holds:

$$\int_{-\infty}^{\infty} d\nu \sum_{L'M''\overline{L}\overline{M}'} \left| \begin{matrix} \lambda L' \\ \nu(M'') \\ \overline{L} \end{matrix} \right\rangle_{(\overline{M}')(\overline{M}')} \left\langle \begin{matrix} \lambda L' \\ \nu(M'') \\ \overline{L} \end{matrix} \right| = \mathbf{1}. \tag{4.6}$$

This formula allows us to decompose any function in $\mathcal{L}^2(SO_n)$ in terms of the noncanonical basis functions (4.4) as a sum over discrete labels and an integral over the continuous label ν. We stress the fact that the "redundant" labels enter in the noncanonical chain, that is, relations (4.5) and (4.6) in the *same* way as they do for the canonical chain. This means that the $SO_{n,1}$ labels L' and the redundant labels $\overline{M}'$ not only appear in block diagonal form with the corresponding subgroup chain, but also appear in block diagonal form in the overlap functions computed explicitly in the next section. For fixed $SO_{n,1}$ and "redundant" labels, we obtain an irreducible subspace and a basis in this subspace is given by the remaining labels, including the "missing" labels, which are essential. All of the discrete labels, of course, are constrained by the branching rules. In particular, for the "missing" labels $M'' \equiv J''_{n-2}$ we have

$$\min(J_{n-1,k}, J'_{n-1,k}) \geq J''_{n-2,k} \geq \max(J_{n-1,k+1}, J'_{n-1,k+1}), \tag{4.7a}$$

for $k = 1, \ldots, [n/2] - 2$, while

$$J''_{n-2,[n/2]-1} \geq 0 \quad \text{for } n \text{ odd} \tag{4.7b}$$

and

$$\min(J_{n-1,[n/2]-1}, J'_{n-1,[n/2]-1}) \geq |J''_{n-2,[n/2]-1}| \quad \text{for } n \text{ even}. \tag{4.7c}$$

The number of possible "missing" labels J''_{n-2} for a fixed $SO_{n,1}$ UIR and for a fixed $SO_{1,1} \otimes SO_{n-1}$ UIR gives the *multiplicity* of the decomposition.

5. THE OVERLAP FUNCTIONS

As both the canonical and noncanonical bases functions are orthogonal and complete, one can easily obtain an integral representation for the overlap function

$$\begin{pmatrix}\lambda L_1' \\ \nu(M'') \\ \overline{L}_1\end{pmatrix}_{(\overline{M_1'})} \Bigg| \begin{matrix}\lambda L_2' \\ J \\ \overline{L}_2\end{matrix} \Bigg\rangle_{(\overline{M_2'})} = \delta_{L_1'L_2'}\delta_{\overline{L}_1\overline{L}_2}\delta_{\overline{M}_1'\overline{M}_2'}2^{-\lambda-3/2}$$
$$\times\left(\pi^{-3/2}\,\frac{\Gamma(\tfrac{1}{2}n)}{\Gamma(\tfrac{1}{2}[n-1])}\,\frac{\dim J\,\dim M''}{\dim L_1\,\dim L_1'}\right)^{1/2}$$
$$\times\int_0^{\pi} d\theta\,\sin^{-\lambda-1-i\nu}\tfrac{1}{2}\theta\,\cos^{-\lambda-1+i\nu}\tfrac{1}{2}\theta\,d^{J}_{L_1',M''L_1'}(\theta), \qquad (5.1)$$

where we have used the orthogonality relations for SO_{n-1}, (2.8) and (2.9). Equation (5.1) contains much information. The appearance of the "missing" labels M'' is explicitly in the d function inside the integral while the result does not depend on the "redundant" labels $\overline{M'}$ (which appear only in the δ's) nor on the UIR labels of SO_p, $p \leq n-2$. The singularity structure[13] and asymptotic behavior in ν can be examined noting that the d functions can be written as polynomials in $\sin^{k'}\tfrac{1}{2}\theta\,\cos^{k''}\tfrac{1}{2}\theta$ where $k', k'' \geq 0$, and run over a finite range. A typical term occurring in the integral (5.1) yields an integral representation for the Beta function[14]

$$\int_0^{\pi} d\theta\,\sin^{-\lambda-1-i\nu+k'}\tfrac{1}{2}\theta\,\cos^{-\lambda-1+i\nu+k''}\tfrac{1}{2}\theta$$
$$= B(\tfrac{1}{2}[-\lambda-i\nu+k'],\tfrac{1}{2}[-\lambda+i\nu+k'']), \qquad (5.2)$$

from which we can see that the overlap functions exhibit simple poles at the points $\nu = \pm i(\lambda - 2k)$, where $k = 0, 1, 2, \cdots$. In particular cases there may be zeroes cancelling some poles due to the influence of the d function. There are no other singularities, however. Moreover, from (5.2) the asymptotic behavior of (5.1) can be obtained from Stirling's formula[14] to be $\sim|\nu|^{\gamma}$ $\exp(-\tfrac{1}{2}\pi|\nu|)$, for some fixed γ, which is typical of many such overlap functions and assures the convergence of the decomposition.

When changing bases, the integration contour over ν runs along the real axis and we see that none of the poles of either (5.1) nor its complex conjugate interfere with the integration. If we analytically continue the $SO_{n,1}$ UIRs to the supplementary series[15] by allowing λ to take values $0 > \lambda > -(n-1)$ [or equivalently τ to lie on the imaginary axis between $-i\tfrac{1}{2}(n-1)$ and $i\tfrac{1}{2}(n-1)$, not including the endpoints], we see that still none of the above poles interfere with the integration contour. Thus our decomposition remains valid for the supplementary series of $SO_{n,1}$ as well.

6. OUTLOOK

We have discussed the example $SO_{n,1} \supset SO_{1,1} \otimes SO_{n-1}$ for its relative simplicity. The corresponding unitary groups can be worked out using the results of Ref. 16 and for the linear groups we can point to Ref. 17. Future work[18] should provide the framework for the reduction $SO_{n,1} \supset SO_{n-k} \otimes SO_{k,1}$ and $SO_{n,k} \supset SO_n \otimes SO_k$ and their unitary and symplectic counterparts. This is due to the relative ease in constructing multiplier representations of noncompact groups. The compact groups should be treatable through analytic continuation and, indeed, the solution of the multiplicity problem does not seem to depend on the noncompact nature of the example presented.

ACKNOWLEDGMENTS

One of us (K. B. W.) would like to thank Professor R. T. Sharp, of McGill University, Montréal, for his hospitality last fall, and Professor P. Winternitz and Dr. E. Kalnins for the stimulating conversations which started this work.

APPENDIX A: $SO_{3,1} \supset SO_{1,1} \otimes SO_2$

In order to find the overlap coefficients, we apply the formula (5.1) keeping in mind that, as $n = 3$, all SO_{n-2} labels disappear and there are no missing nor redundant labels. A straightforward calculation yields, for $\lambda = -1 - i\tau$,

$$\left\langle \begin{matrix}\lambda M \\ \nu \\ m\end{matrix} \middle| \begin{matrix}\lambda M \\ l \\ m\end{matrix} \right\rangle = 2^{-1+i\tau}\pi^{-1/2}(2l+1)^{1/2}$$
$$\times\int_0^{\pi} d\theta\,\sin^{i(\tau-\nu)}\tfrac{1}{2}\theta\,\cos^{i(\tau+\nu)}\tfrac{1}{2}\theta d^{l}_{Mm}(\theta)$$
$$= 2^{-\lambda-1}(2\pi)^{-1/2}i^{\mu_s-\mu_i}\,(2l+1)^{1/2}$$
$$\times\left(\frac{(l+\mu_s)!\,(l-\mu_i)!}{(l-\mu_s)!\,(l+\mu_i)!}\right)^{1/2}\frac{\Gamma(\tfrac{1}{2}[\mu_s+\mu_i-\lambda+i\nu])}{(\mu_s-\mu_i)!}$$
$$\times\frac{\Gamma(\tfrac{1}{2}[\mu_s-\mu_i-\lambda-i\nu])}{\Gamma(\mu_s-\lambda)}$$
$$\times{}_3F_2\left[\begin{matrix}\mu_s-l & \mu_s+l+1 & \tfrac{1}{2}(\mu_s-\mu_i-\lambda-i\nu) \\ \mu_s-\mu_i & \mu_s-\lambda & \end{matrix};1\right], \qquad (A1)$$

where $\mu_s \equiv \max(m, M)$ and $\mu_i = \min(m, M)$, and which can be compared with Kuznetzov et al.[3] We see that the poles occur at $\nu = \pm i[\lambda - (\mu_s \mp \mu_i) - 2k]$, $k = 0, 1, 2, \cdots$.

APPENDIX B: $SO_{4,1} \supset SO_{1,1} \otimes SO_3$

Using the SO_4 d-functions as given, e.g., in Ref. 11, one finds for $SO_{4,1}$ that

$$\begin{pmatrix}\lambda L' \\ \nu(M'') \\ L \\ M\end{pmatrix}_{(M')} \Bigg| \begin{matrix}\lambda L' \\ J_1 J_2 \\ L \\ M\end{matrix} \Bigg\rangle_{(M')}$$
$$= 2^{-\lambda}(2\pi)^{-1}\left(\frac{\dim(J_1,J_2)}{\dim L\,\dim L'}\right)^{1/2}\int_0^{\pi} d\theta\,\sin^{-\lambda-1-i\nu}\tfrac{1}{2}\theta$$
$$\cos^{-\lambda-1+i\nu}\tfrac{1}{2}\theta\,d^{J_1J_2}_{L'M''L}(\theta)$$
$$= 2^{-\lambda}(2\pi)^{-1}\left(\frac{(J_1+1)^2-J_2^2}{(2L+1)(2L'+1)}\right)^{1/2}\sum_m C(\tfrac{1}{2}[J_1+J_2],$$
$$\tfrac{1}{2}[J_1-J_2], L';\ \tfrac{1}{2}[M''+m],\tfrac{1}{2}[M''-m], M'')$$
$$\times C(\tfrac{1}{2}[J_1+J_2],\tfrac{1}{2}[J_1-J_2], L;$$
$$\tfrac{1}{2}[M''+m],\tfrac{1}{2}[M''-m], M'')\sum_n i^n\binom{n}{2m}$$
$$\times B(\tfrac{1}{2}[n-\lambda-i\nu],\tfrac{1}{2}[2m-n-\lambda+i\nu]), \qquad (B1)$$

which can be seen to be independent of the SO_2 and redundant labels M and M'. There are poles due to the B function at $\nu = \pm i(\lambda - 2k)$ $(k = 0, 1, 2, \cdots)$. Notice that the "missing" label M'' appears in a rather "geometric" fashion through the entries of the Clebsch-Gordan coefficient C. If a given UIR of $SO_{1,1} \otimes SO_3$ (in the noncanonical chain) given by (ν, L) appears in the decomposition of a $(\lambda L')$ UIR of $SO_{4,1}$, the multiplicity is given by the possible values of M''. This is constrained by the minimum of $2L+1$ and $2L'+1$, hence the multiplicity is $2\min(L, L') + 1$.

The overlap coefficients for $SO_{5,1}$ can be obtained in a similar fashion from the SO_5 d^1s as given by Holman.[19]

[1]M. Moshinsky, *Group Theory and the Many Body Problem* (Gordon and Breach., New York, 1967).
[2]C. R. Hagen and A. J. Macfarlane, J. Math. Phys. **6**, 1366 (1965); V. Syamala Devi and T. Venkatarayudu, J. Math. Phys. **9**, 1057 (1968); V. Syamala Devi and T. Venkatarayudu, J. Math. Phys. **11**, 169 (1970); R. T. Sharp and C. S. Lam, J. Math. Phys. **10**, 2033 (1969); R. T. Sharp and D. Lee, Rev. Mex. Física **20**, 203 (1970); R. T. Sharp, Proc. Camb. Phil. Soc. **68**, 571 (1970); K. Ahmed and R. T. Sharp, Ann. Phys. **71**, 421 (1971); R. T. Sharp, J. Math. Phys. **13**, 183 (1971); J. Mickelsson, J. Math. Phys. **11**, 2803 (1970); J. Mickelsson, J. Math. Phys. **12**, 2378 (1971).
[3]G. I. Kuznetsov, M. A. Liberman, A. A. Makarov, and Ya. A. Smorodinskii, Sov. J. Nucl. Phys. **10**, 370 (1970).
[4]E. G. Kalnins, J. Math. Phys. **13**, 1304 (1972); **14**, 654 (1973).
[5]M. K. F. Wong, J. Math. Phys. **11**, 1489 (1970).
[6]W. H. Klink, J. Math. Phys. **9**, 1669 (1968).
[7]L. S. Pontryagin, *Topological Groups* (Gordon and Breach, New York, 1966).
[8]R. L. Anderson and K. B. Wolf, J. Math. Phys. **11**, 3176 (1970).
[9]M. A. Melvin, Bull. Am. Phys. Soc. **7**, 493 (1962); **8**, 356 (1963); A. Sankaranarayanan, Nuovo Cimento **38**, 1441 (1965); Y. Dothan, M. Gell-Mann, and Y. Ne'eman, Phys. Lett. **17**, 148 (1965); J. Rosen and P. Roman, J. Math. Phys. **7**, 2072 (1966); R. Gilmore, J. Math. Phys. **13**, 883 (1972).
[10]These deformations are intimately related to multiplier representations as constructed by V. Bargmann, Ann. Math. **48**, 568 (1947) and I. M. Gel'fand and M. A. Naĭmark, *Unitäre Darstellungen der Klassischen Gruppen* (Akademie Verlag, Berlin, 1957); for a recent review where the connection between the two approaches is established, see C. P. Boyer, *Symposium on Symmetry in Nature*, México D. F., June 1973 (to appear in Rev. Mex. Física).
[11]K. B. Wolf, J. Math. Phys. **12**, 197 (1971).
[12]C. P. Boyer, J. Math. Phys. **12**, 1599 (1971); C. P. Boyer and F. Ardalan, J. Math. Phys. **12**, 2070 (1971).
[13]I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic, New York, 1964), Vol. 1.
[14]A. Erdélyi et al., *Higher Transcendental Functions* (McGraw-Hill, New York 1953), Vol. 1.
[15]C. P. Boyer, J. Math. Phys. **14**, 609 (1973).
[16]K. B. Wolf, J. Math. Phys. **13**, 1634 (1972).
[17]C. P. Boyer and K. B. Wolf, J. Math. Phys. (to appear).
[18]C. P. Boyer and K. B. Wolf, work in progress.
[19]W. J. Holman III, J. Math. Phys. **10**, 1710 (1969).

Expansions in Breit–Wigner amplitudes and biorthogonal functions

William H. Klink

Department of Physics and Astronomy, The University of Iowa, Iowa City, Iowa 52240
(Received 22 December 1972; final revised manuscript received 19 July 1973)

The problem of finding the expansion coefficients for a finite superposition of Breit-Wigner amplitudes is discussed. Since such amplitudes are not mutually orthogonal, another set of functions orthogonal to the Breit–Wigner functions are computed. Advantages and disadvantages of this biorthogonal set of functions are compared with conventional orthogonal functions and, in particular, the problem of reading off expansion coefficients when the modulus squared of the amplitude is known—rather than the amplitude itself—is discussed.

I. INTRODUCTION

If a complete set of orthonormal functions is given in a Hilbert space $\mathcal{H}$, it is of basic importance in quantum mechanics that any square integrable function in $\mathcal{H}$ can be expanded in this complete set, and the expansion coefficients are then given by the inner product of the function with the relevant orthonormal function.

Now generally complete sets of orthonormal functions arise as eigenfunctions of Hermitian operators; sometimes, however, it is desirable to be able to expand functions in a complete set of functions that are not orthogonal. Such is the case with Breit–Wigner amplitudes; [1] one would like to be able to fit at least the resonant parts of cross section data as an appropriate superposition of Breit–Wigner amplitudes, calculating both the expansion coefficients and masses and widths of the resonances from the given data. To be able to carry out such an expansion of scattering amplitudes as an appropriate superposition of Breit–Wigner amplitudes requires, since the Breit–Wigner functions are not orthogonal, that certain criteria to be discussed in Sec. II be met.

Once these criteria are satisfied it can be shown that there exists another set of functions, generating a so-called biorthogonal set of functions, with the property that they are orthogonal to the Breit–Wigner amplitudes. Once the biorthogonal set is known it is possible to compute the expansion coefficients as a suitable inner product, in much the way as is done for an orthonormal set. Section III will be devoted to the problem of finding the functions orthogonal to the Breit–Wigner amplitudes, using some properties of Blaschke products and Stieltjes transforms. Finally, Sec. IV will discuss an example, to show why the problem of computing the complex expansion coefficients of an amplitude from the real coefficients obtained from experimental data involves simpler algebraic equations than is the case with expansions arising from group symmetries, such as partial-wave expansions and phase shift analysis.

II. BIORTHOGONAL FUNCTIONS

We wish to investigate the possibility of expanding a function $f(x)$ as a superposition of Breit–Wigner functions,

$$B_i(x) \equiv \frac{1}{x + \lambda_i}, \tag{II. 1}$$

where x is a (dimensionless) energy and λ_i has as its real and imaginary parts, a (dimensionless) mass and width, respectively; that is, we wish to ask under what conditions it is possible to write

$$f(x) = \sum_i c_i B_i(x), \qquad 0 \le x \le \infty, \tag{II. 2}$$

where f is assumed, for the moment, to be square integrable on a Hilbert space $\mathcal{H}$ with norm

$$\|f\|^2 = \int_0^\infty dx\, |f(x)|^2 < \infty, \qquad f \in \mathcal{H}. \tag{II. 3}$$

In the analysis that follows it would be possible to include a weight function in (II. 3), but by a suitable redefinition of the functions orthogonal to the Breit–Wigner functions, it is possible to absorb the weight function and so it will not be discussed further.

Now with the norm given by (II. 3) it is immediately clear that the Breit–Wigner functions, Eq. (II. 1), do not form an orthonormal set; that is,

$$\begin{aligned}(B_i, B_j) &= \int_0^\infty \frac{dx}{(x + \lambda_i^*)(x + \lambda_j)} \\ &= M_{ij} \\ &\neq \delta_{ij}.\end{aligned} \tag{II. 4}$$

The question then arises under what circumstances the set of Breit–Wigner functions for a given $\{\lambda_i\}$ does form a complete set and when there exists a set orthogonal to the Breit–Wigner functions.

The answer to this latter question depends on the matrix M_{ij}, for a necessary and sufficient condition that the expansion (II. 2) define a biorthogonal set is that M be positive definite.[2] For if M is positive definite, there then exists a set of functions $\{\varphi_i\}$ with the property that

$$(\varphi_i, B_j) = \delta_{ij}. \tag{II. 5}$$

The goal of this section will be to discuss the meaning of M being positive definite for Breit–Wigner functions while the next section will show how to obtain the functions φ_i. And once the functions φ_i are known, it is possible to compute the coefficients c_i of Eq. (II. 2) as

$$c_i = (\varphi_i, f). \tag{II. 6}$$

The meaning of the positive definiteness of M can be seen by noting that for any set of functions M is Hermitian. Hence there exists a suitable unitary transformation U that diagonalizes M:

$$M = U^\dagger \Lambda U, \tag{II. 7}$$

where Λ is a diagonal real matrix giving the eigenvalues of M. If all the eigenvalues are nonnegative, then $\Lambda = \Lambda^{1/2}\Lambda^{1/2}$ and

$$\begin{aligned}M &= U^\dagger \Lambda^{1/2}\Lambda^{1/2} U \\ &= (\Lambda^{1/2}U)^\dagger(\Lambda^{1/2}U) \\ &= P^\dagger P,\end{aligned} \tag{II. 8}$$

which is the definition of a nonnegative definite matrix. If the eigenvalues are in fact positive, then P is nonsingular.

Now if the functions φ_i satisfying $(\varphi_i, B_j) = \delta_{ij}$ exist, they too can be expanded as

$$\begin{aligned}\varphi_i &= \sum_j N_{ji} B_j, \\ (\varphi_k, \varphi_i) &= \sum_j N_{ji}(\varphi_k, B_j) \\ &= N_{ki}.\end{aligned} \tag{II. 9}$$

Also,

$$\begin{aligned}(B_k, \varphi_i) &= \sum_j N_{ji}(B_k, B_j) \\ &= \sum_j N_{ji} M_{kj} \\ &= \delta_{ki}\end{aligned} \tag{II. 10}$$

so that $N = M^{-1}$. Thus, if a set $\{\varphi_i\}$ of functions exist, it follows that M has an inverse, which then naturally has the same eigenvectors and inverse eigenvalues as M. Hence a biorthogonal set has common eigenvectors and inverse eigenvalues. For such a biorthogonal set to exist then there can be no zero eigenvalues, so that M must be positive definite.

The usual probabilistic interpretation of quantum mechanics is gotten by writing

$$\begin{aligned}\|f\|^2 &= \sum_{i,j} (d_i \varphi_i, c_j B_j) \\ &= \sum_i d_i^* c_i,\end{aligned} \tag{II. 11}$$

where $c_i = (\varphi_i, f)$ and $d_i = (B_i, f)$. But

$$\begin{aligned}d_i &= (B_i, f) \\ &= (B_i, \sum c_j B_j) \\ &= \sum_j M_{ij} c_j\end{aligned} \tag{II. 12}$$

and

$$\begin{aligned}\|f\|^2 &= \sum_i d_i c_i^* \\ &= \sum_{i,j} M_{ij} c_j c_i^* \\ &= c^\dagger M c \\ &= c^\dagger U^\dagger \Lambda^{1/2} \Lambda^{1/2} U c \\ &= (\Lambda^{1/2} U c)^\dagger (\Lambda^{1/2} U c) \\ &= \tilde{c}^\dagger \tilde{c} \\ &= \sum_i |\tilde{c}_i|^2,\end{aligned} \tag{II. 13}$$

where $\tilde{c} = \Lambda^{1/2} U c$. Since Λ and U are fixed for a given set of functions such as the Breit–Wigner functions, it is clear that while the expansion coefficients c_i cannot be used in the computation of probabilities, the coefficients $\tilde{c}_i$ can. Also, it is to be noted that if M were not positive definite $\|f\|^2$ would not necessarily be positive.

As an example, let $\mathcal{H}$ consist of a two-dimensional real space V_2, with nonorthogonal unit vectors v_1 and v_2 chosen so that

$$\begin{aligned}&(v_1, v_1) = (v_2, v_2) = 1, \\ &(v_1, v_2) = \cos\theta_{12};\end{aligned} \tag{II. 14}$$

then

$$M = \begin{pmatrix} 1 & \cos\theta_{12} \\ \cos\theta_{12} & 1 \end{pmatrix} \tag{II. 15}$$

which has positive eigenvalues unless $\theta_{12} = 0, \pi$. Clearly, in this case the vectors φ_i correspond to vectors satisfying.

$$(\varphi_1, v_2) = (\varphi_2, v_1) = 0, \tag{II. 16}$$

which has a simple geometrical interpretation. Note that φ_i is not of unit length, for $(\varphi_i, v_i) = 1$,

$$\cos\theta_i = \frac{(\varphi_i, v_i)}{(\varphi_i, \varphi_i)^{1/2}(v_i, v_i)^{1/2}} = \frac{1}{(\varphi_i, \varphi_i)^{1/2}}.$$

Returning now to the Breit–Wigner functions, we have

$$\begin{aligned}M_{ij} &= \int_0^\infty \frac{dx}{(x + \lambda_i^*)(x + \lambda_j)} \\ &= \frac{1}{\lambda_i^* - \lambda_j} \ln \frac{\lambda_i^*}{\lambda_j},\end{aligned} \tag{II. 17}$$

$$M_{ii} = \theta_i / \mathrm{Im}\lambda_i > 0, \qquad |\lambda_i| < \infty, \qquad \theta_i = \arg\lambda_i.$$

If M is a finite-dimensional matrix, corresponding to a choice $\{\lambda_i\}_{i=1}^N$, then, as can be seen from Eq. II. 13, M is positive definite. But if M is infinite dimensional, then the possibility arises that the eigenvalues of M may cluster to zero. It can be shown that the possibility of interpolating with Breit–Wigner amplitudes is tied to those choices of $\{\lambda_i\}_{i=1}^\infty$ for which the eigenvalues of M do not have zero as an accumulation point. In the following sections we will always assume that the expansions are finite expansions.

III. BIORTHOGONAL FUNCTIONS FOR THE BREIT-WIGNER AMPLITUDES

We wish to show that if a function f is expanded as

$$f(x) = \sum_i c_i B_i = \sum_i \frac{c_i}{x + \lambda_i}, \tag{III. 1}$$

then there are functions φ_i satisfying $(\varphi_i, B_i) = \delta_{ij}$, that is

$$\int_0^\infty \frac{\varphi_i^* dx}{x + \lambda_j} = \delta_{ij}. \tag{III. 2}$$

Such a set of functions can be found by making use of the Stieltjes transform,[3] written as

$$\int_0^\infty \frac{\varphi(x) dx}{x + \lambda} = F(\lambda) \tag{III. 3}$$

which has the property that if the integral exists, the function $F(\lambda)$ is analytic in the cut complex λ plane, the cut going from zero to minus infinity. To use the Stieltjes transform we notice that it can be rewritten as

$$\int_0^\infty \frac{\varphi_i^* dx}{x + \lambda} = F_i(\lambda). \tag{III. 4}$$

If a set of functions $F_i(\lambda)$ can be found with appropriately located zeroes, i.e., so that

$$\begin{aligned}&F_i(\lambda_j) = 0, \qquad i \neq j, \\ &F_i(\lambda_i) \neq 0,\end{aligned} \tag{III. 5}$$

then our problem is solved.[4] For, as shown in Ref. 3, if a function $F(\lambda)$ satisfying Eq. (III. 3) is given, it is merely necessary to evaluate the discontinuity across the cut to evaluate the function $\varphi(x)$. This can be seen by noting that

$$\begin{aligned} 2\pi i\varphi(x) &= \lim_{\epsilon\to 0}\left[F(-|x|-i\epsilon)-F(-|x|+i\epsilon)\right] \\ &= \lim_{\epsilon\to 0}\int_0^\infty dx'\left(\frac{\varphi(x')}{x'-|x|-i\epsilon}-\frac{\varphi(x')}{x'-|x|+i\epsilon}\right) \\ &= \lim_{\epsilon\to 0}\int_0^\infty \varphi(x')dx'\,\frac{2i\epsilon}{(x'-|x|)^2+\epsilon^2} \\ &= \int_0^\infty \varphi(x')dx'2\pi i\delta(x'-|x|). \end{aligned} \tag{III. 6}$$

To find functions $F_i(\lambda)$ having zeroes at $\lambda=\lambda_j$, $j\neq i$, we make use of the fact that in the Hardy space H^2 of functions analytic in the open unit disk, there are functions called Blaschke products that are determined solely by the location of their zeroes[5]:

$$B(z)=\prod_n \frac{|a_n|}{a_n}\frac{a_n-z}{1-a_n^*z},\qquad \sum_n(1-|a_n|)<\infty\cdot \tag{III. 7}$$

By defining a new product $B_i(z)$, in which the ith zero is left out of the product, we get a function analytic on the open unit disk which has the property that it is zero for $z=a_n, n\neq i$:

$$\begin{aligned} B_i(z) &= \prod_{\substack{n\\ n\neq i}}\frac{|a_n|}{a_n}\frac{a_n-z}{1-a_n^*z} \\ &= \prod_n{}'\frac{|a_n|}{a_n}\frac{a_n-z}{1-a_n^*z}, \qquad B_i(z)=0, \quad z=a_n, \quad n\neq i. \end{aligned} \tag{III. 8}$$

Now since $B_i(z)$ is analytic on the open unit disk, the conformal map that carries the open unit disk to the cut plane,

$$z=(1-\lambda^{1/2})/(1+\lambda^{1/2}), \tag{III. 9}$$

will generate a set of functions $F_i(\lambda)$ with correctly located zeroes. $B_i(z)$ is thus transformed into

$$\tilde F_i(\lambda)=\beta_i\prod_n{}'\frac{\lambda_n^{1/2}-\lambda^{1/2}}{\lambda_n^{1/2*}+\lambda^{1/2}}, \tag{III. 10}$$

where β_i is a normalization constant to be determined. It is clear that by construction $\tilde F_i$ has the correct analytic behaviour and further has zeroes satisfying Eq. (III. 5).

However, before evaluating the discontinuity across the cut of $\tilde F_i(\lambda)$, it is necessary to check that $\tilde F_i(\lambda)$ has the correct asyomptotic properties, so that the biothogonal set indeed have the property expressed by Eq. (III. 2). The asymptotic properties given in Ref. 3 are

$$\begin{aligned} &\lim_{x\to\infty}F_i^{(m)}(x)x^m=0, \\ &\lim_{x\to 0}F^{(m)}(x)x^{m+1}=0, \qquad m=0,1,2,\cdots, \end{aligned} \tag{III. 11}$$

where $F^{(m)}(x)$ means the mth derivative.

Now it is immediately clear that the functions given by Eq. (III. 10) do not have the proper behavior at infinity, for they go to a constant, not zero. But if the factor $\lambda^{-1/2}$, also having the correct cut structure, is joined to the $\tilde F_i(\lambda)$ of Eq. (III. 10), there results a set of functions

$$F_i(\lambda)=\beta_i\lambda^{-1/2}\prod_n{}'\frac{\lambda_n^{1/2}-\lambda^{1/2}}{\lambda_n^{1/2*}+\lambda^{1/2}} \tag{III. 12}$$

that do have the correct asymptotic behavior. This can be seen by writing

$$\begin{aligned} F_i(x) &= \beta_i\frac{1}{\sqrt{x}}\frac{(\sqrt{\lambda_1}-\sqrt{x})(\sqrt{\lambda_2}-\sqrt{x})\cdots}{(\sqrt{\lambda_1^*}+\sqrt{x})(\sqrt{\lambda_2^*}+\sqrt{x})\cdots}, \\ F_i(x)\sqrt{x}(\sqrt{\lambda_1^*}+\sqrt{x})(\sqrt{\lambda_2^*}+\sqrt{x})\cdots &\sim (\sqrt{\lambda_1}-\sqrt{x})(\sqrt{\lambda_2}-\sqrt{x})\cdots \end{aligned} \tag{III. 13}$$

and differentiating both sides with respect to x m times. What results is a sum of terms expressing $F_i^{(m)}(x)$ as a function of all the lower derivatives. By taking the appropriate limits $x\to\infty$ or $x\to 0+$, it can be seen (although it is rather tedious) that indeed Eq. (III. 11) is satisfied. For example,

$$F_i(x)\to 1/\sqrt{x}, \qquad x\to\infty, \qquad x\to 0+. \tag{III. 14}$$

Then

$$\begin{aligned} F_i'(x)\sqrt{x}(\sqrt{\lambda_1^*}+\sqrt{x})\cdots + F_i(x)\frac{1}{\sqrt{x}}(\sqrt{\lambda_1^*}+\sqrt{x})\cdots+\cdots \\ \approx\frac{1}{\sqrt{x}}(\sqrt{\lambda_2}-\sqrt{x})\cdots+\cdots. \end{aligned} \tag{III. 15}$$

For $x\to 0+$, the various terms go like

$$\begin{aligned} &F_i'(x)\sqrt{x}+F_i(x)(1/\sqrt{x})+F_i(x)\approx(1/\sqrt{x}) \\ &F_i'(x)\approx(1/x)+(1/x^{3/2}), \end{aligned} \tag{III. 16}$$

so that

$$\begin{aligned} &\lim_{x\to 0+}x^2F_i'(x)=0, \\ &\lim_{x\to\infty}xF_i'(x)=0. \end{aligned} \tag{III. 17}$$

Since the functions $F_i(\lambda)$ defined in Eq. (III. 12) now have the correct asymptotic properties, it is possible to evaluate their discontinuities, in order to get the functions $\varphi_i^*(x)$:

$$\begin{aligned} 2\pi i\varphi_i^*(x) &= F_i(xe^{-i\pi})-F_i(xe^{i\pi}) = \beta_i/\sqrt{xe^{-i\pi}}\prod_n{}'\frac{\lambda_n^{1/2}-\sqrt{xe^{-i\pi}}}{\lambda_n^{1/2*}+\sqrt{xe^{-i\pi}}}-\beta_i/\sqrt{xe^{+i\pi}}\prod_n{}'\frac{\lambda_n^{1/2}-\sqrt{xe^{+i\pi}}}{\lambda_n^{1/2*}+\sqrt{xe^{+i\pi}}} \\ &= \frac{i\beta_i}{\sqrt{x}}\left(\prod_n{}'\frac{\lambda_n^{1/2}+i\sqrt{x}}{\lambda_n^{1/2*}-i\sqrt{x}}+\prod_n{}'\frac{\lambda_n^{1/2}-i\sqrt{x}}{\lambda_n^{1/2*}+i\sqrt{x}}\right)=i\beta_i\left(\frac{(1/\sqrt{x})\prod_n'(|\lambda_n|-x+2i\sqrt{x}\,\mathrm{Re}\lambda_n^{1/2})+(1/\sqrt{x})\prod_n'(|\lambda_n|-x-2i\sqrt{x}\mathrm{Re}\lambda_n^{12})}{\prod_n'(x+\lambda_n^*)}\right). \end{aligned} \tag{III. 18}$$

The computation of the biorthogonal functions is the main result of this paper, for it gives the explicit form of $\varphi_i(x)$, so constructed to be orthogonal to the Breit–Wigner functions. It is to be noted that the denominator of $\varphi_i(x)$ contains a product of Breit–Wigner functions, with the ith term in the product missing. Only when $F_i(\lambda_j)$ is evaluated at the zero $\lambda_j=\lambda_i$ is it nonzero; what its value is depends on the normalization factor β_i. From considerations to be discussed in the next section it is best to leave the normalization coefficient unspecified.

IV. A SIMPLE EXAMPLE—COMPARISON WITH ORTHOGONAL EXPANSIONS OVER GROUPS

As a simple example of the general results obtained in Sec. III, consider a function $f(x)$ that can be written as a superposition of two Breit–Wigner amplitudes,

$$f(x) = \sum_{i=1}^{2} c_i B_i = \frac{c_1}{x + \lambda_1} + \frac{c_2}{x + \lambda_2}, \tag{IV. 1}$$

where c_1 and c_2 are to be determined. For such a function the relevant biorthogonal functions are

$$\varphi_1(x) = \frac{\beta_1}{\sqrt{x}} \frac{|\lambda_2| - x}{x + \lambda_2},$$
$$\varphi_2(x) = \frac{\beta_2}{\sqrt{x}} \frac{|\lambda_1| - x}{x + \lambda_1}. \tag{IV. 2}$$

To fix the normalization constants β_i, we demand that

$$(\varphi_1, B_1) = 1 = \beta_1 \int_0^\infty \frac{dx(1/\sqrt{x})(|\lambda_2| - x)}{(x + \lambda_2^*)(x + \lambda_1)} = \frac{\beta_1 \pi}{\lambda_1^{1/2}} \frac{\lambda_2^{1/2} - \lambda_1^{1/2}}{\lambda_2^{1/2*} + \lambda_1^{1/2}},$$

so that

$$\beta_1 = \frac{\lambda_1^{1/2}}{\pi} \frac{\lambda_2^{1/2*} + \lambda_1^{1/2}}{\lambda_2^{1/2} - \lambda_1^{1/2}}$$

with a similar result for β_2. The coefficients c_i of Eq. (IV. 1) are then

$$c_i = (\varphi_i, f), \qquad i = 1, 2. \tag{IV. 3}$$

Of more interest in quantum mechanics is the computation of c_i when $|f(x)|^2$ rather than $f(x)$ is given. We wish to show that Breit–Wigner expansions may have some advantages over expansions in orthonormal sets, especially those generated by function over groups, such as partial-wave expansions.

For the above example we have

$$|f|^2 = \frac{|c_1|^2}{(x + \lambda_1)(x + \lambda_1^*)} + \frac{c_1 c_2^*}{(x + \lambda_1)(x + \lambda_2^*)} + \frac{c_2 c_1^*}{(x + \lambda_2)(x + \lambda_1^*)} + \frac{|c_2|^2}{(x + \lambda_2)(x + \lambda_2^*)}$$
$$= \frac{A_{11}}{(x + \lambda_1)(x + \lambda_1^*)} + \frac{A_{12}}{(x + \lambda_1)(x + \lambda_2^*)} + \frac{A_{21}}{(x + \lambda_2)(x + \lambda_1^*)} + \frac{A_{22}}{(x + \lambda_2)(x + \lambda_2^*)}. \tag{IV. 4}$$

To find the coefficients A_{ij}, define the symbol $(\lambda_1, \lambda_2 \ldots, \lambda_N)$ as the numerator of the functions $\varphi_i^*(x)$ (up to normalization), noting that it is symmetric in all the λ_i. For example,

$$(\lambda_1) = \frac{1}{\sqrt{x}}(|\lambda_1| - x),$$
$$(\lambda_1, \lambda_2) = (1/\sqrt{x})[(|\lambda_1| - x)(|\lambda_2| - x) - 4x \operatorname{Re}\lambda_1^{1/2} \operatorname{Re}\lambda_2^{1/2}]. \tag{IV. 5}$$

Then the following equations in A_{ij} can be obtained using the biorthogonality properties of the φ_i^*:

(1) Multiply by $[(\lambda_1, \lambda_2)/(x + \lambda_2^*)]$ and integrate

$$\int \frac{dx(\lambda_1, \lambda_2)}{x + \lambda_2^*} |f|^2 = A_{12} \int \frac{dx(\lambda_1, \lambda_2)}{(x + \lambda_1)(x + \lambda_2^*)^2} + A_{22} \int \frac{dx(\lambda_1, \lambda_2)}{(x + \lambda_2)(x + \lambda_2^*)^2}.$$

(2) Multiply by $[(\lambda_1, \lambda_2)/(x + \lambda_1^*)]$ and integrate

$$\int \frac{dx(\lambda_1, \lambda_2)|f|^2}{(x + \lambda_1^*)} = A_{11} \int \frac{dx(\lambda_1, \lambda_2)}{(x + \lambda_1)(x + \lambda_1^*)^2} + A_{21} \int \frac{dx(\lambda_1, \lambda_2)}{(x + \lambda_2)(x + \lambda_1^*)^2}.$$

(3) Multiply by (λ_1) and integrate

$$\int dx(\lambda_1)|f|^2 = A_{12} \int \frac{(\lambda_1)dx}{(x + \lambda_1)(x + \lambda_2^*)} + A_{21} \int \frac{(\lambda_1)dx}{(x + \lambda_2)(x + \lambda_1^*)} + A_{22} \int \frac{(\lambda_1)dx}{(x + \lambda_2)(x + \lambda_2^*)}.$$

(4) Multiply by (λ_2) and integrate

$$\int dx(\lambda_2)|f|^2 = A_{11} \int \frac{(\lambda_2)dx}{(x + \lambda_1)(x + \lambda_1^*)} + A_{12} \int \frac{(\lambda_2)dx}{(x + \lambda_1)(x + \lambda_2^*)} + A_{21} \int \frac{(\lambda_2)dx}{(x + \lambda_2)(x + \lambda_1^*)}. \tag{IV. 6}$$

There are four complex equations in four unknowns A_{ij} in Eq. (IV. 6). The integrals are all readily evaluated if $|f|^2$ is known so that it should be possible to solve for the A_{ij}, subject to the constraints that $|A_{12}| = |A_{21}| = \sqrt{|A_{11}||A_{22}|}$ and $A_{12} = A_{21}^*$, since the overall phase cannot be computed.

This procedure obviously generalizes to superpositions of N Breit–Wigner resonance amplitudes, so that, if

$$f(x) = \sum_{i=1}^{N} \frac{c_i}{x + \lambda_i}, \tag{IV. 7}$$

where the c_i are unknown and to be determined from knowledge of $|f|^2$, there results a set of N^2 equations in N^2 unknowns A_{ij}, where $A_{ii} = |c_i|^2$ and $A_{ij} = A_{ji}^* = c_i c_j^*$, $i \neq j$. The A_{ij} are thus obtained from $|f|^2$, but then directly yield the magnitudes of c_i from the diagonal A_{ii} and the phases of c_i from the off diagonal A_{ij}. If $|f|^2$ is given from experimental data, then the fact that A_{ij} must be a rank one Hermitian matrix puts constraints on the allowed fits. Presumably it should be possible to vary the values of λ_i about some "eyeballed" values and see when, for different values of λ_i, the constraints on A_{ij} are satisfied; the numerical analysis required for such a program will be discussed in a future publication.

This entire procedure is to be compared with data fitting that is done in terms of expansions generated by group matrix elements. If $F(g)$ is a function over a (compact) group G and $\mathfrak{D}^{\chi}_{MM'}(g)$ an irreducible matrix element, then

$$F(g) = \sum_{\chi MM'} C_{\chi MM'} \mathfrak{D}^{\chi}_{MM'}(g),$$
$$|F(g)|^2 = \sum_{\substack{\chi_1 M_1 M_1' \\ \chi_2 M_2 M_2'}} C_{\chi_1 M_1 M_1'} C^*_{\chi_2 M_2 M_2'} \mathfrak{D}^{\chi_1}_{M_1 M_1'}(g) \mathfrak{D}^{\chi_2 *}_{M_2 M_2'}(g) = \sum_{\chi_3 M_3 M_3'} C_{\chi_3 M_3 M_3'} \mathfrak{D}^{\chi_3}_{M_3 M_3'}(g), \tag{IV. 8}$$

where the coefficients $C_{\chi_3 M_3 M_3'}$ are related to $C_{\chi_1 M_1 M_1'}$ and $C_{\chi_2 M_2 M_2'}$ by appropriate Clebsch–Gordan coefficients. For example, if G is the rotation group $O(3)$ and F is a function over the sphere, then $F(\hat{n}) = \sum_{LM} C_{LM} Y_{LM}(\hat{n})$. Though it is in general straightforward to obtain the coefficients $C_{\chi_3 M_3 M_3'}$ from $|F|^2$ by making use of the orthogonality relations of the $\mathfrak{D}$ functions, it is in general quite difficult to get back to the coefficients $C_{\chi_1 M_1 M_1'}$ because these coefficients are nonlinearly related to the coefficients $C_{\chi_3 M_3 M_3'}$. This difficulty is well known in phase shift analysis where one imposes many other constraints in order to get a unique $C_{\chi_1 M_1 M_1'}$ from the $C_{\chi_3 M_3 M_3'}$.

CONCLUSION

It has been shown that for any finite superposition of Breit–Wigner amplitudes a biorthogonal set of functions including the $\varphi_i(x)$ exist which allow one to compute the expansion coefficients in the superposition. Several problems remain, however, of both a mathematical and physical character.

Although it has been shown (Eq. (II. 11)ff.) that the matrix M is positive definite in the finite-dimensional case, it remains to be shown that in the limit as M becomes infinite dimensional, the eigenvalues do not cluster to zero; that is, it remains to be shown that when M is infinite dimensional, it remains positive definite. The proof of M being positive definite even in the infinite-dimensional case would presumably involve some conditions on the λ_i, how they are separated[4] and how they go to infinity. There is already a condition on the $\{\lambda_i\}$, given with respect to the zeroes $\{a_i\}$ of Eq. (III. 7), but presumably other conditions would also be needed for the positive definiteness of M

From a physical point of view there is no reason to demand that scattering amplitudes are square integrable in the energy. In fact, it is well known that amplitudes may grow at infinity roughly like E^2; [6] this indicates that a Hilbert space does not provide the correct setting for the expansion of general scattering amplitudes;[7] further, since the Breit–Wigner amplitudes all go to zero like E^{-2} at infinity, some sort of clustering of the $\{\lambda_i\}$ at infinity would be needed to reproduce the E^2 dependence. But such a clustering would violate the Hilbert space constraint and undoubtedly also violate the positive definiteness of M. For this reason the positive definiteness of M in the infinite-dimensional case was not investigated. For finite energies, with finite superposition of Breit–Wigner amplitudes, no such problem arises and one can compute the expansion coefficients with aid of the φ_i functions. In this light one can regard the Hilbert space as providing a convenient means for reading off the expansion coefficients. The real problem arises when one wants to reproduce the "background" by a superposition of Breit–Wigner amplitudes; then it might be necessary to use an infinite number of Breit–Wigner amplitudes, depending on how "flat" the background is. It is to be noted that the functions $\{\varphi_i\}$ are not square integrable on $(0, \infty)$, at both limits going like $\ln x$. But when integrated against the Breit–Wigner functions the results are all well defined.

Although there are problems arising when the energy becomes infinite, it is to be hoped that the analysis presented in Sec. IV will be useful in a practical sense. By viewing experimental data and seeing how many resonances there are that need to be fitted, it should be possible to use the appropriate φ_i functions and develop a computer program that takes advantage of the manner in which it is possible to read off the expansion coefficients from the data. Since the position and widths of resonances get shifted when several resonances overlap significantly, it will be necessary to develop a routine that allows for variations in λ_i about some "eyeballed" mean values, consistent with the constraints that must hold for the magnitudes and phases of the expansion coefficients.

The author wishes to thank Dr. Paul Muhly for many helpful mathematical discussions.

[1]There is a long history, both experimental and theoretical, of using Breit-Wigner amplitudes to fit data. Typical examples are given in: J. B. Marion, P. H. Nettles, C. L. Cocke, and G. J. Stephenson, Jr., Phys. Rev. **157**, 847 (1967); E. Norbeck, L. L. Gadeken, and F. D. Ingram, Phys. Rev. **C3**, 2073 (1971).

[2]H. S. Shapiro and A. L. Shields, Amer. J. Math. **83**, 513 (1961).

[3]D. Widder, *The Laplace Transform* (Princeton U. P., Princeton, N. J., 1941), Chap. VIII.

[4]P. L. Duren, *Theory of H^p Spaces* (Academic, New York, 1970). Chap. 9.

[5]Reference 4, Chap. 2.

[6]See, for example, R. J. Eden, *High Energy Collisions of Elementary Particles* (Cambridge U. P., Cambridge, 1967), Chap. 6.

[7]C. Cronström and W. H. Klink, Ann. Phys. **69**, 218 (1972).

Spheroidal analysis of Coulomb scattering

Ming Chiang Li

Department of Physics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
(Received 10 September 1973)

In studying the spheroidal problem of a charged particle scattered by two charged centers, we have had to deal with a differential equation. Its solution was complicated. In this paper, we study a very similar differential equation, which appears in the Coulomb problem. The solution for this equation, however, can be put in a simple form. For completeness the spheroidal analysis of the Coulomb scattering amplitude is also discussed.

1. INTRODUCTION

In a series of papers[1] we have studied spheroidal potential scattering. The importance of such a study is two fold. First, it denotes a class of real physical problems, such as the scattering of electrons by diatomic molecules,[2] and of deformed nuclei.[3] Second, it describes the scattering between a particle and the simplest composite system which can be formed by two particles separated at a fixed distance.

The spheroidal scattering is less understood than the spherical scattering because of its complexity. In spherical scattering, after the separation of variables, only the radial equation depends on the potential. For normal cases, the equation contains two singularities, one at $r = 0$ and the other at $r = \infty$. But in spheroidal scattering, both the radial and angle equations can depend on the potential. For example, the Coulomb potential of a pair of fixed unequal charges appears in both the radial and angle equations. Thus, we have to solve two ordinary differential equations instead of just one as in the case of spherical scattering. These equations, even without the potential term, have two regular and one irregular singularities, and are in the class of Lamé differential equations.[4] In general, they are more difficult to solve.

In spheroidal scattering, the scattering amplitude is expressed in terms of spheroidal angle functions. This decomposition may be referred to as spheroidal analysis,[5] as compared with the partial wave analysis in spherical scattering. These angle functions do not obey conventional recurrence relations.[6] The so-called "recurrence relations" for these functions are actually the integral relations among them. These functions cannot be simply expressed, and their numerical construction is also troublesome. The trouble leads to a dilemma on the numerical treatment of spheroidal problems. The dilemma is on making a choice between the one-center or two-center approaches. This is a choice between spherical and spheroidal expansions of the scattering amplitude. The former expansion is an unnatural one and leads to a very slow convergent series; but each term in the series is easy to calculate. The later expansion is a natural one and leads to a very fast convergent series, but each term in the series is difficult to calculate. The difficulty is mainly due to the troublesome numerical construction of spheroidal angle functions, which makes the spheroidal analysis in the numerical approach unattractive.

The complexity of the spheroidal problem has led to a theoretical effort in a different direction.[7] Before treating the spheroidal scattering problem, one often uses some kind of approximations and tries to make problems simple and manageable. A good example is the scattering theory of electrons by polar molecules, in which the dipole is replaced by a point dipole.[7] According to its definition, a point dipole is a dipole without a spatial extent. The main idea behind the effort is to reduce the spheroidal problem into a spherical problem.

Despite its complexity, the spheroidal problem is exactly solvable.[8] It is a problem describing scattering by a composite system. Most problems in this category are not solvable without some kind of approximations. The well-founded scattering theories, whether they are potential or quantum field types, are basically theories for pointlike particles. Therefore, a fully understood scattering theory for any composite system, such as a spheroidal one, is of great value.

Our interest here is in the spheroidal aspect of the Coulomb problem. The importance of spheroidal scattering and its encountered difficulties suggest that we first deal with similar problems with fewer complications. The Coulomb problem is just one of them. Our interest in the Coulomb problem has a stronger basis. In studying the spheroidal scattering,[1] we have had to deal with a differential equation describing a charged particle scattered by two-charged centers. It was treated by constructing an integro-differential equation which, in turn, was solved by means of a substracted iteration method.[1] The final expression was messy. It was observed later that a very similar differential equation also appeared in the Coulomb problem. It is obtained through a spheroidal separation of the wave equation. This has led us to the interesting question, as to whether this equation can be solved differently and simply.

In this paper, along with the problem of main interest to us, we present a detailed spheroidal analysis of the Coulomb problem. In Sec. 2, the spheroidal symmetry of the Coulomb problem is discussed and the resulting angle function is considered. In Sec. 3, we treat the solution of the radial equation. This equation is very similar to that obtained from a charged particle scattered by two charged centers. In Sec. 4, the spheroidal analysis of the Coulomb scattering amplitude is given.

2. COULOMB EQUATION AND ANGLE FUNCTIONS

A fixed charge Q is located on the z axis with coordinates $\mathbf{r}_a = (0, 0, d/2)$, where d is the interfocal distance of the prolate spheroidal coordinates[4]:

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

with $1 \le \xi < \infty$, $-1 \le \eta \le 1$, $0 \le \phi \le 2\pi$.
The Coulomb potential V at distance $\mathbf{r}$ has the form

$$V = \frac{Q}{|\mathbf{r}-\mathbf{r}_a|} = \frac{2Q}{d}\,\frac{\xi+\eta}{\xi^2-\eta^2}. \tag{2.2}$$

This is a potential with prolate spheroidal symmetry. The Schrödinger equation for describing a charged particle scattered by the potential in Eq. (2.2) has the form

$$-(\hbar^2/2\mu)\nabla^2\psi + V\psi = (\hbar^2k^2/2\mu)\psi, \tag{2.3}$$

where μ is the mass, q the charge, and k the momentum of the incident particle. The above equation may be expressed simply in the prolate spheroidal coordinates as

$$\left[\frac{\partial}{\partial\eta}(1-\eta^2)\frac{\partial}{\partial\eta} + \frac{\partial}{\partial\xi}(\xi^2-1)\frac{\partial}{\partial\xi} + \left(\frac{1}{1-\eta^2}+\frac{1}{\xi^2-1}\right)\frac{\partial^2}{\partial\phi^2} + (\tfrac{1}{2}kd)^2(\xi^2-\eta^2) - \frac{\mu dQq}{\hbar^2}(\xi+\eta)\right]\psi = 0. \tag{2.4}$$

The equation is separable and its solution can be obtained in the form of the Lamé products:

$$\psi_{mn} = R_{mn}(A, C; \xi)\, S_{mn}(A, C; \eta) \begin{cases} \cos m\phi \\ \sin m\phi. \end{cases} \tag{2.5}$$

Functions $R_{mn}(A, C; \xi)$ and $S_{mn}(A, C; \eta)$ satisfy ordinary differential equations

$$\frac{d}{d\xi}\left((\xi^2-1)\frac{d}{d\xi}R_{mn}(A, C; \xi)\right) - \left(\Lambda_{mn}(A, C) - C^2\xi^2 + AC\xi + \frac{m^2}{\xi^2-1}\right)R_{mn}(A, C, \xi) = 0, \tag{2.6}$$

$$\frac{d}{d\eta}\left((1-\eta^2)\frac{d}{d\eta}S_{mn}(A, C; \eta)\right) + \left(\Lambda_{mn}(A, C) - C^2\eta^2 - AC\eta - \frac{m^2}{1-\eta^2}\right)S_{mn}(A, C; \eta) = 0, \tag{2.7}$$

where

$$C = \tfrac{1}{2}kd, \qquad A = (2\mu Qq)/k\hbar^2 \tag{2.8}$$

and constant m is an integer, which comes from the single-value requirement of the wavefunction in Eq. (2.1). The separation constant $\Lambda_{mn}(A, C)$ is chosen so as to ensure that the solution $S_{mn}(A, C; \eta)$ of Eq. (2.7) which is free of logarithmic terms is regular in the entire range $-1 \le \eta \le 1$. The functions $S_{mn}(A, C; \eta)$ also appear in scattering of a charged particle by a dipole. For convenience we will refer to $S_{mn}(A, C; \eta)$ as the Coulomb spheroidal angle function and $R_{mn}(A, C; \eta)$ as the Coulomb spheroidal radial function.

When C vanishes, the differential Eq. (2.7) becomes the one which is satisfied by the associated Legendre functions. It follows that the angle functions must reduce to the associated Legendre functions of the integral order and degree, as C goes to zero. Therefore,

$$\Lambda_{mn}(0, 0) = n(n+1). \tag{2.9}$$

When C is not zero, Eq. (2.7) differs from the associated Legendre equation by having an irregular singularity at infinity. This suggests for the angle functions, an infinite sum of the form

$$S_{mn}(A, C; \eta) = \sum_{r=0}^{\infty} D_r^{mn}(A, C) P_{m+r}^m(\eta). \tag{2.10}$$

Substitution of Eq. (2.10) in Eq. (2.7), with the subsequent use of the associated Legendre differential equation and of the recursion formulas for the associated Legendre functions, yields the following recursion formulas for the coefficients $D_r^{mn}(A, C)$:

$$\begin{aligned}&\frac{(2m+r+2)(2m+r+1)C^2}{(2m+2r+3)(2m+2r+5)}D_{r+2}^{mn}(A, C) \\ &+ \frac{(2m+r+1)AC}{(2m+2r+3)}D_{r+1}^{mn}(A, C) \\ &+ \Big((m+r)(m+r+1) - \Lambda_{mn}(A, C) \\ &+ \frac{2(m+r)(m+r+1) - 2m^2 - 1}{(2m+2r-1)(2m+2r+3)}C^2\Big)D_r^{mn}(A, C) \\ &+ \frac{rAC}{2m+2r-1}D_{r-1}^{mn}(A, C) \\ &+ \frac{r(r-1)C^2}{(2m+2r-3)(2m+2r-1)}D_{r-2}^{mn}(A, C) = 0.\end{aligned} \tag{2.11}$$

The recursion formula for the expansion coefficients constitutes a linear homogeneous difference equation of the fourth order. When A vanishes, the recursion formula reduces to one which is satisfied by the expansion coefficients $d_r^{mn}(C)$ of the spheroidal angle functions in terms of the associated Legendre functions. We now substitute in Eq. (2.11) the expansions

$$D_r^{mn}(A, C) = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} B_{r(j,k)}^{mn} C^{2j}(AC)^k, \tag{2.12}$$

$$\Lambda_{mn}(A, C) = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} l_{(j,k)}^{mn} C^{2j}(AC)^k \tag{2.13}$$

and use the perturbation method to calculate expansion coefficients $B_{r(j,k)}^{mn}$ and $l_{(j,k)}^{mn}$. The lower-order coefficients are found to be

$$\begin{aligned}
&l_{(0,0)}^{mn} = n(n+1), \\
&l_{(1,0)}^{mn} = \tfrac{1}{2}\left(1 - \frac{(2m-1)(2m+1)}{(2n-1)(2n+3)}\right), \\
&l_{(2,0)}^{mn} = \tfrac{1}{2}\left(\frac{(n-m-1)(n-m)(n+m-1)(n+3)}{(2n-3)(2n-1)^3(2n+1)} - \frac{(n-m+1)(n-m+2)(n+m+1)(n+m+2)}{(2n+1)(2n+3)^3(2n+5)}\right) \\
&l_{(0,1)}^{mn} = 0, \\
&l_{(0,2)}^{mn} = \frac{(n-m)(n+m)}{2n(2n-1)(2n+1)} - \frac{(n+m+1)(n-m+1)}{(2n+3)^2(2n+1)}, \\
&l_{(1,1)}^{mn} = 0, \\
&B_{r(0,0)}^{mn} = \delta_{n-m,r}, \\
&B_{r(1,0)}^{mn} = \tfrac{1}{2}\delta_{n-m,r+2}\frac{(2m+r+2)(2m+r+1)}{(2m+2r+3)^2(2m+2r+5)} - \tfrac{1}{2}\delta_{n-m,r-2}\frac{r(r-1)}{(2m+2r-3)(2m+2r-1)}, \\
&B_{r(0,1)}^{mn} = \tfrac{1}{2}\delta_{n-m,r+1}\frac{(2m+2r+1)}{(2m+2r+3)(m+r+1)} - \tfrac{1}{2}\delta_{n-m,r-1}\frac{r}{(2m+2r-1)(m+r)}.
\end{aligned} \tag{2.14}$$

We have normalized the coefficients so that each Coulomb spheroidal angle function reduces exactly to the corresponding associated Legendre function when C becomes zero. From the general theory of Sturm-Liouville differential equations it follows that the functions $S_{mn}(A, C; \eta)$ form a complete orthogonal set on the interval $(-1, 1)$. Thus

$$\int_{-1}^{1} S_{mn}(A, C; \eta) S_{mn'}(A, C; \eta) d\eta = \delta_{nn'} N_{mn}(A, C), \quad (2.15)$$

where $N_{mn}(A, C)$ is easily found with the use of the normalization factor of the associated Legendre functions to be

$$N_{mn}(A, C) = 2 \sum_{r=0}^{\infty} \frac{(r + 2m)!}{(2r + 2m + 1)r!} [D_r^{mn}(A, C)]^2. \quad (2.16)$$

Near the end points $\eta = \pm 1$, the associated Legendre functions $P_n^m(\eta)$ have the following behavior:

$$\begin{aligned} P_n^m(\eta) &= O[(1 - \eta^2)^{m/2}] \quad \text{for } (n - m) \text{ even}, \\ P_n^m(\eta) &= O[\eta(1 - \eta^2)^{m/2}] \quad \text{for } (n - m) \text{ odd}. \end{aligned} \quad (2.17)$$

From these relations, one obtains the following behavior of the Coulomb spheroidal wavefunction near $\eta = \pm 1$ with the aid of Eq. (2.10):

$$S_{mn}(A, C; \eta) = O[(1 - \eta^2)^{m/2}]. \quad (2.18)$$

3. RADIAL FUNCTION

The differential equation (2.6) satisfied by the Coulomb spheroidal radial function is very similar to one encountered in the scattering of two charged centers.[1] The difference between these two is in the parameter dependence. The differential equation (2.6) depends on the Coulomb spheroidal eigenvalue $\Lambda_{mn}(A, C)$. The other one depends on the spheroidal eigenvalue $\lambda_{mn}(C) \equiv \Lambda_{mn}(0, C)$. We have developed a method for solving the latter equation. In the method, an integro-differential equation is used to express the solution of the differential equation. After removal of the second-order derivative, the integro-differential equation is solved through an iteration procedure. The series is convergent under a proper substraction. The same method can also be directly applied to the present equation (2.6). Due to the complexity of the method and the messiness of the expression, we shall approach the present problem through a different path.

From the general theory of integral representations of solutions of differential equations, it follows that the function defined by the integral

$$\int_a^b K_m(\xi, \eta) S_{mn}(A, C; \eta) d\eta \quad (3.1)$$

is a solution of the Coulomb radial differential equation (2.6), provided that the limits a and b are so chosen that the bilinear concomitant,

$$(1 - \eta^2) \left(\frac{\partial K_m(\xi, \eta)}{\partial \eta} S_{mn}(A, C; \eta) - K_m(\xi, \eta) \frac{\partial}{\partial \eta} S_{mn}(A, C; \eta) \right), \quad (3.2)$$

vanishes at both limits, and that the kernel $K_m(\xi, \eta)$ satisfies the differential equation

$$\begin{aligned} &\left(\frac{\partial}{\partial \eta} (1 - \eta^2) \frac{\partial}{\partial \eta} + \frac{\partial}{\partial \xi} (\xi^2 - 1) \frac{\partial}{\partial \xi} \right) \\ &\quad - m^2 \left(\frac{1}{1 - \eta^2} + \frac{1}{\xi^2 - 1} \right) + C^2(\xi^2 - \eta^2) \\ &\quad - AC(\xi + \eta)] K_m(\xi, \eta) = 0. \end{aligned} \quad (3.3)$$

By comparing Eq. (3.3) with Eq. (2.4) we observe that the kernel $K_m(\xi, \eta)$ is actually a solution of the Coulomb scattering equation[9] with the azimuthal angle ϕ dependence factored out. If a fixed charge Q is at the origin of a coordinate system, one of the solutions for the Coulomb scattering equation has the form

$$e^{(ik/2)(\xi_0 - \eta_0)} \xi_0^{m/2} \eta_0^{m/2} \, {}_1F_1(-\tfrac{1}{2}Ai; m + 1; ik\eta_0), \quad (3.4)$$

where the function ${}_1F_1(x; y; z)$ is a confluent hypergeometric function,[10] ξ_0 and η_0 are the parabolic coordinates of the above coordinate system. In the present investigation we used a different coordinate system and placed the fixed charge Q on the z axis with coordinates $r_a = (0, 0, \frac{1}{2}d)$. The relations of these two coordinate systems are as follows:

$$\eta_0 = \tfrac{1}{2}d(\xi + 1)(1 - \eta), \qquad \xi_0 = \tfrac{1}{2}d(\xi - 1)(1 + \eta). \quad (3.5)$$

By using the above relations the solution in Eq. (3.4) may be expressed as

$$\begin{aligned} (\tfrac{1}{2}d)^m e^{ic(\xi\eta - 1)} (\xi^2 - 1)^{m/2} (1 - \eta^2)^{m/2} \, {}_1F_1(-\tfrac{1}{2}Ai; m \\ + 1; ic(\xi + 1)(1 - \eta)). \end{aligned} \quad (3.6)$$

It may be verified directly that the function in Eq. (3.6) satisfies the differential equation in Eq. (3.3). We denote this function as kernel $K_m^{(1)}(\xi, \eta)$. The two obvious limits a and b, at which the bilinear concomitant in Eq. (3.2) vanishes, are

$$a = -1, \qquad b = 1. \quad (3.7)$$

The Coulomb spheroidal radial function based on the kernel $K_m^{(1)}(\xi, \eta)$ will be called $R_{mn}^{(1)}(A, C; \xi)$:

$$R_{mn}^{(1)}(A, C; \xi) \equiv \int_{-1}^{1} K_m^{(1)}(\xi, \eta) S_{mn}(A, C; \eta) d\eta. \quad (3.8)$$

The function $R_{mn}^{(1)}(A, C; \xi)$ is regular at $\xi = 1$, at which the differential equation (2.6) has a regular singularity. In the following we will discuss its asymptotic forms, which has a special interest in the scattering theory. After the substitution of the expansion of Coulomb spheroidal angle function $S_{mn}(A, C; \eta)$ in Eq. (2.10) to Eq. (3.8), we have

$$\begin{aligned} R_{mn}^{(1)}(A, C; \xi) &= \sum_{r=0}^{\infty} D_r^{mn}(A, C) \int_{-1}^{1} K_m^{(1)}(\xi, \eta) P_{m+r}^m(\eta) d\eta \\ &= \sum_{r=0}^{\infty} D_r^{mn}(A, C) \left(\frac{d}{2} \right)^m (\xi^2 - 1)^{m/2} \int_{-1}^{1} d\eta \, P_{m+r}^m(\eta) \\ &\quad \times e^{ic(\xi\eta - 1)} (1 - \eta^2)^{m/2} \\ &\quad \times {}_1F_1(-\tfrac{1}{2}Ai; m + 1; ic(\xi + 1)(1 - \eta)). \end{aligned} \quad (3.9)$$

The confluent hypergeometric function ${}_1F_1(x; y; z)$ has the integral representation

$$\begin{aligned} {}_1F_1(\alpha; \gamma; z) = -\frac{1}{2\pi i} \frac{\Gamma(1 - \alpha)\Gamma(\gamma)}{\Gamma(\gamma - \alpha)} \\ \times \oint_{c'} e^{tz} (-t)^{\alpha - 1} (1 - t)^{\gamma - \alpha - 1} dt. \end{aligned} \quad (3.10)$$

If γ is a positive integer and $\mathrm{Re}(\gamma - \alpha) > 0$, contour c' can be any contour which passes around both the points $t = 0$ and $= 1$. We recall the integration

$$\int_{-1}^{1} e^{i\eta z} (1 - \eta^2)^{m/2} P_{m+r}^m(\eta) d\eta = \frac{2i^n (2m + r)!}{r!} \frac{j_{m+r}(z)}{z^m}, \quad (3.11)$$

where $j_{m+r}(z)$ is a spherical Bessel function. After

using Eqs. (3.10) and (3.11) the function $R^{(1)}_{mn}(A, C; \xi)$ in Eq. (3.9) becomes

$$R^{(1)}_{mn}(A, C; \xi) = 2\left(\frac{d}{2}\right)^m (\xi^2 - 1)^{m/2} \frac{\Gamma(1 + \frac{1}{2}Ai)\Gamma(m + 1)}{\Gamma(m + 1 + \frac{1}{2}Ai)} \times \sum_{r=0}^{\infty} D^{mn}_r(A, C) \frac{i^r(2m + r)!}{r!} K^m_r(A, C; \xi), \quad (3.12)$$

where

$$K^m_r(A, C; \xi) \equiv -\frac{1}{2\pi i} \oint_{c'} dt \{C[\xi(1 - t) - t]\}^{-m} \times j_{m+r}\{c[\xi(1-t)-t]\}(-t)^{-(Ai/2)-1}(1 - t)^{m+Ai/2} \times e^{ic[t(\xi+1)-1]}. \quad (3.13)$$

The spherical Bessel function is related to the confluent hypergeometric function

$$z^{-n} j_n(z) = \frac{\sqrt{\pi}}{2^{n+1}\Gamma(n + \frac{3}{2})} e^{-iz}\ {}_1F_1(n + 1; 2n + 2; 2iz). \quad (3.14)$$

The confluent hypergeometric function satisfies the Kummer transformation

$${}_1F_1(\alpha; \gamma; z) = e^z\ {}_1F_1(\gamma - \alpha; \gamma; -z). \quad (3.15)$$

With the help of Eqs. (3.14) and (3.15), we rewrite Eq. (3.13) as

$$K^m_r(A, C; \xi) = -\frac{1}{2\pi i} \frac{\sqrt{\pi}\, 2^{-(m+r+1)}}{\Gamma(m + r + \frac{3}{2})} e^{ic(\xi-1)} \times \oint_{c'} dt[C\{\xi(1 - t) - t\}]^r(-t)^{-(Ai/2)-1}(1 - t)^{m+(Ai/2)} \times {}_1F_1(m + r + 1; 2m + 2r + 2; -2ic[\xi(1 - t) - t]). \quad (3.16)$$

The confluent hypergeometric function has a series expansion

$${}_1F_1(\alpha; \gamma; z) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)} \sum_{k=0}^{\infty} \frac{\Gamma(\alpha + k)}{\Gamma(\gamma + k)\Gamma(k + 1)} z^k. \quad (3.17)$$

The hypergeometric function has a contour integral representation

$${}_2F_1(\alpha, \beta; \gamma; z) \equiv -\frac{1}{2\pi i} \frac{\Gamma(1 - \alpha)\Gamma(\gamma)}{\Gamma(\gamma - \alpha)} \times \oint_{c'} dt(-t)^{\alpha-1}(1 - t)^{\gamma-\alpha-1}(1 - tz)^{-\beta}. \quad (3.18)$$

From Eqs. (3.17) and (3.18) we may express Eq. (3.16) as

$$K^m_r(A, C; \xi) = \frac{2^{m+r}\Gamma(m + 1 + Ai/2)}{\Gamma(1 + Ai/2)\Gamma(m + 1)} e^{ic(\xi-1)} \times \sum_k \frac{\Gamma(m + r + 1 + k)}{\Gamma(2m + 2r + 2 + k)\Gamma(k + 1)} (c\xi)^{r+k}(-2i)^k \times {}_2F_1(-Ai/2, -r - k; m + 1; (\xi + 1)/\xi). \quad (3.19)$$

In reaching Eq. (3.19) these identities are also used:

$$\Gamma(\tfrac{1}{2}) = \pi,$$
$$\Gamma(n + \tfrac{1}{2}) = \frac{1.3.5\ldots(2n - 1)}{2^n}\Gamma(\tfrac{1}{2}). \quad (3.20)$$

Equation (3.12) together with Eq. (3.19) yields a complete series solution of the Coulomb spheroidal radial function $R^{(1)}_{mn}(A, C; \xi)$. If we used the general method,[1] which was previously suggested to solve the integro-differential equation arising from the spheroidal scattering problem we would obtain another series expansion for the function $R^{(1)}_{mn}(a, C, \xi)$. However the latter expansion has more complicated form. The asymptotic form of the function $R^{(1)}_{mn}(A, C; \xi)$ has considerable interests in the scattering theory. This form follows directly from the series expansion presented here.

The hypergeometric function satisfies a relation

$${}_2F_1(\alpha, \beta; \gamma; z) = \frac{\Gamma(\gamma)\Gamma(\gamma - \alpha - \beta)}{\Gamma(\gamma - \alpha)\Gamma(\gamma - \beta)} {}_2F_1(\alpha, \beta; \alpha + \beta + 1 - \gamma; 1 - z) + \frac{\Gamma(\gamma)\Gamma(\alpha + \beta - \gamma)}{\Gamma(\alpha)\Gamma(\beta)} (1 - z)^{\gamma-\alpha-\beta} \times {}_2F_1(\gamma - \alpha, \gamma - \beta; \gamma + 1 - \alpha - \beta; 1 - z). \quad (3.21)$$

Equation (3.21) leads us to the following asymptotic form function $K^m_r(A, C; \xi)$ in Eq. (3.19):

$$K^m_r(A, C; \xi) \underset{\xi\to\infty}{\longrightarrow} \frac{2^{m+r}\Gamma(m + r + 1 + Ai/2)}{\Gamma(1 + Ai/2)\Gamma(2m + 2r + 2)} (c\xi)^r e^{ic(\xi-1)} \times {}_1F_1(m + r + 1 + Ai/2; 2m + 2r + 2; -2ic\xi). \quad (3.22)$$

In arriving at Eq. (3.22), we also used Eq. (3.17) and the equation

$$1/\Gamma(-n) = 0, \quad n = 0, 1, 2, \cdots. \quad (3.23)$$

The spherical regular Coulomb wave function $F_L(\frac{1}{2}A, \rho)$ has the form

$$F_L(A/2, \rho) = C_L(A/2)\rho^{L+1}e^{i\rho}\,{}_1F_1(L + 1 + Ai/2, 2L + 2; -2i\rho), \quad (3.24)$$

where

$$C_L\left(\frac{A}{2}\right) = \frac{2^L e^{-\pi A/4}|\Gamma(L + 1 + iA/2)|}{\Gamma(2L + 2)}. \quad (3.25)$$

From Eqs. (3.12), (3.22), and (3.24) we obtain the asymptotic form of the Coulomb spheroidal radial function

$$R^{(1)}_{mn}(A, C; \xi) \underset{\xi\to\infty}{\longrightarrow} \frac{2\Gamma(m + 1)}{C\xi k^m\Gamma(m + 1 + Ai/2)} \exp\left(-ic + \frac{\pi A}{4}\right) \times \sum_{r=0} D^{mn}_r(A, C) \frac{i^r(2m + r)!}{r!} e^{i\sigma_{m+r}} F_{m+r}\left(\frac{A}{2}, C\xi\right), \quad (3.26)$$

where

$$\sigma_{m+r} = \arg\Gamma(m + r + 1 + Ai/2) \quad (3.27)$$

is the spherical Coulomb phase shift. The spherical regular Coulomb wave function $F_L(A/2, \rho)$ has the asymptotic form

$$F_L(\tfrac{1}{2}A, \rho) \underset{\rho\to\infty}{\longrightarrow} \sin(\rho - \tfrac{1}{2}A\ln 2\rho - \tfrac{1}{2}L\pi + \sigma_L). \quad (3.28)$$

The final asymptotic form of the Coulomb spheroidal radial function follows directly from Eqs. (3.26) and (3.28);

$$R^{(1)}_{mn}(A, C; \xi) \underset{c\xi\to\infty}{\longrightarrow} \frac{2\Gamma(m + 1)}{c\xi k^m\Gamma(m + 1 + \frac{1}{2}Ai)} \exp\left(-ic + \frac{\pi A}{4}\right) \times (H^2_{mn} + I^2_{mn})^{\frac{1}{2}} \sin[c\xi - \tfrac{1}{2}A\ln(2c\xi) - \tfrac{1}{2}n\pi + \sigma_n + \Sigma^m_n], \quad (3.29)$$

where

$$H_{mn} = \sum_{r=0} D_r^{mn}(A, C)\, \frac{i^r(2m + r)!}{r!}\, e^{i\sigma_{m+r}} \times \cos[(n - m - r)\pi/2 + \sigma_{m+r} - \sigma_n],$$

$$I_{mn} = \sum_{r=0} D_r^{mn}(A, C)\, \frac{i^r(2m + r)!}{r!}\, e^{i\sigma_{m+r}} \times \sin[(n - m - r)\pi/2 + \sigma_{m+r} - \sigma_n], \tag{3.30}$$

$$\Sigma_n^m = \tan^{-1}[I_{mn}/H_{mn}].$$

4. SCATTERING AMPLITUDE

In this section the spheroidal expression of the Coulomb scattering amplitude is discussed. The complete Coulomb wavefunction (incident plus scattered wave) is known. This wavefunction is the starting point of the present discussion. If the fixed charge Q is at the origin of the spherical coordinate system, the complete Coulomb wavefunction has the form

$$U_c = v^{-1/2}\,\Gamma(1 + \tfrac{1}{2}Ai)e^{-A\pi/4} \times e^{ikr_0\cos\Theta}\; {}_1F_1(-\tfrac{1}{2}iA; 1; 2ikr_0 \sin^2\tfrac{1}{2}\Theta), \tag{4.1}$$

where v is the velocity of the incident particle, Θ is the angle between the position vector $\mathbf{r}_0$ and the incident momentum vector $\mathbf{k}$. In Eq. (4.1) the incident beam is normalized to unit flux. In the present prolate spheroidal coordinate system, the fixed charge is placed on the z axis with coordinates $\mathbf{r}_a = (0, 0, \frac{1}{2}d)$. The relations between these two coordinate systems are as follows:

$$r_0 = \tfrac{1}{2}d(\xi - \eta), \tag{4.2}$$

$$z_0 = \tfrac{1}{2}d(\xi\eta - 1). \tag{4.3}$$

The Coulomb wavefunction in Eq. (4.1) is regular at $\xi = 1$, and satisfies Eq. (2.4). This function has the following asymptotic forms:

$$U_c \underset{\xi\to\infty}{\longrightarrow} v^{-1/2}\left[\exp i\{c\xi\cos\Theta + \tfrac{1}{2}A\ln[c\xi(1 - \cos\Theta)]\} \times \left(1 + \frac{A^2}{4ic\xi(1 - \cos\Theta)}\right)U(\Theta) + \frac{k}{c\xi}f_c(\Theta)\exp i\left(c\xi - \frac{A}{2}\ln(2c\xi)\right)\right] \tag{4.4}$$

or

$$U_c \underset{\xi\to\infty}{\longrightarrow} v^{-1/2}\frac{1}{2ic\xi}\sum_{n=0}^{\infty}\{\exp i(c\xi - \tfrac{1}{2}A\ln(2c\xi) + 2\sigma_n) - \exp(-i)[c\xi - n\pi - \tfrac{1}{2}A\ln(2c\xi)]\}(2n + 1)P_n(\cos\Theta), \tag{4.5}$$

where $U(\Theta)$ is a step function

$$U(\Theta) = \begin{cases} 1, & \text{for } \Theta > 0, \\ 0, & \text{for } \Theta \le 0, \end{cases} \tag{4.6}$$

and $f_c(\Theta)$ is the Coulomb scattering amplitude

$$f_c(\Theta) = \frac{A}{4k\sin^2\frac{1}{2}\Theta}\exp(-i)[\tfrac{1}{2}A\ln(\sin^2\tfrac{1}{2}\Theta) - \pi - 2\sigma_0] = \frac{1}{2ik}\sum_{n=0}^{\infty}(2n + 1)e^{2i\sigma_n}P_n(\cos\Theta). \tag{4.7}$$

The angle Θ can be expressed as

$$\cos\Theta = \cos\theta_0\cos\theta' + \sin\theta_0\sin\theta'\cos(\phi - \phi'), \tag{4.8}$$

where the incident momentum $\mathbf{k}$ and the position vector $\mathbf{r}_0$ have the spherical coordinates θ', ϕ' and $\theta_0, \phi_0 \equiv \phi$, respectively. In the asymptotic region there is no difference between the spheroidal coordinate η and $\cos\theta_0$:

$$\cos\Theta \underset{\xi\to\infty}{\longrightarrow} \eta\cos\theta' + (1 - \eta^2)^{1/2}\sin\theta'\cos(\phi - \phi'). \tag{4.9}$$

Now we wish to express the Coulomb scattering amplitude $f_c(\Theta)$ in terms of the spheroidal analysis. To start with the complete Coulomb wavefunction U_c in Eq. (4.1) is expanded as

$$U_c = \sum_{m=0}^{\infty}\sum_{n=m}^{\infty} B_{mn}(c, \theta')R_{mn}^{(1)}(A, C; \xi)\,S_{mn}(A, C, \eta)\cos[m(\phi - \phi')]. \tag{4.10}$$

In writing down Eq. (4.10), we use the fact that the Coulomb wavefunction can be expressed in terms of the Lamé products in Eq. (2.5) and is rotational invariant along the z axis. The expansion coefficients have the form

$$B_{mn}(c, \theta')R_{mn}^{(1)}(A, C, \xi) = [2\pi N_{mn}(A, C)]^{-1}(2 - \delta_{0m}) \times \int_{-1}^{1}d\eta\int_0^{2\pi}d\phi\, U_c S_{mn}(A, C, \eta)\cos[m(\phi - \phi')]. \tag{4.11}$$

To find function $B_{mn}(c, \theta')$, we would like to discuss the asymptotic behavior of the above equation on both sides. The asymptotic form on the left-hand side of Eq. (4.11) is easy to find and follows directly from Eqs. (3.26) and (3.28):

$$\frac{2\Gamma(m + 1)}{c\xi k^m\Gamma(m + 1 + \frac{1}{2}Ai)}B_{mn}(c, \theta')\exp\left(-ic + \frac{\pi A}{4}\right) \times \sum_{r=0}^{\infty}D_r^{mn}(A, C)\frac{i^r(2m + r)!}{r!}e^{\sigma_{m+r}} \times \sin[c\xi - \tfrac{1}{2}A\ln(2c\xi) - \tfrac{1}{2}(m + r)\pi + \sigma_{m+r}]. \tag{4.12}$$

The asymptotic form on the right-hand side of Eq. (4.11) deserves some careful consideration. The complete Coulomb wavefunction in Eq. (4.1) can be expanded as

$$U_c = v^{-1/2}e^{-A\pi/4}\sum_{m=0}^{\infty}\sum_{n=m}^{\infty}\frac{\Gamma(n + 1 + \frac{1}{2}iA)}{(2n)!}(2ikr_0)^n \times e^{ikr_0}\,{}_1F_1(n + 1 + \tfrac{1}{2}iA; 2n + 2; -2ikr_0) \times (2 - \delta_{0m})\frac{(n - m)!}{(n + m)!}P_n^m(\cos\theta_0)P_n^m(\cos\theta') \cos[m(\phi' - \phi_0)]. \tag{4.13}$$

In the asymptotic region $\xi \to \infty$, we have

$$r_0 \to \tfrac{1}{2}d\xi, \qquad \cos\theta_0 \to \eta_0. \tag{4.14}$$

By using Eqs. (4.13) and (4.14), the asymptotic form on the right-hand side of Eq. (4.11) can be written as

$$(2\pi N_{mn}(A, C)]^{-1}(2 - \delta_{0m})v^{-1/2}e^{-A\pi/4} \times \sum_{m'=n'}^{\infty}\sum_{n'=0}^{\infty}\frac{\Gamma(n' + 1 + \frac{1}{2}iA)}{(2n')!}(2ic\xi)^{n'}e^{ic\xi}(2 - \delta_{0m'}) \times \frac{(n' - m')!}{(n' + m')!}P_{n'}^{m'}(\cos\theta')\,{}_1F_1(n' + 1 + \tfrac{1}{2}iA; 2n' + 2; -2ic\xi) \times \int_{-1}^{1}d\eta\int_0^{2\pi}d\phi\, S_{mn}(A, C; \eta)P_n^{m'}(\eta)\cos m'(\phi' - \phi) \times \cos m(\phi' - \phi). \tag{4.15}$$

The associated Legendre functions are orthogonal:

$$\int_{-1}^{1} d\eta P_n^m(\eta) P_{n'}^m(\eta) = \delta_{nn'} \frac{2(n+m)!}{(2n+1)(n-m)!}. \quad (4.16)$$

We carry out the integrations in Eq. (4.15) through the help of Eqs. (2.10) and (4.16) and obtain

$$\frac{2(2-\delta_{0m})}{N_{mn}(A,C)} v^{-1/2} e^{-A\pi/4} \sum_{n'=m}^{\infty} \frac{\Gamma(n'+1+iA/2)}{(2n')!} \times (2ic\xi)^{n'} e^{ic\xi} \frac{1}{2n'+1} D_{n'-m}^{mn}(A,C) P_{n'}^m(\cos\theta') \times {}_1F_1(n'+1+\tfrac{1}{2}iA; 2n'+2; -2ic\xi). \quad (4.17)$$

From Eqs. (3.24) and (3.28) we obtain the final asymptotic form on the right-hand side of Eq. (4.11):

$$\frac{2(2-\delta_{0m})}{c\xi N_{mn}(A,C)} v^{-1/2} \sum_{r=0} i^{r+m} e^{i\sigma_{r+m}} D_r^{mn}(A,C) \times P_{r+m}^m(\cos\theta') \sin[c\xi - \tfrac{1}{2}A\ln(2c\xi) - \tfrac{1}{2}(r+m)\pi + \sigma_{m+r}]. \quad (4.18)$$

The function $B_{mn}(c,\theta')$ is then found by matching Eqs. (4.12) and (4.18):

$$B_{mn}(c,\theta') = \frac{(ik)^m \Gamma(m+1+\frac{1}{2}iA)}{\Gamma(m+1)} \exp\left(ic - \frac{A\pi}{4}\right) \frac{(2-\delta_{0m})}{N_{mn}(A,C)} \times v^{-1/2}\left(\sum_{r=0}^{\infty} (-1)^r \frac{(2m+r)!}{r!} D_r^{mn}(A,C)\right)^{-1} \times S_{mn}(A,C;-\cos\theta'). \quad (4.19)$$

The spherical expression of the Coulomb scattering amplitude in Eq. (4.7) is obtained through the comparison of the exponential dependence on the factor $\exp i[c\xi - \frac{1}{2}A\ln(2c\xi)]$ between the asymptotic form in Eq. (4.4) and the spherical asymptotic expansion in Eq. (4.5) of the complete Coulomb wavefunction. By the same token, we obtain the spheroidal expression of the Coulomb scattering amplitude through the comparison between the asymptotic form in Eq. (4.4) and the spheroidal asymptotic expansion, which follows from Eqs. (4.10) and (4.12):

$$f_c(\theta) = v^{1/2}(-i) \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} \frac{\Gamma(m+1)}{k^{m+1}\Gamma(m+1+\frac{1}{2}Ai)} \times \exp\left(-ic + \frac{\pi A}{4}\right)\left(\sum_{r=0}^{\infty} D_r^{mn}(A,C) \frac{(-i)^m(2m+r)!}{r!} \times e^{2i\sigma_{m+r}}\right) B_{mn}(c,\theta') S_{mn}(A,C;\eta) \cos m(\phi-\phi') = \frac{1}{ik} \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} \frac{2-\delta_{0m}}{N_{mn}(A,C)} e^{2i\tilde{\sigma}_{mn}} S_{mn}(A,C;-\cos\theta') \times S_{mn}(A,C;\eta) \cos m(\phi-\phi'), \quad (4.20)$$

where

$$e^{2i\tilde{\sigma}_{mn}} = \sum_{r=0}^{\infty} \frac{(2m+r)!}{r!} D_r^{mn}(A,C) e^{2i\sigma_{m+r}} \Big/ \sum_{r=0}^{\infty} (-)^r \frac{(2m+r)!}{r!} D_r^{mn}(A,C). \quad (4.21)$$

The quantity $\tilde{\sigma}_{mn}$ may be called the spheroidal Coulomb phase shift and is not real. For a real spheroidal potential we always have a real spheroidal phase shift.[1] The complexedness of the spheroidal Coulomb phase shift $\tilde{\sigma}_{mn}$ may be the cause of the dual spherical[11] and spheroidal nature for the Coulomb potential

ACKNOWLEDGMENT

The author wishes to express his appreciation to Professor D. Kaplan for several discussions.

[1] Ming Chiang Li, J. Math. Phys. **12**, 936 (1971); **13**, 1381 (1972); **14**, 1358 (1973).
[2] H. Stier, Z. Physik **16**, 439 (1932); J. B. Fisk, Phys. Rev. **49**, 167 (1936); K. Takayanagi, Progr. Theoret. Phys. (Kyoto) Suppl. **40**, 216 (1967); S. Hara, J. Phys. Soc. Japan **27**, 1009 (1969); M. H. Mittleman and R. E. Von Holdt, Phys. Rev. **140**, A726 (1965).
[3] J. Rainwater, Phys. Rev. **79**, 432 (1950); S. Granger and R. D. Spence, Phys. Rev. **83**, 460 (1951).
[4] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge U. P., London, England, 1962); C. Flammer, *Spheriodal Wave Function* (Stanford U. P., Palo Alto, California, 1957); J. Meixner and F. W. Schafke, *Mathieusche Functionen und Spharoid-function* (Springer-Verlag, Berlin, 1954).
[5] M. Kotani, Proc. Phys.-Math. Soc. **15**, 30 (1955); C. J. Bouwkamp, J. Math. Phys. **26**, 79 (1947); J. Meixner and W. Andrejewski, Ann. Phys. (N. Y.) **1**, 157, (1950); C. Flammer, J. Appl. Phys. **24**, 1218, 1224 (1953).
[6] I. Marx, J. Math. and Phys. **32**, (1954); Mich. Math. Journal **2**, 45 (1953); **2**, 99 (1954); N. Chako, J. Math. and Phys. **36**, 62 (1957).
[7] S. Altshuler, Phys. Rev. **107**, 114 (1957); O. H. Crawford, A. Dalgarno, and P. B. Hays, Mol. Phys. **13**, 181 (1967); C. Bottcher, Mol. Phys. **19**, 193 (1970); O. H. Crawford, Mol. Phys. **20**, 585 (1971); W. R. Garrett, Mol. Phys. **20**, 751 (1971); C. Bottcher, Chem. Phys. Lett. **9**, 57 (1971); F. S. Stephens, R. M. Diamond, and R. de Boer, Phys. Rev. Lett. **27**, 1151 (1971); W. Bertozzi, T. Cooper, N. Ensslin, J. Heisenberg, S. Kowalski, M. Mills, T. Turchinetz, C. Williamson, S. P. Fivojinsky, J. W. Lightbody, Jr., and S. Penner, Phys. Rev. Lett. **28**, 1711 (1972).
[8] O. Burrau, Kgl. Danske Videnskab Selskab, Mat.-Fys. Medd. **7**, 14 (1927).
[9] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968).
[10] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge U. P., London, 1962); L. D. Landau and E. M. Lifshitz *Quantum Mechanics* (Pergamon, New York, 1965); M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Natl. Bur. Stds., Washington, D. C., 1964).
[11] M. Bander and C. Itzykson, Rev. Mod. Phys. **38**, 330 (1966); M. J. Englefield, *Group Theory And The Coulomb Problem* Interscience, New York, 1972).

Spectral properties of phase operators

J. Dombrowski

Department of Mathematics, Wright State University, Dayton, Ohio 45431
(Received 7 August 1973)

This paper studies the spectral properties of phase operators associated with the phase of the harmonic oscillator. It is shown that all such phase operators have an absolutely continuous part and that those of a certain subclass are absolutely continuous.

INTRODUCTION

Let $\mathcal{H}$ be a separable Hilbert space with orthonormal basis $\{\phi_n\}_{n=1}^{\infty}$. For any bounded linear operator A on $\mathcal{H}$, let $\mathrm{Sp}(A)$ denote the spectrum of A. If A is self-adjoint with spectral resolution $A = \int \lambda dE_\lambda$, denote by $\mathcal{H}_a(A)$ the set of elements x in $\mathcal{H}$ for which $\|E_\lambda x\|^2$ is an absolutely continuous function of λ. It can be shown that $\mathcal{H}_a(A)$ is a subspace of $\mathcal{H}$ which reduces A.[1] The restriction of A to $\mathcal{H}_a(A)$ is called the absolutely continuous part of A, and, if $\mathcal{H}_a(A) = \mathcal{H}$, the operator A is said to be absolutely continuous. In particular, an absolutely continuous operator has no point spectrum.

Consider now the following operators on $\mathcal{H}$:

$$C = \tfrac{1}{2}\begin{bmatrix} 0 & a_1 & 0 & 0 & \cdot & \cdot & \cdot \\ a_1 & 0 & a_2 & 0 & \cdot & \cdot & \cdot \\ 0 & a_3 & 0 & a_3 & \cdot & \cdot & \cdot \\ 0 & 0 & a_3 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix},$$

$$S = \tfrac{1}{2i}\begin{bmatrix} 0 & a_1 & 0 & 0 & \cdot & \cdot & \cdot \\ -a_1 & 0 & a_2 & 0 & \cdot & \cdot & \cdot \\ 0 & -a_2 & 0 & a_3 & \cdot & \cdot & \cdot \\ 0 & 0 & -a_3 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

where $\{a_n\}_{n=1}^{\infty}$ is a sequence of positive real numbers converging monotonically to 1, so chosen that $\mathrm{Sp}(C) = \mathrm{Sp}(S) = [-1, 1]$. Such operators, referred to as phase operators, have been studied in conjunction with the phase of the harmonic oscillator.[2–4]

Ifantis[4] conjectured that the spectrum of every phase operator is purely continuous or, equivalently, that the point spectrum is empty. The eigenvalue problem was studied directly by Eswaran.[5] It will be shown below that a somewhat stronger result holds for those phase operators for which the corresponding sequence $\{a_n\}_{n=1}^{\infty}$ increases monotonically to 1. Such operators are, in fact, absolutely continuous. Also, it will be shown that every phase operator for which the corresponding sequence $\{a_n\}_{n=1}^{\infty}$ decreases monotonically to 1 has an absolutely continuous part whose spectrum is the interval $[-1, 1]$.

THE DEFINITION OF A PHASE OPERATOR

If V denotes the unilateral shift operator on $\mathcal{H}$, so that $V\phi_n = \phi_{n+1}$, and if A is defined by $A\phi_n = a_{n-1}\phi_n$ with $a_0 = 0$, then

$$C = \tfrac{1}{2}\ (V^*A + AV),$$
$$S = (1/2i)\,(V^*A - AV).$$

Thus C and S are the real and imaginary parts, respectively, of the operator $T = V^*A$. Since C and S are unitarily equivalent,[4] it suffices to consider the spectral properties of C.

Infantis[4] has observed that

$$C = \tfrac{1}{2}\,(V + V^*) + \tfrac{1}{2}\,[(A - I)\,V + V^*\,(A - I)],$$

where the operator $\frac{1}{2}[(A - I)V + V^*(A - I)]$ is compact, self-adjoint and the operator $\frac{1}{2}(V + V^*)$ has a purely continuous spectrum which contains the interval $[-1, 1]$. Hence it follows from Weyl's theorem that $\mathrm{Sp}(C) \supset [-1, 1]$.

The final restriction on the sequence $\{a_n\}_{n=1}^{\infty}$ needed to guarantee that $\mathrm{Sp}(C) \subset [-1, 1]$, and hence that $\mathrm{Sp}(C) = [-1, 1]$, is given by the following proposition due to Lerner, Huang and Walters[3,6]:

Proposition: $\mathrm{Sp}(C) \subset [-1, 1]$ if and only if $\{\frac{1}{4}a_n^2\}_{n=1}^{\infty}$ is a "chain sequence." That is

$$\tfrac{1}{4}a_n^2 = (1 - g_{n-1})g_n, \quad \text{where } 0 \le g_0 < 1, \quad 0 < g_n < 1 \quad (n > 0). \tag{1}$$

Hence a phase operator C can be precisely defined as an operator of the form

$$C = \tfrac{1}{2}\begin{bmatrix} 0 & a_1 & 0 & 0 & \cdot & \cdot & \cdot \\ a_1 & 0 & a_2 & 0 & \cdot & \cdot & \cdot \\ 0 & a_2 & 0 & a_3 & \cdot & \cdot & \cdot \\ 0 & 0 & a_3 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix},$$

where $\{a_n\}_{n=1}^{\infty}$ is a sequence of positive real numbers monotonically converging to 1 and satisfying the chain sequence condition (1).

THE SPECTRUM OF A PHASE OPERATOR

It can be shown that any sequence of positive numbers dominated term by term by a chain sequence is itself a chain sequence.[3] Since the sequence $\{a_n = 1\}_{n=1}^{\infty}$ satisfies the chain sequence condition for $g_n = \frac{1}{2}$ $(n = 0, 1, 2, \cdots)$, it follows that every sequence of positive real numbers increasing monotonically to 1 gives rise to a phase operator.

The following proposition follows from the results of Putnam.[7]

Proposition: For any sequence $\{a_n\}_{n=1}^{\infty}$ of positive real numbers increasing monotonically to 1, the corresponding phase operator C is absolutely continuous.

Proof: Consider the operator $T = C + iS$. It is easily verified that $T^*T - TT^* = K$ is a diagonal matrix with diagonal elements

$$\vdash a_1^2, a_1^2 - a_2^2, a_2^2 - a_3^2, \cdots .$$

Then $CS - SC = -(i/2)K$, where $K \leq 0$. It follows that $\mathcal{H}_a(C)$ contains the smallest subspace of $\mathcal{H}$ reducing C and S and containing the range of K. Since $K\phi_1 = -a_1^2\phi_1$, $\phi_1 \in \mathcal{H}_a(C)$. Also, $C\phi_1 = \frac{1}{2}a_1\phi_2$ implies that $\phi_2 \in \mathcal{H}_a(C)$. Now suppose it has been shown that $\phi_1, \phi_2, \cdots, \phi_n \in \mathcal{H}_a(C)$ for $n \geq 2$. Since

$$C\phi_n = \tfrac{1}{2}(a_{n-1}\phi_{n-1} + a_n\phi_{n+1}),$$

it follows that $\phi_{n+1} \in \mathcal{H}_a(C)$. Therefore, $\mathcal{H}_a(C) = \mathcal{H}$, as was to be shown.

Consider now a phase operator C corresponding to a sequence $\{a_n\}_{n=1}^{\infty}$ of positive real numbers decreasing monotonically to 1. The operator K, defined by $CS - SC = -(i/2)K$, no longer satisfies $K \geq 0$ or $K \leq 0$ and the previous proof fails. A different approach, however, yields a somewhat weaker result. The following lemma is needed.

Lemma: Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of positive real numbers decreasing monotonically to 1 and satisfying the chain sequence condition. That is,

$$\tfrac{1}{4}a_n^2 = (1 - g_{n-1})g_n, \quad \text{where } 0 \leq g_0 < 1, 0 < g_n < 1 \ (n > 0).$$

Then $\sum_{n=1}^{\infty}(a_n - 1) < \infty$.

Proof: It is convenient to make the following observations.

(1) $g_n < 1$ implies $g_{n-m} < (m+2)/(2m+2)$ for $n \geq 1$, $0 \leq m \leq n-1$.

Proof: [By induction on m]. Let $m = 1$. Since

$$g_n = \tfrac{1}{4}a_n^2/(1 - g_{n-1}) < 1,$$

it follows that $1 \leq a_n^2 < 4(1 - g_{n-1})$ and hence that $g_{n-1} < \frac{3}{4}$. Now suppose $g_{n-m} < (m+2)/(2m+2)$ $(1 \leq m \leq n-1)$. Since

$$g_{n-m} = \tfrac{1}{4}a_{n-m}^2/(1 - g_{n-m-1}) < (m+2)/(2m+2),$$

it follows that $1 \leq a_{n-m}^2 < 4[(m+2)/(2m+2)](1 - g_{n-m-1})$ and hence that $g_{n-m-1} < (m+3)/(2m+4)$,

(2) The sequence $\{(m+2)/(2m+2)\}$ monotonically decreases to $\frac{1}{2}$.

(3) $g_n \leq \frac{1}{2}$, $n = 0, 1, 2, \cdots$.

Proof: Suppose $g_N > \frac{1}{2}$. Then there exists an m^* such that

$$g_N > (m^* + 2)/(2m^* + 2) > \tfrac{1}{2}.$$

But this contradicts (1) since $g_{N+m^*} < 1$.

(4) $g_n \geq 1/4(1 - g_{n-1})$, $n = 1, 2, 3, \cdots$.

Proof: $g_n = a_n^2/4(1 - g_{n-1}) \geq 1/4(1 - g_{n-1})$.

(5) $\{g_n\}_{n=0}^{\infty}$ is a monotonically increasing sequence.

Proof: For $n \geq 1, g_n - g_{n-1} \geq 1/4(1 - g_{n-1}) - g_{n-1}$. Therefore $g_n - g_{n-1} \geq (1 - 2g_{n-1})^2/4(1 - g_{n-1}) \geq 0$.

(6) $\lim_{n\to\infty} g_n = \frac{1}{2}$.

Proof: (3) and (5) imply that $g = \lim_{n\to\infty} g_n$ exists. It then follows from the fact that

$$\tfrac{1}{4}a_n^2 = (1 - g_{n-1})g_n, \quad n = 1, 2, 3, \cdots$$

that $g = \frac{1}{2}$.

These facts can now be used to show that

$$\sum_{n=1}^{\infty}(a_n^2 - 1) < \infty.$$

In particular,

$$\begin{aligned}\sum_{n=1}^{\infty}(\tfrac{1}{4}a_n^2 - \tfrac{1}{4}) &= \sum_{n=1}^{\infty}[(1 - g_{n-1})g_n - \tfrac{1}{4}] \\ &= \sum_{n=1}^{\infty}[(\tfrac{1}{2}g_n - \tfrac{1}{4}) + (\tfrac{1}{2} - g_{n-1})g_n] \\ &\leq \sum_{n=1}^{\infty}[\tfrac{1}{2}(g_n - \tfrac{1}{2}) + \tfrac{1}{2}(\tfrac{1}{2} - g_{n-1})] \\ &\leq \sum_{n=1}^{\infty}\tfrac{1}{2}(g_n - g_{n-1}) \\ &\leq \tfrac{1}{4}.\end{aligned}$$

Finally, since $a_n \geq 1$,

$$\sum_{n=1}^{\infty}(a_n - 1) \leq \sum_{n=1}^{\infty}(a_n^2 - 1) < \infty.$$

Proposition: For any sequence $\{a_n\}_{n=1}^{\infty}$ of positive real numbers decreasing monotonically to 1 and satisfying the chain sequence condition (1), the corresponding phase operator C has an absolutely continuous part whose spectrum is the interval $[-1, 1]$.

Proof: As noted above $C = \frac{1}{2}(V + V^*) + \frac{1}{2}[(A - I)V + V^*(A - I)]$, where V is the unilateral shift operator on $\mathcal{H}$ and A is defined by $A\phi_n = a_{n-1}\phi_n$ with $a_0 = 0$. By the previous lemma, A, and hence also $\frac{1}{2}[(A - I)V + V^*(A - I)]$, is of trace class. It then follows by a theorem of Kato[8] that C has an absolutely continuous part unitarily equivalent to $\frac{1}{2}(V + V^*)$.

GENERAL REMARKS

The previous proposition suggests the possibility that all phase operators are absolutely continuous. No proof, however, has yet been achieved.

[1]P. R. Halmos, *Introduction to Hilbert Space and the Theory of Spectral Multiplicity* (Chelsea, New York, 1957), p. 104.
[2]E. C. Lerner, Nuovo Cimento B **56**, 183 (1968).
[3]E. C. Lerner, H. W. Huang, and G. E. Walters, J. Math. Phys. 11, 1679 (1970).
[4]E. K. Ifantis, J. Math. Phys. 12, 1021 (1971).
[5]K. Eswaran, Nuovo Cimento B **70**, 1 (1970).
[6]See also H. Wall, *Analytic Theory of Continued Fractions* (Van Nostrand, New York, 1948), p. 79.
[7]C. R. Putnam, *Commutation Properties of Hilbert Space Operators and Related Topics* (Springer, New York 1967), p. 146.
[8]T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1957), p. 540.

Brownian motion in assemblies of coupled harmonic oscillators*

Sungwoon Kim†

Theoretical Physics Institute, Faculty of Mathematics and Sciences, The University of Nijmegen, Nijmegen, The Netherlands

(Received 24 October 1973)

The generalized Langevin equation is derived for a particle of arbitrary mass in an assembly of harmonic oscillators with general interaction matrix and with an external force acting on the particle. The reduction of the equation to the ordinary Langevin equation is studied in various limits. The reduction to the Langevin equation is achieved apart from the distribution of the bath particles, and the results thus obtained are valid whether the bath is in equilibrium or not. It is found that if the Ford–Kac–Mazur interaction is assumed, the particle achieves Brownian motion regardless of its mass ratio to the bath particle. The weak coupling limit is effected by scaling the equation with the mass ratio. In the weak coupling limit, the explicit formula for the friction coefficient is obtained assuming a general interaction matrix. It is demonstrated that the generalized Langevin equation is a very convenient starting point for the study of Brownian motion.

1. INTRODUCTION

For the study of time dependent phenomena, the assembly of coupled harmonic oscillators is one of the very few models which are mathematically tractable and yet exhibit salient features. The model has been studied extensively in connection with Brownian motions.

Rubin,[1] Hemmer,[2] Turner,[3] and Ullersma[4] have studied the motion of a heavy particle with nearest neighbor interactions and found that a sufficiently heavy particle achieves a free Brownian motion. On the other hand, Toda[5] and Takeno and Hori,[6] pointing out that the nature of the approximation procedure in the calculations is not clear, have concluded that a heavy particle behaves as a Brownian particle if the force constant becomes large as well as the mass of the particle.

Ford, Kac, and Mazur[7] have studied the motion of a particle of mass equal to the bath particle and found that the particle attains Brownian motion when the interaction is a very special long range type with a limiting cutoff frequency. Mazur and Braun[8] have studied the motion of a heavy particle with arbitrary interactions. They have found that a heavy particle performs Brownian motion with arbitrary interactions provided that the spectral density of eigenvalues, ω^2, of the interaction matrix is proportional to ω^{-1} for $\omega \to 0$.

Most of these works did not study the equation of motion itself. However, we find that the direct appeal to the exact equation of motion—the generalized Langevin equation—is very profitable. By studying the equation in various limits we not only rediscover the results obtained previously in a unified manner, but also generalize them further. This is all possible only because, as first shown by Deutch and Silbey,[9] the generalized Langevin equation is exact for an assembly of coupled harmonic oscillators. This approach is very similar to the procedure adopted in the general microscopic theory of Brownian motions[10] and thus sheds more light on it.

In Sec. 2, we define the model and explain some of its properties. In Sec. 3, we derive the exact generalized Langevin equation for a particle of arbitrary mass subject to an external force with arbitrary interaction matrix. We obtain the explicit expressions for the Laplace transform of the random force and the kernel in the generalized Langevin equation.

We then study the reduction of the generalized Langevin equation to the ordinary Langevin equation in various limits. When the Langevin equation is reduced directly from the generalized Langevin equation, the special interaction form of Ford-Kac-Mazur is recovered. It is found that if their special interaction form is assumed, the particle achieves Brownian motion regardless of its mass. This surprising result is due to the very special nature of the interaction. This is presented in Sec. 4.

In Sec. 5, we scale the generalized Langevin equation with the mass ratio λ^2 and study the equation in the weak coupling limit. The connection between the weak coupling limit and others is pointed out. We also derive, in the weak coupling limit, the explicit expression for the friction coefficient for general interactions.

In Sec. 6, we find that the random force has proper stochastic properties for a bath in equilibrium. The direct calculation of the force autocorrelation function is sketched. In the last section, 7, we briefly discuss the Toda-Takeno-Hori limit. The advantage of the generalized Langevin equation for the study of Brownian motion problems is emphasized.

2. THE MODEL

We consider a system which consists of a particle of mass M and $2N$ bath particles of mass m interacting with each other with harmonic forces. The Hamiltonian of the system is given by

$$H = p_0^2/2M + H_b \tag{2.1}$$

and

$$H_b = \tfrac{1}{2}\sum_{j\neq 0}(p_j^2/m) + \tfrac{1}{2}\sum_{j,k} q_j A_{jk} q_k.^{11} \tag{2.2}$$

Here, $p_0(p_j)$ is the momentum of the zeroth (jth) particle of mass $M(m)$, $q_0(q_j)$ is the displacement of the zeroth (jth) particle from its equilibrium position, and A_{jk} is the jk-element of the interaction matrix A characterizing the interactions of the particles.

The interaction matrix A is assumed to be completely general except that it has no negative eigenvalues. We assume that all particles are subject to the same interaction, and also impose periodic boundary conditions. This implies that A is a cyclic and symmetric matrix. The matrix element A_{jk} depends only on $|j-k|$,

$$A_{j,k} = A_{|j-k|} \equiv A_l. \tag{2.3}$$

Under these conditions, the matrix element, A_l can be expressed as[12]

$$A_l = \frac{1}{2N+1} \sum_n \omega_n^2 \exp\left(\frac{2\pi i}{2N+1} nl\right) \tag{2.4}$$

and

$$\omega_n^2 = \sum_l A_l \exp\left(-\frac{2\pi i}{2N+1} nl\right). \tag{2.5}$$

The quantity ω_n^2 is the eigenvalue of the interaction matrix A corresponding to the eigenvector $\xi^{(n)}$;

$$A\xi^{(n)} = \omega_n^2 \xi^{(n)}. \tag{2.6}$$

The kth element of $\xi^{(n)}$ is given by

$$\xi_k^{(n)} = \frac{1}{(2N+1)^{1/2}} \exp\left(\frac{2\pi i}{2N+1} nk\right). \tag{2.7}$$

The conservation of the total momentum of the system further requires that

$$\sum_j A_{j,k} = 0 \quad \text{for all } k \tag{2.8}$$

which is a very important identity for our discussion. From Eqs. (2.4) and (2.8) one notices that

$$\omega_0^2 = 0. \tag{2.9}$$

Eventually, we have to deal with infinite systems. In the limit $N \to \infty$, the summation in Eq. (2.4) is approximated by an integral

$$A_l = \frac{1}{2\pi} \int_{-\pi}^{\pi} \omega^2(\theta) e^{il\theta} d\theta \tag{2.10}$$

where

$$\omega^2(\theta) = \omega_n^2, \quad \left(n = \frac{2N+1}{2\pi}\theta\right). \tag{2.11}$$

We assume that ω_n^2 changes slowly enough to warrant the approximation. In passing we note that symmetry of A yields symmetry of ω_n^2;

$$\omega_n^2 = \omega_{-n}^2 \text{ and } \omega^2(\theta) = \omega^2(-\theta). \tag{2.12}$$

3. THE GENERALIZED LANGEVIN EQUATION

To be general, we assume that an external force,

$$F_{\text{ext}}(q_0(t)) = -\frac{\partial V(q_0(t))}{\partial q_0} \tag{3.1}$$

acts on the zeroth particle in addition to the harmonic interactions of the particles. The equation of motion of the jth particle is given by

$$[m + (M-m)\delta_{j,0}]\ddot{q}_j(t) = -\sum_k A_{jk} q_k(t) + F_{\text{ext}}(q_0(t)) \cdot \delta_{j,0}. \tag{3.2}$$

The equation may be solved by introducing a generating function $G_n(t)$:

$$G_n(t) \equiv \frac{1}{(2N+1)^{1/2}} \sum_j \exp\left(\frac{2\pi i}{2N+1} jn\right) q_j(t) \tag{3.3}$$

and

$$q_j(t) = \frac{1}{(2N+1)^{1/2}} \sum_n \exp\left(-\frac{2\pi i}{2N+1} jn\right) G_n(t). \tag{3.4}$$

Differentiating $G_n(t)$ twice and using Eqs. (3.2) and (2.5), we obtain

$$\ddot{G}_n(t) = -\frac{\omega_n^2}{m} G_n(t) + \frac{1}{(2N+1)^{1/2}} \left(1 - \frac{M}{m}\right)\ddot{q}_0(t) + \frac{1}{(2N+1)^{1/2}} \frac{1}{m} F_{\text{ext}}(q_0(t)). \tag{3.5}$$

We take the Laplace transform of the equation to obtain

$$\tilde{G}_n(\epsilon) = \frac{1}{\epsilon^2 + \omega_n^2/m} \left([\epsilon G_n(0) + \dot{G}_n(0)] - \frac{(M/m-1)}{(2N+1)^{1/2}} [\epsilon^2 \tilde{q}_0(\epsilon) - \epsilon q_0(0) - \dot{q}_0(0)] + \frac{1}{(2N+1)^{1/2}} \frac{1}{m} \tilde{F}_{\text{ext}} \right). \tag{3.6}$$

Summing over n and using Eq. (3.4), we have

$$\tilde{q}_j(\epsilon) + (M/m-1)\epsilon^2 \tilde{f}_j(\epsilon)\tilde{q}_0(\epsilon) = \sum_k [\epsilon q_k(0) + \dot{q}_k(0)]\tilde{f}_{k-j}(\epsilon) + \{(M/m-1)[\epsilon q_0(0) + \dot{q}_0(0)] + (1/m)\tilde{F}_{\text{ext}}\}\tilde{f}_j(\epsilon), \tag{3.7}$$

where

$$\tilde{f}_j(\epsilon) \equiv \frac{1}{2N+1} \sum_n \frac{1}{\epsilon^2 + \omega_n^2/m} \exp\left(\frac{2\pi i}{2N+1} nj\right). \tag{3.8}$$

In particular, setting $j = 0$, we obtain

$$\tilde{q}_0(\epsilon) = \frac{1}{1 + (M/m-1)\epsilon^2 \tilde{f}_0(\epsilon)} \times \left(\sum_{k \neq 0} [\epsilon q_k(0) + \dot{q}_k(0)]\tilde{f}_k(\epsilon) + \frac{M}{m} [\epsilon q_0(0) + \dot{q}_0(0)]\tilde{f}_0(\epsilon) + \frac{1}{m} \tilde{F}_{\text{ext}} \tilde{f}_0(\epsilon) \right). \tag{3.9}$$

Equations (3.7) and (3.9) are the solutions for $q_j(t)$. The $p_j(t)$ may be obtained from

$$\tilde{p}_j(\epsilon) = m[\epsilon \tilde{q}_j(\epsilon) - q_j(0)]. \tag{3.10}$$

We write Eq. (3.9) in the form of the generalized Langevin equation,

$$\dot{p}_0(t) = F_{\text{ext}}(t) + F(t) - \lambda^2 \int_0^t d\tau\, K(\tau) p_0(t-\tau) \tag{3.11}$$

by identifying the terms in Eq. (3.9) with the Laplace transform of Eq. (3.11) which is

$$\tilde{q}_0(\epsilon) = \frac{1}{\epsilon} q_0(0) + \frac{1}{\epsilon^2 + \lambda^2 \epsilon \tilde{K}(\epsilon)} \left(\dot{q}_0(0) + \frac{1}{M} \tilde{F}_{\text{ext}} + \frac{1}{M} \tilde{F} \right). \tag{3.12}$$

Here,

$$\lambda^2 = m/M. \tag{3.13}$$

The results are

$$\tilde{K}(\epsilon) = [1/\epsilon f_0(\epsilon) - \epsilon] \tag{3.14}$$

and

$$\tilde{F}(\epsilon) = -mq_0(0)\tilde{K}(\epsilon) + \frac{m}{\tilde{f}_0(\epsilon)} \sum_{j \neq 0} [\epsilon q_j(0) + \dot{q}_j(0)]\tilde{f}_j(\epsilon). \tag{3.15}$$

Using Eq. (2.9) and the fact that

$$\sum_j \tilde{f}_j(\epsilon) = 1/\epsilon^2, \tag{3.16}$$

we can rewrite Eq. (3.15) in more convenient form,

$$\tilde{F}(\epsilon) = \frac{1}{\tilde{f}_0(\epsilon)} \sum_{j \neq 0} \{m\epsilon[q_j(0) - q_0(0)] + p_j(0)\}\tilde{f}_j(\epsilon). \tag{3.17}$$

For later convenience, we also define g_j by

$$\tilde{g}_j(\epsilon) \equiv \epsilon \tilde{f}_j(\epsilon). \tag{3.18}$$

Later on we will also need the expression for $\tilde{f}_j(\epsilon)$ in the limit $N \to \infty$;

$$\tilde{f}_j(\epsilon) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i\theta j}}{\epsilon^2 + \omega^2(\theta)/m} d\theta. \tag{3.19}$$

In the following sections the presence of F_{ext} is irrelevant to our discussion and we will neglect the term.

A few remarks are due on the generalized Langevin equation, Eq. (3.11). First of all, we stress the fact that the equation is exact and valid for all mass ratio λ^2 and for all time. This was first pointed out by Deutch and Silbey.[10] Its dependence on the mass of the zeroth particle, M is very simple; M enters the equation only through λ^2. This remarkably simple dependence on M may not be expected for other systems in general. These simple properties and the close resemblance of the equation to the ordinary Langevin equation greatly facilitate the study of the equation for Brownian motion. Furthermore, we note that all initial conditions appear only in $F(t)$.

The $K(t)$ depends only on the mass of the bath particle and the properties of the interaction matrix. Therefore the reduction of the equation to the Langevin equation can be discussed separately from the distribution of the bath particles. The results thus obtained may be valid whether or not the bath particles are in equilibrium. However, one must realize that, because the equation is exact for all time, it reduces to the Langevin equation only in appropriate limits, despite its close resemblance to the Langevin equation.

4. FORD, KAC, AND MAZUR LIMIT

Let us now try to reduce the Langevin equation directly from the generalized Langevin equation, Eq. (3.11), by requiring that

$$K(t) = \gamma\delta(t) \tag{4.1}$$

or, equivalently,

$$\tilde{K}(\epsilon) = \gamma. \tag{4.2}$$

Using Eqs. (3.8), (3.14), and (3.18), we obtain

$$g_0(\epsilon) = 1/(\gamma + \epsilon). \tag{4.3}$$

In view of Eqs. (3.8) and (3.18), it is obvious that the requirement in Eq. (4.3) cannot be met with a finite number of particles. Therefore, we consider the limit of infinitely many particles.

In the limit $N \to \infty$, the requirement Eq. (4.3) takes the following form:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\epsilon}{\epsilon^2 + \omega^2(\theta)/m} d\theta = \frac{1}{\gamma + \epsilon} \tag{4.4}$$

or, expressed in terms of t,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \cos\left(\frac{t\omega^2(\theta)}{m}\right) d\theta = e^{-\gamma t}. \tag{4.5}$$

This is the integral equation for $\omega^2(\theta)$ which was first found by Ford, Kac, and Mazur[7] in the study of Brownian motion of a particle of mass $M = m$. They found the solution of the equation to be

$$\omega^2(\theta) = \gamma^2 \tan^2\theta/2. \tag{4.6}$$

As they pointed out, when this expression is inserted into Eq. (2.10), the matrix element A_l diverges. To avoid this difficulty, a cutoff value θ_L is introduced, such that

$$\omega_L^2(\theta) = \begin{cases} \gamma^2 \tan^2\theta/2 & \text{for } |\theta| < \theta_L \\ 0 & \text{for } \theta_L \leq |\theta| \leq \pi. \end{cases} \tag{4.7}$$

By taking the limit $\theta_L \to \pi$, it has been argued that a particle (of mass $M = m$) can achieve Brownian motion with the interaction given by Eq. (4.7).

For more detailed discussions, we refer to the original article of Ford, Kac, and Mazur.[7] However, one must note here that if the special interaction form, Eq. (4.7), with the limiting cutoff frequency is realized, the particle achieves Brownian motion regardless of the mass ratio λ^2 (including $M < m$). This somewhat surprising result is due to the special nature of the interaction. It does not contradict the finding of Cukier and Mazur,[13] which states that the kinetic energy of a particle of mass $M < m$ is a nonergodic function. Their study is limited to nearest-neighbor interaction and does not apply to the special interaction, Eq. (4.7).

5. SCALING AND THE WEAK-COUPLING LIMIT

Instead of the straight reduction of the Langevin equation discussed in the previous section, we first scale the generalized Langevin equation and then reduce it to the Langevin equation in appropriate limits. Although other choices are possible, we choose λ^2 to be the scaling parameter and set

$$s = \lambda^2 t \quad \text{and} \quad \sigma = \lambda^2 \tau \tag{5.1}$$

in Eq. (3.11). The generalized Langevin equation now takes the following form:

$$\frac{d\pi(s)}{ds} = E(s) - \int_0^s C(\sigma)\pi(s - \sigma)d\sigma, \tag{5.2}$$

where

$$\pi(s) = p_0(s/\lambda^2), \tag{5.3}$$

$$E(s) = F(s/\lambda^2)/\lambda^2, \tag{5.4}$$

$$C(s) = K(s/\lambda^2)/\lambda^2. \tag{5.5}$$

We now reduce Eq. (5.2) to the Langevin equation by requiring

$$C(s) = \gamma\delta(s) \tag{5.6}$$

or, equivalently,

$$\tilde{K}(\lambda^2\epsilon) = \gamma \tag{5.7}$$

in the limit $\lambda^2 \to 0$. When the condition of Eq. (5.6) is met, the solution of Eq. (5.2) is given by

$$\pi(s) = \pi(0)e^{-\gamma s} \tag{5.8}$$

or

$$p_0(t) = p_0(0)e^{-\lambda^2\gamma t}. \tag{5.9}$$

Therefore, to have a meaningful solution, we must require that

$$\lambda^2\gamma t = \text{finite} \tag{5.10}$$

in any limiting process. Thus, we are led to the weak coupling limit:

$$\lambda^2\gamma \to 0, \quad t \to \infty \tag{5.11}$$

and $\lambda^2\gamma t = \text{finite}$.

In this limit, Eq. (5.7) reduces to

$$\lim_{\xi\to 0}\tilde{K}(\xi) = \gamma. \tag{5.12}$$

It is clearly seen here that the limiting process taken by many workers[1-4] exactly corresponds to the weak coupling limit. The weak coupling limit plays an important role in the general microscopic theory of Brownian motions.[10] Mazur and Braun[8] have pointed out the connection between their limiting process and the weak coupling limit.

Thanks to the simplicity of Eq. (5.12), one can obtain an explicit expression for the friction coefficient, γ in the weak coupling limit. With Eqs. (3.8), (3.14), and (3.18), the requirement, Eq. (5.12), may be written as

$$\lim_{\xi\to 0}\frac{1}{\pi}\int_0^{\pi}\frac{\xi}{\xi^2+\omega^2(\theta)/m}\,d\theta = \frac{1}{\gamma}. \tag{5.13}$$

Here we have used the symmetry property of $\omega^2(\theta)$. It is important to remember that by assumption $\omega^2(\theta)$ cannot be negative. One can easily see that only the neighborhoods of the zeros of $\omega^2(\theta)$ contribute to the integral. We recall here that $\omega^2(\theta)$ has at least one zero, $\theta = 0$, from Eq. (2.9). The integral depends strongly on the behavior of $\omega^2(\theta)$ at its zeros, and this enables us to evaluate the integral explicitly. Let us suppose that $\omega^2(\theta)$ behaves like

$$\omega^2(\theta) \approx \alpha_i|\theta-\theta_i|^{\beta_i} \tag{5.14}$$

as the curve $\omega^2(\theta)$ emerges[14] from one of its zeros, θ_i. The contribution of this branch to the integral is

$$I_{\beta_i} = \lim_{\xi\to 0}\frac{1}{\pi}\int_0^{\delta}\frac{\xi}{\xi^2+(\alpha_i/m)x^{\beta i}}\,dx. \tag{5.15}$$

One can show that

$$I_{\beta_i} = \begin{cases} 0 & \text{if } 0 \le \beta_i < 2 \\ \frac{1}{2}(m/\alpha_i)^{1/2} & \text{if } \beta_i = 2 \\ \infty & \text{if } 2 < \beta_i. \end{cases} \tag{5.16}$$

Therefore, we have

$$\frac{1}{\gamma} = \sum_{[[\omega^2(\theta_i)=0]]}\Gamma_i, \tag{5.17}$$

where

$$\Gamma_i = \begin{cases} 0 & \text{if } \omega^2(\theta)\approx\alpha_i|\theta-\theta_i|^{\beta_i}, \quad 0\le\beta_i<2 \text{ at } \theta_i \\ \frac{1}{2}\sqrt{m/\alpha_i} & \text{if } \omega^2(\theta)\approx\alpha_i|\theta-\theta_i|^2 \quad \text{at } \theta_i \\ \infty & \text{if } \omega^2(\theta)\approx\alpha_i|\theta-\theta_i|^{\beta_i}, \quad 2<\beta_i \quad \text{at } \theta_i. \end{cases} \tag{5.18}$$

The summation is performed for all zeros, θ_i, and the ingoing and outgoing branches at the zero points must be counted separately.[14]

If $\omega^2(\theta)$ has at least one zero at which the function $\omega^2(\theta)$ changes more slowly than $|\theta-\theta_i|^2$ (that is, $2<\beta_i$), the friction coefficient γ is zero. On the other hand, a zero at which $\omega^2(\theta)$ changes faster than $|\theta-\theta_i|^2$ (that is, $0<\beta_i<2$) does not contribute to the integral. If $\omega^2(\theta)$ changes faster than $|\theta-\theta_i|^2$ at all its zeros, the friction coefficient is infinite.

If $\omega^2(\theta)$ is an analytic function, we have the simpler expression for γ:

$$\frac{1}{\gamma} = \sum_{[\omega^2(\theta_i)=0]}\frac{1}{1+\delta_{0,\theta_i}+\delta_{\pi,\theta_i}}\left(\frac{2m}{[\omega^2(\theta)]''}\right)^{1/2}. \tag{5.19}$$

In this expression, each zero has been counted twice already for the ingoing and outgoing branches and the summation should be taken only once over all zeros of $\omega^2(\theta)$.

6. STATISTICAL PROPERTIES OF $F(t)$

As stated earlier, all initial conditions are included in $F(t)$ and the discussions in the previous sections on the properties of $K(t)$ are independent of the conditions of the bath. The results are valid whether the bath particles are in equilibrium or not. To complete the discussion of Brownian motion, it remains to show that the random force, $E(s)$ has the proper stochastic properties for a bath in equilibrium. The properties of $E(s)$ are determined from $F(t)$ by Eq. (5.4).

Let us assume that the initial values of q_j and p_j of the bath particles are distributed according to the canonical distribution in the potential field created by the fixed zeroth particle;

$$D(q_j,p_j) = e^{-\beta H_b}/\int e^{-\beta H_b}d\mathbf{p}^{2N}d\mathbf{q}^{2N}, \tag{6.1}$$

where

$$d\mathbf{p}^{2N} = dp_1\cdots dp_N dp_{-1}\cdots dp_{-N}$$

and

$$d\mathbf{q}^{2N} = dq_1\cdots dq_N dq_{-1}\cdots dq_{-N}.$$

The average value of x may be defined by

$$\langle x\rangle = \int e^{-\beta H_b}x\,d\mathbf{p}^{2N}d\mathbf{q}^{2N}/\int e^{-\beta H_b}d\mathbf{p}^{2N}d\mathbf{q}^{2N}. \tag{6.2}$$

It is convenient to change the variable q_j to y_j by

$$y_j = q_j - q_0 \quad \text{for } j \neq 0. \tag{6.3}$$

With the aid of Eq. (2.8), one can write

$$H_b = \tfrac{1}{2}\sum_{j\neq 0}\frac{p_j^2}{m} + \tfrac{1}{2}\sum_{j,k\neq 0}y_jA_{jk}y_k, \tag{6.4}$$

$$\tilde{F}(\epsilon) = \frac{1}{\tilde{f}_0(\epsilon)}\sum_{l\neq 0}[m\epsilon y_l(0)+p_l(0)]\tilde{f}_l(\epsilon), \tag{6.5}$$

and

$$F(0) = -\sum_{m\neq 0}A_{0,m}y_m(0). \tag{6.6}$$

We note that the distribution is Gaussian and $F(t)$ is a linear combination of $y_j(0)$ and $p_j(0)$. This is sufficient[15] to have the desired stochastic properties of $E(s)$;

$$\langle E(s_1)E(s_2)\cdots E(s_{2n+1})\rangle = 0 \tag{6.7}$$

and

$$\langle E(s_1)E(s_2)\cdots E(s_{2n})\rangle = \sum\prod_{i<j}^{2n}\langle E(s_i)E(s_j)\rangle, \tag{6.8}$$

where the summation is performed over all different ways in which $s_1,\ldots,s_{2n}$ may be divided in n pairs.

One may calculate $\langle F(t)F(0)\rangle$ directly using Eqs. (6.5) and (6.6). With the aid of Eq. (2.8), we obtain

$$\langle\tilde{F}(\epsilon)F(0)\rangle = \frac{m\epsilon}{\tilde{f}_0(\epsilon)}\sum_{l\neq 0}\tilde{f}_l(\epsilon)\sum_{i,m\neq 0}A_{i,m}\langle y_l(0)y_m(0)\rangle. \tag{6.9}$$

The result of the Gaussian integral is well known:

$$\int\exp\left(-\frac{\beta}{2}\sum_{j,k\neq 0}y_jA_{jk}y_k\,dy^{2N}\right) = \frac{\pi^N}{\sqrt{|A'|}}\frac{1}{(\beta/2)^N} \tag{6.10}$$

where $|A'|$ is the $2N\times 2N$ determinant defined by

$$|A'| = \begin{vmatrix} A_{1,1}A_{1,2}\cdots A_{1,N}A_{1,-1}\cdots A_{1,-N} \\ A_{2,1}A_{2,2}\cdots A_{2,N}A_{2,-1}\cdots A_{2,-N} \\ \cdot \quad\quad\quad\quad\quad\quad\quad\quad \cdot \\ \cdot \quad\quad\quad\quad\quad\quad\quad\quad \cdot \\ \cdot \quad\quad\quad\quad\quad\quad\quad\quad \cdot \\ A_{-N,1}A_{-N,2}\cdots A_{-N,N}A_{-N,-1}\cdots A_{-N,-N} \end{vmatrix} . \quad (6.11)$$

Here, we assume that $|A'| > 0$. Thus, we obtain

$$\langle y_l(0) y_m(0) \rangle = (1/\beta)(1/|A'|) c_{l,m} \quad (6.12)$$

where $c_{l,m}$ is the cofactor of $A_{l,m}$ in the determinant $|A'|$.

With the aid of the well-known identity

$$\sum_{m \neq 0} A'_{im} c_{l,m} = |A'| \delta_{i,l} \quad (6.13)$$

and Eq. (3.16), we obtain

$$\langle \tilde{F}(\epsilon) F(0) \rangle = (m/\beta) \tilde{K}(\epsilon). \quad (6.14)$$

Therefore, only in appropriate limits, the correlation function, $\langle F(t)F(0) \rangle$ has short memory and can be represented by a δ function.

7. DISCUSSIONS

We wish to point out that, in addition to the limits we have discussed, there are various other limits in which Brownian motion can be realized. For example, we may take the following limit; $\lambda^2 \to 0$, $\gamma \to \infty$, and $\lambda^2 \gamma =$ finite. This is the limit Toda[5] and Takeno and Hori[6] employed in their study of the system with nearest neighbor interaction.[16] In this limit, it is notable that the particle achieves Brownian motion in a finite time.

We have seen that the generalized Langevin equation provides a natural starting point for the study of Brownian motion. Since its structure is very similar to the Langevin equation, the reduction of the equation to the Langevin equation can be effected in a direct and simple manner. The distribution of the bath particles does not enter directly in the process of reduction. Furthermore, since the generalized Langevin equation is valid for any mass ratio and for all time, the equation can be used for any limiting case one wants to study.

We believe that this approach to the study of Brownian motion based on the generalized Langevin equation will be equally powerful for other similar problems such as more-than-one-defect particles, quantum systems, etc.

Note added on proof: After the submission of the manuscript for publication, Professor P. Mazur kindly alerted the author to another relevant work on the subject: P. Mazur, *Proceedings of the International Symposium on Statistical Mechanics and Thermodynamics, Aachen, June 1964*, edited by J. Meixner (North-Holland, Amsterdam), p. 69.

ACKNOWLEDGMENTS

The author would like to express his appreciation to the members of the Institute, in particular to Dr. A. Weyland, for pleasant and stimulating discussions.

*Supported by Fundamenteel Onderzoek der Materie, Nederland.

†Present address: Department of Mathematics, Korea University, Seoul, Korea.

[1] R. J. Rubin, *Proceedings of the International Symposium on Transport Processes in Statistical Mechanics, August 1956*, edited by I. Prigogine, (International Publishers, New York, 1958), p. 155. J. Math. Phys. **1**, 309 (1960); **2**, 373 (1961).
[2] P. C. Hemmer, "Dynamics and Stochastic Types of Motion in the Linear Chain," Thesis, (Trondheim, 1959).
[3] R. E. Turner, Physica **26**, 269 (1960).
[4] P. Ullersma, Physica **32**, 27, 56, 74 (1966).
[5] M. Toda, Comments at the Third Symposium on the Theory of Lattice Vibrations of Imperfect Crystals, June 1962 (Hokkaido University).
[6] S. Takeno and J. Hori, Suppl. Prog. Theor. Phys. (Kyoto) **23**, 177 (1962).
[7] G. W. Ford, M. Kac and P. Mazur, J. Math. Phys. **6**, 504 (1965).
[8] P. Mazur and E. Braun, Physica **30**, 1973 (1964).
[9] J. M. Deutch and R. Silbey, Phys. Rev. A **3**, 2049 (1971).
[10] P. Mazur and I. Oppenheim, Physica **50**, 241 (1970). S. Kim and I. Oppenheim, Physica **54**, 593 (1971); **57**, 469 (1972).
[11] Throughout the article, the summation indices are understood to run from $-N$ to $+N$ including 0 unless explicitly noted (e.g., $j \neq 0$).
[12] See Ref. 7 and the references quoted therein.
[13] R. I. Cukier and P. Mazur, Physica **53**, 157 (1970).
[14] To be general, we treat separately the ingoing and outgoing branches of $\omega^2(\theta)$ at its zeros. The discussion equally applies to the ingoing branches. For the zero at $\theta = 0$, only the outgoing branch contributes to the integral, while only the ingoing branch contributes to the integral if the zero is located at $\theta = \pi$.
[15] M. C. Wang and G. E. Uhlenbeck, Rev. Mod. Phys. **17**, 323 (1945).
[16] For nearest-neighbor interaction, we have
$\gamma = \sqrt{4k/m}$, $K(t) = (\gamma/t) J_1(\gamma t)$,
$C(s) = (\gamma/\lambda^2)[1/(\gamma s/\lambda^2)] J_1(\gamma s/\lambda^2)$.

A novel approach to the exact calculation of correlation functions of a one-dimensional random Ising chain

A. K. Rajagopal and G. S. Grest*

Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803
(Received 30 August 1973; revised manuscript received 11 October 1973)

An exact result for the partition function for a general one-dimensional Ising chain of N spin-1/2 particles described by the Hamiltonian $H_N = -\Sigma_{i=1}^{N} J_i\sigma_i\sigma_{i+1} - \Sigma_{i=1}^{N} H_i\sigma_i$ is given. For an open strand, $J_N = 0$; for a closed chain, $\sigma_{N+1} = \sigma_1$, $J_N \neq 0$. The novelty of the trick used enables one to obtain the partition function and all the spin correlation functions for open and closed chains with equal ease. Special cases of this model have been discussed before to elucidate certain features of some biological systems. New expressions for parallel and perpendicular susceptibilities for this model are also derived. When J_i and H_i are treated as random variables, the above Hamiltonian describes a one-dimensional Ising spin glass. In this case some simple models and formal averaging procedures are discussed.

I. INTRODUCTION

Various types of regular arrays of scalar spins capable of only "up" or "down" spin orientations but with mutual interactions between nearest neighbors only have been studied extensively for almost half a century.[1] The two-dimensional version of this model has been the only exactly soluble model for phase transitions until recently. Even though the one-dimensional model does not show any phase transition, it is of intrinsic interest as it can also be solved exactly. Variations of this one-dimensional model have recently found enormous applications in elucidating some features of certain long chain molecules of great biological importance.[1] Pioneering work in this area has been done by Goel[2] and Montroll.[3] McCoy[4] has recently written an account of investigations on a two-dimensional Ising net of spins with a special random arrangement. Random systems in general are of significance in discussing amorphous substances, such as glass.

In this paper, we obtain an exact expression for the partition function of N spins arranged on a straight line but obeying the Hamiltonian

$$H_N = -\sum_{i=1}^{N} J_i\sigma_i\sigma_{i+1} - \sum_{i=1}^{N} H_i\sigma_i. \tag{1}$$

Here the set $\{\sigma_i\}$ take on the values + 1 or − 1 and the various σ_i's commute. J_i is the interaction strength between the spin i with its nearest neighbor to the right. H_i can be thought of as a local magnetic field on site i. Clearly for an open strand, $J_N = 0$, whereas for a closed chain, $\sigma_{N+1} = \sigma_1$, $J_N \neq 0$. The Hamiltonian given by (1) for an open strand can also be considered as an Ising model with nearest and next nearest neighbor interactions, for $N + 1$ spins,

$$H'_{N+1} = -\sum_{i=1}^{N-1} J_i\sigma_i\sigma_{i+2} - \sum_{i=1}^{N} H_i\sigma_i\sigma_{i+1}, \tag{2}$$

if we introduce a new "spin"[5]

$$t_0 = \sigma_1, \quad t_i = \sigma_i\sigma_{i+1}, \quad i = 1, 2, \ldots, N \text{ (open chain)}. \tag{3}$$

The partition function is given by

$$Z_N = \sum_{\sigma_1=\pm1} \cdots \sum_{\sigma_N=\pm1} \prod_{i=1}^{N} A_i(\sigma_i, \sigma_{i+1}), \tag{4}$$

where

$$A_i(\sigma_i, \sigma_{i+1}) = \exp(\mathcal{J}_i\sigma_i\sigma_{i+1} + h_i\sigma_i) \tag{5}$$

and

$$\beta J_i = \mathcal{J}_i, \quad \beta H_i = h_i. \tag{5'}$$

A_i is usually known as the "transfer matrix."

Falk[6] calculated Z_N for an *open* regular array where all the J_i are equal (= J) and all the H_i are equal (= H). (We will have occasion to refer to this work later.) Earlier Thompson[1] had discussed completely the special cases when all $H_i = 0$ and unequal J_i, and $J_i = J, H_i = H$ (all i), for a *closed chain*. Goel and Montroll, in their work on the application of one-dimensional Ising model to the study of certain biological molecules, considered the case of *infinite closed chain* with unequal H_i's and all J_i's equal (= U).

In order to focus attention of the reader to the novelty of our method, we will briefly outline in the Introduction itself a transfer matrix method used by all these authors, suitably modified here, for the general Hamiltonian given by Eq. (1). Both open and closed chains can be treated in this way.[1,6]

If $\lambda_j(i)$ $(j = 1, 2)$ are the two eigenvalues of the (2×2) transfer matrix, $A_i(\sigma_i, \sigma_{i+1})$ and $\{\psi_{j_i}(\sigma_i)\}$ are the associated normal eigenvectors, then one obtains the exact result for Z_N in the case of *closed chain* $(\sigma_{N+1} = \sigma_1)$:

$$Z_N^{(\text{closed})} = \sum_{j_1=1}^{2} \cdots \sum_{j_N=1}^{2} \lambda_{j_1} \cdots \lambda_{j_N}(j_1, j_2) \cdots (j_N, j_1), \tag{6}$$

where

$$(j_l, j_m) = \sum_{\sigma=\pm1} \psi_{j_l}(\sigma)\psi_{j_m}(\sigma) \tag{7}$$

is really the "inner product" between two eigenvectors of different transfer matrices representing the "overlap" between the neighboring spins.

Our technique simplifies the expression (6) to a much greater extent. In essence, we write A_i in a "standard" representation[1]

$$A_i(\sigma_i, \sigma_{i+1}) = a_i + r_i\hat{n}_i \cdot \boldsymbol{\tau} \equiv A_i^{\tau}, \tag{8}$$

where $\boldsymbol{\tau}$ is a Pauli spin vector matrix, $\hat{n}_i$ is a unit vector. a_i, r_i are given by

$$\begin{aligned} a_i &= \tfrac{1}{2}\,\mathrm{tr}_{\tau}\, A_i^{\tau} = e^{\mathcal{J}_i}\cosh h_i, \\ a_i^2 - r_i^2 &= \det A_i^{\tau} = 2\sinh 2\mathcal{J}_i. \end{aligned} \tag{9}$$

Of course, a_i and r_i are related to the eigenvalues $\lambda_j(i)$ via

$$a_i = \tfrac{1}{2}[\lambda_1(i) + \lambda_2(i)], \quad r_i = \tfrac{1}{2}[\lambda_1(i) - \lambda_2(i)], \tag{10}$$

if $\lambda_1 > \lambda_2$. The unit vector $\hat{n}_i$ is given by

$$\hat{n}_i = (e^{-\mathcal{J}_i}\cosh h_i/r_i, ie^{-\mathcal{J}_i}\sinh h_i/r_i, e^{\mathcal{J}_i}\sinh h_i/r_i) \tag{11}$$

and has the interpretation that it *specifies the spatial orientation of the ith spin (or ith bond)*. We then show that

$$Z_N = \mathrm{tr}_\tau \left(\prod_{i=1}^{N} (a_i + r_i \hat{n}_i \cdot \boldsymbol{\tau}) \right). \tag{12}$$

The only difference between the open and closed chains is that $J_N = 0$ for the open chain in Eq. (12). With this observation expression (12) can be used for both the open and the closed chains. The trace here is only over the τ matrices. Similar expressions with only trace over τ are obtained for all the spin correlation functions.

The representation (12) is the central result of the present work. A derivation of it and further evaluation of the trace over τ are given in the second section. The expression (6) is equivalent to our expression even though to prove the equivalence is quite complicated. However, we have verified this in special cases. Expressions for parallel and perpendicular susceptibilities for this model are also expressed as traces only over τ matrices. The results for the various special cases (simple regular chain, and a general periodic chain with a basis) mentioned earlier will all be obtained in the third section. The structure of these expressions will appear slightly different from those found in the literature, partly because our method leads to the end result directly. The fourth section deals with the problem of random J_i, H_i and is a variant of a one-dimensional version of McCoy's work. It serves as a model for an amorphous spin system. The last section summarizes the results obtained.

II. COMPLETE SOLUTION OF THE GENERAL MODEL

We first demonstrate the result

Lemma 1:

$$\sum_{\sigma_2=\pm 1} \cdots \sum_{\sigma_N=\pm 1} A_1(\sigma_1,\sigma_2) A_2(\sigma_2,\sigma_3) \cdots A_N(\sigma_N,\sigma_{N+1}) \equiv A_1^\tau \cdots A_N^\tau, \tag{13}$$

where A_i^τ is as defined by Eqs. (8)–(11).

The proof is straightforward. We have already shown in the Introduction that $A_i(\sigma_i,\sigma_{i+1})$ can indeed be represented by A_i^τ. The next step is to show that

$$\sum_{\sigma_{i+1}=\pm 1} A_i(\sigma_i,\sigma_{i+1}) A_{i+1}(\sigma_{i+1},\sigma_{i+2}) = A_i^\tau A_{i+1}^\tau. \tag{14}$$

This is easily verified as follows. We first rewrite A_i in the usual form:

$$\begin{aligned} A_i(\sigma_i,\sigma_{i+1}) &= e^{\mathcal{J}_i \sigma_i \sigma_{i+1} + h_i \sigma_i} \\ &\equiv \cosh\mathcal{J}_i \cosh h_i + \sigma_i \cosh\mathcal{J}_i \sinh h_i \\ &\quad + \sigma_{i+1} \sinh\mathcal{J}_i \sinh h_i + \sigma_i\sigma_{i+1} \sinh\mathcal{J}_i \cosh h_i. \end{aligned} \tag{15}$$

Then the left side of (14) can be explicitly computed and the resulting expression recast as a 2×2 matrix. By writing out the product, $A_i^\tau A_{i+1}^\tau$ as a 2×2 matrix, it can be verified that these two are identical. By repeating this procedure, the identity given by (13) is in general established. An important corollary of this is

$$\sum_{\sigma_{i+1}=\pm 1} A_i(\sigma_i,\sigma_{i+1}) \sigma_{i+1} A_{i+1}(\sigma_{i+1},\sigma_{i+2}) = A_i^\tau \tau_z A_{i+1}^\tau. \tag{14'}$$

This follows at once by differentiating both sides of Eq. (14) with respect to h_{i+1} and using the definitions (5) and (8). This identity is useful in computing spin correlation functions in general.

Now, it is quite easy to prove

Lemma 2:

$$Z_N^{(\mathrm{closed})} = \mathrm{tr}_\tau \left\{ \prod_{i=1}^{N} A_i^\tau \right\}, \tag{16a}$$

$$Z_N^{(\mathrm{open})} = \mathrm{tr}_\tau \left\{ \left(\prod_{i=1}^{N-1} A_i^\tau \right) A_N^\tau (J_N = 0; H_N) \right\}; \tag{16b}$$

tr_τ here stands for trace on τ.

The proof is immediate

$$\begin{aligned} Z_N^{(\mathrm{closed})} &= \sum_{\sigma_1=\pm 1} \cdots \sum_{\sigma_N=\pm 1} A_1(\sigma_1,\sigma_2) \cdots A_N(\sigma_N,\sigma_1) \\ &= \mathrm{tr}_\tau \left\{ \prod_{i=1}^{N} A_i^\tau \right\}, \end{aligned}$$

where we used the fact that, for a closed chain, $\sigma_{N+1} = \sigma_1$. To prove (16b), we note that there is *zero coupling* ($J_N = 0$) between (N) and (1) and so A_N is represented by

$$A_N(\sigma_N,\sigma_1) = \begin{pmatrix} e^{h_N} & e^{h_N} \\ e^{-h_N} & e^{-h_N} \end{pmatrix}$$

or

$$A_N^\tau(J_N = 0; H_N) = a_N(J_N = 0; H_N) + \tau \cdot \mathbf{r}_N (J_N = 0; H_N). \tag{17}$$

Thus,

$$\begin{aligned} Z_N^{(\mathrm{open})} &= \sum_{\sigma_1=\pm 1} \cdots \sum_{\sigma_N=\pm 1} A_1(\sigma_1,\sigma_2) \cdots A_{N-1}(\sigma_{N-1},\sigma_N) \\ &\quad \times A_N(\sigma_N,\sigma_1) \\ &= \mathrm{tr}_\tau \left[\left(\prod_{i=1}^{N-1} A_i^\tau \right) A_N^\tau (J_N = 0; H_N) \right]. \end{aligned}$$

We have thus established (16a, b). We now establish similar expressions for spin correlation functions of all orders for both closed and open chains.

Lemma 3:

$$\langle \sigma_{l_1} \rangle = \frac{1}{Z_N} \mathrm{tr}_\tau \{ A_1^{(l_1-1)\tau} \tau_z A_{l_1}^{(N)\tau} \}, \tag{18a}$$

$$(l_2 \geq l_1) \langle \sigma_{l_1} \sigma_{l_2} \rangle = \frac{1}{Z_N} \mathrm{tr}_\tau \{ A_1^{(l_1-1)\tau} \tau_z A_{l_1}^{(l_2-1)\tau} \tau_z A_{l_2}^{(N)\tau} \} \tag{18b}$$

and quite generally,

$(l_k \geq l_{k-1} \geq \cdots \geq l_1)$,

$$\begin{aligned} &\langle \sigma_{l_1} \sigma_{l_2} \cdots \sigma_{l_k} \rangle \\ &\quad = \frac{1}{Z_N} \mathrm{tr}_\tau \{ A_1^{(l_1-1)\tau} \tau_z A_{l_1}^{(l_2-1)\tau} \tau_z \cdots A_{l_{k-1}}^{(l_k-1)\tau} \tau_z A_{l_k}^{(N)\tau} \}. \end{aligned} \tag{18c}$$

Here

$$\begin{aligned} A_{l_i}^{(0)\tau} &= 1 \quad (2 \times 2 \text{ unit matrix}), \\ A_{l_i}^{(l_i)\tau} &= A_{l_i}^\tau \equiv A_{l_i}(\sigma_{l_i}, \sigma_{l_i+1}), \\ A_{l_i}^{(l_j)\tau} &\equiv A_{l_i}^\tau A_{l_i+1}^\tau \cdots A_{l_j}^\tau \quad (l_j \geq l_i; l_j - l_i + 1 \text{ factors}). \end{aligned} \tag{19}$$

In all the above, $l_1 \cdots l_k$ take on the values $1, 2, \ldots, N$.

These are established by a straightforward application of (14'). We will illustrate it for closed chains and the results for the open chains follow if the trick used in (16b) is employed. By definition,

$$\langle \sigma_{l_i} \rangle^{(\text{closed})} = \frac{1}{Z_N^{(\text{closed})}} \sum_{\sigma_1=\pm 1} \cdots \sum_{\sigma_N=\pm 1} A_1(\sigma_1, \sigma_2) \times \cdots A_{l_i-1}(\sigma_{l_i-1}, \sigma_{l_i}) \sigma_{l_i} A_{l_i}(\sigma_{l_i}, \sigma_{l_i+1}) \cdots A_N(\sigma_N, \sigma_1).$$

Using (14′) and (14), we obtain

$$\langle \sigma_{l_i} \rangle^{(\text{closed})} = \frac{1}{Z_N^{(\text{closed})}} \operatorname{tr}_\tau \{ A_1^\tau \cdots A_{l_i-1}^\tau \tau_z A_{l_i}^\tau \cdots A_N^\tau \}.$$

Similarly, for $l_2 \geq l_1$,

$$\begin{aligned}\langle \sigma_{l_1} \sigma_{l_2} \rangle^{(\text{closed})} &= \frac{1}{Z_N^{(\text{closed})}} \sum_{\sigma_1=\pm 1} \cdots \sum_{\sigma_N=\pm 1} A_1(\sigma_1, \sigma_2) \\ &\times \cdots A_{l_1-1}(\sigma_{l_1-1}, \sigma_{l_1}) \sigma_{l_1} A_{l_1}(\sigma_{l_1}, \sigma_{l_1+1}) \\ &\times \cdots A_{l_2-1}(\sigma_{l_2-1}, \sigma_{l_2}) \sigma_{l_2} A_{l_2}(\sigma_{l_2}, \sigma_{l_2+1}) \cdots A_N(\sigma_N, \sigma_1) \\ &= \frac{1}{Z_N^{(\text{closed})}} \operatorname{tr}_\tau \{ A_1^{(l_1-1)\tau} \tau_z A_{l_1}^{(l_2-1)\tau} \tau_z A_{l_2}^{(N)\tau} \}.\end{aligned}$$

Similarly $\langle \sigma_{l_1} \cdots \sigma_{l_k} \rangle$ can be derived following the same procedures. These expressions are seen to be the ones given in (18a) and (18b) and the introduction of the notation (19) becomes evident. We must point out that these correlation functions could have been obtained by a direct differentiation of the expressions for Z_N appropriately with respect to magnetic fields $\{h_i\}$ or the interaction strengths $\{\mathcal{J}_i\}$ but in practice the above traces are found to be easier to evaluate in the absence of the fields, $\{h_i = 0\}$.

We must next evaluate the traces on τ. To do this, we use repeatedly the well-known identity that if **A**, **B** are two three-dimensional vectors, then, in the usual notation,

$$(\boldsymbol{\tau} \cdot \mathbf{A})(\boldsymbol{\tau} \cdot \mathbf{B}) = \mathbf{A} \cdot \mathbf{B} + i\boldsymbol{\tau} \cdot (\mathbf{A} \times \mathbf{B}) \tag{20}$$

and the fact that $\operatorname{tr}_\tau \tau = 0$.

It is clear that by employing (20), we can write

$$A_{l_i}^{(l_j)\tau} = \mathfrak{A}_{l_i}^{(l_j)} + \boldsymbol{\tau} \cdot \mathfrak{R}_{l_i}^{(l_j)} \tag{21}$$

where $\mathfrak{A}_{l_i}^{(l_j)}$ and $\mathfrak{R}_{l_i}^{(l_j)}$ are multinomials of $\{a_i, r_i, \hat{n}_i\}$:

$$\begin{aligned}\mathfrak{A}_{l_i}^{(l_j)} &= \left(\prod_{i=l_i}^{l_j} a_i \right) \Bigg(1 + \sum_{(i_1,i_2)}{}' s_{i_1} s_{i_2} (\hat{n}_{i_1} \cdot \hat{n}_{i_2}) \\ &+ i \sum_{(i_1,i_2,i_3)}{}' s_{i_1} s_{i_2} s_{i_3} [(\hat{n}_{i_1} \times \hat{n}_{i_2}) \cdot \hat{n}_{i_3}] \\ &+ \sum_{(i_1,i_2,i_3,i_4)}{}' s_{i_1} s_{i_2} s_{i_3} s_{i_4} [(\hat{n}_{i_1} \cdot \hat{n}_{i_2})(\hat{n}_{i_3} \cdot \hat{n}_{i_4}) \\ &- (\hat{n}_{i_1} \times \hat{n}_{i_2}) \cdot (\hat{n}_{i_3} \times \hat{n}_{i_4})] + i \sum_{(i_1 \cdots i_5)}{}' s_{i_1} s_{i_2} s_{i_3} s_{i_4} s_{i_5} \\ &\times [(\hat{n}_{i_1} \cdot \hat{n}_{i_2})[(\hat{n}_{i_3} \times \hat{n}_{i_4}) \cdot \hat{n}_{i_5}] + [(\hat{n}_{i_1} \times \hat{n}_{i_2}) \cdot \hat{n}_{i_5}](\hat{n}_{i_3} \cdot \hat{n}_{i_4}) \\ &- [(\hat{n}_{i_1} \times \hat{n}_{i_2}) \times (\hat{n}_{i_3} \times \hat{n}_{i_4})] \cdot \hat{n}_{i_5}] + \cdots \Bigg) \\ &\equiv \left(1 + \sum_{k=2}^{(l_j - l_i + 1)} \mathfrak{A}^{(k)} \right) \left(\prod_{i=l_i}^{l_j} a_i \right),\end{aligned} \tag{21a}$$

$$\begin{aligned}\mathfrak{R}_{l_i}^{(l_j)} &= \left(\prod_{i=l_i}^{l_j} a_i \right) \Bigg(\sum_{(i_1)} s_{i_1} \hat{n}_{i_1} + i \sum_{(i_1,i_2)}{}' s_{i_1} s_{i_2} (\hat{n}_{i_1} \times \hat{n}_{i_2}) \\ &+ \sum_{(i_1,i_2,i_3)}{}' s_{i_1} s_{i_2} s_{i_3} [(\hat{n}_{i_1} \cdot \hat{n}_{i_2}) \hat{n}_{i_3} - (\hat{n}_{i_1} \times \hat{n}_{i_2}) \times \hat{n}_{i_3}] + \cdots \Bigg) \\ &\equiv \left(\sum_{k=1}^{(l_j - l_i + 1)} \mathfrak{R}^{(k)} \right) \left(\prod_{i=l_i}^{l_j} a_i \right),\end{aligned} \tag{21b}$$

where we have introduced a new symbol

$$s_i = r_i / a_i. \tag{22}$$

In both (21a, b) $\sum'_{(i_1 \cdots i_k)}$ denotes that $i_1 \cdots i_k$ are ordered ($i_1 < i_2 < \cdots < i_k$), that no two of $(i_1 \cdots i_k)$ are equal, and they all go over $l_i \cdots l_j$. The expressions for $\mathfrak{A}_{l_i}^{(l_j)}$ and $\mathfrak{R}_{l_i}^{(l_j)}$ are remarkable in that $\mathfrak{A}^{(2k)}$ involves only scalars, $\mathfrak{A}^{(2k+1)}$ only pseudoscalars, whereas $\mathfrak{R}^{(2k)}$ involves only vectors, and $\mathfrak{R}^{(2k+1)}$ only pseudovectors formed out of the unit vectors $\{\hat{n}_i\}$. The interpretation that $\hat{n}_i$ specifies the spatial orientation of the *ith spin* or *bond* makes these expressions physically appealing as one can perhaps motivate further schemes of approximations for the "amorphous" system. We will return to this question in Sec. IV. We thus obtain

$$\begin{aligned}Z_N^{(\text{closed})} &= 2\,\mathfrak{A}_1^{(N)}, \\ Z_N^{(\text{open})} &= 2[\mathfrak{A}_1^{(N-1)} a_N(J_N = 0; H_N) + \mathfrak{R}_1^{(N-1)} \\ &\quad \cdot \mathbf{r}_N(J_N = 0; H_N)].\end{aligned} \tag{23}$$

This completes the derivation of the partition functions and spin correlation functions for the *finite, open,* and *closed* Ising chains for the Hamiltonian given by (1). We will now compute parallel and perpendicular susceptibilities of this system. This generalizes the work of Fisher[7] as well as provides a new method of computing the perpendicular susceptibility using our formalism.

(*a*) *Parallel susceptibility:* The parallel susceptibility is given

$$\begin{aligned}\chi_\parallel(\{J_i, H_i\}) &\equiv \frac{g^2 \mu_B^2}{NkT} \sum_{l,m=1}^{N} [\langle \sigma_l \sigma_m \rangle - \langle \sigma_l \rangle \langle \sigma_m \rangle] \\ &= \frac{g^2 \mu_B^2}{NkT} \Bigg(\sum_{l=1}^{N} (1 - \langle \sigma_l \rangle^2) + 2 \sum_{l=2}^{N} \sum_{m=1}^{l-1} (\langle \sigma_m \sigma_l \rangle \\ &\quad - \langle \sigma_m \rangle \langle \sigma_l \rangle) \Bigg).\end{aligned} \tag{24}$$

Using the expressions (18a, b) and (21) and the cyclic property of the trace operations, we may express this in an elegant form:

$$\langle \sigma_l \rangle = \frac{1}{Z_N} \operatorname{tr}_\tau [C_{(l;1)}^{(N;\,l-1)\tau} \tau_z], \tag{25a}$$

$$(m \geq l) \quad \langle \sigma_l \sigma_m \rangle = \frac{1}{Z_N} \operatorname{tr}_\tau [C_{(m;1)}^{(N;\,l-1)\tau} (\tau_z A_l^{(m-1)\tau} \tau_z)], \tag{25b}$$

where we have introduced the notation

$$A_m^{(N)\tau} A_1^{(l-1)\tau} \equiv C_{(m;1)}^{(N;\,l-1)\tau} \equiv [K_{(m;1)}^{(N;\,l-1)} + \mathbf{L}_{(m;1)}^{(N;l-1)} \cdot \boldsymbol{\tau}]. \tag{25c}$$

This leads us to the expressions

$$\langle \sigma_l \rangle = \frac{2}{Z_N} L_{(l;1)z}^{(N;l-1)}, \tag{26a}$$

$$(m \geq l) \quad \langle \sigma_l \sigma_m \rangle = \frac{2}{Z_N} [K_{(m;1)}^{(N;l-1)} \mathfrak{A}_l^{(m-1)} + \mathbf{L}_{(m;1)}^{(N;l-1)} \cdot \mathfrak{R}_l^{(m-1)'}], \tag{26b}$$

where $\mathfrak{R}_l^{(m-1)'} \equiv (-\mathfrak{R}_{lx}^{(m-1)}, -\mathfrak{R}_{ly}^{(m-1)}, \mathfrak{R}_{lz}^{(m-1)})$ with components of $\mathfrak{R}_l^{(m-1)}$ defined as in expression (21). No further simplification occurs in the general case. When all H_i are set equal to zero but with general J_i, some more simplification obtains and the result is discussed in Sec. IV.

(*b*) *Perpendicular susceptibility:* Fisher[7] developed a diagram summation method for computing the perpendicular susceptibility $\chi_\perp$. For the special case of all J_i

equal and all H_i equal to zero, by summing all the relevant diagrams associated with the one-dimensional Ising model he obtained an exact expression for it for both open and closed chains. We here employ our method to obtain $\chi_\perp$ for the general Hamiltonian given by Eq. (1). We apply the standard Kubo formula for the linear response function and re-express this in terms of our formulation. Cabib and Mahanti[8] have recently used the Kubo formula to derive $\chi_\perp$ when all H_i are zero for an open chain. We begin with the expression

$$\chi_\perp(\{J_i, H_i\}) = \frac{g^2\mu_B^2}{NZ_N} \sum_{l,m=1}^{N} \int_0^\beta dy \, \mathrm{Tr}_{\{\sigma\}}(e^{-\beta H_N} \sigma_{lx}(y)\, \sigma_{mx}(0)). \tag{27}$$

Here $\mathrm{Tr}_{\{\sigma\}}$ implies trace over the N Pauli spin vectors σ_k. Clearly, we have used the fact that $\langle \sigma_{lx} \rangle = 0$ in (27), which is trivially established. Also,

$$\sigma_{lx}(y) = e^{yH_N}\, \sigma_{lx}(0)\, e^{-yH_N}. \tag{28}$$

We now express

$$S_{lm}(y) = \mathrm{Tr}_{\{\sigma\}}\, [e^{-\beta H_N}\, \sigma_{lx}(y)\, \sigma_{mx}(0)] \tag{29}$$

in terms of trace over only the standard Pauli spin vector τ. Now

$$S_{lm}(y) = \mathrm{Tr}_{\{\sigma\}} \{e^{-H_N(\beta-y)}\, \sigma_{lx}(0)\, e^{-H_N y}\, \sigma_{mx}(0)\}$$
$$= \mathrm{Tr}_{\{\sigma\}} [A_1^{(\beta-y)} A_2^{(\beta-y)} \cdots A_N^{(\beta-y)}\, \sigma_{lx} A_1^{(y)} \cdots A_N^{(y)} \sigma_{mx}(0)].$$

We notice that σ_{lx} commutes with all A_i except the ones with $i = l$ and $i = l-1$; and of course all the A_i's commute among themselves. In view of these observations, we can write (we have here used the cyclic property of the trace)

$$S_{lm}(y) = \mathrm{Tr}_{\{\sigma\}} [A_1^{(\beta)} \cdots \tilde{B}_{m-1}^{(\beta;\, y)}\, \tilde{B}_m^{(\beta;\, y)} A_{m+1}^{(\beta)} \cdots \sigma_{mx}$$
$$\cdot \sigma_{lx}\, \tilde{B}_{l-1}^{(\beta;\, y)}\, \tilde{B}_l^{(\beta;\, y)} A_{l+1}^{(\beta)} \cdots A_N^{(\beta)}],$$

where

$$\tilde{B}_l^{(\beta;\, y)} = \sigma_{lx} A_l^{(\beta-y)}\, \sigma_{lx} A_l^{(y)}. \tag{30}$$

We now make several observations.

(a) $\sigma_{lx} A_l^{(\beta-y)}\, \sigma_{lx}$ is independent of σ_{lx} as can be checked directly by using, for instance, the representation (15).

(b) $S_{lm}(y) = 0 \quad \text{if } l \neq m.$ (31)

This is because σ_{lx} and σ_{my} occur linearly and traces on them are zero.

(c) $S_{ll}(y) = \mathrm{tr}_\tau \{A_{(1)}^{(l-2)\tau}\, \tilde{B}_{l-1}^{\tau}(\beta; y)\, \tilde{B}_l^{\tau}(\beta; y)\, A_{(l+1)}^{(N)\tau}\}$, (32)

where we have used the notation introduced in expressions (19) and (21). The procedure to go over to the τ-representation from expression (32) is the same as before. One thus obtains the general result, using the new symbol introduced in Eq. (25c):

$$\chi_\perp(\{J_i, H_i\}) = \frac{g^2\mu_B^2}{NZ_N} \sum_{l=1}^{N} \mathrm{tr}_\tau [C_{(l+1;1)}^{(N;\, l-2)\tau} \int_0^\beta \tilde{B}_{l-1}^{\tau}(\beta; y) \tilde{B}^{\tau}(\beta; y) dy]. \tag{33}$$

The integration over y can be done under the trace operation quite easily for both the open and closed cases. These will be illustrated in the special case of general J_i and all $H_i = 0$ in Sec. IV. Elsewhere these formulas are employed to study the case where all J_i's are equal to J and all H_i's are equal to H.[10]

III. ILLUSTRATIVE EXAMPLES-REGULAR CHAINS

We will obtain the known results for a few regular chains in this section.

(a) *Simple regular chains:* Here all the J_i's are equal ($= J$) and all the H_i's are equal ($= H$) (except $J_N = 0$ for an open chain). Then all the A_i's become equal. Then $\mathfrak{A}_1^{(N)}$ and $\mathfrak{R}_1^{(N)}$ become very simple expressions:

$$\mathfrak{A}_1^{(N)} = a^N \{1 + \binom{N}{2} s^2 + \binom{N}{4} s^4 + \cdots\}$$
$$= \tfrac{1}{2}(\lambda_1^N + \lambda_2^N), \tag{34}$$

after using (10) to simplify $(a \pm r)$ in terms of the eigenvalues of A. Similarly

$$\mathfrak{R}_1^{(N)} = \hat{n}\, a^N \{\binom{N}{1} s + \binom{N}{3} s^3 + \cdots\} = (\hat{n}/2)(\lambda_1^N - \lambda_2^N). \tag{35}$$

Hence

$$Z_N^{(\mathrm{closed})} = \lambda_1^N + \lambda_2^N, \tag{36a}$$

$$Z_N^{(\mathrm{open})} = \lambda_1^{N-1} [\cosh h + (e^{-\vartheta} + e^{\vartheta} \sinh^2 h)/|\mathbf{r}|]$$
$$+ \lambda_2^{N-1} [\cosh h - (e^{-\vartheta} + e^{\vartheta} \sinh^2 h)/|\mathbf{r}|]. \tag{36b}$$

Expression (36a) is given in Thompson's book[1] while (36b) was given by Falk.[6]

When $H = 0$, $|\mathbf{r}| = e^{-\vartheta}$ and the results for both open and closed chains agree with the corresponding ones given by Thompson.[1]

(b) *A regular chain with a basis:* By a regular chain with a basis we mean

$$J_{p+k} = J_k, \qquad H_{p+k} = H_k, \tag{37}$$

where p is some fixed integer and k is any integer. When $p = 1$, we have the simple chain. This means that $(J_1 \cdots J_p), (H_1 \cdots H_p)$ are unequal among themselves and that this "unit" repeats itself along the chain. In this case, it is clear that we obtain (N = total length of the chain and so N/p is an integer)

$$Z_N = \mathrm{tr}_\tau \{(A_1^\tau \cdots A_p^\tau)^{N/p}\}. \tag{38}$$

We can write as before

$$A_1^\tau \cdots A_p^\tau \equiv A_1^{(p)\tau} = \mathfrak{A}_1^{(p)} + \tau \cdot \mathfrak{R}_1^{(p)} \tag{39}$$

as in expression (21). By using the same procedures as in case (a) we obtain the result

$$Z_N^{(\mathrm{closed})} = (\lambda_{p1}^{N/2} + \lambda_{p2}^{N/2}). \tag{40}$$

The corresponding expression for the open chain may be obtained as before and has the structure of Eq. (36b). Here $\lambda_{p1}, \lambda_{p2}$ are the two eigenvalues of $A_1^{(p)\tau}$. The case where $p = 2$, with $J_1 = J_2 (= U)$ in the thermodynamic limit, for the closed chain, is discussed in Refs. 2 and 3. Our results agree with theirs after some considerable manipulation.

IV. ONE-DIMENSIONAL ISING GLASS[8,9]

We mean by a one-dimensional Ising glass a random array of spin 1/2 particles on a straight line, in analogy to a one-dimensional model of an amorphous system of particles moving in a random one-dimensional potential. Since our model incorporates only randomness in the strengths (J_i, H_i), it is more akin to the model of an amorphous system where only the strengths of the potentials are random and not their location (viz. Kronig-Penney model with random strength parameters but keeping the periodicity intact). Since $\{J_i\}$ represent the

interaction between neighboring spins and $\{H_i\}$ represent the "energy" of the spin at site i, their statistical fluctuations can be of many different varieties corresponding to different situations. We will here discuss a few of these. For the sake of illustration, we consider only the case of open chains.

Case (i): $H_i = 0$ for all i: To begin with, let us treat all J_i's as random variables;[9] let $P_m(J_1 \cdots J_m)$ represent the probability distribution function for m-"bonds" without regard to the sequence of bonds $(1 \cdots m)$. Quite generally one has a sequence of probabilities $P_1, P_2 \cdots P_N$ if N is the total number of "bonds" with the usual relations among themselves

$$\int_{-\infty}^{\infty} P_m(J_1 \cdots J_m) dJ_m = P_{m-1}(J_1 \cdots J_{m-1}) \quad (m = 1, 2, \ldots, N), \qquad P_0 = 1. \tag{41}$$

The free energy of the system is given by

$$F_N(\{J_i\}) = -kT \ln Z_N(\{J_i\}) \tag{42}$$

and it is this quantity that one must average over all $\{J_i\}$, when one is considering a macroscopic system. For "small" systems such as biopolymers, it is more appropriate to average first over Z_N. The results of these two operations are of course, very different. Thus

$$\bar{F}_N = \int F_N(\{J_i\}) P_N(\{J_i\}) dJ_1 \cdots dJ_N. \tag{43}$$

When $H_i = 0$, we know from (11) that $\hat{n}_i = (1, 0, 0)$ and from (10), (21a), and (23)

$$Z_N^{(\text{open})}(\{J_i\}) = 2 \prod_{i=1}^{N-1} (2 \cosh J_i/kT) \tag{44}$$

since $\lambda_i = 2 \cosh J_i/kT$, $J_N = 0$. Thus using (4), we obtain

$$\bar{F}_N^{(\text{open})} = -(N-1)kT \int_{-\infty}^{\infty} P_1(J) \ln(2 \cosh J/kT) \times dJ - kT \ln 2. \tag{45}$$

This scheme clearly *assumes* equal *a priori* probabilities for *all i*. From (21) we note that the open strand is simpler to analyze for the present purposes.

One could introduce a more general scheme of probabilities involving *different probabilities for different bonds i* in which case $P_m(J_1 \cdots J_m)$ describes the probability distribution of the sequence of bonds $J_1 \cdots J_m$. This leads to more complexities, but is perhaps applicable to biological problems where we may have an Ising system with a *basis*. Then the expression for $\bar{F}_N^{(\text{open})}$ becomes

$$\bar{F}_N^{(\text{open})} = -kT \sum_{i=1}^{N-1} \int_{-\infty}^{\infty} P_i(J_i) \ln(2 \cosh J_i/kT) \times dJ_i - kT \ln 2 \tag{45'}$$

with

$$\int_{-\infty}^{\infty} P_i(J_i) dJ_i = 1 \quad (i = 1, \ldots, N-1), \qquad P_N(J_N) = \delta(J_N).$$

It is of interest to point out that the result (45) could have been obtained by first assuming all J_i's to be equal (except that $J_N = 0$) and then averaging over J. For, then

$$F_N^{(\text{open})} = -(N-1)kT \ln(2 \cosh J/kT) - kT \ln 2 \tag{46}$$

and averaging over different J's leads to (45). This is only a restatement of the well known equivalence principle concerning the free energy, that if all the *a priori* probabilities of the random variables are the same and the random variables are statistically independent, $P_m(J_1 \cdots J_m) = \prod_{i=1}^{m} P_i(J_i)$, then an average over different J_i is the same as the "ensemble" average over (46). The calculation of susceptibilities should be done carefully. Using (18a, b) we can compute very easily the averages $\langle \sigma_l \rangle$ and $\langle \sigma_m \sigma_l \rangle$ for $\{H_i \neq 0\}$. For $\{H_i = 0\}$, one then notes

$$A_i^{\tau}(J_i, H_i = 0) = (a_i + r_{ix} \tau_x) \tag{47}$$

and, further, (21a, b) simplify.

Clearly then,

$$\langle \sigma_l \rangle \Big|_{\{H_i = 0\}} = 0, \qquad \text{for both open and closed chains.} \tag{48}$$

Similarly,

$$\langle \sigma_l \sigma_m \rangle^{(\text{open})} \Big|_{\{H_i = 0\}} = \prod_{i=1}^{(m-1)} [\lambda_2(i)/\lambda_1(i)]$$

$$\langle \sigma_l \sigma_m \rangle^{(\text{closed})} \Big|_{\{H_i = 0\}} = \frac{1}{2} \times \frac{\prod_{i=1}^{N} \lambda_1(i) \prod_{k=l}^{m-1} [\lambda_2(k)/\lambda_1(k)] + \prod_{i=1}^{N} \lambda_2(i) \prod_{k=l}^{m-1} [\lambda_1(k)/\lambda_2(k)]}{\prod_{i=1}^{N} \lambda_1(i) + \prod_{i=1}^{N} \lambda_2(i)}. \tag{49}$$

We thus see that it is easier to compute $\chi_{\parallel}$ for open, finite strands $[\lambda_1(i) = 2 \cosh J_i/kT, \lambda_2(i) = 2 \sinh J_i/kT]$. One obtains, in the general case, for the average susceptibility

$$\bar{\chi}_{\parallel}^{(\text{open})} = \frac{g^2 \mu_B^2}{NkT} \left\{ N + 2 \left[\sum_{i=1}^{(N-1)} \left(\overline{\tanh \frac{J_i}{kT}} \right)_1 + \sum_{i=1}^{(N-2)} \left(\overline{\tanh \frac{J_i}{kT} \tanh \frac{J_{i+1}}{kT}} \right)_2 + \cdots + \left(\tanh \frac{J_1}{kT} \cdots \tanh \frac{J_{N-1}}{kT} \right)_{N-1} \right] \right\}, \tag{50}$$

where

$$\left(\overline{\tanh \frac{J_1}{kT} \cdots \tanh \frac{J_k}{kT}} \right)_k \equiv \int \cdots \int dJ_1 \cdots dJ_k \, P_k(J_1 \cdots J_k) \prod_{i=1}^{k} \tanh \frac{J_i}{kT}. \tag{51}$$

Thus, we find all the higher-order joint probability functions appearing in computing $\bar{\chi}_{\parallel}^{(\text{open})}$. If we *assume*[8] statistical independence of these variables $\{J_i\}$, we have

$$\left(\overline{\tanh \frac{J_1}{kT} \cdots \tanh \frac{J_k}{kT}} \right)_k = \left(\overline{\tanh \frac{J}{kT}} \right) \equiv t^k. \tag{52}$$

Expression (50) can then be simplified further:

$$\bar{\chi}_{\parallel}^{(\text{open})} = \frac{g^2 \mu_B^2}{NkT} \left(N \left(\frac{1+t}{1-t} \right) + 2t \, \frac{(1-t^N)}{(1-t)^2} \right). \tag{53}$$

In the thermodynamic limit $(N \to \infty)$, and since $|t| < 1$ in general, we obtain a simple result

$$\bar{\chi}_{\parallel(N \to \infty)} = \frac{g^2 \mu_B^2}{kT} \left(\frac{1+t}{1-t} \right) = \frac{g^2 \mu_B^2}{kT} \left(\frac{1 + \overline{\tanh J/kT}}{1 - \overline{\tanh J/kT}} \right). \tag{54}$$

Note that if we had proceeded by first calculating $\chi_{N \to \infty_{\parallel}}^{(\text{open})}$, we would have obtained

$$\chi_{(N\to\infty)\parallel} = \frac{g^2\mu_B^2}{kT}\left(\frac{1+\tanh J/kT}{1-\tanh J/kT}\right) = \frac{g^2\mu_B^2\, e^{2J/kT}}{kT}, \qquad (55)$$

whose average over J is nowhere similar to (54).

The perpendicular susceptibility can also be calculated when all the H_i are zero. For, performing the y integration in (33), and recasting the result in the τ representation, we have

$$\int_0^\beta dy\, \tilde{B}^{\tau}_{l-1}(\beta;y)\,\tilde{B}^{\tau}_{l}(\beta;y)\Big|_{\{H_i=0\}} = 2\beta\left(\frac{\sinh(\mathcal{J}_l+\mathcal{J}_{l+1})}{(\mathcal{J}_1+\mathcal{J}_{l+1})} + \tau_x\,\frac{\sinh(\mathcal{J}_l-\mathcal{J}_{l-1})}{(\mathcal{J}_l-\mathcal{J}_{l-1})}\right).$$

And

$$C^{(N;\,l-2)\tau}_{(l+1;\,1)} = [(\mathcal{A}^{(N-1)}_{l+1} + \mathcal{R}^{(N-1)}_{l+1x}\tau_x)(a_N + r_{Nx}\tau_x) \times (\mathcal{A}^{(l-2)}_1 + \mathcal{R}^{(l-2)}_{1x}\tau_x)].$$

Hence, we obtain for the open strand

$$\chi_\perp^{(\mathrm{open})}(\{J_i\}) = \frac{g^2\mu_B^2}{NkT}\sum_{l=1}^{N}\frac{1}{(\mathcal{J}_l^2-\mathcal{J}_{l-1}^2)}(\mathcal{J}_l\tanh\mathcal{J}_l - \mathcal{J}_{l-1}\tanh\mathcal{J}_{l-1}) \qquad (56)$$

and for the closed chain

$$\chi_\perp^{(\mathrm{closed})}(\{J_i\}) = \frac{g^2\mu_B^2}{NkT}\sum_{l=1}^{N}\frac{1}{(\mathcal{J}_l^2-\mathcal{J}_{l-1}^2)}\left[(\mathcal{J}_l\tanh\mathcal{J}_l - \mathcal{J}_{l-1}\tanh\mathcal{J}_{l-1}) + \prod_{i=1}^{N}(\tanh\mathcal{J}_i)\left(\frac{\mathcal{J}_l}{\tanh\mathcal{J}_l} - \frac{\mathcal{J}_{l-1}}{\tanh\mathcal{J}_{l-1}}\right)\right]\frac{1}{(1+\Pi_{i=1}^{N}\tanh\mathcal{J}_i)}. \qquad (57)$$

For a uniform system, all the $\mathcal{J}_l$'s are the same; one then resorts to a limiting procedure and the results obtained coincide with those derived by Fisher.[7] One must also note that again the averaging process over random $\{\mathcal{J}_i\}$ can be more easily accomplished with the open chain as in the parallel susceptibility.

Case (ii): $J_i = U$ for all i: This is the model used frequently by Goel and Montroll for describing some biopolymers. When all the J_i's are equal, from (23a) we can show quite easily that all the *pseudoscalars vanish* because

$$\hat{n}_i = (e^{-\mathcal{J}}\cosh h_i/r_i,\, ie^{-\mathcal{J}}\sinh h_i/r_i,\, e^{\mathcal{J}}\sinh h_i/r_i),$$

$$\hat{n}_1\times\hat{n}_2 = (0,\sinh(h_1-h_2)/r_1r_2, -ie^{-2\mathcal{J}}\sinh(h_1-h_2)/r_1r_2),$$

and hence

$$(\hat{n}_{i_1}\times\hat{n}_{i_2})\cdot\hat{n}_{i_3} = 0 \quad \text{etc.} \quad (\text{for } i_1\neq i_2\neq i_3).$$

This is a consequence of "isotropy" (nongyrotropic medium!) of such a model. This appears to be an amusing consequence of our formalism. In (25a) therefore only the terms $\mathcal{A}^{(2k)}$ are nonzero. One at once finds that in the present model, the averaging procedure for $\ln Z_N(U;\{H_i\})$ would be quite involved, and much more so for the even more general model where we have $\ln Z_N(\{J_i\};\{H_i\})$. One therefore must resort to approximation schemes. We shall not discuss these questions any further.

V. SUMMARY

In this paper we have presented a novel method of calculating the partition function and spin correlation functions of a generalized Ising system of N spins (spin $\frac{1}{2}$) in a magnetic field and arranged on a one-dimensional chain described by the Hamiltonian (1). This general problem, to the best of our knowledge, has not been discussed before in the literature and is of considerable interest in itself. Besides, it is of interest in dealing with some models of long chain molecules of biological interest where the assumption of the existing model discussed in Case (ii) of Sec. IV may not hold. It can also serve as a useful model for discussing averaging procedures in the theory of amorphous systems. For amorphous systems of interest in solid state, it is believed that averaging over the thermodynamic quantities such as free energy, susceptibility, specific heat, etc. is appropriate, because, on such a macroscopic scale in both time and spatial extent, the system tends to be in a state of minimum free energy and minimum fluctuations of all the correlations. This is the scheme we have examined in Sec. IV and we have explicitly worked out the case of a simple model where all $H_i = 0$. For biological applications, it is perhaps more meaningful to study random chains of fixed, finite length and average over the partition function itself and from it, deduce the free energy, etc. The final results of these calculations differ a great deal and a simple model such as the one discussed here brings out these differences quite transparently. The formalism developed here has also been applied to study the parallel and perpendicular susceptibilities of (a) an Ising chain (nearest neighbor model) in the presence of a uniform magnetic field and (b) an Ising chain with nearest and next-nearest neighbor interactions. These results will be discussed elsewhere.[10] We have also used this method to study the finite size and impurity effects on the properties of model DNA chains.[11]

* NDEA Fellow.

[1] See for instance, C. J. Thompson, *Mathematical Statistical Mechanics* (MacMillan, New York, 1972), Chaps. 5, 6, and 7. The "standard" representation given by Eq. (12) in the text is a slight generalization of the one given in Appendix D of this book. The present work exploits this representation to the full.

[2] N. S. Goel, in *Progress of Theoretical Biology*, edited by R. Rosen and F. M. Snell (Academic, New York, 1972), Vol. 2. References to original works are contained in this and the companion review article, Ref. 3.

[3] R. M. Wartell and E. W. Montroll, in *Advances in Chemical Physics*, edited by I. Prigogine and S. A. Rice (Interscience, New York, 1972), Vol. 22.

[4] B. McCoy, in *Phase Transitions and Critical Phenomena*, edited by O. Domb and M. S. Green (Academic, New York, 1972), Vol. 2.

[5] H. Falk showed this trick to one of us (AKR) and also gave us the references: J. F. Dobson, J. Math. Phys. **10**, 40 (1969), and J. Stephenson, Can. J. Phys. **48**, 1724 (1970). This trick is useful only for a finite open chain; for a finite closed chain one has the extra constraint $t_1 t_2 \ldots t_N = 1$ which seems to make the trick less useful. The thermodynamic limit ($N\to\infty$), of course, does not depend on the boundary conditions.

[6] H. Falk, Am. J. Phys. **40**, 1453 (1972).

[7] M. E. Fisher, J. Math. Phys. **4**, 124 (1963).

[8] D. Cabib and S. D. Mahanti (preprint), to be published (1974). This section is an outgrowth of an attempt to build simple models of "spin glass" in collaboration with Dr. S. D. Mahanti, Michigan State University, East Lansing, Michigan. AKR thanks him for communicating some of his unpublished results.

[9] C. Fan and B. M. McCoy, Phys. Rev. **182**, 614 (1969). Note that our model differs from the model of these authors in that we allow for statistical dependence of J_i's on different sites. This new aspect of our model makes the calculation of the various thermodynamic quantities more interesting, as the calculation of $\chi_\parallel$ illustrates, for example.

[10] G. S. Grest and A. K. Rajagopal, J. Math. Phys. **15**, 589 (1974).

[11] G. S. Grest and A. K. Rajagopal, J. Theor. Biol. (1974) (to be published).

Perpendicular susceptibility of two one-dimensional Ising chains

G. S. Grest* and A. K. Rajagopal

Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803
(Received 30 August 1973; revised manuscript received 11 October 1973)

Exact, closed expressions are obtained for the perpendicular and parallel susceptibilities of the one-dimensional spin-1/2 Ising model (i) with nearest neighbor interactions in a uniform magnetic field and (ii) with next-nearest neighbor interactions but without the magnetic field, in the thermodynamic limit. Graphs of the susceptibilities as a function of temperature for various magnetic field strengths in (i), and as a function of different ratios of the interaction strengths in (ii), are discussed.

I. INTRODUCTION

This is a sequel to the companion paper on an approach to an exact treatment of a one-dimensional spin-1/2 Ising system. The first paper[1] contains the general theory while this paper will consider its applications to two model Hamiltonians. The first model will be a system with the nearest neighbor (nn) interaction in a magnetic field. The case of $H = 0$ with random nn interaction strengths has previously been considered in Ref. 1. The second system to be considered will be the next nearest neighbor (nnn) interaction in a zero magnetic field. We obtain exact formulas in closed forms for both perpendicular ($\chi_\perp$) and parallel ($\chi_\parallel$) susceptibilities for these systems.

These two problems were chosen because of their special significance. The first problem is of intrinsic interest because it enables one to study the behavior of the system as one increases the strength of the magnetic field. There is a similar study of the effect of magnetic field on the magnetization of a linear chain of classical spins by Blume and co-workers.[2] These studies reveal, for instance, how the magnetization approaches its equilibrium value as the field strength is increased, for a fixed temperature. The present study enables one to study the behavior of two types of second order spin correlations as a function of field strength.

$\chi_\perp$ for the case of nn interactions was first considered by Fisher[3] for zero magnetic field. For a nonzero magnetic field, Marsh, [4] by calculating a perturbation series for the free energy, deduced an expression for $\chi_\perp$. He seems to have missed an essential feature of $\chi_\perp$ for $H = 2|J|, J < 0$. We calculate $\chi_\perp$ and $\chi_\parallel$ and plot both of these as a function of temperature for various values of the magnetic field.

It is shown by Dobson[5] that the partition function for an open chain Ising system with nnn interaction in zero magnetic field is related to that of the nn interaction with a magnetic field by a simple spin variable transformation. It is of interest to examine how far this observation of Dobson is applicable when one considers higher order spin correlation functions of this system. Stephenson[6] has studied this question to some extent; he has used the Dobson trick to calculate the pair correlation function and hence also $\chi_\parallel$. He noted some very interesting behavior of the system when the ratio of the strength of the second neighbor to the first neighbor interaction, α, is varied. For instance, he showed that $\chi_\parallel$ has an essential singularity at zero temperature for the special value of $\alpha = 0.5$ with antiferromagnetic nn coupling. We find this feature persisting in $\chi_\perp$. Similar behavior has been noted for the same Hamiltonian but with Heisenberg interaction for the corresponding value of $\alpha = 0.5$.[7] $\chi_\perp$ also has the interesting property of being independent of the sign of the nn interaction. We calculate $\chi_\perp$ and $\chi_\parallel$ for this model in the thermodynamic limit using our method and their dependence on temperature are plotted for various values of α and different signs of the coupling strengths.

In Sec. II, we will calculate $\chi_\perp$ and $\chi_\parallel$ for the nn model in a magnetic field while Sec. III will contain the derivation of the same functions for the nnn model. Concluding remarks are given in Sec. IV.

II. NEAREST NEIGHBOR INTERACTION IN A MAGNETIC FIELD

Consider the Hamiltonian for a closed chain,

$$\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_{\text{ext}}, \tag{1}$$

where

$$\mathcal{H}_1 = -J \sum_{i=1}^{N} \sigma_i^z \sigma_{i+1}^z - h \sum_{i=1}^{N} \sigma_i^z , \tag{2a}$$

$$\mathcal{H}_{\text{ext}} = -h_z \sum_{i=1}^{N} \sigma_i^z - h_x \sum_{i=1}^{N} \sigma_i^x , \tag{2b}$$

with $\sigma_{N+1} = \sigma_1$ and $h = g\mu_B H$, where H is the magnetic field. We consider here a closed chain and in the next section study the second problem for an open chain for the explicit purpose of displaying the power of the method of Ref. 1. Of course, in the thermodynamic limit, the boundary conditions do not matter. $\mathcal{H}_{\text{ext}}$ will be considered as a perturbation. $\chi_\perp$ is derived by letting $h_z = 0$ and determining the linear response of the system to h_x in the limit $h_x \to 0$, while $\chi_\parallel$ is determined by letting $h_x = 0$ first. For $h_x = h_z = 0$, the partition function is given by ($\beta = 1/kT$, T stands for temperature)

$$Z_N = \mathrm{Tr}_{\{\sigma_i^z\}} \left[\exp\left(\beta J \sum_{i=1}^{N} \sigma_i^z \sigma_{i+1}^z + \beta h \sum_{i=1}^{N} \sigma_i^z \right) \right]$$
$$= \mathrm{Tr}_{\{\sigma_i^z\}} \left[A_1^{(\beta)}(\sigma_1^z, \sigma_2^z) \cdots A_N^{(\beta)}(\sigma_N^z, \sigma_1^z) \right]. \tag{3}$$

In Ref. 1, it is shown that the trace over $\{\sigma_i^z\}$ can be converted into a trace on only the three Pauli matrices, τ_x, τ_y, and τ_z. For purposes of introducing the notation used in this paper, this partition function will be determined. Write $A_i^{(\beta)}$ in the τ representation, [1]

$$A_i^{(\beta)}(\sigma_i^z, \sigma_{i+1}^z) \Rightarrow A^T(\beta) = a + \boldsymbol{r} \cdot \boldsymbol{\tau} = \begin{pmatrix} e^{\beta J + \beta h} & e^{-\beta J + \beta h} \\ e^{-\beta J - \beta h} & e^{+\beta J - \beta h} \end{pmatrix} \tag{4}$$

where τ are the Pauli matrices and

$$a = e^{\beta J} \cosh \beta h, \tag{5a}$$

$$\boldsymbol{r} = (e^{-\beta J} \cosh \beta h, i e^{-\beta J} \sinh \beta h, e^{\beta J} \sinh \beta h). \tag{5b}$$

The partition function can then be expressed as

$$Z_N = \mathrm{Tr}_\tau \{(A^\tau)^N\} = \mathrm{Tr}_\tau \{ (a + \boldsymbol{r}\cdot\tau)^N \}.$$

It can easily be shown that

$$(a + \boldsymbol{r}\cdot\tau)^N = \mathfrak{A}^{(N)} + \mathfrak{R}^{(N)}\, \boldsymbol{r}\cdot\tau/|r|, \tag{6a}$$

where

$$\mathfrak{A}^{(n)} = (\lambda_1^N + \lambda_2^N)/2, \qquad \mathfrak{R}^{(n)} = (\lambda_1^N - \lambda_2^N)/2, \tag{6b}$$

and therefore

$$Z_N = \lambda_1^N + \lambda_2^N. \tag{7}$$

Here $\lambda_{1,2}$ are the eigenvalues, $(a \pm |\boldsymbol{r}|)$, of A^τ. This agrees with the standard results.[8]

A. Perpendicular susceptibility

By means of standard perturbation techniques, in the limit $h_x \to 0$ we have the expression

$$\chi_\perp(J,H) = \frac{g^2\mu_B^2}{N} \sum_{i,j=1}^{N} \int_o^\beta dy\, [\langle \sigma_i^x(y)\, \sigma_i^x(o)\rangle - \langle\sigma_i^x\rangle\langle\sigma_i^x\rangle] \tag{8}$$

for the perpendicular susceptibility.

Here

$$\sigma_i^x(y) = e^{y\mathcal{H}_1}\sigma_i^x(o)\, e^{-y\mathcal{H}_1} \tag{9}$$

and $\langle\cdots\rangle$ is the usual thermodynamic average. But $\langle\sigma_i^x\rangle$ and $\langle\sigma_i^x\sigma_j^x\rangle$ $(i \neq j)$ are obviously zero for the Hamiltonian considered. Hence,

$$\chi_\perp(J,H) = \frac{g^2\mu_B^2}{N} \sum_{i=1}^{N} \int_o^\beta dy\, \langle e^{y\mathcal{H}_1}\sigma_i^x(o)\, e^{-y\mathcal{H}_1}\sigma_i^x(o)\rangle. \tag{10}$$

Now

$$\langle\sigma_i^x(y)\sigma_i^x(0)\rangle = \frac{\mathrm{Tr}\{e^{-\beta\mathcal{H}_1+y\mathcal{H}_1}\sigma_i^x(0)\, e^{-y\mathcal{H}_1}\sigma_i^x(0)\}}{\mathrm{Tr}\{e^{-\beta\mathcal{H}_1}\}} = \frac{\mathrm{Tr}_{\{\sigma_1^z\}}\{A_1^{(\beta-y)}(\sigma_1^z,\sigma_2^z)\cdots A_N^{(\beta-y)}(\sigma_N^z,\sigma_1^z)\sigma_i^x A_1^{(y)}(\sigma_1^z,\sigma_2^z)\cdots A_N^{(y)}(\sigma_N^z,\sigma_1^z)\sigma_i^x\}}{\mathrm{Tr}_{\{\sigma_i^z\}}\{A_1^{(\beta)}\cdots A_N^{(\beta)}\}}. \tag{11}$$

Since σ_i^x commutes with $\sigma_j^x (i \neq j)$, $\sigma_i^{x2} = 1$ and using the cyclic property of the trace, one can show that

$$\begin{aligned}&\mathrm{Tr}_{\{\sigma_i^z\}}\{A_1^{(\beta-y)}\cdots A_{i-2}^{(\beta-y)}\sigma_i^x A_{i-1}^{(\beta-y)}\sigma_i^x\sigma_i^x A_i^{(\beta-y)}\sigma_i^x A_{i+1}^{(\beta-y)} \\ &\quad\times\cdots A_N^{(\beta-y)}A_1^{(y)}\cdots A_N^{(y)}\}/\mathrm{Tr}_{\{\sigma_i^z\}}\{(A^{(\beta)})^N\} \\ &= \mathrm{Tr}_{\{\sigma_i^z\}}\{A_1^{(\beta)}A_2^{(\beta)}\cdots A_{i-2}^{(\beta)}(A_{i-1}^{(y)}\sigma_i^x A_{i-1}^{(\beta-y)}\sigma_i^x) \\ &\quad\times(A_i^{(y)}\sigma_i^x A_i^{(\beta-y)}\sigma_i^x)A_{i+1}^{(\beta)}\cdots A_N^{(\beta)}\}/\mathrm{Tr}_{\{\sigma_i^z\}}\{(A^{(\beta)})^N\}.\end{aligned} \tag{12}$$

If one uses the fact that

$$A_i^{(y)}(\sigma_i^z,\sigma_{i+1}^z) = \cosh yJ\,\cosh yh + \sigma_i^z\cosh yJ\,\sinh yh + \sigma_{i+1}^z\sinh yJ\,\sinh yh + \sigma_i^z\sigma_{i+1}^z\sinh yJ\,\cosh yh$$

and defines

$$B_{i-1}(\beta;y) = A_{i-1}^{(y)}\sigma_i^x A_{i-1}^{(\beta-y)}\sigma_i^x, \tag{13a}$$

$$B_i(\beta;y) = A_i^{(y)}\sigma_i^x A_i^{(\beta-y)}\sigma_i^x, \tag{13b}$$

then in the τ representation as shown in Ref. 1, we obtain

$$\langle\sigma_i^x(y)\sigma_i^x(0)\rangle = \mathrm{Tr}_\tau\{[A^\tau(\beta)]^{N-2}B_{i-1}^\tau(\beta;y)B_i^\tau(\beta;y)\}/\times\mathrm{Tr}_\tau\{[A^\tau(\beta)]^N\}. \tag{14}$$

If one now notes that only B_{i-1}^τ and B_i^τ depend on y, the integral can easily be done as follows

$$C_i^\tau(\beta) = \int_0^\beta dy\, B_{i-1}^\tau(\beta;y)\, B_i^\tau(\beta;y)$$

which in the τ representation is

$$C^\tau(\beta) = b + s\cdot\tau, \tag{15}$$

where

$$b = e^{\beta h}\frac{\sinh\beta(2J+h)}{2J+h} + e^{-\beta h}\frac{\sinh\beta(2J-h)}{2J-h}, \tag{16a}$$

$$s = \left(2\frac{\sinh\beta h\,\cosh\beta h}{h}, 2i\frac{\sinh^2\beta h}{h}, e^{\beta h}\frac{\sinh\beta(2J+h)}{2J+h} - e^{-\beta h}\frac{\sinh\beta(2J-h)}{2J-h}\right) \tag{16b}$$

and

$$\int_0^\beta \langle\sigma_i^x(y)\sigma_i^x(0)\rangle\, dy = \mathrm{Tr}_\tau\{[A^\tau(\beta)]^{N-2}C^\tau(\beta)\}/\mathrm{Tr}_\tau\{[A^\tau(\beta)]^N\} = 2\,(\mathfrak{A}^{(N-2)}b + \mathfrak{R}^{(N-2)}\boldsymbol{r}\cdot s/|r||s|)/(\lambda_1^N + \lambda_2^N). \tag{17}$$

Therefore

$$\begin{aligned}\chi_\perp(J,H) &= \frac{g^2{\mu_B}^2}{\lambda_1^N+\lambda_2^N}\left[(\lambda_1^{N-2}+\lambda_2^{N-2})\left(e^{\beta h}\frac{\sinh\beta(2J+h)}{2J+h} + e^{-\beta h}\frac{\sinh\beta(2J-h)}{2J-h}\right) + \frac{(\lambda_1^{N-2}-\lambda_2^{N-2})e^{\beta J}\sinh\beta h}{(e^{-2\beta J}+e^{2\beta J}\sinh^2\beta h)^{1/2}}\right. \\ &\quad\left.\times\left(e^{\beta h}\frac{\sinh\beta(2J+h)}{2J+h} - e^{-\beta h}\frac{\sinh\beta(2J-h)}{2J-h} + \frac{2e^{-2\beta J}}{h}\right)\right].\end{aligned} \tag{18}$$

In the thermodynamic limit $(N \to \infty)$, we obtain

$$\begin{aligned}\chi_\perp(J,H) &\underset{N\to\infty}{=} \frac{g^2\mu_B^2}{\lambda_1^2}\left[\left(e^{\beta H}\frac{\sinh\beta(2J+h)}{2J+h} + e^{-\beta H}\frac{\sinh\beta(2J-h)}{2J-h}\right) + \frac{e^{\beta J}\sinh\beta h}{(e^{-\beta J}+e^{2\beta J}\sinh^2\beta h)^{1/2}}\left(e^{\beta H}\frac{\sinh\beta(2J+h)}{2J+h}\right.\right. \\ &\quad\left.\left. - e^{-\beta h}\frac{\sinh\beta(2J-h)}{2J-h} + \frac{2e^{-2\beta J}}{h}\right)\right].\end{aligned} \tag{19}$$

This reduces to Fisher's[3] formula in the case $H = 0$,

$$\chi_\perp(J,H=0) \underset{N\to\infty}{=} \frac{g^2\mu_B^2}{2J}(\tanh\beta J + \beta J\,\mathrm{sech}^2\beta J). \tag{20}$$

Figure 1 shows results for $\chi_\perp(J,H)$ as a function of temperature for various values of magnetic field in thermodynamic limit. Note that as the magnetic field increases, nothing of interest happens for the ferromagnetic interaction but, for the antiferromagnetic nn

coupling, a critical value is reached namely $H = 2|J|$, for which $\chi_\perp$ has an essential singularity at $T = 0$. This curve is also an upper bound for all other $\chi_\perp$ for different values of H. For $H > 2|J|$, $\chi_\perp$ decreases until the magnetic is so strong, the nn interactions is negligible and $\chi_\perp$ becomes independent of the sign of J.

B. Parallel susceptibility

Because of the cyclic properties of a closed chain, it is obvious that the magnetization is independent of position and

$$\langle \sigma_i^z \rangle = \mathrm{Tr}_{\{\sigma_i^z\}}\{\sigma_i^z A_1(\sigma_1^z, \sigma_2^z)\cdots A_N(\sigma_1^z, \sigma_2^z)\}/Z_N = 2\Re^{(N)} r_z/|\boldsymbol{r}|(\lambda_1^N + \lambda_2^N) \ . \tag{21}$$

Hence the magnetization is given by

$$\frac{M}{M_{\mathrm{sat}}} = \left(\frac{\lambda_1^N - \lambda_2^N}{\lambda_1^N + \lambda_2^N}\right)\frac{r_z}{|\boldsymbol{r}|} \underset{N\to\infty}{\to} \frac{r_z}{|\boldsymbol{r}|} . \tag{22}$$

The parallel susceptibility is given by

$$\chi_\parallel(J,H) = \beta\frac{g^2\mu_B^2}{N}\sum_{i,j=1}^{N}(\langle\sigma_i^z\sigma_j^z\rangle - \langle\sigma_i^z\rangle\langle\sigma_j^z\rangle) = \frac{\beta g^2\mu_B^2}{N}\left(\sum_{i=1}^{N}(1 - \langle\sigma_i^z\rangle^2) + 2\sum_{i=2}^{N}\sum_{j=1}^{i-1}(\langle\sigma_j^z\sigma_i^z\rangle - \langle\sigma_i^z\rangle\langle\sigma_j^z\rangle)\right). \tag{23}$$

We now require $(i > j)$ and we express the result in $\boldsymbol{r}$ representation following Ref. 1.; using the cyclic property of the trace, one has

$$\langle\sigma_j^z\sigma_i^z\rangle = [1/2(\lambda_1^N + \lambda_2^N)]\{[\lambda_1^N + \lambda_2^N + \lambda_1^N(\lambda_2/\lambda_1)^{i-j} + \lambda_2^N(\lambda_1/\lambda_2)^{i-j}] + [\lambda_1^N + \lambda_2^N - \lambda_1^N(\lambda_2/\lambda_1)^{i-j} - \lambda_2^N(\lambda_1/\lambda_2)^{i-j}](\mathbf{r}\cdot\mathbf{r}'/|\boldsymbol{r}|^2)\} \tag{24}$$

where $\boldsymbol{r}' = (-r_x, -r_y, r_z)$.

By doing the indicated sums in (2) it can be shown that

$$\chi_\parallel(J,H) = \frac{\beta g^2\mu_B^2}{N}\left[N\left(1 - r_z^2/|\boldsymbol{r}|^2\right)\left(\frac{(\lambda_1 + \lambda_2)(\lambda_1^N - \lambda_2^N)}{(\lambda_1 - \lambda_2)(\lambda_1^N + \lambda_2^N)}\right) - N^2\frac{r_z^2}{|\boldsymbol{r}|^2}\left(\frac{(\lambda_1^N - \lambda_2^N)^2}{(\lambda_1^N + \lambda_2^N)^2} - 1\right)\right] . \tag{25}$$

In thermodynamic limit, we obtain

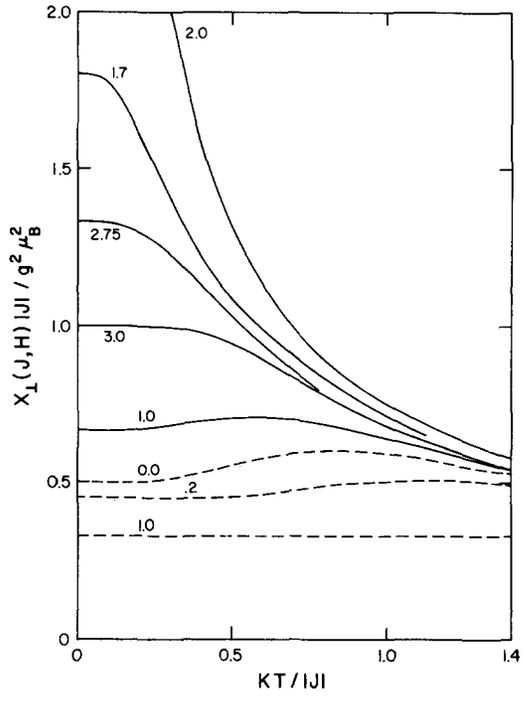
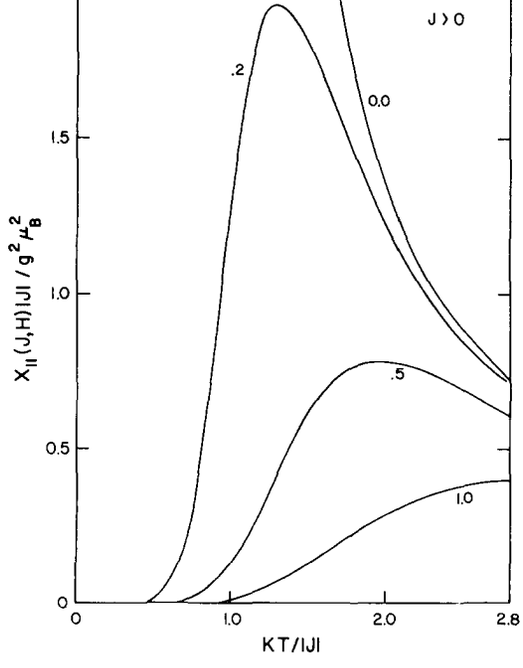


FIG. 1: Perpendicular susceptibility as a function of temperature for various values of $|H/J|$ $(N \to \infty)$. Solid curves are for $J < 0$ and dashed curves for $J > 0$. $\chi_\perp(J,H)$ is finite as $T \to 0$ for all values of $|H/J|$ except 2.00, $J < 0$. For $H = 0$, $\chi_\perp$ is independent of the sign of J.



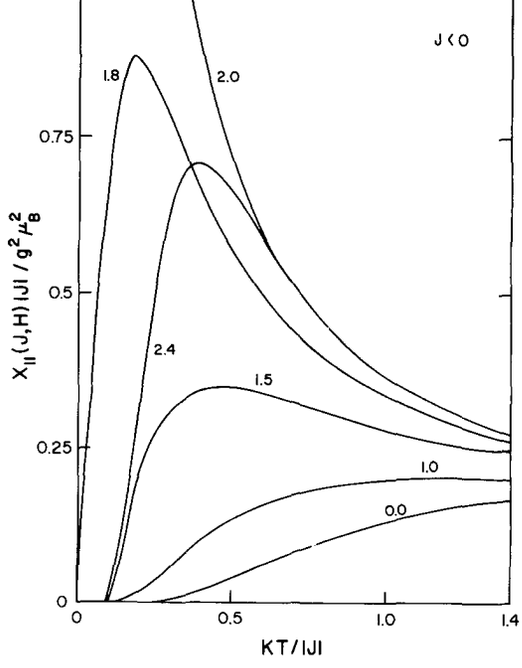


FIG. 2: Parallel susceptibility as a function of temperature for various values of $|H/J|$ $(N \to \infty)$.

$$\chi_{\parallel}(J,H) \underset{N\to\infty}{=} \frac{\beta g^2\mu_B^2}{N}\left(1-\gamma_z^2/|\boldsymbol{\gamma}|^2\right)\left(\frac{\lambda_1+\lambda_2}{\lambda_1-\lambda_2}\right). \quad (26)$$

Figures 2(a) and 2(b) show $\chi_{\parallel}(J,H)$ in the thermodynamic limit as a function of temperature for various values of H. Figure 2(a) is for $J>0$, which shows that for all $H\neq 0$, $\chi_{\parallel}\to 0$ as $T\to 0$. For the case $J<0, H=2|J|$ is the critical value of magnetic field for which $\chi_{\parallel}\to\infty$ as $T\to 0$.

III. NEXT NEAREST NEIGHBOR INTERACTION

Consider now the nnn Hamiltonian for an open chain:

$$\mathcal{H}_2 = -J_1\sum_{i=1}^{N-1}\sigma_i^z\sigma_{i+1}^z - \alpha J_1\sum_{i=1}^{N-2}\sigma_i^z\sigma_{i+2}^z. \quad (27)$$

We consider the open chain because Dobson's trick can be successfully applied here. The partition function is given by

$$Z_N = \sum_{\sigma_i=\pm 1}\cdots\sum_{\sigma_N=\pm 1}\exp\left(\beta J_1\sum_{i=1}^{N-1}\sigma_i^z\sigma_{i+1}^z + \beta\alpha J_1\sum_{i=1}^{N-2}\sigma_i^z\sigma_{i+2}^z\right). \quad (28)$$

By the use of the transformation, [5]

$$t_0=\sigma_1^z,\quad t_i=\sigma_i^z\sigma_{i+1}^z,\quad i=1,2,\cdots,N-1,$$

we can express $\sigma_i^z = t_{i-1}t_{i-2}\cdots t_1t_0$, so that

$$\sigma_i^z\sigma_{i+1}^z = t_i,\quad \sigma_i^z\sigma_{i+2}^z = t_it_{i+1}$$

and so the partition function is given by

$$Z_N(J_1,\alpha) = \sum_{t_0=\pm 1}\cdots\sum_{t_{N-1}=\pm 1}\exp\left(\beta J_1\sum_{i=1}^{N}t_i + \alpha\beta J_1\sum_{i=1}^{N}t_it_{i+1}\right). \quad (29)$$

Hence [5]

$$Z_N(J_1,\alpha) = 2\ \bar{Z}_{N-1}(J=\alpha J_1, H=J_1). \quad (30)$$

where $\bar{Z}_{N-1}(J,H)$ is the partition for $(N-1)$ particles with nn interaction J in magnetic field H. The factor 2 in (30) appears because of sum over t_0. Note that this can be done only for open chains because the transformation leads to an extra constraint in the case of closed chain which is not easily incorporated in such a calculation. [1] Since we are interested only in the thermodynamic limit, this is really not important.

A. Perpendicular susceptibility

As before,

$$\chi_{\perp}(J_1,\alpha) = \frac{g^2\mu_B^2}{N}\sum_{i=1}^{N}\int_0^{\beta}dy\ \langle\sigma_i^x(y)\sigma_i^x(0)\rangle. \quad (31)$$

Now,

$$\langle\sigma_i^x(y)\sigma_i^x(0)\rangle = \mathrm{Tr}_{\{\sigma_i^z\}}\{A_1^{(\beta-y)}\cdots A_{N-1}^{(\beta-y)}\sigma_i^xA_1^{(y)}\times\cdots A_{N-1}^{(y)}\sigma_i^x\}/\mathrm{Tr}_{\{\sigma_i^z\}}\{A_1^{(\beta)}\cdots A_{N-1}^{(\beta)}\} \quad (32)$$

where

$$A_i^{(\beta)} \equiv \exp(\beta J_1\sigma_i^z\sigma_{i+1}^z + \beta\alpha J_1\sigma_i^z\sigma_{i+2}^z)$$

and

$$A_{N-1}^{(\beta)} \equiv \exp(\beta J\sigma_{N-1}^z\sigma_N^z).$$

By applying the same type of tricks as in case (i) and remembering that $A_i^{(\beta)}$ depends on σ_i^z, σ_{i+1}^z and σ_{i+2}^z, one finds that

$$\langle\sigma_i^x(y)\sigma_i^x(0)\rangle = \mathrm{Tr}_{\{\sigma_i^z\}}\{A_1^{(\beta)}\cdots A_{i-3}^{(\beta)}B_{1i}(\beta;y)A_{i+1}^{(\beta)}\cdots A_{N-1}^{(\beta)}\}/\mathrm{Tr}_{\{\sigma_i^z\}}\{A_1^{(\beta)}\cdots A_{N-1}^{(\beta)}\}, \quad (33)$$

where we have introduced

$$B_{1i}(\beta;y) = (A_{i-2}^{(y)}\sigma_i^xA_{i-2}^{(\beta-y)}\sigma_i^x)(A_{i-1}^{(y)}\sigma_i^xA_{i-1}^{(\beta-y)}\sigma_i^x)\times(A_i^{(y)}\sigma_i^xA_i^{(\beta-y)}\sigma_i^x). \quad (34)$$

Let

$$C_{1i}(\beta) = \int_0^{\beta}dy\,B_{1i}(\beta;y). \quad (35)$$

Then one can write going over to the τ representation

$$\chi_{\perp}(J_1,\alpha) = \frac{g^2\mu_B^2}{N}\times\sum_{i=1}^{N}\frac{\mathrm{Tr}_{\tau}\{[A_1^{\tau}(\beta)]^{i-3}C_{1i}^{\tau}(\beta)[A_1(\beta)]^{N-i-2}A_{N-1}(\beta)\}}{\mathrm{Tr}_{\tau}\{A_1^{\tau}(\beta)]^{N-2}A_{N-1}^{\tau}(\beta)\}}. \quad (36)$$

In the thermodynamic limit,

$$\chi_{\perp}(J_1,\alpha) \underset{N\to\infty}{=} g^2\mu_B^2\ \frac{\mathrm{Tr}_{\tau}\{(A_1^{\tau})^{N-3}C_1^{\tau}(\beta)}{\mathrm{Tr}_{\tau}\{(A_1^{\tau})^N\}}. \quad (37)$$

Writing

$$A_1^{\tau} = \mathfrak{A}_1 + \boldsymbol{\gamma}_1\cdot\boldsymbol{\tau},\qquad \mathfrak{A}_1 = e^{\beta\alpha J_1}\cosh\beta J_1, \quad (38a)$$

$$\boldsymbol{\gamma}_1 = (e^{-\beta\alpha J_1}\cosh\beta J_1, ie^{-\beta\alpha J_1}\sinh\beta J_1, e^{\beta\alpha J_1}\sinh\beta J_1), \quad (38b)$$

and

$$C_1^{\tau} = 2(b_1+\mathbf{s}_1\cdot\boldsymbol{\tau}), \quad (38c)$$

where

$$b_1 = \beta e^{-\beta\alpha J_1}\cosh\beta J_1 + \frac{1}{2}e^{\beta\alpha J_1}\times\left(e^{\beta J_1}\frac{\sinh2\beta(J_1+\alpha J_1)}{2(J_1+\alpha J_1)} + e^{-\beta J}\frac{\sinh2\beta(J_1-\alpha J_1)}{2(J_1-\alpha J_1)}\right), \quad (39a)$$

$$s_{1x} = \left(e^{\beta\alpha J_1}\frac{\sinh2\beta J_1}{2J_1} + e^{-\beta\alpha J_1}\frac{\sinh2\beta\alpha J_1}{2\beta\alpha J_1}\right)\cosh\beta J_1, \quad (39b)$$

$$s_{1y} = i\left(e^{\beta\alpha J_1}\frac{\sinh2\beta J_1}{2J_1} + e^{-\beta\alpha J_1}\frac{\sinh2\beta\alpha J_1}{2\beta\alpha J_1}\right)\sinh\beta J_1, \quad (39c)$$

and

$$s_{1z} = \beta e^{-\beta\alpha J_1}\sinh\beta J_1 + \frac{e^{\beta\alpha J_1}}{2}\times\left(e^{\beta J_1}\frac{\sinh2\beta(J_1+\alpha J_1)}{2(J_1+\alpha J_1)} - e^{-\beta J_1}\frac{\sinh2\beta(J_1-\alpha J_1)}{2(J_1-\alpha J_1)}\right), \quad (39d)$$

we have

$$\chi_{\perp}(J_1,\alpha) \underset{N\to\infty}{=} (g^2\mu_B^2/\lambda_1^3)\,2\,(b_1+\boldsymbol{\gamma}_1\cdot\mathbf{s}_1/|\boldsymbol{\gamma}_1|). \quad (40)$$

Here

$$\lambda_1 = \mathfrak{A}_1 + |\boldsymbol{\gamma}_1| = e^{\alpha\beta J_1}\cosh\beta J_1 + \left(e^{-2\beta\alpha J_1} + e^{2\beta\alpha J_1}\sinh^2\beta J_1\right)^{1/2}.$$

Further, this reduces to Fisher's result for $\alpha=0$.

Figure 3 shows $\chi_{\perp}(J,\alpha)$ as a function of temperature for various values of α. Note that $\chi_{\perp}$ is independent of the sign of nn interaction, and depends only on the sign

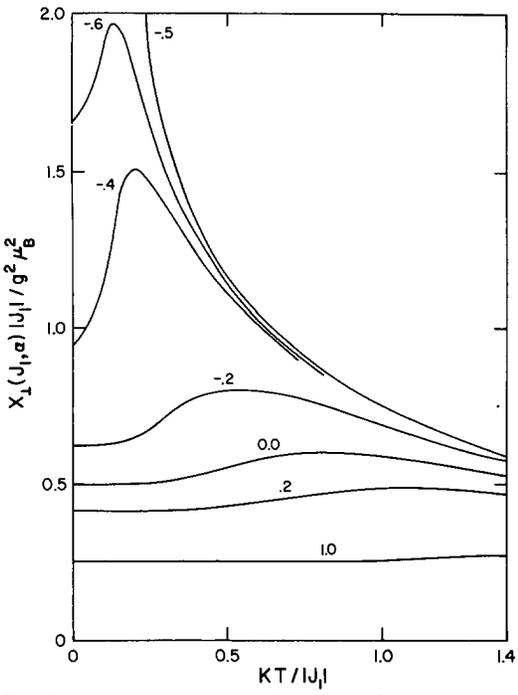


FIG. 3: Perpendicular susceptibility as a function of temperature for various values of $\alpha J_1/|J_1|$. $\chi_\perp(J_1, \alpha)$ is independent of ths sign of J_1 and is finite as $T \to 0$ for $\alpha J_1/|J_1| \neq -0.5$.

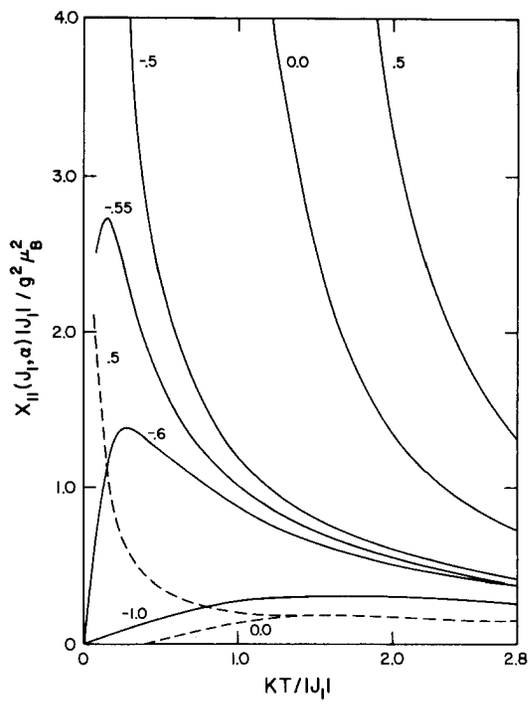


FIG. 4: Parallel susceptibility as a function of temperature for various values of α. Solid curves are for $J_1 > 0$ and dashed curves for $J_1 < 0$.

of α. The value $\alpha J_1/|J_1| = -0.5$ is special and can be associated with the corresponding special value of the magnetic field in case (i), $H = 2|J|$, $J < 0$. This curve also forms an upper bound on $\chi_\perp$ for all other values of α. In this respect, this feature is similar to case (i).

$$\langle \sigma_j^z \sigma_i^z \rangle = \langle \sigma_j^z \sigma_{j+1}^z \sigma_{j+1}^z \cdots \sigma_{i-1}^z \sigma_{i-1}^z \sigma_i^z \rangle = \langle t_j t_{j+1} \cdots t_{i-1} \rangle = \mathrm{Tr}_{\{t_i\}} \frac{\{A_1^{(\beta)} \cdots A_{j-1}^{(\beta)} (t_j A_j^{(\beta)}) \cdots (t_{i-1} A_{i-1}^{(\beta)}) A_i^{(\beta)} \cdots A_{N-1}^{(\beta)}\}}{\mathrm{Tr}_{\{t_i\}} \{A_1^{(\beta)} \cdots A_{N-1}^{(\beta)}\}}. \quad (42)$$

Transforming to the τ representation, [1] in the thermodynamic limit, one gets

$$\langle \sigma_j^z \sigma_i^z \rangle = \tfrac{1}{2} \{(\lambda_1/\mu_1)^{i-j} (1 + \mathfrak{A}_1 r_{1z}/|\boldsymbol{r}_1 \| \boldsymbol{r}_1'|) + (\lambda_1/\mu_2)^{i-j} (1 - \mathfrak{A}_1 r_{1z}/|\boldsymbol{r} \| \boldsymbol{r}_1|)\}, \quad (43)$$

where $\mu_{1,2} = \mathfrak{A}_1' \pm |\boldsymbol{r}_1'|$, $\mathfrak{A}_1' = r_{1z}, \boldsymbol{r}_1' = (-ir_{1y}, ir_{1x}, \mathfrak{A}_1)$. Finally,

$$\chi_\parallel(J_1, \alpha) \underset{N\to\infty}{=} \beta g^2 \mu_B^2 \left[1 + (1 + \mathfrak{A}_1 r_{1z}/|\boldsymbol{r}_1 \| \boldsymbol{r}_1'|) \frac{\mu_1/\lambda_1}{(1 - \mu_1/\lambda_1)} + \left(1 - \frac{\mathfrak{A}_1 r_{1z}}{|\boldsymbol{r}_1 \| \boldsymbol{r}_1'|}\right) \frac{\mu_2/\lambda_1}{(1 - \mu_2/\lambda_1)}\right]. \quad (44)$$

If one lets $\alpha = 0$, $\chi_\parallel$ goes to the known limit

$$\chi_\parallel(J_1, \alpha = 0) \underset{N\to\infty}{=} \beta g^2 \mu_B^2 \frac{(1 + \tanh\beta J_1)}{(1 - \tanh\beta J_1)} \quad (45)$$

Figure 4 shows $\chi_\parallel$ as a function of temperature for various values of α. For $J > 0$, ferromagnetic nnn coupling only causes $\chi_\parallel$ to diverge faster as $T \to 0$. If $\alpha < 0$ and $J > 0$, $\chi_\parallel$ diverges at $T \to 0$ for all $\alpha > -0.5$, but, for values of α smaller than -0.5, $\chi_\parallel$ has a maximum and goes to zero as T approaches zero. For $J < 0$, $\alpha = 0$, $\chi_\parallel$ does not diverge, and in fact $\chi_\parallel$ diverges as $T \to 0$ for only one value of α, equal to 0.5.

IV. CONCLUSION

By means of the formalism developed in Ref. 1 for a random Ising system in one dimension, we derive closed formulas for $\chi_\perp$ and $\chi_\parallel$ for two one-dimensional systems. It has been noted elsewhere[5] that the partition function for the nnn interaction in a zero magnetic field is related to the nn interaction system with a nonzero magnetic field. It should be clear from the calculated forms of $\chi_\perp$ and $\chi_\parallel$ that there is no apparent direct relationship between higher spin correlation functions. Correspondingly, Figs. 1–4 adequately bring out this aspect in a transparent way. Moreover, the behavior of $\chi_\parallel$ for $J > 0$ as a function of the magnetic field displayed in Fig. 2 in no way resembles the corresponding figure, Fig. 4, showing $\chi_\parallel$ as a function of α. The striking behavior of $\chi_\parallel$ and $\chi_\perp$ for $H = 2J$, $J < 0$ remains for $\alpha = \frac{1}{2}$ in the corresponding expressions in the second model.

B. Parallel susceptibility

One has by transformation of spin variables

$$\langle \sigma_i^z \rangle = \mathrm{Tr}_{\{t_i\}} \{(t_{i-1} t_{i-2} \cdots t_1 t_0) e^{-\beta \mathcal{H}_2}\}/Z_N = 0, \quad (41)$$

because the sum over t_0 gives zero. And

*NDEA Fellow.

[1] A. K. Rajagopal and G. S. Grest, J. Math. Phys. 15, 583 (1974).
[2] M. Blume, M. Lurie, and P. Heller, in preparation (1973) (private communication); M. Blume and A. K. Rajagopal, unpublished work (1972); A. K. Rajagopal, Proc. Ind. Acad. Sci. (A) 78, 13 (1973).
[3] M. E. Fisher, J. Math. Phys. 4, 124 (1963).
[4] J. S. Marsh, Phys. Rev. 145, 251 (1966).
[5] J. F. Dobson, J. Math. Phys. 10, 40 (1969).
[6] J. Stephenson, Can. J. Phys. 48, 1724 (1970).
[7] C. K. Majumdar, J. Math. Phys. 10, 177 (1969); C. K. Majumdar J. Phys. C: Solid State Phys. 3, 911 (1970); C. K. Majumdar, K. Krishan and V. Mubayi, J. Phys. C: Solid State Phys. 5, 2896 (1972).
[8] C. J. Thompson, *Mathematical Statistical Mechanics* (MacMillan, New York, 1972).

The general solution to Einstein–Maxwell equations with plane symmetry

P. S. Letelier*

Instituto de Física, Universidad Católica de Chile, Casilla 114-D, Santiago, Chile

R. R. Tabensky

Departamento de Física, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile
(Received 26 June 1973)

The general solution to Einstein–Maxwell equations with plane symmetry is obtained. This solution has an extra Killing vector ξ^a not assumed *a priori*. The solution may be classified in two classes depending on the sign of an integration constant. The first class depends only on the time; ξ^a is spacelike. In the second class ξ^a is spacelike, lightlike, and timelike in different regions like Schwarzschild's metric.

1. INTRODUCTION

Einstein–Maxwell field equations in the absence of charge are given by[1]

$$G_{ab} = -kT_{ab}, \tag{1}$$

$$T_{ab} = F_a{}^c F_{cb} + \tfrac{1}{4} g_{ab} F_{cd} F^{cd}, \tag{2}$$

$$F^{ab}{}_{;b} = 0, \tag{3}$$

$$F_{[ab;c]} = 0, \tag{4}$$

where all symbols have the usual meaning.

The line element with plane symmetry can be written in the form[2]

$$ds^2 = e^{w(u,v)} du dv - e^{\mu(u,v)}(dx^2 + dy^2). \tag{5}$$

When the plane symmetry condition is used, it is found that the only nonvanishing components of F_{ab} are given by[3]

$$F_{23} = C_1, \qquad F_{01} = \tfrac{1}{2} C_2 e^{w-\mu}, \tag{6}$$

where $x^0 = u, x^1 = v, x^2 = x, x^3 = y$ and C_1, C_2 are arbitrary constants. The nonvanishing components of the energy–momentum tensor are

$$T_0{}^0 = T_1{}^1 = -T_2{}^2 = -T_3{}^3 = \tfrac{1}{2}(C_1{}^2 + C_2{}^2) e^{-2\mu}). \tag{7}$$

From (1), (5), and (7) it is found that

$$\mu,_{00} + \tfrac{1}{2}\mu,_0^2 - w,_0 \mu,_0 = 0, \tag{8}$$

$$\mu,_{11} + \tfrac{1}{2}\mu,_1^2 - w,_1 \mu,_1 = 0, \tag{9}$$

$$\mu,_{01} + w,_{01} + \tfrac{1}{2}\mu,_0 \mu,_1 = \tfrac{1}{4} k e^{w-2\mu}(C_1{}^2 + C_2{}^2), \tag{10}$$

$$\mu,_{01} + \mu,_0 \mu,_1 = -\tfrac{1}{4} k e^{w-2\mu}(C_1{}^2 + C_2{}^2), \tag{11}$$

$$\mu,_{01} + \tfrac{3}{4}\mu,_0 \mu,_1 + \tfrac{1}{2} w,_{01} = 0. \tag{12}$$

Equation (12) expresses the fact that the Ricci scalar vanishes. It is the purpose of this paper to find and to discuss some aspects of the general solution to equations (8) through (12).

In Sec. 2 we solve the differential equations. In Sec. 3 we use the Hawking–Penrose energy condition[4] and we show that the solution is physically reasonable. In Sec. 4 we exhibit two classes of solutions depending on the sign of an integration constant. The first class is a homogeneous solution that depends only on time. The second class has a singularity that separates two regions. In the first region the solution is homogeneous and in the second it is a static one. The static solution was found by Patnaik.[3] In Sec. 5 we show that Einstein–Maxwell equations imply an extra Killing vector ξ^a not assumed *a priori*. Finally in Sec. 6 we describe a general similarity with the Schwarzschild solution.

2. THE SOLUTION

From (8) and (9) we get

$$w - \tfrac{1}{2}\mu = \ln[\mu,_0 G,_1(v)], \tag{13a}$$

$$w - \tfrac{1}{2}\mu = \ln[\mu,_1 F,_0(u)], \tag{13b}$$

where F and G are arbitrary functions of their arguments. Hence, from (13)

$$\mu,_0 G,_1 = \mu,_1 F,_0.$$

This condition tells us that μ is a function of u and v of the form

$$\mu = \mu[F(u) + G(v)]. \tag{14}$$

So, from (14) and (13) we get

$$e^w = e^{\mu/2} \mu' F,_0 G,_1, \tag{15}$$

where $(\,)'$ means differentation with respect to $t = F(u) + G(v)$. When we put (15) and (14) into equations (11), (10), and (12), we get

$$\mu'' + \mu'^2 = -\tfrac{1}{4} k \mu' e^{-3\mu/2}(C_1{}^2 + C_2{}^2), \tag{16}$$

$$\mu'' + \tfrac{1}{2}\mu'^2 + w'' = \tfrac{1}{4} k \mu' e^{-3\mu/2}(C_1{}^2 + C_2{}^2), \tag{17}$$

$$\mu'' + \tfrac{3}{4}\mu'^2 = -\tfrac{1}{2} w''. \tag{18}$$

From (15) and (16) we have

$$w'' = -\tfrac{1}{2}\mu'' + \tfrac{3}{8} k \mu' e^{-3\mu/2}(C_1{}^2 + C_2{}^2) \tag{19}$$

and from (19) and (16) follows (17) and (18). So we only need solve (16) which may be cast into the form

$$(e^{\mu})'' = \tfrac{1}{2} k (C_1{}^2 + C_2{}^2)(e^{-\mu/2})',$$

whence

$$2e^{\mu}(e^{\mu/2})' - \tfrac{1}{2} k (C_1{}^2 + C_2{}^2) + C_3 e^{\mu/2} = 0. \tag{20}$$

This ordinary differential equation is solved by

$$(e^{\mu/2} + A)^2 + \ln(e^{\mu/2} - A) = -C_3 t + C_4, \tag{21}$$

where the C_i, $i = 1, \ldots, 4$, are arbitrary constants and $A = \frac{1}{2}k(C_1^2 + C_2^2)/C_3$.

From (15) and (5) we obtain

$$ds^2 = 2(e^{\mu(t)/2})'dFdG - e^{\mu(t)}(dx^2 + dy^2), \tag{22}$$

and letting $z = G - F$, $t = G + F$ (as before) and $\Omega = \frac{1}{2}(e^{\mu(t)/2})'$, we get

$$ds^2 = \Omega(t)(dt^2 - dz^2) - e^{\mu(t)}(dx^2 + dy^2), \tag{23}$$

where $\mu(t)$ is the solution to (21).

3. THE ENERGY CONDITION

The Hawking–Penrose condition tells us in this case that $T_{ab}\eta^a\eta^b \geq 0$ for any timelike vector η^a, i.e., $g_{ab}\eta^a\eta^b > 0$. From (5) and (7) this condition gives us only that e^{μ} is positive. Therefore, the sign of e^{w} is not restricted.

4. THE TWO CLASSES OF SOLUTIONS

It is clear from (23) that t and z are temporal and spatial coordinates respectively if Ω is positive. From (20) it is clear that $C_3 \leq 0$ makes Ω always positive, but when $C_3 = 0$ solution (21) must be changed to

$$e^{\mu} = [\tfrac{3}{2}k(C_1^2 + C_2^2)t + C_4]^{2/3}.$$

Therefore, when $C_3 \leq 0$, we have an homogeneous non-static solution. If $C_3 > 0$, we have two types of regions because from (20) we see that Ω may be positive, negative, or zero depending on whether $e^{\mu/2} < A$, $e^{\mu/2} > A$, or $e^{\mu/2} = A$, respectively. When $e^{\mu/2} < A$ we have that the solution is homogeneous. When $e^{\mu/2} > A$, Ω is negative and in this case t and z represent space and time respectively. We have a static solution which is the Patnaik solution.[3]

5. THE EXTRA KILLING VECTOR

From (23) it is clear that $\xi^a = \delta^a_1$ is a Killing vector, where now $x^0 = t, x^1 = z$. The metric (23) also gives us

$$\xi^a\xi_a = -\tfrac{1}{2}(e^{\mu/2})' = -\Omega.$$

Thus ξ^a is spacelike when $C_3 \leq 0$ or $C_3 > 0$ and $e^{\mu/2} < A$. ξ^a is timelike when $e^{\mu/2} > A$ and $C_3 > 0$ and it is null when $e^{\mu/2} = A$.

6. DISCUSSION

In some general aspects the solution found is similar to Schwarzschild's solution. The Schwarzschild solution has an extra Killing vector implied by the field equations (Birkhoff's theorem[5]). The solution is always static when $m < 0$. When $m > 0$, we also have two regions, the first when $r > 2m$ and the second when $r < 2m$. The Killing vector changes from timelike to spacelike and it is null on $r = 2m$.

ACKNOWLEDGMENT

One of us R. T. would like to acknowledge the financial support extended by Instituto de Física, Universidad Católica.

*Present address: Department of Physics, Boston University.

[1]A. Einstein, *The Meaning of Relativity* (Princeton U. P., Princeton, N. J., 1955).
[2]A. H. Taub, Ann. Math. 53, 472 (1951).
[3]S. Patnaik, Proc. Camb. Phil. Soc. 67, 127 (1970).
[4]S. W. Hawking, and R. Penrose, Proc. Roy. Soc. A 314, 529 (1970).
[5]G. D. Birkhoff, *Relativity and Modern Physics* (Harvard U. P., Cambridge, Mass., 1923).

On the moments of the distribution of widths corresponding to an ensemble of random matrices*

J. F. McDonald

Department of Mathematics, University of Windsor, Windsor, Ontario, Canada
(Received 14 August 1973)

In this work, we discuss the moments of the width distribution corresponding to an ensemble of random matrices. The ensemble considered is one where the real and imaginary parts of the Hamiltonian matrices have Gaussian distributions with different widths. In particular, we give exact expressions for the moments when $N = 2$ ($N =$ dimension of the Hamiltonian submatrix with a fixed set of quantum numbers) and give a method for obtaining a power series representation when N is arbitrary.

1. INTRODUCTION

Of recent interest is the question of how a small time-reversal odd term in the Hamiltonian of a complex system (such as a heavy nucleus) affects the statistical properties of the system's energy spectrum.[1-5] This has led to the study of ensembles of random matrices which are not representationally invariant.[6]

An ensemble which has received much attention is one in which both the real and imaginary parts of the matrix elements have a Gaussian distribution. However, the width of these Gaussians are assumed to be different for the real and imaginary parts of the off diagonal matrix elements.

In this work we consider the moments of the distribution of widths corresponding to such an ensemble of random matrices. In particular, we give below exact expressions for these moments when $N = 2$ ($N =$ dimension of the Hamiltonian submatrix with a fixed set of quantum numbers), and a method for finding a power series expansion for these moments when N is arbitrary.

2. THE ENSEMBLE AND THE DISTRIBUTION OF WIDTHS

The distribution of matrix elements is chosen to be

$$p_N(H, \gamma, \alpha) = \eta \exp(-\alpha \mathrm{Tr} H^2) \exp\left(-2\gamma \sum_{i>j} S_{ij}^2\right), \quad (1)$$

where $\alpha > 0$, $\alpha + \gamma > 0$,

$$\eta = 2^{N(N-1)/2}(\alpha/\pi)^{N(N+1)/4}[(\alpha + \gamma)/\pi]^{N(N-1)/4}, \quad (2)$$

and

$$H_{ij} = R_{ij} + iS_{ij}. \quad (3)$$

Since the Hamiltonian matrix, H, is Hermitian,

$$R_{ij} = R_{ji}, \quad (4)$$

and

$$S_{ij} = -S_{ji}, \quad (5)$$

where R and S are both real matrices. Here, all of the R_{ij}, $i \geq j$, and S_{ij}, $i > j$, are assumed to vary independently between $-\infty$ and $+\infty$. Note that, for $\gamma \to +\infty$, p_N becomes the orthogonal Gaussian distribution and for $\gamma \to 0^+$ it becomes the unitary Gaussian distribution.[7]

If certain assumptions are made, the distribution of widths can be shown to be given by[8]

$$P_N(X, \gamma, \alpha) = \int \delta(X - NA_{11}A_{11}^*) p_N(H, \gamma, \alpha) d\mathbf{H}_N, \quad (6)$$

where

$$X = \Gamma/\langle\Gamma\rangle, \quad (7)$$

Γ being the width and $\langle\Gamma\rangle$ the average width,

$$d\mathbf{H}_N = \prod_{i \geq j} dR_{ij} \prod_{k>l} dS_{kl}, \quad (8)$$

and A is the unitary matrix which diagonalizes H, i.e.,

$$E = A^+ HA, \quad (9)$$

where E is diagonal.

The moments of the distribution of widths are given by

$$\langle X^n\rangle = \int_0^N P_N(X, \gamma, \alpha) X^n dX, \quad (10)$$

where n is a nonnegative integer. It easily follows from (6) and (10) that

$$\langle X^n\rangle = N^n \langle |A_{11}|^{2n}\rangle, \quad (11)$$

where

$$\langle |A_{11}|^{2n}\rangle = \int |A_{11}|^{2n} p_N(H, \gamma, \alpha) d\mathbf{H}_N. \quad (12)$$

It is these moments which shall be discussed in the remainder of this work.

3. THE SPECIAL CASE $N = 2$

When $N = 2$, $|A_{11}|^2$ can be expressed explicitly in terms of the R_{ij} and S_{ij}. In particular,

$$|A_{11}|^2 = \tfrac{1}{2}\{1 + (R_{11} - R_{22})/[(R_{11} - R_{22})^2 + 4(R_{21}^2 + S_{21}^2)]^{1/2}\}. \quad (13)$$

Further, from (1), (2), and (12) it easily follows that

$$\langle |A_{11}|^{2n}\rangle = \frac{2\alpha^{3/2}\hat{\gamma}^{1/2}}{\pi^2} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-[\alpha(R_{11}^2 + R_{22}^2 + 2R_{21}^2) + \hat{\gamma}S_{21}^2]} \times |A_{11}|^{2n} d\mathbf{H}_2, \quad (14)$$

where

$$\hat{\gamma} \equiv \gamma + \alpha. \quad (15)$$

If one lets

$$\xi \equiv R_{11} + R_{22}, \quad -\infty < \xi < +\infty, \quad (16)$$

$$r \equiv [(R_{11} - R_{22})^2 + 4R_{21}^2 + 4S_{21}^2]^{1/2}, \quad r \geq 0, \quad (17)$$

$$\tan\phi \equiv 2R_{21}/(R_{11} - R_{22}), \quad 0 \leq \phi < 2\pi, \quad (18)$$

and

$$\cos\theta \equiv 2S_{21}/r, \quad 0 \leq \theta \leq \pi, \quad (19)$$

Eq. (14) can be rewritten as

$$\langle |A_{11}|^{2n}\rangle = \frac{\alpha^{3/2}\hat{\gamma}^{1/2}}{2^{n-1}\pi^2} \int_{-\infty}^{+\infty} e^{-\alpha\xi^2/2} d\xi \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta\, d\theta \times \int_0^{+\infty} r^2 dr\, e^{-(\alpha\sin^2\theta + \hat{\gamma}\cos^2\theta)r^2/2} (1 + \sin\theta\cos\phi)^n. \quad (20)$$

596

If the integration with respect to ξ and r is carried out explicitly, one obtains

$$\langle|A_{11}|^{2n}\rangle = \frac{\alpha\hat{\gamma}^{1/2}}{\pi 2^{n+2}} \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta\, d\theta\, (1 + \sin\theta\cos\phi)^n \times (\alpha \sin^2\theta + \hat{\gamma}\cos^2\theta)^{-3/2}. \quad (21)$$

From the binomial theorem we have

$$(1 + \sin\theta\cos\phi)^n = \sum_{k=0}^{n} \binom{n}{k} \sin^k\theta \cos^k\phi. \quad (22)$$

Further, it is easily verified by mathematical induction that[9]

$$\int_0^{2\pi} \cos^k\phi d\phi = [1 + (-1)^k]\,\Gamma\left(\frac{k+1}{2}\right)\Big/\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{k+2}{2}\right) = \pi[1 + (-1)^k](k-1)!!/k!!, \quad (23)$$

if k is a nonnegative integer. Thus, from (21), (22), and (23) it follows that

$$\langle|A_{11}|^{2n}\rangle = \frac{\epsilon}{2^{n+1}} \sum_{l=0}^{[n/2]} \binom{n}{2l} \frac{(2l-1)!!}{(2l)!!} \int_0^{\pi} \frac{\sin^{2l+1}\theta\, d\theta}{(1 - \mu\sin^2\theta)^{3/2}}, \quad (24)$$

where $[n/2]$ is the largest integer less than or equal to $n/2$,

$$\epsilon \equiv \alpha/\hat{\gamma}, \quad (25)$$

and

$$\mu \equiv 1 - \epsilon. \quad (26)$$

Note that $\mu < 1$, since $\epsilon > 0$.

Now,

$$\int_0^{\pi} \frac{\sin^{2l+1}\Theta}{(1 - \mu\sin^2\Theta)^{3/2}} = 2\int_0^{\pi/2} \frac{\sin^{2l+1}\Theta}{(1 - \mu\sin^2\Theta)^{3/2}} d\Theta = \int_0^1 \frac{t^l}{\sqrt{1-t}\,(1-\mu t)^{3/2}} dt, \quad (27)$$

where $t \equiv \sin^2\Theta$. Further,

$$\int_0^1 \frac{t^l}{\sqrt{1-t}\,(1-\mu t)^{3/2}} dt = \frac{\Gamma(l+1)\Gamma(\frac{1}{2})}{\Gamma(l+\frac{3}{2})} \times F(\tfrac{3}{2}, l+1; l+\tfrac{3}{2}; \mu), \quad (28)$$

where F is the hypergeometric function.[10] Thus, (24) can now be rewritten as

$$\langle|A_{11}|^{2n}\rangle = \frac{1-\mu}{2^n} \sum_{l=0}^{[n/2]} \binom{n}{2l} \frac{1}{2l+1} F(\tfrac{3}{2}, l+1; l+\tfrac{3}{2}; \mu), \quad (29)$$

or, alternatively,

$$\langle|A_{11}|^{2n}\rangle = \frac{1}{2^n} \sum_{l=0}^{[n/2]} \binom{n}{2l} \frac{1}{2l+1} F(\tfrac{1}{2}, l; l+\tfrac{3}{2}; \mu). \quad (30)$$

When $0 < \mu < 1$, it is easily shown that

$$F(\tfrac{1}{2}, l; l+\tfrac{3}{2}; \mu) = 1, \quad l = 0,$$

$$= \frac{(2l+1)!!}{2l-2)!!} \left[\frac{1}{2\sqrt{\mu}} \left(\sum_{r=0}^{l} A^l_r \mu^{-r}\right) \ln\left(\frac{1+\sqrt{\mu}}{1-\sqrt{\mu}}\right) - \sum_{r=1}^{l} A^l_r \mu^{-r} \sum_{j=0}^{r-l} \frac{\mu^j}{2j+1}\right], \quad l \geq 1, \quad (31)$$

where

$$A^l_r = \frac{(2l-2r-1)!!\,(2r-1)!!}{(2r)!!\,(2l-2r)!!\,(2l-2r-1)}. \quad (32)$$

It follows by mathematical induction with respect to k that

$$C^l_k \equiv \sum_{r=0}^{k} A^l_r = \frac{(2l-2k-3)!!(2k+1)!!}{2^l l!(l-k-l)!!}, \quad 0 \leq k \leq l-1. \quad (33)$$

From this one finds immediately that $\sum_{r=0}^{l} A^l_r \equiv 0$. It is then easily verified that

$$\sum_{r=0}^{l} A^l_r \mu^{-r} = \frac{\mu-1}{\mu^l} \sum_{k=0}^{l-1} C^l_k \mu^{l-1-k}. \quad (34)$$

Combining (30), (31), and (34), we obtain

$$\langle|A_{11}|^{2n}\rangle = \frac{1}{2^n}\left(1 + \frac{\mu-1}{2\sqrt{\mu}} f_1(\mu) \ln\frac{1+\sqrt{\mu}}{1-\sqrt{\mu}} - f_2(\mu)\right), \quad (35)$$

where

$$f_1(\mu) \equiv \sum_{l=1}^{[n/2]} \binom{n}{2l} \frac{(2l-1)!!}{(2l-2)!!} \mu^{-l} \sum_{k=0}^{l-1} C^l_k \mu^{l-1-k}, \quad (36)$$

and

$$f_2(\mu) \equiv \sum_{l=1}^{[n/2]} \binom{n}{2l} \frac{(2l-1)!!}{(2l-2)!!} \sum_{r=l}^{l} A^l_r \mu^{-r} \sum_{j=0}^{r-l} \frac{\mu^j}{2j+1}. \quad (37)$$

Thus, for example,

$$\langle|A_{11}|^4\rangle = \frac{1}{4}\left[1 + \frac{1}{2\mu} + \frac{\mu-1}{4\mu^{3/2}} \ln\left(\frac{1+\sqrt{\mu}}{1-\sqrt{\mu}}\right)\right], \quad (38)$$

and

$$\langle|A_{11}|^6\rangle = \frac{1}{8}\left[1 + \frac{3}{2\mu} + \frac{3(\mu-1)}{4\mu^{3/2}} \ln\left(\frac{1+\sqrt{\mu}}{1-\sqrt{\mu}}\right)\right], \quad (39)$$

when $0 < \mu < 1$.

Let us next consider the above moments in the limiting case $\mu \to 1^-$ (i.e., orthogonal). The case when $\mu \cong 1$ (i.e., $\hat{\gamma} >> \alpha$) corresponds to the situation, mentioned above, where the Hamiltonian contains a small part that is odd with respect to time reversal. Thus this is the case which is of most interest.

From (30) it follows that

$$\lim_{\mu\to 1^-} \langle|A_{11}|^{2n}\rangle = \frac{1}{2^n} \sum_{l=0}^{[n/2]} \binom{n}{2l} \frac{\Gamma(l+\frac{1}{2})}{\Gamma(l+1)\Gamma(\frac{1}{2})}. \quad (40)$$

The sum on the right-hand side is easily evaluated by use of the identity

$$\Gamma(l+\tfrac{1}{2})\Gamma(\tfrac{1}{2}) = \Gamma(l+1)\int_0^{\pi} \cos^{2l}\psi\, d\psi. \quad (41)$$

Equation (40) then becomes

$$\lim_{\mu\to 1^-} \langle|A_{11}|^{2n}\rangle = \Gamma(n+\tfrac{1}{2})/[\Gamma(n+1)\Gamma(\tfrac{1}{2})], \quad (42)$$

which is the well-known orthogonal result.[11]

From (35) it is easily seen that for $\mu \cong 1$

$$\langle|A_{11}|^{2n}\rangle = \lim_{\mu\to 1} \langle|A_{11}|^{2n}\rangle + \tfrac{1}{2}f_1(1)(\mu-1)\ln\left(\frac{1+\sqrt{\mu}}{1-\sqrt{\mu}}\right) + O\left[(\mu-1)^2 \ln\left(\frac{1+\sqrt{\mu}}{1-\sqrt{\mu}}\right)\right] = \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+1)\Gamma(\frac{1}{2})} + \tfrac{1}{2}f_1(1)\,\epsilon \ln\frac{\epsilon}{4} + O[\epsilon^2 \ln\epsilon]. \quad (43)$$

Note that $\epsilon = 0$ is a branch point. Note also that

$$\lim_{\epsilon \to 0^+} \frac{\partial \langle |A_{11}|^{2n} \rangle}{\partial \epsilon} = -\infty, \tag{44}$$

since $f_1(1) > 0$.

Finally, let us consider a power series representation for the moments.

From (30) it follows in a straightforward manner that

$$\langle |A_{11}|^{2n} \rangle = \sum_{l=0}^{+\infty} B_l^{(n)} \mu^l , \tag{45}$$

where

$$B_l^{(n)} \equiv \frac{(2l-1)!!}{2^n (2l)!!} \sum_{k=0}^{[n/2]} \binom{n}{2k} \frac{(2k-1)!!\,(2l+2k-2)!!}{(2l+2k+1)!!\,(2k-2)!!} . \tag{46}$$

For large l,

$$B_l^{(n)} \sim b_n l^{-3/2} , \tag{47}$$

where b_n is independent of l. Thus, the series given by (45) converges for $|\mu| \leq 1$. Below we shall show how this power series can be derived when N is arbitrary.

4. SERIES EXPANSION FOR ARBITRARY N

For N arbitrary,

$$\langle |A_{11}|^{2n} \rangle = 2^{N(N-1)/2} \pi^{-N^2/2} \alpha^{N(N+1)/4} \hat{\gamma}^{N(N-1)/4} \pi^{-N^2/2} \times \int e^{-\alpha \mathrm{Tr} R^2} \exp\left(-2\hat{\gamma} \sum_{i>j} S_{ij}^2\right) d\mathbf{H}_N . \tag{48}$$

This can be rewritten as

$$\langle |A_{11}|^{2n} \rangle = 2^{N(N-1)/2} \pi^{-N^2/2} \hat{\gamma}^{N(N+1)/4} (1-\mu)^{N(N+1)/4} \int e^{-\hat{\gamma} \mathrm{Tr} H^2} e^{(\hat{\gamma}-\alpha)\mathrm{Tr} R^2} |A_{11}|^{2n} d\mathbf{H}_N . \tag{49}$$

Since $\hat{\gamma} > \hat{\gamma} - \alpha$, it is easily shown that the second exponential in the integrand can be expanded and the resultant series integrated term by term. Thus, for $\mu < 1$, (49) can be rewritten as

$$\langle |A_{11}|^{2n} \rangle = \sum_{l=0}^{+\infty} (-1)^l \binom{\beta_N}{l} \mu^l \sum_{k=0}^{+\infty} \frac{1}{k!} \times \langle |A_{11}|^{2n} (\hat{\gamma} \mathrm{Tr} R^2)^k \rangle_u \mu^k , \tag{50}$$

where

$$\langle Q \rangle_u \equiv 2^{N(N-1)/2} \pi^{-N^2/2} \hat{\gamma}^{\beta_N} \int e^{-\hat{\gamma} \mathrm{Tr} H^2} Q \, d\mathbf{H}_N \tag{51}$$

(i.e., the average with respect to the unitary ensemble),

$$\binom{\beta_N}{l} = \frac{\Gamma(\beta_N + 1)}{\Gamma(\beta_N + 1 - l)\Gamma(l+1)} , \tag{52}$$

and

$$\beta_N \equiv N(N+1)/4. \tag{53}$$

Note that $\langle |A_{11}|^{2n} (\hat{\gamma} \mathrm{Tr} R^2)^k \rangle_u$ is independent of $\hat{\gamma}$. Equation (50) can be rewritten as

$$\langle |A_{11}|^{2n} \rangle = \sum_{k=0}^{+\infty} B_k^{(n)} \mu^k , \tag{54}$$

where

$$B_k^{(n)} \equiv \sum_{l=0}^{k} \frac{(-l)^l}{(k-l)!} \binom{\beta_N}{l} \langle |A_{11}|^{2n} (\hat{\gamma} \mathrm{Tr} R^2)^{k-l} \rangle_u . \tag{55}$$

Thus, to obtain the coefficients in the expansion given by (54), one must evaluate averages of the form $\langle |A_{11}|^{2n} (\hat{\gamma} \mathrm{Tr} R^2)^p \rangle_u$, where p is a nonnegative integer.

The average with $p = 0$ is well known, and is given by[11]

$$\langle |A_{11}|^{2n} \rangle_u = \Gamma(N)\,\Gamma(n+1)/\Gamma(N+n). \tag{56}$$

For $p \geq 1$, one can express $\mathrm{Tr} R^2$ in terms of the eigenvalues E_i and the rotation matrix A_{ij}. One can then reduce the problem to evaluation of averages of the form $\langle E_1^{m_1} E_2^{m_2} \cdots E_r^{m_r} \rangle_u$, where $\sum m_i = 2p$, and averages of the form

$$\langle |A_{11}|^{2n} A_{i_1 j_1}^{s_1} A_{i_1 j_1}^{*t_1} \cdots A_{i_r j_r}^{s_r} A_{i_r j_r}^{*t_r} \rangle_u ,$$

where $\sum (s_i + t_i) \leq 4p$. Procedures for evaluating both types of averages for any value p were given previously by the author.[12,13] Unfortunately, the number of averages required and the difficulty of obtaining these averages explicitly increases rapidly as p increases. Thus, the difficulty of evaluating $B_k^{(n)}$ explicitly increases drastically as k increases. Thus, for definiteness we shall evaluate $B_k^{(n)}$ only for $k = 0, 1$, and 2.

From (55) it follows immediately that

$$B_0^{(n)} = \langle |A_{11}|^{2n} \rangle_u , \tag{57}$$

$$B_1^{(n)} = \langle |A_{11}|^{2n} \hat{\gamma} \mathrm{Tr} R^2 \rangle_u - \beta_N B_0^{(n)}, \tag{58}$$

and

$$B_2^{(n)} = \tfrac{1}{2} \langle |A_{11}|^{2n} (\gamma \mathrm{Tr} R^2)^2 \rangle_u - \beta_N B_1^{(n)} - \tfrac{1}{2} \beta_N (\beta_N + 1) B_0^{(n)} . \tag{59}$$

Now, if (9) is written in terms of matrix elements, we have

$$H_{ij} = \sum_{k=1}^{N} A_{ik} A_{jk}^* E_k . \tag{60}$$

It follows directly from this that

$$\mathrm{Tr} R^2 = \tfrac{1}{2} (\mathrm{Tr} H^2 + \delta), \tag{61}$$

where

$$\delta = \sum_{i,j} H_{ij}^2 = \sum_{i,j} (R_{ij}^2 - S_{ij}^2) = \sum_{i,j,k,l} A_{ik} A_{jk}^* A_{il} A_{jl}^* E_k E_l . \tag{62}$$

Further, it is easily verified for the unitary ensemble that

$$\langle Q \mathrm{Tr} H^2 \rangle_u = \frac{N^2}{2\hat{\gamma}} \langle Q \rangle_u + \left\langle \frac{\partial Q}{\partial \hat{\gamma}} \right\rangle_u - \frac{\partial}{\partial \hat{\gamma}} \langle Q \rangle_u , \tag{63}$$

and

$$\langle Q (\mathrm{Tr} H^2)^2 \rangle_u = \frac{N^2}{2\hat{\gamma}^2} \left(\frac{N^2}{2} + 1 \right) \langle Q \rangle_u + \frac{N^2}{\hat{\gamma}} \left\langle \frac{\partial Q}{\partial \hat{\gamma}} \right\rangle_u - \frac{N^2}{\hat{\gamma}} \frac{\partial}{\partial \hat{\gamma}} \langle Q \rangle_u - 2 \frac{\partial}{\partial \hat{\gamma}} \left\langle \frac{\partial Q}{\partial \gamma} \right\rangle_u + \left\langle \frac{\partial^2 Q}{\partial \hat{\gamma}^2} \right\rangle_u + \frac{\partial^2}{\partial \hat{\gamma}^2} \langle Q \rangle_u , \tag{64}$$

where Q is any function of the matrix elements H_{ij} (or equivalently of the eigenvalues E_i and rotation matrix A_{ij}) and $\hat{\gamma}$.

Using (61), (63), and (64), Eq. (58) and (59) can be rewritten as

$$B_1^{(n)} = \tfrac{1}{2} \hat{\gamma} \langle |A_{11}|^{2n} \delta \rangle_u - \tfrac{1}{4} N B_0^{(n)}, \tag{65}$$

$$B_n^{(n)} = \tfrac{1}{8} \hat{\gamma} \langle |A_{11}|^{2n} \delta^2 \rangle_u - \tfrac{1}{8} \hat{\gamma} (N-2) \langle |A_{11}|^{2n} \delta \rangle_u - \tfrac{1}{8} N (\tfrac{1}{4} N + 1) B_0^{(n)} . \tag{66}$$

Now,

$$\langle |A_{11}|^{2n}\delta\rangle_u = \sum_{i,j,k,k'} \langle A_{ik}A_{jk}^* A_{ik'}A_{jk'}^* |A_{11}|^{2n}\rangle_u \langle E_k E_{k'}\rangle_u$$
$$= \{\langle E_1^2\rangle_u - \langle E_1 E_2\rangle_u\}\Delta_0 + N\langle E_1 E_2\rangle_u \langle |A_{11}|^{2n}\rangle_u, \quad (67)$$

where

$$\Delta_0 = \sum_{i,j,k} \langle A_{ik}^2 A_{jk}^{*2} |A_{11}|^{2n}\rangle_u. \quad (68)$$

Note that the statistical independence of the E_i and A_{ij}, as well as the symmetry of the unitary ensemble with respect to the E_i and A_{ij}, has been used in the derivation of (67).

From the results given in Refs. 12 and 13 it follows in a straightforward manner that

$$\langle E_1^2\rangle_u - \langle E_1 E_2\rangle_u = (N+1)/2\gamma, \quad (69)$$

$$\langle E_1 E_2\rangle_u = -1/(2\hat{\gamma}), \quad (70)$$

and

$$\Delta_0 = [\Gamma(N)\Gamma(n+1)/N\Gamma(N+n+2)][(n+1)(n+2)N + 4N(N-1) + 2(N-1)(N+n)(N+n-1)]. \quad (71)$$

From these results it follows directly that

$$\langle |A_{11}|^{2n}\delta\rangle = \frac{1}{2\hat{\gamma}}\left(\frac{n(n-1)(N+2)(N-1)}{N(N+n)(N+n+1)} + N\right) B_0^{(n)}, \quad (72)$$

and thus

$$B_1^{(n)} = \frac{n(n-1)(N+2)(N-1)}{4N(N+n)(N+n+1)} B_0^{(n)}. \quad (73)$$

In a similar manner, one can show that

$$\langle \delta^2 |A_{11}|^{2n}\rangle_u = (1/4\hat{\gamma}^2)[N(N+1)(2\delta_1 + 7\delta_2 + \delta_4 + 4\delta_5 + \delta_6) - 2(N-2)(N+1)\delta_3 - N(N+4) B_0^{(n)}], \quad (74)$$

where the definitions and the explicit expressions for the δ_i are given in Table I.

Inserting the explicit results for the δ_i into (74) and then inserting the resulting expression into (66) yields

$$B_2^{(n)} = \frac{n(n-1)\Gamma(N+n)}{16N(N+2)\Gamma(N+n+4)}[4N^5 + (n^2+3n+30)N^4 + (5n^2+23n+58)N^3 + 2(3n^2+13n-6)N^2 - 8(n^2+7n+16)N - 16(n+2)(n+3)] B_2^{(0)}. \quad (75)$$

Note that, for large N,

$$B_k^{(n)} \sim [n(n-1)/4N] B_0^{(n)}, \quad k = 1, 2. \quad (76)$$

Thus for small μ and large N,

$$\langle |A_{11}|^{2n}\rangle \cong \langle |A_{11}|^{2n}\rangle_u \{1 + [n(n-1)/4N](\mu + \mu^2)\}. \quad (77)$$

At this point one can see that $B_k^{(n)}$ can be calculated for any particular k in a relatively straightforward manner. However, it is also clear that the labor involved increases drastically as k increases. Our original hope was to find $B_k^{(n)}$ for arbitrary k, at least in the limit that N is large. However, we were unable to do so. It appears that an approximation for $B_k^{(n)}$ when N is large is nearly as difficult to obtain as an exact expression. For example, from (55) it would appear that for large N

$$B_k^{(n)} \sim B_0^{(n)}[a_{0k}^{(n)} N^{2k} + a_{1k}^{(n)} 2k-1 + \cdots + a_{2k,k}^{(n)} + b_{1k}^{(n)} N^{-1} + b_{2k}^{(n)} N^{-2} + \cdots], \quad (78)$$

where the $a_{ik}^{(n)}$ and $b_{ik}^{(n)}$ are independent of N. However, as shown above, for $k = 1$ and $k = 2$ all of the a's vanish. Thus it would seem that, to approximate $B_k^{(n)}$, one must approximate all contributions to $B_k^{(n)}$ at least to order $B_0^{(n)}/N$.

TABLE I. Definitions and explicit forms of the δ_i.

Quantity	Definition	Explicit form
δ_1	$\sum_{i,k} \langle |A_{11}|^{2n} |A_{ik}|^8\rangle_u$	$\frac{\Gamma(N)\Gamma(n+1)}{\Gamma(N+n+4)}\left(\frac{\Gamma(n+5)}{\Gamma(n+1)} + 48(N-1) + \frac{24(N-1)\Gamma(N)\Gamma(N+n+3)}{\Gamma(N+3)\Gamma(N+n-1)}\right)$
δ_2	$\sum_{\substack{i\neq i'\\ k}} \langle |A_{11}|^{2n} |A_{ik}|^4 |A_{i'k}|^4\rangle_u$	$\frac{4(N-1)\Gamma(N)\Gamma(n+1)}{\Gamma(N+n+4)}\left[(n+1)(n+2) + (N-2) + \frac{(N+n)(N+n-1)}{N}\left(\times 2 + \frac{(N-2)(N+n+1)(N+n+2)}{(N+1)(N+2)}\right)\right]$
δ_3	$\sum_{i,l} \langle |A_{11}|^{2n} |A_{il}|^4\rangle_u$	$\frac{\Gamma(N)\Gamma(n+1)}{\Gamma(N+n+2)}\left((n+1)(n+2) + 4(N-1) + \frac{2(N-1)(N+n)(N+n-1)}{N}\right)$
δ_4	$\sum_{\substack{i\neq i'\\ k\neq k'}} \langle |A_{11}|^{2n} |A_{ik}|^4 |A_{i'k'}|^4\rangle_u$	$\frac{4(N-1)\Gamma^2(N)\Gamma(n+1)}{\Gamma(N+3)\Gamma(N+n+4)}\left((n+1(n+2)(N+1)(N+2)(N+n+2)(N+n+1) + 4(N-2)^2 + 2(N+1)^2(N+2)^2 + 4(N-2)(N+1)(N+2)(N+n+2)(N+n+1) + \frac{(N-2)N(N+1)(N+n+2)(N+n-1)(N+n+2)(N+n+3)}{N-1}\right)$
δ_5	$\sum_{\substack{i\neq j\\ k\neq l}} \langle |A_{11}|^{2n} |A_{ik}|A_{jk}|^2 |A_{il}|^2\rangle_u$	$\frac{\Gamma^2(N)\Gamma(n+1)}{\Gamma(N+3)\Gamma(N+n+4)}\{4(n+1)(N-1)(N+1)(N+2)[4(n+2) + (N-2)(N+n+3)] + 4(N-1)(N-2)(N+1)(N+2)[8 + (N+n-2)(N+n+3)] + (N-2)[16(N-1)(N-2) + 4(N+n+3)(N+n-2)(N-1)(N-2) + (N+n+2)(N+n-3)(N+n-2)(N+n-1)(N^2-N+2)]\}$
δ_6	$\sum_{\substack{i\neq j\\ i'\neq j'\\ k\neq k'}} \langle |A_{11}|^{2n} A_{ik}^2 A_{jk}^{*2} A_{i'k'}^2 A_{j'k'}^{*2}\rangle_u$	$\frac{8\Gamma^2(N)\Gamma(n+1)}{\Gamma(N+3)\Gamma(N+n+4)}\{2(n+1)(n+2)(N-1)(N+1)(N+2) + 4(N-1)(N-2)(N+1)(N+2) + (N-2)[2(N-1)(N-2) + (N+n-2)(N+n-1)(N+n+2)(N+n+3) - 2(N-1)(N+n-2)(N+n+3)]\}$

The original intent of this work, assuming that the $B_k^{(n)}$ could be found, was to use the resulting expansion to obtain information about the moments when $\mu \cong 1$. Since we can obtain explicitly only the first few terms of the expansion, we are not in a position to derive information in this region. However, we can use what we have obtained to test results derived by other methods. An example of such an application is given in the next section.

5. ON A POSSIBLE APPROXIMATION OF THE WIDTH DISTRIBUTION FOR LARGE N

In a previous article, [5] we gave an approximation for the distribution of widths which seems to be a good approximation when N is large and ϵ is small. In particular, we found that the distribution

$$P_N(X,\gamma,\alpha) = C\{E[\beta\nu/(1+\beta\nu)]/(1+\beta\nu)^{1/2}\} \times (1-X/N)^{A(N-2)}, \quad (79)$$

where

$$\nu = X(1-X/N), \quad (80)$$

$$\beta = B\gamma/\alpha, \quad (81)$$

and

$$E(m) \equiv \int_0^{\pi/2} (1 - m\sin^2\Theta)^{1/2} d\Theta \quad (82)$$

(i.e., a complete elliptic integral of the second kind)[14] seems to approximate quite well the numerical results of Rosenzweig, Monahan, and Mehta[1] if

$$B = 1/0.3N, \quad (83)$$

and A and C are determined by the conditions

$$\int_0^N X^n P_N(X,\gamma,\alpha)\, dX = 1, \qquad n = 0, 1. \quad (84)$$

Since one would expect the assumption made in deriving (79) to be valid also when $\epsilon \cong 1$ (i.e., $\mu \cong 0$), one might conjecture that the distribution given by (79) might be used as an approximation for the actual width distribution for $0 < \epsilon < 1$ if one allows B to be a function of ϵ. We shall now examine this possibility.

In particular, let us consider the moments of the distribution given by (79), where we assume that

$$\beta = B\gamma/\alpha = B\mu/(1-\mu), \quad (85)$$

$$B = \sum_{k=0}^{\infty} B_k \mu^k, \quad (86)$$

and

$$A = A_0 + \sum_{k=1}^{\infty} [a_k/(N-2)]\mu^k. \quad (87)$$

These assumptions seem reasonable since we know from above that the moments of the actual distribution have a power series representation with respect to μ.

It is well known that[15]

$$P_N(X,0,\alpha) = [(N-1)/N](1-X/N)^{N-2} \quad (88)$$

(i.e., the unitary result). Thus, it follows that $A_0 = 1$.

At this point our procedure will be to expand the moments of the above distribution in a power series in μ. Then by equating the coefficients of this series with those given in the previous section for the actual distribution we will determine the a_k and B_k.

Let us define

$$I_n \equiv \int_0^N \frac{E[\beta\mu/(1+\beta\mu)]}{\sqrt{1+\beta\mu}} \left(1-\frac{X}{N}\right)^{A(N-2)} X^n\, dX. \quad (89)$$

If we let $t = X/N$, then

$$I_n = N^{n+1}\int_0^1 \frac{E[y/(1+y)]}{\sqrt{1+y}} (1-t)^{A(N-2)} t^n dt, \quad (90)$$

where $y \equiv \beta N t(1-t)$. It is easily shown that

$$E[y/(1+y)]/\sqrt{1+y} = \tfrac{1}{2}\pi\{1 - \tfrac{3}{4}t(1-t)b_1\mu + [\tfrac{45}{64}b_1^2 t^2(1-t)^2 - \tfrac{3}{4}t(1-t)b_2]\mu^2 + \cdots\}, \quad (91)$$

and

$$(1-t)^{A(N-2)} = (1-t)^{N-2} e^{[1n(1-t)](a_1\mu + a_2\mu^2 + \cdots)} = (1-t)^{N-2}(1 + a_1 1n(1-t)\mu + \{a_2 1n(1-t) + \tfrac{1}{2}a_1^2[1n(1-t)]^2\}\mu^2 + \cdots), \quad (92)$$

where

$$b_1 \equiv NB_0, \quad (93)$$

and

$$b_2 \equiv N(B_0 + B_1). \quad (94)$$

Thus,

$$I_n = N^{n+1}\tfrac{1}{2}\pi(\xi_0^{(n)} + \xi_1^{(n)}\mu + \xi_2^{(n)}\mu^2 + \cdots), \quad (95)$$

where

$$\xi_0^{(n)} = J_0(n, N-2), \quad (96)$$

$$\xi_1^{(n)} = -\tfrac{3}{4}b_1 J_0(n+1, N-1) + a_1 J_1(n, N-2), \quad (97)$$

$$\xi_2^{(n)} = a_2 J_1(n, N-2) + \tfrac{1}{2}a_1^2 J_2(n, N-2) - \tfrac{3}{4}a_1 b_1 J_1(n+1, N-1) + \tfrac{45}{64}b_1^2 J_0(n+2, N) - \tfrac{3}{4}b_2 J_0(n+1, N-1), \quad (98)$$

and

$$J_l(z,w) \equiv \int_0^1 [1n(1-t)]^l t^z (1-t)^w dt. \quad (99)$$

Now, $J_0(Z, W)$, when $Rz > -1$ and $Rw > -1$, is just an integral representation of the Beta function. [16] Thus,

$$J_0(z,w) = \Gamma(z+1)\Gamma(w+1)/\Gamma(z+w+2). \quad (100)$$

Further, if z is a nonnegative integer, it can be shown using mathematical induction that

$$J_1(z,w) = -J_0(z,w) j_1(z,w), \quad (101)$$

where

$$j_1(z,w) = \sum_{k=1}^{z+1} \frac{1}{w+k}, \quad (102)$$

and that

$$J_2(z,w) = 2J_0(z,w) j_2(z,w), \quad (103)$$

where

$$j_2(z,w) = \sum_{k=1}^{z+1} \frac{1}{w+k} \sum_{l=k}^{z+1} \frac{1}{w+l}. \quad (104)$$

Now, from (11) we have

$$\langle |A_{11}|^{2n}\rangle = I_n/N^n I_0. \quad (105)$$

Combining (95)-(105), one obtains

$$\langle |A_{11}|^{2n}\rangle = (1 + \hat{B}_k^{(n)}\mu + \hat{B}_k^{(n)}\mu^2 + \cdots, \quad (106)$$

where

$$\hat{B}_1^{(n)} \equiv [\bar{\xi}_1^{(n)} - \bar{\xi}_1^{(0)}]\langle |A_{11}|^{2n}\rangle_u, \quad (107)$$

$$\hat{B}_2^{(n)} \equiv \{\bar{\xi}_2^{(n)} - \bar{\xi}_2^{(0)} - \bar{\xi}_1^{(0)}[\bar{\xi}_1^{(n)} - \bar{\xi}_1^{(0)}]\} \times \langle |A_{11}|^{2n}\rangle_u, \quad (108)$$

$$\bar{\xi}_1^{(n)} = -\frac{3(n+1)(N-1)b_1}{4(N+n+1)(N+n)} - a_1 j_1(n, N-2), \quad (109)$$

and

$$\bar{\xi}_2^{(n)} = -a_2 j_1(n, N-2) + a_1^2 j_2(n, N-2) + \tfrac{3}{4} a_1 b_1 j_1 \times (n+1, N-1)\frac{(n+1)(N-1)}{(N+n)(N+n+1)} + \tfrac{45}{64} b_1^2 \frac{(n+2)(n+1)N(N-1)}{(N+n+3)(N+n+2)(N+n+1)(N+n)} - \frac{3b_2}{4}\frac{(n+1)(N-1)}{(N+n+1)(N+n)}. \quad (110)$$

If one now equates the coefficients $\hat{B}_1^{(n)}$ to the $B_1^{(n)}$, given by (73), for $n = 1$ and $n = 2$, one obtains two simultaneous linear algebraic equations for a_1 and b_1. The solutions are found to be

$$a_1 = -(N+1)N-2)/6N, \quad (111)$$

and

$$b_1 = 2(N+1)^2(N+2)/9N(N-1). \quad (112)$$

Inserting these results into (109), we find that

$$\bar{\xi}_1^{(n)} = -\frac{(n+1)(N+1)^2(N+2)}{6(N+n)(N+n+1)N} + \frac{(N+1)(N-2)}{6N} \times \sum_{k=0}^{n} \frac{1}{N-1-k}. \quad (113)$$

If we now insert (113) into (108), we find for $N > 2$ that $\hat{B}_1^{(n)} = B_1^{(n)}$ iff $n = 0, 1$, or 2. However, for large N, we find that

$$\hat{B}_1^{(n)} \sim [n(n-1)/4N]B_0^{(n)}. \quad (114)$$

Thus, for large N, $\hat{B}_1^{(n)} \cong B_1^{(n)}$ for all n.

If next one equates $\hat{B}_2^{(n)}$ to $B_2^{(n)}$ for $n = 1$, and $n = 2$, one obtains two simultaneous equations for a_2 and b_2. If one solves these, and then evaluates $\hat{B}_2^{(n)}$, it is found that if $N > 2$, $B_2^{(n)} = \hat{B}_2^{(n)}$ iff $n = 0, 1, 2$. Further, it is found that, for large N,

$$\hat{B}_2^{(n)} \sim -[n(n-1)(5n-37)/108N]\langle |A_{11}|^{2n}\rangle_u \quad (115)$$

Thus, unfortunately, we see that the distribution given by (79), with A and B given by (86) and (87) respectively, cannot be used to approximate the actual distribution for $0 < \epsilon < 1$ even when N is very large.

6. DISCUSSION

We found above that when $N = 2$,

$$\lim_{\epsilon \to 0^+} \frac{\partial}{\partial \epsilon} \langle |A_{11}|^{2n}\rangle = -\infty.$$

The numerical results of Rosenzweig, Monahan and Mehta[1] suggest that this is probably the case for all N. Thus, it seems that it will in fact be extremely difficult, at best, to derive a valid approximation when N is arbitrary and ϵ is small, for width distribution and/or its moments. Certainly the usual types of perturbation theory will be of little help.

On the other hand, it is a relatively simple matter to approximate the moments when $\epsilon \cong 1$. In fact, as shown above, a series representation, in powers of $1 - \epsilon$, exists, and the coefficients can be found to any particular order. We feel that it may somehow be possible to approximate the general coefficient when N is large (although to date we have been unable to do so). If this can be done, one would then have a representation of the moments for large N when $0 < \epsilon < 1$. Hopefully, one could then obtain much information about the moments in the region of physical interest, namely when ϵ is small.

*Supported in part by an N.R.C. grant (Grant No. A 4625).

[1]N. Rosenzweig, J. E. Monahan, and M. L. Mehta, Nucl. Phys. A 109, 437 (1968).
[2]M. L. Mehta, Nuovo Cimento B 65, 107 (1970).
[3]M. L. Mehta, *The International Conference on Statistical Properties of Nuclei*, Albany, New York, August 1971, see *Statistical Properties of Nuclei*, edited by J. B. Garg (Plenum, New York, 1972).
[4]L. D. Favro and J. F. McDonald, Phys. Rev. Letters 19, 1254 (1967).
[5]J. F. McDonald, J. Math. Phys. 10, 1191 (1969).
[6]See for example, the introductory review of C. E. Porter, *Statistical Theories of Spectra: Fluctuations* (Academic, New York, 1915). See also M. L. Mehta, *Random Matrices and the Statistical Theory of Energy Levels* (Academic, New York, 1967).
[7]F. J. Dyson, J. Math. Phys. 3, 140 (1962).
[8]See, for example, Ref. 1.
[9]The double factorial is defined by the recursion relation $n!! = n\{(n-2)!!\}$, with $0!! \equiv 1$, $1!! \equiv 1$. Note that $\frac{1}{(-2n)!!} = 0$, $n > 0$ while $[-(2n+1)]!! = \frac{(-1)^n(2n+1)}{(2n+1)!!}$, $n \geq 0$.
[10]See, for example, *Applied Mathematics Series, No. 55: Handbook of Mathematical Functions* (National Bureau of Standards, Washington, D.C., 1964), pp. 556-66. All of the properties of the hypergeometric functions required for our discussion can be found here.
[11]N. Ullah, Nucl. Phys. 58, 65 (1964).
[12]J. F. McDonald, J. Math. Phys. 12, 542 (1971).
[13]J. F. McDonald, J. Math. Phys. 13, 1299 (1971).
[14]Ref. 10, p. 590.
[15]N. Ullah, J. Math. Phys. 4, 1279 (1963).
[16]See, for example, Ref. 10, p. 258.

On dynamical groups: Classification of Lie algebras with Galilei subalgebras

H. D. Doebner* and J. Hennig†

International Centre for Theoretical Physics, Trieste, Italy
(Received 1 June 1973)

Because of the pure group theoretical approach to the free nonrelativistic particle through an integrable irreducible representation of the quantum mechanical Galilei (Lie) algebra G, it is reasonable to construct so-called dynamical (Lie) algebras D which possess an integrable representation describing interacting (nonrelativistic) systems. Such dynamical algebras D should contain the geometrical subalgebra G_0 of G, spanned by the mass operator, the momentum, the angular-momentum and the position operators. Furthermore, the relation between the free and the interacting system is simplified if the one-dimensional Lie algebra T generating free time translations is a subalgebra of D. Hence preferred candidates for D are those Lie algebras L which possess a subalgebra isomorphic to G or to G_0, i.e., those L for which an injective homomorphism $\epsilon: G \to \epsilon(G) \subset L$ or $G_0 \to \epsilon(G_0) \subset L$ exists. ϵ is called an embedding of G or of G_0 in L.
Our main result is a complete classification of
(i) all nonsemisimple L with Levi decomposition $L = S \overset{+}{\subset} F$ with G or G_0 embedding ϵ such that $S \subset \epsilon(G)$ or $S \subset \epsilon(G_0)$
(ii) all complex semisimple $\bar{L}$ with an embedding of the complex extension of G,
(iii) all real simple L being real forms or realifications of the lowest dimensional $\bar{L}$ (i.e., A_5, B_3, C_4) with G-embedding.
The result gives a fairly complete list of all candidates for dynamical nonrelativistic algebras. The physical aspects of two of them, the conformal Galilei algebra G_c and a limitable dynamical algebra D^t, are discussed. A method for the construction of physically useful integrable representations for G_c and D^t is given. Some general properties of nonrelativistic and of relativistic dynamical algebras with G-embeddings are considered.

1. INTRODUCTION

1. One of the more attractive methods in a group theoretical approach to particle physics is the concept of dynamical (Lie) algebras D which possess an irreducible and integrable representation U_{ph}, such that the properties of an interacting system are completely determined by $U_{ph}(D)$, similarly as an irreducible and integrable representation of the Galilei (Lie) algebra G gives a description for the free nonrelativistic particle.

The completeness of the description through $U_{ph}(D)$ implies for nonrelativistic systems that the position, momentum, and angular momentum operators q_k, p_k, and m_k, $k = 1, 2, 3$ (three space dimensions) are generators of D and that the quantum mechanical commutation relations are valid for q_k, p_k, and m_k. Hence the mass operator c is also a generator of D. This gives

$$G_0 \subset D$$

as an essential condition for a dynamical algebra, with G_0 being the geometrical subalgebra of G spanned by q_k, p_k, m_k, and c.

Furthermore, also the Hamiltonian h generating the dynamics of the system has to be a generator of D, making a calculation of bound and scattering states possible.[1] If $[H, G_0]$ is not contained in $G_0 + H$ (H being the linear hull of h), which is the general case, then D contains further generators, some of them being again Hamiltonians. The set of all these Hamiltonians is called the interaction type of D. D is a finite dimensional Lie algebra. Hence there is only a finite number of linearly independent Hamiltonians in D.

2. An additional and restrictive requirement is the assumption that the Hamiltonian t of the free system also belongs to D. It implies

$$G_0 \subset G \subset D.$$

A justification can be given through a somewhat implicit physical argument. A nonrelativistic interaction can generally be turned off. This (noncontinuous) process connects interacting and free systems. If the interacting system is described via $U_{ph}(D)$ and the free one via $U'_{ph}(G)$ it should be possible to define a mathematical procedure which gives $U'_{ph}(G)$ as a "limit" of $U_{ph}(D)$. This procedure is a mathematical translation of the physical turning off process. Following Segal,[2] it is reasonable to assume that the "limit" can be realized as a "contraction" between integrable representations of D and G. However, D and G have different dimensions and the definition of contractions becomes a complicated mathematical problem. Quite different methods are known.[3] The simplest one, used in Ref. 4, is to restrict a unitary representation of D to one of its subalgebras and can be applied only if $G \subset D$. This relates the assumption $t \in D$ to a special limiting procedure. Note that also if $t \notin D$ a group theoretical method should exist, which turns the interacting system into a free one.

3. To simplify our notation we call an injective homomorphism ϵ mapping a Lie algebra A into a Lie algebra B an A-embedding into B, $\epsilon: A \to \epsilon(A) \subset B$. Two A-embeddings into B, ϵ_1 and ϵ_2, are said to be equivalent, if there exist automorphisms ρ_A of A and ρ_B of B such that $\epsilon_2 = \rho_B \epsilon_1 \rho_A$ holds. Then the result of the discussion can be summarized as:

Let G be the Galilei algebra and G_0 its geometrical subalgebra. Lie algebras with G- or with G_0-embeddings are preferable candidates for nonrelativistic dynamical algebras.

This approach was frequently used (see also Sec. 5). A Lie algebra which contains G, the conformal Galilei algebra G_c, was derived in Ref. 5 and is a maximal kinematical invariance algebra of the free massive Schrödinger operator. The generators in G_c, not contained in $G \subset G_c$, were identified as infinitesimal dila-

tations and as a kind of infinitesimal conformal transformations.

Two Lie algebras D^t and D^t_s with G-embeddings were given in Refs. 4 and 6 together with a representation theory and an identification of the physical systems described through D^t and D^t_s.

Lie algebras with G-embeddings can also be used in the relativistic case. The Poincaré algebra $\hat{P}$ itself has an embedding of the Galilei algebra in two space dimensions, $G^{(2)}$, which was used as Galilei structure in the light-cone formalism and in phenomenological parton models.[7] Similarly, the conformal algebra $so(4, 2)$ contains an embedding of $G^{(2)}_c$, being the conformal Galilei algebra in two space dimensions.[8] A "relativistic" Galilei algebra G_r, with embeddings $\epsilon(G) \subset G_r$ and $\epsilon'(\hat{P}) \subset G_r$ such that $\epsilon(G) \cap \epsilon'(\hat{P})$ is the Euclidean algebra E_3 in three space dimensions, was outlined in Ref. 9.

4. It is the purpose of this paper to give a classification of all Lie algebras L possessing a G-embedding or a G_0-embedding (Sec. 2, 3, 4) and to discuss their physical properties (Sec. 5).

Because a classification of solvable Lie algebras and their semidirect sums with semisimple ones is not yet available, we restrict the classification to those $L = S \in F$ (Levi decomposition) with a G-, G_0-embedding such that the radical S is contained in $\epsilon(G)$ or $\epsilon(G_0)$, respectively, i.e., $S \subset \epsilon(G) \subset L$, or $S \subset \epsilon(G_0) \subset L$. We call such radicals minimal.[4]

5. The material is organized as follows:

The notation is explained in Appendix A and the Galilei algebra (in three space dimensions) is described in Appendix B.

The first step of the classification is Lemma 1 and its corollary, which list the minimal radicals for $\epsilon(G) \subset L$ and $\epsilon(G_0) \subset L$, respectively. The next step is a determination of the Levi factors. The result, collected in Sec. 3. 1, is a classification of all nonsemisimple $L \supset \epsilon(G)$ (lemma 2) and $L \supset \epsilon(G_0)$ (Lemma 3) with minimal radicals in both cases. The proofs are given in Secs. 3. 2 and 3. 3. for lemma 2 and Lemma 3, respectively. A method to construct all real simple $L \supset \epsilon(G)$ via an embedding of the complex extension $\overline{G}$ of G in complex simple Lie algebras $\overline{L}$ is explained in Sec. 4. 1. The classification theorems are presented in Sec. 4. 2. We give all complex simple $\overline{L}$ with $\overline{L} \supset \epsilon(\overline{G})$ in Theorem 1, all real simple L being realifications with $L \supset \epsilon(G)$ in Theorem 2 and all real simple $L \supset \epsilon(G)$ being real forms of the lowest dimensional Lie algebras $\overline{L}$ with $\overline{G}$-embeddings in Theorem 3. The mathematical background of the proofs is outlined in Secs. 4. 3 and 4. 4.

A discussion of the results and examples are given in Sec. 5. The problem of identifying the physical system described through a dynamical algebra is outlined and partially solved in Sec. 5. 1, which contains also a representation theory, based on the method of Nelson extensions, for some $L \supset \epsilon(G)$. This method is also applicable to G_c. Some examples are treated in Sec. 5. 2.

2. PROPERTIES OF LIE ALGEBRAS WITH G-, G_0-EMBEDDINGS

1. Let L be a real Lie algebra and $L = S \in {}_\sigma F$ its Levi decomposition (S = radical, L = Levi factor). Consider $G = R \in M$ and $G_0 = R_0 \in M$ with $M \approx so(3)$, $R = (C \oplus T \oplus P) \in Q$ and $R_0 = (C \oplus P) \in Q$, respectively. Then those L with a G- or a G_0-embedding ϵ such that $S \subset \epsilon(G)$ or $S \subset \epsilon(G_0)$ (minimal radical) have the form given in Lemma 1 (for G-embeddings) and its corollary (for G_0-embeddings).

We use the notation $A \subsetneq B$ if there exists $A' \approx A$, $A' \subset B$.

Lemma 1: Let ϵ be an embedding of G into L.

A. Non-semisimple L with $S \subset \epsilon(G)$

a. If $L = S \in {}_\sigma F$, σ being an injective homomorphism $\sigma: F \to \mathrm{Der}\, S$, then L has the structure
(I) $S = \epsilon(C + T + P)$, $M + Q \subsetneq F$ or
(II) $S = \epsilon(C + P + Q)$, $M + T \subsetneq F$ or
(III) $S = \epsilon(C + T + P + Q)$, $M \subsetneq F$.

b. If $L = S \in {}_\sigma F$, σ being an arbitrary homorphism $\sigma: F \to \mathrm{Der}\, S$, then L can be written as $L = L' \oplus F_2$, $L' = S \in {}_\rho F_1$, $F = F_1 \oplus F_2$, $\rho = \sigma/F_1$ injective, with $G \subsetneq L'$ and L' having the structure of L in part a.

B. Semisimple L

If L is decomposed into $L = \oplus\, {}^r_1 L_i$, L_i simple, then $G \subsetneq L_i$ for at lease one i.

Corollary: Let ϵ be an embedding of G_0 into L.

A. Non-semisimple L with $S \subset \epsilon(G_0)$

a. If $L = S \in {}_\sigma F$, σ being an injective homorphism $\sigma: F \to \mathrm{Der}\, S$, then L has the structure
(I) $S = \epsilon(C + P)$, $M + Q \subsetneq F$ or
(I′) $S = \epsilon(C + Q)$, $M + P \subsetneq F$ or
(II) $S = \epsilon(C + P + Q)$, $M \subsetneq F$.

b. (See lemma 1 with G_0 inserted for G).

B. Semisimple L (See Lemma 1).

2. *Proof of Lemma 1:* For simplicity we shall identify $\epsilon(G)$ with G; no confusion should arise. Consider the projection operators P_S, P_F in the vector space $S + F$ with $P_F/S = P_S/F = 0$ and $P_S + P_F = \mathbb{1}$. Then P_F is a (Lie algebra) homorphism $L \to F$, because for $x, x' \in L$ $[x, x'] = [(P_S + P_F)(x), (P_S + P_F)(x')] = y + [P_F(x), P_F(x')]$ with $y \in S$, i.e., $P_F([x, x']) = [P_F(x), P_F(x')]$.

Note that P_S is not necessarily a homomorphism. For P_F we have $\mathrm{Ker}\, P_F \cap G = J_\nu$ with J_ν, $\nu = 1, \ldots, 6$, being one of the six solvable ideals of G given in Appendix B. Hence, at most, those S with $S = J_\nu$ will be minimal. We check for which J_ν an embedding is possible. Take $S = C$. Then $[(P_F + P_S)(P), (P_F + P_S)(Q)] = [P_F(P), P_F(Q)] = P_F([P, Q]) = 0$, which contradics $[P, Q] = C$.

The case $S = C + P$ is also forbidden: $[T, Q] \subset C + [P_F(T), P_F(Q)] = C + P_F([T, Q]) = C + P_F(P) = C$ contradicts $[T, Q] = P$. The theorem of Malcev and Harish-Chandra[10] implies that, if $G \subset L$, then there is always an automorphism τ of L such that $\tau(M) \subset F$, $\tau(S) = S$. Moreover, for $S = C + T + P$ an automorphism ξ of G with $\xi/C + T + P + M = \mathbb{1}$ and $\xi(q_i) = \tau^{-1}(P_F(q_i))$ exists, i.e., $\tau\xi(Q + M) \subset F$, thus $Q + M \subsetneq F$. A similar result holds for $S = C + P + Q$. The case $S = R$ is trivial.

Let $L = S \in {}_\sigma F$ be decomposed into $L = L' \oplus F_2$, $L' = S \in {}_\rho F_1$, $F_2 = \ker \sigma$, $\rho = \sigma/F_1$, $F = F_1 \oplus F_2$, then ρ is an injective homomorphism $F_1 \to \mathrm{Der}\, S$, and $\epsilon' = P_{L'}\epsilon$ is injective because all nonzero ideals in G contain C and $\epsilon(C) \subset S$ for all $J_\nu \neq 0$ ($P_{L'}$ is the projection operator on L'). Hence ϵ' is an embedding of G in L'. For $S = 0$, L is decomposed into simple ideals

$L = \oplus_1^r L_i$. Every projection operator P_i is a homomorphism $L \to L_i$ and the kernel of P_i is an ideal in G. If $P_i(G) \not\subset G$ for $i = 1, \ldots, r$, then $C \subset \ker P_i$ for all i and hence $C \not\subset L$ in contradiction to $G \subset L$. So there is at least one simple ideal L_i in L with G-embedding.

3. The proof of the corollary is the same as for Lemma 1, except the part in which $S = C + P$ is excluded because this part uses T. P and Q have to be treated symmetrically.

4. We mention that because of the results given in Secs. 3 and 4, all types of L with G-embedding given in Lemma 1 and its corollary are possible

3. NONSEMISIMPLE LIE ALGEBRAS WITH G-, G_0-EMBEDDINGS

We calculate a complete list of all nonsemisimple $L = S \dot{\oplus}_\sigma F$ of type I, II, III for G and of type I, II for G_0. The results are given in Lemmas 2 and 3 in Sec. 3. The proofs for G and G_0 are outlined in Secs. 3. 2 and 3. 3, respectively. In the proofs, first Der S is determined, then properties of $\sigma(F) = \mathrm{ad}_S F$ are derived such that F can be identified using the classification of irreducible real and comple representations of simple Lie algebras. Note that $\mathrm{ad}_S F$ gives the commutators between S and F.

A. The results for minimal radicals

Lemma 2: Let G be the Galilei algebra $G = R \dot{\oplus} M$, $R = (C \oplus T \oplus P) \dot{\oplus} Q$, $M \approx so(3)$. Let ϵ be an embedding of G into $L = S \dot{\oplus}_\sigma F$ with σ being an injective homomorphism $\sigma: F \to \mathrm{Der}\, S$, and let S be a minimal radical and not zero.

Then L has the following form:

(I) $S = \epsilon(C + T + P)$

- (a) $F \approx sl(5,R)$, $\mathrm{ad}_S F = sl(5,R)$ or
- (b) $F \approx so(4,1)$, $\mathrm{ad}_S F = so(4,1)$

(II) $S = \epsilon(C + P + Q)$

- (a) $F \approx sp(3,R)$, $\mathrm{ad}_{P+Q} F = sp(3,R)$ and $\mathrm{ad}_C F = 0$ or
- (b) $F \approx sp(1,R) \oplus so(3)$, $\mathrm{ad}_{P+Q} F = (\mathbb{1}_3 \times sp(1,R)) \oplus (so(3) \times \mathbb{1}_2)$ and $\mathrm{ad}_C F = 0$

(III) $S = \epsilon(C + T + P + Q)$, $F \approx so(3)$.

Lemma 3: Let G_0 be the geometrical Galilei algebra $G_0 = R_0 \dot{\oplus} M, R_0 = (C \oplus P) \dot{\oplus} Q$, $M \approx so(3)$. Let ϵ be an embedding of G_0 into $L = S \dot{\oplus}_\sigma F$ with σ being an injective homomorphism σ: $F \to \mathrm{Der}\, S$ and let S be minimal but not trivial.

Then L has the following form:

(I) $S = \epsilon(C + P)$ or $S = \epsilon(C + Q)$, $F \approx sl(4,R)$, $\mathrm{ad}_S F = sl(4,R)$ or

(II) $S = \epsilon(C + P + Q)$

- (a) (see IIa in Lemma 2) or
- (b) (see IIb in Lemma 2) or
- (c) $F \approx so(3)$, $\mathrm{ad}_{P+Q} F = so(3) + so(3)$, $\mathrm{ad}_C F = 0$ or
- (d) $F \approx sl(3,R)$, $\mathrm{ad}_C F = 0$ $\mathrm{ad}_{P+Q} F = sl(3,R) \dot{+} sl(3,R)$.

We add the following remarks:

Case III in Lemma 2 shows that there exists no embedding of G in $L = R \dot{\oplus} F$ except for $L \approx G$. The corresponding case for a G_0-embedding, i.e., $S = R_0$, has a trivial ($L \approx G_0$) and nontrivial solutions (as II in Lemma 3).

$sp(3,R)$ is the largest simple Lie algebra which can be coupled to $C + P + Q$. Any semisimple Lie algebra for which each simple ideal has nontrivial commutators with the radical $C + P + Q$ is a subalgebra of $sp(3,R)$, i.e., $sp(1,R) \oplus so(3) \subset sp(3,R)$, $so(3) \subset sp(3,R)$ or $sl(3,R) \subset sp(3,R)$, and the corresponding $\mathrm{ad}_S F$ can be calculated from $\mathrm{ad}_{P+Q} sp(3,R) = sp(3,R)$ and $\mathrm{ad}_C F = 0$

It was mentioned already that any L with $\epsilon(G) \subset L$ has also a G_0-embedding, $\epsilon(G_0) \subset L$. But there exist embeddings ϵ_0 of G_0 into some L which cannot be extended to embeddings of G into L (e.g. Lemma 3 case I); in these cases and if L has a minimal radical the Levi factor of L must be enlarged in order to get extensions of ϵ_0 to G-embeddings (except for $L \approx G_0$).

B. Proof of the classification lemma for G-embeddings

1. Embeddings with radical $S = \epsilon(C + T + P)$ (Type I): $\epsilon(G) \equiv G$ for simplicity.

(1) S is a five-dimensional Abelian subalgebra; hence Der $S = sl(5,R) \oplus T_1$ with T_1 spanned by $\mathbb{1}_5$. A suitable basis in S is

$$\mathfrak{L}_S = \{c, t, p_i / i = 1, 2, 3\}.$$

(2) We have $\sigma(F) \subset sl(5,R)$. $\mathrm{ad}_S F$ is real irreducible (i.e., because of dim $S = 5$, also complex irreducible). To show this, suppose the contrary. Then $\mathrm{ad}_S F$ reduces into a 3- or a 4-dimensional real irrep. and a two- or a one-dimensional rep. because of $\mathrm{ad}_S M = 0_2 \dot{+} so(3)$. Then $S \approx T_5$ implies that G decomposes into $G = G_1 \oplus G_2$ with dim $G_2 = 2$ or 1 which is not possible.

(3) F must be simple, because otherwise Schur's lemma would imply that F contains a simple ideal commuting with S, which contradicts σ being injective. There are only real forms of A_3, A_4, B_2, D_2 as simple subalgebras $X \subset sl(5,R)$ with dim $X > 6$. A_3 is excluded since it has no five-dimensional irrep. For $A_4 \approx sl(5,C)$, only the real form $sl(5,R)$ is isomorphically contained in Der S. $so(5,C)$ has three real forms, $so(5)$, $so(4,1)$, and $so(3,2)$. Because $so(5)$ is the compact real form of B_2, the maximal dimension of its abelian subalgebras is equal to the rank of B_2, i.e., it is 2 and $F \approx so(5)$ is excluded.

The same holds for $F \approx so(3,2)$. In this case consider the following $so(3,2)$ basis $\mathfrak{L}' = \{m_i', n_{i4}, n_{i5}, k / i = 1, 2, 3\}$ with $m_i' = -\epsilon_{ijk} E_{jk}$, $n_{i4} = E_{i4} - E_{4i}$, $n_{i5} = E_{i5} + E_{5i}$, $k = E_{45} - E_{54}$. The rep. ad $so(3)$ reduces on $so(3,2)$ as $so(3) \dot{+} so(3) \dot{+} so(3) \dot{+} 0_1$, and $\{m_i\}$, $\{n_{i4}\}$, and $\{n_{i5}\}$ transform under ad $so(3)$ in the same way. Because, up to automorphism of $so(3,2)$, there is only one subalgebra in $so(3,2)$, isomorphic to $so(3)$, we have that $q_j = am_j' + bn_{j4} + dn_{j5}$ with a, b, d real. But $[q_1, q_2] = 0$ implies $a = b = d = 0$. The remaining case $F \approx so(4,1)$ is possible. There exists an embedding of G into L, e.g., with a basis $\mathfrak{L}'' = \{p_1, p_2, p_3, c + p, c - p\}$ for S, with the rep. $so(3) + 0_2$ for $\mathrm{ad}_S M$ and $\mathrm{ad}_S q_i = -E_{i4} + E_{i5} + E_{4i} + E_{5i}$.

2. Embeddings with radical $S = \epsilon(C + P + Q)$ (Type II):

(1) S is isomorphic to the Heisenberg algebra in three space dimensions. We use the basis $\mathfrak{L}_S =$

$\{c, p_i, q_i / i = 1, 2, 3\}$ and infer from Ref. 4 that Der $S \approx T_6 \subsetneq (T_1 \oplus sp(3,R))$. The symplectic algebra $sp(3,R)$ appears in the form

$$O_1 \dot{+} \begin{pmatrix} S_1 & S_2 \\ S_3 & -S_1^T \end{pmatrix}, \; S_i \text{ real } 3\times 3 \text{ matrices}, S_2, S_3 \text{ symmetric}.$$

The generator of T_1 is a diagonal matrix and the six generators of T_6 are $E_{1\nu}$, $\nu = 2, \ldots, 7$, i.e., C is an ideal in L.

(2) We have $\sigma(F) \subsetneq sp(3,R)$. $[C, F] \subset C$ yields $[C, F] = 0$, i.e., $\sigma(F) \subset O_1 \dot{+} sp(3,R)$. Hence we have to calculate all semisimple F with

$$\begin{aligned} &(\alpha)\ \sigma(F) = O_1 \dot{+} \mathrm{ad}_{P+Q}F \subset O_1 \dot{+} sp(3,R) \\ &(\beta)\ \sigma(M) = O_1 \dot{+} so(3) \dot{+} so(3) \subset \sigma(F) \\ &(\gamma)\ \sigma(t) \ = O_1 \dot{+} (1_3 \times \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}) \epsilon O_1 \dot{+} (1_3 \times sp(1,R)) \subset \sigma(F). \end{aligned}$$

$\mathrm{ad}_{P+Q}F$ is complex irreducible in $sp(3,R)$. To show this, suppose the contrary. Then, because of (β), $\mathrm{ad}_{P+Q}F$ reduces into two three-dimensional irreps. which contain a rep. of both M and T which commute. So $\mathrm{ad}_{P+Q}T$ must commute with $\mathrm{ad}_{P+Q}F$ and, because $\mathrm{ad}_{P+Q}F$ has no centre $\mathrm{ad}_{P+Q}T$ must be zero in contradiction to $[T, S] \neq 0$.

Now let $\mathrm{ad}_{P+Q}F$ be a complex irreducible subalgebra of $sp(3,R)$. Then $\overline{\mathrm{ad}_{P+Q}F} = \mathrm{ad}_{P+Q}\overline{F}$ is also irreducible in $sp(3, C)$. A complete list of irreducible subalgebras of $sp(3, C)$ can be found in Ref. 11 and we have (a) $\mathrm{ad}_{P+Q}\overline{F} = sp(3, C)$ or (b) $\mathrm{ad}_{P+Q}\overline{F} \approx (1_3 \times sp(1, C)) \oplus (so(3, C) \times 1_2)$. Hence $\mathrm{ad}_{P+Q}F$ is a real form of $\mathrm{ad}_{P+Q}\overline{F}$ such that $\mathrm{ad}_{P+Q}F$ remains complex irreducible. This implies for case (a) $\mathrm{ad}_{P+Q}F = sp(3,R)$ and, because of (β), for case (b) $\mathrm{ad}_{P+Q}F = (1_3 \times sp(3,R)) \oplus (so(3) \times 1_2)$. There exist embeddings of both types.

3. Embeddings with radical $S = \epsilon(C + T + P + Q)$ (type III):

(1) A straightforward calculation shows $\mathrm{Der}\, S = [((A_P \oplus A_Q) \subsetneq (A_T \oplus T_5)) \subsetneq (T_1 \subsetneq T_2)] \subsetneq A_M$ with $A_P \approx A_Q$ three-dimensional abelian, A_T one-dimensional, $A_M \approx so(3)$ and T_1, T_2, T_5 being one-, two-, five-dimensional Abelian. Furthermore $\mathrm{ad}\, S = A_P + A_Q + A_T, ((A_P \oplus A_Q) \subsetneq A_T) \subsetneq A_M \approx G/C$ and $[A_P, T_2] = A_P, [A_Q, T_2] = A_Q, [A_T, T_2] = A_T$, $[T_2, T_1] = T_1, [T_5, A_M] = T_5, [T_5, A_Q] = A_P$, $[T_5, T_2] = T_5$.[12]

(2) From $\sigma(M) \subset \sigma(F) \subsetneq \mathrm{Der}\, S$ we have immediately $F \approx so(3)$.

C. Proof of the classification lemma for G_0-embeddings

1. Embeddings with radical $S = \epsilon(C + P)$ or $S = \epsilon(C + Q)$ (type I, I′):

(1) $[\epsilon(G_0) \equiv G_0]\, S$ is a four-dimensional abelian subalgebra; hence $\mathrm{Der}\, S = sl(4,R) \oplus T_1$ with T_1 spanned by 1_4.

(2) We have $\mathrm{ad}_S F \subset sl(4,R)$. In the same way as in Sec. 3B1 one can show that F is isomorphic to a simple subalgebra of $sl(4,R)$ which contains $Q + M$ (or $P + M$) isomorphically. Only $sl(4,R)$ itself has this property.

2. Embeddings with radical $S = \epsilon(C + P + Q)$ (type II):

The proof is the same as for G-embeddings of type II up to condition (γ) which must be cancelled here. If $\mathrm{ad}_{P+Q}F$ is complex reducible, then it reduces into two three-dimensional irreps. containing $so(3) + so(3)$. This gives (c) $\mathrm{ad}_{P+Q}F = so(3)$ or (d) $\mathrm{ad}_{P+Q}F = sl(3,R)$ because the only semisimple matrix subalgebras in $sp(3,R)$ containing $so(3) + so(3)$ and leaving P and Q separately invariant, are $so(3) + so(3)$ and $sl(3,R) + sl(3,R)$. If $\mathrm{ad}_{P+Q}F$ is complex irreducible in $sp(3,R)$ we have, as before, (a) $\mathrm{ad}_{P+Q}F = sp(3,R)$ or (b) $\mathrm{ad}_{P+Q}F = (1_3 \times sp(1,R)) \oplus (so(3) \times 1_2)$. There exist embeddings of all these types.

4. SIMPLE LIE ALGEBRAS WITH G-EMBEDDINGS

The technique applied to classify real simple L with G-embeddings, which uses as first step embeddings of the complex Galilei algebra $\overline{G}$ into complex simple Lie algebras, is sketched in Sec. 4A. The results are presented in Sec. 4B. A guideline for the proofs is given in Secs. 4C and 4D for complex and real embeddings, respectively.[13]

A. Real and complex embeddings

1. The set of all real simple Lie algebras L splits[14,15] into two disjoint sets $\mathfrak{M}_1$ and $\mathfrak{M}_2$: The set $\mathfrak{M}_1$ contains all real forms, i.e., all real simple L with a complex extension $\overline{L}$ which is also simple. $\mathfrak{M}_2$ contains all realifications, i.e., all real simple L with a complex extension $\overline{L}$ such that $\overline{L} = \overline{L'} \oplus \overline{L'}$ with $\overline{L'}$ being a complex simple Lie algebra.[16]

$$\mathfrak{M}_1 = \{L/\overline{L} \text{ simple}\}, \quad \mathfrak{M}_2 = \{L/\overline{L} = \overline{L'} \oplus \overline{L'}, \overline{L'} \text{ simple}\}.$$

For any complex simple $\overline{L'}$ there exists exactly one real simple L, such that $\overline{L} = \overline{L'} \oplus \overline{L'}$. Let $\mathfrak{L}$ be the set of all complex simple Lie algebras. Then we have a bijective mapping ϕ: $\mathfrak{L} \to \mathfrak{M}_2$ and ϕ puts the well-known classification for $\mathfrak{L}$ into a classification for $\mathfrak{M}_2$.

2. Consider now G-embeddings. Because of Lemma 1 part B one can prove that $\overline{L} \in \mathfrak{L}$ possesses a $\overline{G}$-embedding iff the corresponding $\phi(\overline{L}) \in \mathfrak{M}_2$ has a G-embedding To have all realifications with G-embeddings it is necessary to determine the set $\mathfrak{L}_2$ of all $\overline{L}$ with $\overline{G}$-embedding. So complex embeddings of the complex $\overline{G}$ become a useful tool for the classification of realifications with G-embeddings; they are also needed for the case of G-embeddings into real forms.

B. The results

1. We recall[10] first that the set $\mathfrak{L}$ of complex simple Lie algebras contains four chains: $A_r, B_r, C_r (r \geq 1)$ and $D_r (r \geq 3)$ and the exceptional cases G_2, F_4, E_6, E_7, E_8. There are the isomorphism $A_1 \approx B_1 \approx C_1, B_2 \approx C_2$ and $A_3 \approx D_3$. The following inclusions are valid: $A_r \subset A_{r+1}$, $B_r \subset B_{r+1}, C_r \subset C_{r+1}, D_r \subset D_{r+1}, B_r \subset D_{r+1}, B_3 \subset F_4 \subset E_6 \subset E_7 \subset E_8$ (the list is not complete). Hence $\mathfrak{L}$ can be split into disjoint sets $\mathfrak{L}_1 = \{A_r (r \leq 4), B_r (r \leq 2), C_r (r \leq 3), D_3, G_2\}$ and $\mathfrak{L}_2 = \mathfrak{L} - \mathfrak{L}_1$.

2. Using the techniques indicated in Sec. 4C we show

Lemma 4: The complex Lie algebras A_4, B_2, C_3, and G_2 possess no $\overline{G}$-embeddings.

Lemma 5: $\overline{G}$-embeddings into A_5, B_3, and C_4 exist. Nonequivalent $\overline{G}$-embeddings into A_5 can be labelled with a complex number δ. All $\overline{G}$-embeddings into B_3 are equivalent and all G-embeddings into C_4 are equivalent.

Lemmas 4 and 5 give together with the stated inclusions [each of the Lie algebras in $\mathfrak{L}_2$ has A_5, B_3 and (or) C_4 as subalgebra]:

Theorem 1: Complex simple Lie algebras $\overline{L}$ have a G-embedding iff

$\overline{L} \in \mathfrak{L}_2 = \{A_r(r \geq 5), B_r(r \geq 3), C_r(r \geq 4), D_r(r \geq 4), F_4, E_6, E_7, E_8\}$.

This gives a complete answer for the complex case.

3. As mentioned in Sec. 4A2 a solution for complex $\overline{L}$ gives a solution for those real L which are realifications. So we have:

Theorem 2: Real simple Lie algebras $\phi(L)$ which are realifications have a G-embedding iff $\overline{L} \in \mathfrak{L}_2$ (with $\mathfrak{L}_2$ given in Theorem 1).

The real forms with G-embeddings were determined as described in Sec. 4D for the three lowest-dimensional Lie algebras in the A-, B- and the combined B-D chain in $\mathfrak{L}_2$. A complicated discussion of the involutions of compact real forms and of special properties of G give for A_2-, B_3, and C_4:

Theorem 3: Only the following real forms of A_5, B_3, and C_4 have a G-embedding:

$sl(6,R)$ as a real form of A_5,
$so(5,2)$ as a real form of B_3,
$sp(4,R)$ as a real form of C_4.

4. We add the following remarks: We have not calculated all real forms with G-embeddings but only those for A_5, B_3, C_4. However, the inclusions between real forms give, together with Theorem 3, some information on G-embeddings into higher-dimensional real forms.

In the case of $so(5,2)$ and $sp(4,R)$ there is only one G-embedding up to equivalence (see Sec. 4.4). In general [e.g., in the case of $sl(6,R)$] one has to distinguish between different equivalence classes of embeddings. For some physical applications of embeddings it is useful to know these different equivalence classes.[17]

C. Complex simple Lie algebras with $\overline{G}$-embeddings

1. Suppose that $\overline{L}$ has a $\overline{G}$-embedding. Then every solvable subalgebra of $\overline{G}$ is contained in a maximal solvable subalgebra $\overline{A}$ of $\overline{L}$. This fact can be used in connection with

Morozov's theorem[18]: Let $\overline{A}_1$ and $\overline{A}_2$ be two maximal solvable subalgebras of the complex simple Lie algebra $\overline{L}$. Then there exists an automorphism σ of $\overline{L}$ with $\sigma(\overline{A}_1) = \overline{A}_2$.

One of the maximal subalgebras in $\overline{L}$ can be identified as follows: Let $\overline{H}$ be a Cartan subalgebra of $\overline{L}$, let Σ be the corresponding root system and let Π and Σ_+ be the systems of simple and positive roots. Then the basis $\mathfrak{L}_w = \{h_{\alpha_i}, e_\alpha, e_{-\alpha}/\alpha_i \in \Pi, \alpha \in \Sigma_+\}$ in $\overline{L}$ is called Weyl basis. The h_{α_i}, $\alpha_i \in \Pi$ span $\overline{H}$ and $e_\alpha, e_{-\alpha}$, $\alpha \in \Sigma_+$ span two subalgebras $\overline{N}_+, \overline{N}_-$ and $\overline{L} = \overline{H} + \overline{N}_+ + \overline{N}_-$ holds. With this notation, $\overline{H} + \overline{N}_+$ is a maximal solvable subalgebra in $\overline{L}$.

2. The application of this result to $\overline{G}$ implies that the solvable subalgebra $((\overline{C} \oplus \overline{T} \oplus \overline{P}) \subset \overline{Q}) \subset (\overline{M}_+ + \overline{M}_3) = \overline{S}_{10} \subset \overline{G}$, $\overline{M}_+$ and $\overline{M}_3$ spanned by $im_1 + m_2$ and m_3, respectively must be contained isomorphically in $\overline{H} + \overline{N}_+$ up to automorphism of $\overline{L}$, if $\overline{L}$ has a $\overline{G}$-embedding. Therefore one has to identify $\overline{S}_{10}$ in $\overline{H} + \overline{N}_+$ and, furthermore, identify the remaining element $m_- = im_1 - m_2 \in \overline{G}$ in $\overline{L}$.

D. Realifications and real forms with G-embeddings

1. All realifications with G-embeddings can be simply found from the $\overline{G}$-embeddings into $\overline{L} \in \mathfrak{L}_2$.

2. Real forms of a complex simple n-dimensional $\underline{L}$ can be considered as real n-dimensional subspaces of $\overline{L}$, which close to a real Lie algebra. Let ϵ be a fixed embedding of G into a real Lie algebra, L being a real form of $\overline{L}$. Then we have the following inclusions of real and complex subalgebras of $\overline{L}$: $\epsilon(G) \subset \overline{\epsilon(G)} \subset \overline{L}$, $\epsilon(G) \subset L \subset \overline{L}$. As result of our discussion of the complex case we have in Lemma 5 all embeddings of $\overline{G}$ in A_5, B_3, and C_4 up to automorphism of $\overline{L}$. Because of the special structure of G one can show, moreover, that all real subalgebras $G' \subset \overline{\epsilon(G)} \subseteq \overline{L}$, $G' \approx G$, which can be completed to real forms of $\overline{L}$, are conjugate[19] to $\epsilon(G)$ by inner automorphisms of $\overline{L}$. Furthermore, the parameter δ which labels the non equivalent $\overline{G}$-embeddings into A_5 must be real if there exists a real form L of A_5 with $\epsilon(G) \subset L \subset A_5$. Therefore, for B_3 and C_4 we have to consider only one embedding of $\overline{G}$ into $\overline{L}$, and for A_5 a set of $\overline{G}$-embeddings labelled by a real number. In each case the usual complex transformation $G \to \hat{G}$ ($\hat{G}$ denotes the hyperbolic Galilei algebra, see Appendix B) yields $\epsilon(\hat{G}) \subset L_w$ with L_w being a real form, called Weyl form of $\overline{L}$ (L_w is the real linear hull of the Weyl basis $\mathfrak{L}_w$). Similarly to the methods of Gantmacher,[15] one can prove that all other real forms of $\overline{L}$ can be found from L_w via a complex linear transformation V which fulfills $\overline{V}V^{-1} \in \mathrm{Aut}\ \overline{L}$, with $\overline{V}$ being the complex conjugate of V in the Weyl basis. Hence we must determine those V which (1) transform $\epsilon(\hat{G})$ into $G' \approx \epsilon(G)$ and (2) fulfill $\overline{V}V^{-1} \in \mathrm{Aut}\ \overline{L}$. With V the involution J which gives the real form is known. We have calculated these V for A_5, B_3, and C_4.

In each of the cases B_3 and C_4 there is only one real form with G-embeddings [$so(5,2)$ and $sp(4,R)$, respectively] and for each of these real forms all G-embeddings are equivalent. For A_5 we get only the real form $sl(6,R)$ which has G-embeddings, but there are nonequivalent G-embeddings into $sl(6,R)$ labelled by the real number δ.

5. DISCUSSION AND EXAMPLES

The problem of how to identify a Lie algebra with G-embeddings as dynamical algebra of a physical system is discussed in Sec. 5A. Using some properties of integrable representations, the identification problem is solved for those L listed in case II of Lemma 1. A representation theory for the conformal Galilei algebra is indicated. The Lie algebras with Galilei embeddings already used in particle physics are given in Sec. 5B.

A. The center of $\epsilon(G)$ and $\epsilon(G_0)$

1. If an irreducible and integrable representation U_i of a dynamical algebra D is used for the description of a non relativistic system, one has to identify those generators which are not contained in the Galilei subalgebra $\epsilon(G)$ of D with infinitesimal transformations, e.g., Hamiltonians. The identification can depend on the representation. The generators of $\epsilon(G) \subset D$ should retain their physical interpretation. Applied to the mass operator $U_i(\epsilon(c))$, this implies [with the notation $\epsilon(G) \equiv G$]: If $U_i(c) = im\mathbb{1}$, the system described through U_i has a fixed mass m and corresponds to a particle moving in a potential.

Because U_i is assumed to be faithful a necessary condition for $U_i(c) = im\mathbb{1}$ is that D is not semisimple and that $c \in S$ with S being the radical of D. This condition is always fulfilled for D with G-, G_0-embeddings and minimal radical $S \neq 0$ (see Lemma 1). The mass m is one of the labels which characterize the irreducible representations U_i. The restriction of U_i to the Heisenberg subalgebra $H_3 = (C \oplus P) \subset Q$ yields a quasi-

irreducible representation of H_3, i.e., a degenerated irreducible representation. The irreducible representations of H_3 are uniquely, up to unitary equivalence, determined by m.

If $U_i(c)$ is not a multiple of the identity, then U_i describes, e.g., the relative motion of a two-particle system.

A similar discussion for the physical $so(3)$ subalgebra in $D = (S \subset F_1) \oplus F_2$ (see Sec. 2), which can have projection parts in F_2, shows that U_i yields systems which can have a spin spectrum.

2. We use the results for a solution of the identification problem for nonsemisimple D with minimal radical. Because $U_i|H_3$ is quasi-irreducible, one can express the generators in the Levi factor F of D on a dense set $\mathfrak{D}_D^{an}$ of analytic vectors for U_i as functions of $U_i(q_k), U_k(p_k), k = 1, 2, 3$, and a physical interpretation is (formally) possible. The procedure can be simplified: Let $\epsilon(D)$ and $\epsilon(H_3)$ be the universal enveloping algebras[20] of D and H_3. A basis in $\epsilon(H_3)$ is given by a set of polynomials of the partly noncommuting variables p_k, q_k, c. Let U be an integrable representation of H_3 and $U(\epsilon(H_3))$ be the image of $\epsilon(H_3)$ under U. Assume now that:

(AI) There exists a homomorphism $\eta: \epsilon(D) \to \eta(\epsilon(D)) \subset U(\epsilon(H_3))$ with $\eta(x) = U(x)$ for $x \in \epsilon(H_3)$ and $\eta(x)$ skew symmetric for $x \in D \in \epsilon(D)$.

If the operators $\overline{\eta(x)}, x \in D$, obtained from $U(H_3)$ through η, form an integrable representation $\tilde{U}$ of D, then η yields an identification of the generators of $U_i(D)$ as polynomials in $U(p_k), U(q_k) \in U(H_3)$. To prove the existence of $\tilde{U}$, a second assumption is needed:

(AII) H_3 is an ideal in D.

Then the extension theorem of Nelson[21] can be applied: If AI and AII are fulfilled, then the representation U of H_3 can be extended to an integrable representation $\tilde{U}$ of D, called Nelson extension of U to D. One can show that $\tilde{U}$ is even unique, i.e., all integrable representations of D with ideal H_3 are Nelson extensions of the representation U of H_3. We now check the assumptions. The homomorphism η exists (see Ref. 4) and H_3 is an ideal for all L listed in case II of Lemma 1 and its corollary. Hence: The integrable representations of $H_3 \subset sp(3,R), H_3 \subset (sp(1,R) \oplus so(3)) \approx G_\subset, H_3 \subset so(3) \approx G_0, H_3 \subset sl(3,R)$, with $H_3 = (C \oplus P) \subset Q$ are Nelson extensions of integrable representations of H_3 and the identification of the generators as polynomials in $U(q_k)$, $U(p_k)$ up to unitary equivalence is possible. An irreducible representation of D can be obtained by Nelson extension only from an irreducible representation U of H_3 and describes spinless particles.

3. With the remarks given above the physical situation of the dynamical algebras listed in case II of Lemma 2 and its corollary can be clarified. The other non semisimple Lie algebras have an abelian radical and the Nelson extension cannot be applied.[22]

B. Examples

1. Only the nonsemisimple Lie algebras with G-embeddings listed in case II of Lemma 1 were used up to now.

A dynamical algebra $D^t = H_3 \subset sp(3,R)$ (case II-a) was proposed in Ref. 4 as a Lie algebra of a limitable dynamical group. A representation theory of D^t was constructed using Nelson extensions and it was shown, using the method sketched before, that $U_i(D^t)$ describes spinless one particle systems with Hamiltonians of the type $H = \Sigma_{ij}(a_{ij}p_ip_j + b_{ij}q_iq_j + c_{ij}(p_iq_j + q_ip_j))$ which contain also Hamiltonians with a subalgebra of $sp(3,R)$, e.g., $so(2,1)$, as spectrum generating algebra. With a special method to treat the spin part in unitary representations of the Galilei group, the model was extended in Ref. 6 to particles with spin using $D_s^t = D^t \oplus so(3)_I$ as dynamical algebra. Here the physical $so(3)$ subalgebra has projection parts in D^t and $so(3)_I$

2. A subalgebra of D^t, the conformal Galilei algebra $G_c = H_3 \subset (sp(1,R) \oplus so(3))$ (case II-b), was first derived by (Hagen[5] and later by Niederer[23]) as the maximal kinematical invariance algebra of the Schrödinger operator $(2m\partial_0 - \partial_i\partial_i)$ of a free massive nonrelativistic particle.

The generators in $G_\subset$ not being generators of $\epsilon(G)$ were identified with dilatations and with a kind of conformal transformations. $G_\subset$ was also used in Ref. 5 to discuss scale and conformal transformations in a Galilei invariant field theory.

A representation theory for $G_\subset$ was indicated in Ref. 24 and can be put in a general form via the Nelson extensions of representations of the Heisenberg algebra as shown in Sec. 5A2. The generators in a representation U of $G_\subset$ were identified in Ref. 24 as functions of $U(q_k), U(p_k)$.[25] The $G_\subset$-theory is contained in the D^t-theory because $G_\subset \subset D^t$ and because the Nelson extension of $U(H_3)$ to D^t, restricted to $G_\subset$, gives the Nelson extension of $U(H_3)$ to $G_\subset$. Hence $G_\subset$ describes as D^t a single particle moving in a potential; new internal degrees of freedom are not related to the system.

3. From a more geometrical point of view it would be interesting to construct a dynamical algebra D_γ with a Galilei and a Poincaré embedding $\epsilon(G)$ and $\epsilon'(\hat{P})$, respectively, such that $\epsilon(G) \cap \epsilon'(\hat{P})$ is the Euclidean algebra E_3 in three space dimensions. An algebra of this type, the "relativistic" Galilei algebra $G_\gamma = ((C \oplus T \oplus P_4) \subset Q_4) \subset so(3,1)$, was given in Ref. 9[26] and its representation theory in Ref. 27. The method of Nelson extensions is also applicable to G_γ. The radical in G_γ is not minimal.

One may ask, whether there are more Lie algebras with this mixed G- and $\hat{P}$-embedding. From our classification we find the following result: The Lie algebras $L \approx (C + T + P) \subset sl(5,R)$, $L \approx (C + T + P) \subset so(4,1)$ (Lemma 2, case I) and $L \approx so(5,2)$ (Theorem 3) have the properties: (1) L has a Galilei embedding $\epsilon(G) \subset L$ (2) L has a Poincaré embedding $\epsilon'(\hat{P}) \subset L$ (3) $\epsilon(G) \cap \epsilon'(\hat{P}) = \epsilon(Q \subset so(3))$, i.e., the Euclidean algebra in three space dimensions.

4. Consider now the Galilei algebra $G^{(2)}$ in two space dimensions. Here $G^{(2)} \subset \hat{P}$ and a $G^{(2)}$-embedding into simple Lie algebra $so(4,2)$ was applied[8] to extreme relativistic processes.

6. CONCLUDING REMARKS

The classification and the discussion of the physical properties of possible dynamical algebras D for nonrelativistic systems show, that even in the simple case that D has a G-embedding there are group theoretical difficulties. The main problem is the identification of a given algebra as the dynamical algebra of a single or more particle system. The assumption that G is a subalgebra of D, i.e., that the description of a nonrelativistic free particle is contained in the representation of D gives only a hint but not a general method to identify D. It is necessary to put more physical information in the approach, e.g., a group theoretical ver-

sion of a limiting procedure between interacting and free systems or a general group theoretical approach to the scattering problem.

ACKNOWLEDGMENTS

The authors would like to thank Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste.

APPENDIX A. NOTATION

The semidirect sum and direct sum of two Lie algebras A, B are denoted by $A \mathbin{\dot{+}}_{\rho} B$ (or simply $A \mathbin{\dot{+}} B$) and $A \oplus B$. $\rho : b \to \rho_b$ is a homomorphism mapping B into the Lie algebra Der A of derivations of A, i.e., $\rho_b(a) = (\mathrm{ad}_A b)(a) = [b, a]$, $a \in A$, $b \in B$, if ad_A is the adjoint representation restricted to the ideal A. ρ defines the semidirect sum. $N + M = A$ means that the linear subspace A of a Lie algebra B is the direct vector space sum of the two linear subspaces N and M of B, $+$ means the direct sum of matrices and representations, respectively, and $\times$ the Kronecker product of matrices and representations, respectively.

The n-dimensional zero matrix and unit matrix are written as 0_n and $\mathbb{1}_n$. For fixed i, k the matrix E_{ik} is defined by $(E_{ik})_{lm} = \delta_{il}\delta_{km}$. T_n is the n-dimensional abelian real Lie algebra.

If for Lie algebras A, B there exists $A' \subset B, A' \approx A$, we write $A \subsetneq B$. $\bar{A}$ always denotes the complex extension of A.

APPENDIX B. THE GALILEI ALGEBRA

We define the Galilei algebra G via its Levi decomposition $G = R \mathbin{\dot{+}}_{\sigma} M$ and a decomposition of $R = (C \oplus T \oplus P) \mathbin{\dot{+}}_{\tau} Q$ with $M \approx so(3)$, $P \approx Q \approx T_3$, $T \approx C \approx T_1$ and with standard basis $\mathfrak{L} = \{c, t, p_i, q_i, m_i / i = 1, 2, 3\}$. The semidirect sums in G and R are given by homomorphisms σ: $M \ni m \to \rho_m \in \mathrm{Der}\, R$, $\rho_m = \mathrm{ad}_R m$, and τ: $Q \ni q \to \tau_q \in \mathrm{Der}\,(C + T + P)$ such that the nonzero commutators are $[m_i, p_k] = \epsilon_{ike} p_e$, $[m_i, q_k] = \epsilon_{ike} q_e$, $[p_i q_k] = \delta_{ik} c$, $[q_i, t] = p_i$. The center of G is C. G has the following solvable ideals $J_1 = C$, $J_2 = C + P$, $J_3 = C + T + P$, $J_4 = C + P + Q$, $J_5 = C + T + P + Q = R$, $J_6 = 0$. The maximal solvable ideal, i.e., the radical of G, is J_5.

The complex extension of G is $\bar{G} = \bar{R} \mathbin{\dot{+}}_{\bar{\sigma}} \bar{M}$ with $\bar{R}$ being the complex extension of R and $\bar{M} \approx B_1$. There are two nonisomorphic real forms of $\bar{G}$: the (spherical) Galilei algebra G given above and the hyperbolic Galilei algebra $\hat{G} = \hat{R} \mathbin{\dot{+}}_{\hat{\sigma}} \hat{M}$ with $\hat{R} \approx R$ and $M \approx so(2, 1)$.

The Galilei algebra is the central extension of $\tilde{G} = \tilde{R} \mathbin{\dot{+}} M$ with $\tilde{R} \approx R/C$.[28] We denote $G_0 = R_0 \mathbin{\dot{+}} M$, $R_0 = C + P + Q$, as geometrical Galilei algebra. Galilei algebras $G^{(n)}$ corresponding to Galilei transformations in n space dimensions are defined as $G^{(n)} = R_n \mathbin{\dot{+}} so(n)$, $P_n \approx Q_n \approx T_n$, $T \approx C \approx T_1$, $R_n = (C \oplus T \oplus P_n) \mathbin{\dot{+}} Q_n$, and $G^{(m)} \subset G^{(n)}$ for $m \leqslant n$ holds. G refers to $n = 3$, i.e., $G \equiv G^{(3)}$.

* On leave of absence from Universität Osnabrück, Osnabrück, Fed. Rep. Germany. Permanent address: Technische Universität Clausthal, Clausthal, Fed. Rep. Germany.

† On leave of absence from Technische Universität Clausthal, Clausthal, Fed. Rep. Germany.

[1] If only h and the generators of a degeneracy algebra E ($[h, E] = 0$) are elements of D but *not* p_k, q_k and c, then D is called a spectrum generating algebra for h if D has an integrable representation U such that $h = \Sigma\, a_i U(d_i)$ with $\{ d_i \}$ being a basis of D; a_i real.

[2] I. Segal, Duke Math. J. **18**, 221 (1951).

[3] There are only very few mathematical results concerning contractions between representations. E. J. Saletan, J. Math. Phys. **2**, 742 (1961); J. Michelson and J. Niederle, J. Math. Phys. **13**, 23 (1972). Also the related and much simpler problem of defining contractions between Lie algebras in a suitable way is not yet completely solved [see the different definitions in Ref. 2 and E. Inönü and E. P. Wigner, Proc. Nat. Acad. Sci. USA **39**, 510 (1953); H. D. Doebner and O. Melsheimer, Nuovo Cimento **49**, 306 (1967); R. Hermann, Commun. Math. Phys. 3, 53 (1966); A. Nijenhuis and R. W. Richardson, Bull. Am. Math. Soc. **72**, 1 (1966). The suggestion of Segal, Ref. 2, refers to p-limitations contracting Lie algebras of equal dimension.

[4] H. D. Doebner and O. Melsheimer, J. Math. Phys. **9**, 1638 (1968).

[5] C. R. Hagen, Phys. Rev. D **5**, 377 (1972).

[6] H. D. Doebner and O. Melsheimer, J. Math. Phys. **11**, 1463 (1970).

[7] Ph. Meyer, Cargese Lectures (1972).

[8] G. Burdet, H. Perrin and P. Sorba, preprint, Centre de Phys. Theor. Marseille (Feb. 1973).

[9] J. J. Aghassi, P. Roman and R. M. Santilli, Phys. Rev. D **1**, 2753 (1970).

[10] N. Jacobson, *Lie Algebras* (Wiley, New York, 1962).

[11] E. B. Dynkin, Amer. Math. Soc. Transl. Ser. 2 **6**, 245 (1957).

[12] Der S contains Der $G = [((A_P \oplus A_Q) \mathbin{\dot{+}} A_T) \mathbin{\dot{+}} (T_1 \mathbin{\dot{+}} T_2) \mathbin{\dot{+}} A_M$, which is also calculated very easily [cf. G. Marmo and G. S. Whiston, Int. J. Theor. Phys. **6**, No. 4, 293 (1972)].

[13] A detailed version of the proofs will be given elsewhere.

[14] F. R. Gantmacher, Mat. Sb. **5** (47), 101 (1939).

[15] F. R. Gantmacher, Mat. Sb. **5** (47), 217 (1939).

[16] In contradistinction to real forms, the realifications have not been used in physics, except $\phi(B_1) \approx so(3,1)$. The reason may be, that their lowest dimensions are already high, dim $\phi(\bar{L}) = 2$ dim $\bar{L}$.

[17] H. D. Doebner and J. Hennig, ICTP, Trieste, preprint IC/72/108.

[18] The theorem was stated first by Morozov [V. V. Morozov, thesis, Kazan Gos. Univ. (1943)]. The first exact proof was given by Freudenthal [H. Freudenthal and H. de Vries, *Linear Lie Groups* (Academic, New York, 1969)].

[19] Two subalgebras A_1, A_2 of a Lie algebra A are said to be conjugate if there exists an automorphism σ of A with $A_2 = \sigma(A_1)$.

[20] Let L be a Lie algebra, T_L the tensor algebra of the underlying vector space of L, K the ideal of T_L generated by all elements of the form $[l_1, l_2] - l_1 l_2 + l_2 l_1$, $l_1, l_2 \in L$, then the algebra T_L/K is a realisation of the universal enveloping algebra $\epsilon(L)$ of L.

[21] E. Nelson, Ann. Math. **70**, 572 (1959).

[22] Denote by H_n the Heisenberg algebra in n space dimensions. Then there exists for any L an n such that there is a homomorphism $\xi : L \to \epsilon(H_n)$. However, there are only very few results on finite-dimensional subalgebras in $\epsilon(H_n)$, except, e.g., $sp(n,R) \subset \epsilon(H_n)$.

[23] U. Niederer, Helv. Phys. Acta **45**, 802 (1972).

[24] P. Roman, J. J. Aghassi, R. M. Santilli, and P. L. Huddleston, Nuovo Cimento A **12**, 185 (1972).

[25] In G. Burdet and M. Perrin, Lett. Nuovo Cimento **4**, 651 (1972) G_C is realized in terms of second-order polynomials of $U(q_k)$, $U(p_k)$, $k = 1, \ldots, n$ via an injective homomorphism mapping G_C into $U(\epsilon(H_n))$.

[26] A method for the construction of relativistic algebras, L. Castell, Nuovo Cimento **49**, 285 (1967), is based on those $L \supset \epsilon(\hat{P})$ such that a p-contraction of L yields a Lie algebra $L' \supset \epsilon(G)$.

[27] J. J. Aghassi, P. Roman, and R. M. Santilli, Nuovo Cimento A **5**, 551 (1971).

[28] In some papers $\tilde{G}$ and G are called Galilei algebra and central extension of the Galilei algebra.

New wave-operator identity applied to the study of persistent currents in 1D

Daniel C. Mattis*

Belfer Graduate School of Science, Yeshiva University, New York, New York 10033
(Received 31 October 1973)

We show that a large class of backward-scattering matrix elements involving $\Delta k \sim \pm 2k_F$ vanish for fermions interacting with two-body attractive forces in one dimension. (These same matrix elements are finite for noninteracting particles and infinite for particles interacting with two-body repulsive forces.) Our results demonstrate the possibility of persistent currents in one dimension at $T = 0$, and are a strong indication of a metal-to-insulator transition at $T = 0$ for repulsive forces. They are obtained by use of a convenient representation of the wave operator in terms of density-fluctuation operators.

INTRODUCTION

It is usual to express the density-fluctuation operators[1] $\rho(p)$ as bilinear forms of the wave operator $\Psi(x)$. We have recently succeeded in inverting the process, expressing the wave operator as an exponential form of the density fluctuation operators, in the special case of Luttinger's soluble model of interacting fermions in one dimension[2]. While certain aspects of our procedure could obviously be used in other applications[3] or even adapted to the case of electrons in three dimensions, we limit the present application to the challenging question, of whether persistent currents (i.e., *supercurrents*) can exist in one dimension despite arbitrary random scattering potentials. The surprising result is that, for sufficiently attractive two-body forces, a current-carrying state at $T = 0$ can have infinite lifetime regardless of the strength of the scattering mechanisms. Therefore, it is proved rigorously that superconductivity can exist, at $T = 0$, in one dimension, despite the well-known lack of long range order. We also find the converse, that for sufficiently repulsive two-body forces, the lifetime of a current-carrying state at $T = 0$ tends to zero, and the system acquires the attributes of an insulator. The nontrivial generalization of these results to finite temperature is the subject of an ongoing, separate, study.

DETAILS OF THE MODEL

We first recall certain aspects of the soluble many-fermion model[2] under scrutiny. It consists of right-going particles (labeled 1) having constant velocity v_0, and left-going particles (labeled 2) with velocity $-v_0$, with interactions characterized by a two-body potential $V(x - x')$ and coupling constant λ, obeying a Hamiltonian:

$$\mathcal{H} = v_0 \sum_k k(n_{1k} - n_{2k}) + (\lambda/L) \sum_p U(p)[\rho_1(p) + \rho_2(p)][\rho_1(-p) + \rho_2(-p)], \quad (1)$$

where p, k refer to wave numbers, $U(p)$ is the Fourier transform of $V(x - x')$ and the various operators are

$$n_{ik} = a_{ik}^+ a_{ik}, \qquad \rho_i(p) = \sum_k a_{i\,k+p}^+ a_{ik}$$
$$\Psi_i(x) = L^{-1/2} \sum_k a_{ik} e^{ikx} \quad (2)$$

with L = dimension of the space, for purposes of box normalization. The particle-current operator $j_{\rm op}$ takes the form

$$j_{\rm op} = V_0 \sum_k (n_{1k} - n_{2k}). \quad (3)$$

If, for example, at $T = 0$ we set the coupling constant $\lambda = 0$, we find for the eigenvalue of (3) the value

$$j = v_0 (L/2\pi)(k_{1F} + k_{2F}), \quad (4)$$

for, at $T = 0$, the ground state of the non-interacting particles is described by occupation numbers $n_{1k} = 1$ for $-\infty < k < k_{1F}$ and $n_{1k} = 0$ for $k > k_{1F}$, together with $n_{2k} = 1$ for $k_{2F} < k < +\infty$, and $n_{2k} = 0$ for $k < k_{2F}$. In the ground state, $k_{2F} = -k_{1F}$ and no current flows. In general, however, we can have $k_{1F} \neq -k_{2F}$ and the current eigenvalue j will be nonzero. This conclusion is unaffected by the interactions when $\lambda \neq 0$, for $j_{\rm op}$ commutes with both parts of the Hamiltonian $\mathcal{H}$ separately, and j is therefore a good quantum number until a mechanism for decay of the current is introduced into the Hamiltonian.

Accordingly, we introduce a mechanism allowing electrons to be backward scattered from one branch to the other, in order to test the hypothesis of persistent currents. For definiteness, consider a one-body scattering Hamiltonian $\mathcal{H}'$:

$$\mathcal{H}' = \int dx [W(x) \Psi_2^+(x) \Psi_1(x) + \text{h.c.}], \quad (5)$$

where $W(x)$ is a random potential. Because $j_{\rm op}$ does not commute with $\mathcal{H}'$, j is no longer a constant of the motion, and generally decays exponentially:

$$j(t) = j(0) \exp(-t/\tau), \quad (6)$$

where τ = lifetime of the current, is a measure of the strength of the scattering potential $W(x)$ and of the effective density of one-particle states. To probe the latter, we compute the matrix element:

$$M(i \to f) \equiv \langle f | \Psi_2^+(x) \Psi_1(x) | i \rangle, \quad (7)$$

in which $|i\rangle$ is the initial, exact, eigenstate of both $\mathcal{H}$ and $j_{\rm op}$, and $|f\rangle$ is the final eigenstate of these operators. It will be appreciated that if the initial eigenvalue of $j_{\rm op}$ is j, the final eigenvalue is $j - 2v_0$. The matrix elements M enable us to compute the structure of $\mathcal{H}'$ in the Hilbert space of the eigenstates of $\mathcal{H}$. Some of them could be finite, as for noninteracting particles, and then we would have normal decay. But if we find that the matrix elements connecting low-lying states are *all* zero, then there can be no scattering, and the existence of persistent currents is demonstrated. If, on the other hand, some matrix elements are infinite, then we may conclude either that the lifetime τ of a current is zero, or, more accurately, that the effects of $\mathcal{H}'$ are too profound to be taken into account by perturbation theory (for it causes an insulator phase to replace the metallic phase, and this

should presumably be taken into account before the effects of the two-body forces.) In the following, we shall find all these possibilities to be realizable, depending on the sign and magnitude of $U(0) = \int dx V(x)$.

PRELIMINARY COMPUTATIONS

We start by recalling the unitary transformation S which renders $\exp(iS)\mathcal{H}\exp(-iS)$ diagonal. It has the form[2]

$$S = (2\pi i/L)\sum_{\text{all } p} p^{-1}\varphi(p)\rho_1(p)\rho_2(-p). \tag{8}$$

We recall that, owing to the peculiarities of a filled Fermi sea, the ρ's do not all commute, but obey the commutation relations

$$[\rho_i(p), \rho_j(-p)] = \epsilon_i \delta_{ij} pL/2\pi \tag{9}$$

in which $\epsilon_1 = -1$ and $\epsilon_2 = +1$. The correct value of φ to diagonalize $\mathcal{H}$ is found to be

$$\varphi(p) = -\tfrac{1}{4}\ln[1 + 2\lambda u(p)], \tag{10}$$

where $u(p) \equiv U(p)/\pi v_0$, so that

$$e^{iS}\mathcal{H}e^{-iS} = (2\pi v_0/L)\sum_{p>0}[1 + 2\lambda u(p)]^{1/2} \times [\rho_1(p)\rho_1(-p) + \rho_2(-p)\rho_2(p)] + W_1. \tag{11}$$

Making use of the commutation relations (9), one sees that $\mathcal{H}$ is reduced to a set of noninteracting harmonic oscillators having characteristic energy $E(p) = v_0 p[1 + 2\lambda u(p)]^{1/2}$. W_1 is the vacuum renormalization energy,

$$W_1 = \frac{Lv_0}{2\pi}\int_0^\infty dp\, p\{[1 + 2\lambda u(p)]^{1/2} - 1 - \lambda u(p)\}. \tag{12}$$

To obtain the effect of S on $\mathcal{H}'$, we have found the following new operator identities to be extremely convenient:

$$\Psi_1(x) \langle = \rangle \frac{e^{ik_{1F}x}}{L^{1/2}} \exp\left(\frac{-2\pi}{L}\sum_{p>0} p^{-1}\rho_1(p)e^{-ipx}\right) \times \exp\left(\frac{2\pi}{L}\sum_{p>0} p^{-1}\rho_1(-p)e^{ipx}\right) \tag{13a}$$

and

$$\Psi_2(x) \langle = \rangle \frac{e^{ik_{2F}x}}{L^{1/2}} \exp[i\pi\int dx'\,\Psi_1^\dagger(x')\Psi_1(x')] \times \exp\left(\frac{-2\pi}{L}\sum_{p>0} p^{-1}\rho_2(-p)e^{ipx}\right) \times \exp\left(\frac{2\pi}{L}\sum_{p>0} p^{-1}\rho_2(p)e^{-ipx}\right) \tag{13b}$$

with p = integer $\times 2\pi/L$. The double arrows indicate that the identities hold in a special sense only; supposing $|F,N\rangle$ to be the ground state Fermi sea corresponding to N particles of type 1 and Ω_1 to be an arbitrary function of the $\rho_1(\pm p)$ operators, we have

$$\Psi_1(x)\Omega_1|F,N+1\rangle = \frac{e^{ik_{1F}x}}{L^{1/2}} \exp\left(\frac{-2\pi}{L}\sum_{p>0} p^{-1}\rho_1(p)e^{-ipx}\right) \times \exp\left(\frac{2\pi}{L}\sum_{p>0} p^{-1}\rho_1(-p)e^{ipx}\right)\Omega_1|F,N\rangle \tag{14}$$

and similarly for $\Psi_2(x)$ and for Hermitean conjugate operators $\Psi_i^\dagger(x)$. Since any state in our Hilbert space can be written in the form $\Omega_1|F,N\rangle$, Eqs. (13), and their Hermitean conjugate relations, are operator identities in every practical sense. We note also that they are kinematical identities, independent of the nature of the dynamical interactions or of the magnitude of the coupling constant.

Applying the unitary transformation S to the wave operators,

$$\Psi_i(x) \to \exp(iS)\Psi_i(x)\exp(-iS), \tag{15}$$

we obtain expressions that are readily evaluated using the bosonlike commutation relations, Eq. (9). We cast such expressions into normal ordering [$\rho_1(-p)$ to the right of $\rho_1(+p)$, $\rho_2(p)$ to the right of $\rho_2(-p)$, with $p>0$]. As an example, consider the bilinear form:

$$\begin{aligned}\Psi_1^\dagger(x')\Psi_1(x) \to \frac{1}{L}\exp[ik_{1F}(x-x')] &\exp\left(\frac{2\pi}{L}\sum_{p>0} p^{-1}e^{ip(x'-x)}\right) \\ &\times \exp\left(-\frac{4\pi}{L}\sum_{p>0} p^{-1}[1-\cos p(x-x')]\sinh^2\varphi_p\right) \\ &\times \exp\left(\frac{2\pi}{L}\sum_{p>0} p^{-1}\rho_2(-p)(e^{ipx}-e^{ipx'})\sinh\varphi_p\right) \\ &\times \exp\left(-\frac{2\pi}{L}\sum_{p>0} p^{-1}\rho_2(p)(e^{-ipx}-e^{-ipx'})\cosh\varphi_p\right) \\ &\times \exp\left(-\frac{2\pi}{L}\sum_{p>0} p^{-1}\rho_1(p)(e^{-ipx}-e^{-ipx'})\cosh\varphi_p\right) \\ &\times \exp\left(\frac{2\pi}{L}\sum_{p>0} p^{-1}\rho_1(-p)(e^{ipx}-e^{ipx'})\cosh\varphi_p\right)\end{aligned} \tag{16}$$

This generalizes an earlier result,[4] the calculation of the ground-state expectation value by an entirely different and more laborious technique:

$$\langle F|\Psi_1^\dagger(x')\Psi_1(x)|F\rangle = \frac{1}{L}e^{ik_{1F}(x-x')}\sum(x'-x) \times \exp\left(\frac{-4\pi}{L}\sum_{p>0} p^{-1}[1-\cos p(x-x')]\sinh^2\varphi_p\right). \tag{17}$$

Here, and elsewhere, the following identity proves helpful:

$$\sum(R) \equiv \exp\left[\frac{2\pi}{L}\sum_{p>0} p^{-1}e^{ipR}\right] = \sum_{p>0}^{\infty} e^{ipR} = [1 - e^{i2\pi R/L}]^{-1} \tag{18}$$

We have denoted this quantity $\sum(R)$ for typographical convenience.

SCATTERING MATRIX ELEMENT

The state of lowest energy carrying a current j is denoted the ground state for current j, and symbolized $|F;j\rangle$. At $T = 0$ one may always assume the initial state to be a state of this type.

We therefore calculate the transition matrix element from an initial state, the ground state of current $j > 0$, to a final state, which can be either the ground state of current $j - 2v_0$, or any excited state of the same current. The total rate of decay out of the initial state into the final states, subject to the requirement of conservation of energy, determines the lifetime τ of the current j. It shall, however, not be necessary for us to calculate τ in any detail in cases when $U(0) \neq 0$, for we shall find $\tau = 0$ when $U(0) > 0$ and $\tau = \infty$ when $U(0) < 0$.

We apply (15) to the right-hand sides of Eqs.(13) and their Hermitean conjugates, to obtain

$$\Psi_2^*(x)\Psi_1(x) \to (e^{i(k_{1F}-k_{2F})x}/L)e^{-\alpha}e^{B_2^\dagger}e^{-B_2}e^{-A_1^\dagger}e^{A_1}, \tag{19}$$

in which we note that the phase factor $k_{1F} - k_{2F} \sim 2k_F$ corresponds to backward scattering across the Fermi surface, and

$$\begin{aligned}\alpha &= \frac{2\pi}{L}\sum_{p>0} p^{-1}(e^{2\varphi(p)}-1) = \int_0^\infty dp\, p^{-1}(e^{2\varphi(p)}-1)\\ &= \int_0^\infty dp\, p^{-1}\{[1+2\lambda u(p)]^{-1/2}-1\},\end{aligned} \tag{20}$$

$$A_1 = \frac{2\pi}{L}\sum_{p>0} p^{-1}\rho_1(-p)e^{ipx}e^{\varphi(p)},$$

$$B_2 = \frac{2\pi}{L}\sum p^{-1}\rho_2(p)e^{-ipx}e^{\varphi(p)}.$$

Therefore the ground state-ground state matrix element, which we write $M(F \to F)$ in an obvious notation, takes on the value:

$$M(F \to F) = (1/L)e^{i(k_{1F}-k_{2F})x}e^{-\alpha}. \tag{21}$$

The magnitude of $M(F \to F)$ depends on α, and this in turn depends sensitively on $\varphi(0) \equiv \lim_{p\to 0}\varphi(p)$. A two-body interaction which is attractive on the whole has $U(0) < 0$, hence, by Eq. (10), $\varphi(0) > 0$. Such an interaction implies a positive α which is logarithmically divergent $(+\infty)$, and thus a vanishing matrix element. Similarly, a two-body interaction which is repulsive on the whole implies a negatively divergent value of α, hence an infinite matrix element. When both $\varphi(0)$ and $\varphi(\infty)$ are zero, the integral defining α is well-behaved and the matrix element is finite. It should be noted that any two-body interaction $V(x-x')$, the spatial integral of which is nonzero, corresponds to a Fourier transform $U(p \to 0) \neq 0$, hence to a divergent α (negatively or positively divergent according as to whether the interaction is repulsive or attractive). In all such cases the matrix element $M(F \to F)$ is nonanalytic in the coupling constant λ at $\lambda = 0$, despite the persistence of a "sharp Fermi surface" to finite values of λ (cf. discussion in Ref. 4). In the case of potentials which are neither repulsive nor attractive on the whole, $U(p \to 0) = 0$, α is finite and is a continuous function of λ. In such cases only is the decay of an induced current qualitatively the same as for noninteracting particles.

STRUCTURED FINAL STATES

Concerning the divergence in α arising primarily from long wavelengths $(p \to 0)$, it is legitimate to wonder whether it is not possible to cancel this divergence through an appropriate linear combination of low-lying excited states. We shall examine two typical compound final states in some detail:

$$\langle Q^{(1)}| \equiv \langle F; j - 2v_0 | a_{1k_{1F}} a^\dagger_{1(k_{1F}-Q)}, \tag{22a}$$

$$\langle Q^{(2)}| \equiv \langle F; j - 2v_0 | a^\dagger_{2k_{2F}} a_{2(k_{2F}-Q)}. \tag{22b}$$

These have the advantage of being eigenstates of the free-fermion Hamiltonian ($\lambda = 0$). It is of interest to see whether the matrix elements $M(F \to Q^{(i)})$ vanish or diverge under the same conditions as $M(F \to F)$. We start the analysis under the supposition that the forces are essentially attractive $[\varphi(p) > 0]$; a separate analysis will follow in the case of essentially repulsive forces.

After some elementary manipulations based on Eqs. (2) and (13), we obtain

$$\begin{aligned}M(F \to Q^{(1)}) &= (1/L)e^{-\alpha}e^{i(k_{1F}-k_{2F}-Q)x}\\ &\times (1/L)\int dR e^{-iQR}\sum(R)\\ &\times \exp\left((2\pi/L)\sum_{p>0} p^{-1}e^{ipR}(e^{\varphi(p)}-1)\right)\end{aligned} \tag{23a}$$

and

$$M(F \to Q^{(2)}) = e^{2iQx}M(F \to Q^{(1)}). \tag{23b}$$

It is therefore sufficient to study the behavior of $M(F \to Q^{(1)})$. If $\varphi(0) \neq 0$ the sum in the exponential is logarithmically divergent, and we manipulate it so as to combine it with the divergent expression in α.

Thus

$$\begin{aligned}&\frac{2\pi}{L}\sum_{p>0} p^{-1}e^{ipR}(e^{\varphi(p)}-1)\\ &= \frac{2\pi}{L}\sum_{p>0} p^{-1}(e^{\varphi(p)}-1) - \frac{2\pi}{L}\sum_{p>0} p^{-1}(1-e^{ipR})(e^{\varphi(p)}-1)\end{aligned} \tag{24}$$

Finally,

$$\begin{aligned}M(F \to Q^{(1)}) &= \frac{1}{L}e^{-\alpha'}e^{i(k_{1F}-k_{2F}-Q)x}\ \frac{1}{L}\int dR\ e^{-iQR}\sum(R)\\ &\times \exp\left(-\int_0^\infty dp\, p^{-1}(1-e^{ipR})(e^{\varphi(p)}-1)\right)\end{aligned} \tag{25}$$

$$\equiv \frac{1}{L}e^{-\alpha'}e^{i(k_{1F}-k_{2F}-Q)x}I(Q), \tag{25a}$$

where

$$\alpha' \equiv \int_0^\infty dp\, p^{-1}e^{\varphi(p)}(e^{\varphi(p)}-1). \tag{26}$$

We note that although $\alpha' < \alpha$ [for the case under consideration, viz., $\varphi(p) > 0$], it is nonetheless infinite when $\varphi(0) \neq 0$. It remains only to study the behavior of $I(Q)$, and to verify that all quantities reduce to the appropriate value when the interaction is turned off. For this purpose, it is most convenient to expand the exponential in a power series about $R = 0$, retaining up to quadratic terms. Thus,

$$\int_0^\infty dp\, p^{-1}(1-e^{ipR})(e^{\varphi(p)}-1) = -i\gamma R + \tfrac{1}{2}\delta R^2 + O(R^3), \tag{27}$$

Where

$$\gamma \equiv \int_0^\infty dp(e^{\varphi(p)}-1), \qquad \delta \equiv \int_0^\infty dp\, p(e^{\varphi(p)}-1), \tag{28}$$

both positive quantities in the case under consideration. Then,

$$\begin{aligned}I(Q) &= \frac{1}{L}\int dR\ e^{-iQR}\sum(R)e^{i\gamma R - \delta R^2/2}\\ &= \frac{1}{L}\sum_{p>0}\int_{-\infty}^{\infty} dR\, e^{i(p-Q+\gamma)R}e^{-\delta R^2/2}\\ &= \frac{1}{L}\sum_{p>0}\left(\frac{2\pi}{\delta}\right)^{1/2}e^{-(p+\gamma-Q)^2/2\delta}\\ &= \frac{1}{(2\pi\delta)^{1/2}}\int_0^\infty dp\, e^{-(p+\gamma-Q)^2/2\delta}.\end{aligned} \tag{29}$$

[In the limit $\lambda \to 0$, both γ and δ vanish and, for any finite positive Q, $I = 1$, and $M(F \to Q^{(1)})$ tends to what

is obviously the correct value for free particles.] For any $\lambda > 0$, I is finite and in the case of $\varphi(0) > 0$ *the matrix element* $M(F \to Q^{(i)})$ *vanishes just as did* $M(F \to F)$.

We now verify that, for repulsive forces, the matrix element diverges. It must first be understood that when $\varphi(p) < 0$, the main contribution to the spatial integral in (25) is from a region near $R = \pm\frac{1}{2}L$. Because of periodic boundary conditions, we have

$$e^{ip(R\pm L/2)} = -e^{ipR}, \quad \sum(R\pm L/2) = (1 + e^{i2\pi R/L})^{-1}. \tag{30}$$

Therefore we cast (25) in the form

$$M(F \to Q^{(1)}) = (1/L)\, e^{-\alpha''} e^{i(k_{1F}-k_{2F}-Q)x} J(Q), \tag{25b}$$

where

$$\alpha'' \equiv \int_0^\infty dp\, p^{-1}(e^{\varphi(p)} - 1)(e^{\varphi(p)} + 2)$$

and

$$J(Q) \equiv -\frac{1}{L}\int dR\, e^{-iQR}(1 + e^{i2\pi R/L})^{-1} \times \exp\left(\int_0^\infty dp\, p^{-1}(1 - e^{ipR})(e^{\varphi(p)} - 1)\right) \tag{31}$$

We can evaluate $J(Q)$ by the same methods in (27-29), and show it is finite. *Thus, the divergence* $[\alpha'' \to -\infty$ whenever $\varphi(0) < 0]$ *is again confirmed.*

RECAPITULATION AND FUTURE APPLICATIONS

We have found a representation for fermion wave operators in a specific one-dimensional model, in terms of density fluctuation operators, which enables the exact evaluation of rather complicated matrix elements. In applying this to the problem of persistent current we observed that in the case of repulsive two-body forces, $U(0) > 0$, the scattering matrix elements due to impurities become infinite, and in the case of attractive two-body forces, $U(0) < 0$ they vanish. It should be noted that neither the ground state energy W_1, Eq. (12), nor the sharpness of the Fermi surface[4], are so singularly dependent on $U(0)$.

Our finding of what is tantamount to superconductivity, for electrons interacting with attractive forces, is in harmony with the well-known results of the BCS theory of superconductivity for three-dimensional systems. Recently, Heeger and his collaborators[5] have found anomalously large conductivity in certain linear chain molecules (TTF-TCNQ) near a finite temperature ~ 58°K, followed by a rapid decrease in conductivity as the temperature is further decreased. For these experimental facts to be explained on the basis of any one-dimensional model requires a calculation at finite temperature, and, possibly also, considerations of the electron spin and the electron-phonon interactions.

Note added in proof: We have now succeeded in evaluating τ at finite temperature and in taking the electron-phonon forces explicitly into account. (Full details have been submitted for publication elsewhere).

ACKNOWLEDGMENT

It is a pleasure to thank Mr. R. Peña for his helpful collaboration in the early stages of this work.

*Supported in part by Air Force Contract AFOSR-72-2153.

[1] Also often denoted "current operators" as in "current algebras." We reserve that nomenclature for the quantity defined in Eq. (3).

[2] A model introduced by J. M. Luttinger, J. Math. Phys. **4**, 1154 (1963), given an exact solution by D. Mattis and E. Lieb, J. Math. Phys. **6**, 304 (1965), and which is analogous to the Thirring model; see discussion in E. Lieb and D. Mattis, *Mathematical Physics in One Dimension* (Academic, New York, 1966), Chap. 4.

[3] Such as the Tomonoga model; see discussion, references and reprints in Lieb and Mattis, Ref. 2 (1966) and subsequent work by H. Gutfreund and M. Sehick, Phys. Rev. **168**, 418 (1968).

[4] D. Mattis and E. Lieb, J. Math. Phys. **6**, 304 (1965), Sec. V. Eqs. (5.3) *et seq.*

[5] L. B. Coleman, J. J. Cohen, D. J. Sandman, F. G. Yamagashi, A. F. Garito and A. J. Heeger, Solid State Comm. 12, 1125 (1973).

On the application of quasi-Lagrangian coordinates to random shear flows*

M. L. Thiebaux

National Oceanic and Atmospheric Administration, Boulder, Colorado 80302
(Received 17 April 1973; revised manuscript received 5 November 1973)

An ensemble of homogeneous random shear flows with steady linear mean velocity profiles is considered from a purely kinematical point of view. Quasi-Lagrangian coordinates (advected by the mean flow) are used so that a proper orthogonal decomposition of the fluctuation velocity field is possible and periodic boundary conditions can be imposed. The conditions of stationarity in time and incompressibility take on special forms when applied to wave-vector moments. A simple application of the methodology is presented in the construction of a two-dimensional random shear flow.

1. INTRODUCTION

A turbulent incompressible fluid in a shear flow is a dynamically interesting system because the turbulent energy can be directly drawn from the kinetic energy of the mean flow. The attempt to understand the mechanism of this energy transfer under the simplest conditions has compelled experimental and theoretical investigators to consider the idealized case of turbulence in a homogeneous shear flow.

Nearly homogeneous turbulent shear flows, have been generated and studied in the laboratory.[1,2] One significant feature emerging from the experimental studies is evidence that such flows, although very nearly homogeneous in some ways, are not strictly stationary in time in an Eulerian frame advected by the mean velocity. Also it has been demonstrated analytically that time scales must monotonically increase in homogeneous shear flows.[3] However, such flows may be considered quasistationary if the statistical properties of primary concern are only weakly dependent on time displacements. The notion of temporal stationarity is clearly a useful one for analytical purposes and can be applied to mathematically idealized systems. For example, in studies of the dispersion of passive quantities in randomly generated velocity fields (e.g., see Ref. 4), which already have idealistic properties such as specified probability laws, the assumption of temporal stationarity is not only convenient and harmless, but also may be essential for clarity.

In this paper, we shall consider an ensemble of shear flows with homogeneous, random velocity fluctuations and a steady linear mean velocity profile. We wish to develop a kinematic description of the ensemble that satisfies statistical notions of spatial homogeneity and temporal stationarity. A central feature in this development will be the use of quasi-Lagrangian coordinates. In general, the quasi-Lagrangian frame, as defined by Gifford[5] in connection with atmospheric turbulence, is a time-dependent nonorthogonal coordinate system that is advected (and distorted) by the mean flow. Quasi-Lagrangian coordinates (ξ^1, ξ^2, ξ^3) may then be defined by the requirement that a point in real space moving with the mean velocity maps into a fixed point in ξ-space.

The usefulness and significance of quasi-Lagrangian coordinates in the case of a homogeneous shear flow will be discussed in this paper and are first briefly summarized here: (i) The advection and distortion of the velocity fluctuation field by the large-scale mean flow is an intrinsic property of the quasi-Lagrangian methodology. Thus, the sweeping effect that larger scales of motion have on the smaller scales is partly taken into account by the formalism. (ii) The combined statements of homogeneity, incompressibility, and stationarity in terms of quasi-Lagrangian coordinates take on simple forms that are relatively easy to satisfy simultaneously when synthesizing random realizations of shear flows. An example of such a synthesis is given in Sec. 5. (iii) In both random syntheses and dynamical numerical computations, the simultaneous constraints of homogeneity and periodic boundary conditions are possible with quasi-Lagrangian coordinates.

2. STATIONARY HOMOGENEOUS SHEAR FLOWS

Let $\mathbf{x} = (x_1, x_2, x_3)$ denote coordinates in a Cartesian frame labeled I in which the mean velocity is

$$\langle \mathbf{V} \rangle = \omega x_2 \mathbf{e}_1,$$

where ω is the constant rate of shear and $\mathbf{e}_1$ is the unit vector in the x_1 direction. Now consider a second moving Cartesian frame II, with axes parallel to those of I, but with a different origin. Frame II is moving with the mean velocity $\langle \mathbf{V} \rangle$ at its origin. Hence, the mean flow as seen by an observer in II appears identical to that seen by an observer in I. If all other statistical properties of the ensemble of velocity fields appear identical to the two observers, then the shear flow is homogeneous. In this paper we shall consider only second-order Eulerian moments of the velocity field, i.e., the autocorrelation tensor. If only the autocorrelations are observed and appear identical to the two observers, the shear flow is said to be covariance homogeneous.

To be specific, we write the total velocity $\mathbf{V}$ as the sum of its ensemble mean part and a fluctuating part $\mathbf{v}$,

$$\mathbf{V} = \langle \mathbf{V} \rangle + \mathbf{v},$$

and suppose that measurements of the components of $\mathbf{v}$ are made in I at two space-time points denoted $(I, \mathbf{x}^{(k)}, t^{(k)})$, for $k = 1, 2$. Then the ensemble average of the product of the observed velocity components is the second-order moment, or autocorrelation tensor,

$$\langle v_i(I, \mathbf{x}^{(1)}, t^{(1)}) v_j(I, \mathbf{x}^{(2)}, t^{(2)}) \rangle . \tag{2.1}$$

A similar moment can be determined in II where velocity fluctuations are measured at the same times $t^{(1)}$, $t^{(2)}$ and at points whose locations relative to the origin of II are the same vectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}$. This moment we denote

$$\langle v_i(II, \mathbf{x}^{(1)}, t^{(1)}) v_j(II, \mathbf{x}^{(2)}, t^{(2)}) \rangle . \tag{2.2}$$

The statement of covariance homogeneity is that moments (2.1) and (2.2) are identical.

It is possible to rewrite (2.2) as a moment defined in frame I. If at time t the origin $\mathbf{x}_0$ of II is given in I by the equation

$$\mathbf{x}_0 = \mathbf{a} + \omega t a_2 \mathbf{e}_1,$$

where $\mathbf{a}$ is some constant vector, then (2.2) is the same as

$$\langle v_i(I, \mathbf{a} + \omega t^{(1)} a_2 \mathbf{e}_1 + \mathbf{x}^{(1)}, t^{(1)}) v_j(I, \mathbf{a} + \omega t^{(2)} a_2 \mathbf{e}_1 + \mathbf{x}^{(2)}, t^{(2)}) \rangle. \quad (2.3)$$

The equality of (2.1) and (2.3) is the covariance statement of the extension to shear flows of the definition of instantaneous homogeneity discussed by Batchelor.[6]

A set of quasi-Lagrangian coordinates suitable for the shear flow is given by

$$\xi^1 = x_1 - \omega t x_2, \quad \xi^2 = x_2, \quad \xi^3 = x_3.$$

Then, with the notation

$$v_i(\boldsymbol{\xi}, t) \equiv v_i(I, \mathbf{x}, t),$$

where $\boldsymbol{\xi}$ stands for the three contravariant components ξ^1, ξ^2, ξ^3 of the position vector $\mathbf{x}$ as described in the quasi-Lagrangian frame, the equality of (2.1) and (2.3) is equivalent to

$$\langle v_i(\boldsymbol{\xi}^{(1)}, t^{(1)}) v_j(\boldsymbol{\xi}^{(2)}, t^{(2)}) \rangle = \langle v_i(\boldsymbol{\xi}^{(1)} + \mathbf{a}, t^{(1)}) v_j(\boldsymbol{\xi}^{(2)} + \mathbf{a}, t^{(2)}) \rangle \quad (2.4)$$

for any $\mathbf{a}$.

The great advantage of the form (2.4) (identical to the conventional statement of covariance homogeneity) over the equality of (2.1) and (2.3) is that the expansion of the velocity fluctuation field in orthogonal Fourier modes defined with respect to the quasi-Lagrangian coordinates then results in second-order moments that are diagonal in the corresponding wave-vector space. Furthermore, it is easily seen that the imposition of periodic boundary conditions consistent with homogeneity, of much convenience in numerical and theoretical investigations, is possible in quasi-Lagrangian coordinates, and impossible in Cartesian coordinates.

Another view of this expansion was elucidated by Lumley.[7] It can be shown that the expansion in quasi-Lagrangian Fourier modes is the *proper orthogonal decomposition* in the sense of Loève.[8] Lumley used this decomposition, identified as the set of eigenfunctions of the autocorrelation tensor, as the basis for defining the "big eddies" in a nonhomogeneous flow. In our application, we are giving "proper" treatment to the effect of the largest component of the motion, namely the mean flow. Since the autocorrelation tensor is homogeneous in the components of $\boldsymbol{\xi}$, its eigenfunctions are harmonic functions of $\boldsymbol{\xi}$. A harmonic function (Fourier mode) defined with respect to the quasi-Lagrangian coordinates feels the effect of the mean flow by being advected and distorted by it. In contrast, a conventional Cartesian Fourier mode would be simply superimposed on the mean flow and would ignore its advective effect.

To complete this section, we exhibit the familiar statement of Eulerian covariance stationarity in frame I. For any time interval b,

$$\langle v_i(I, \mathbf{x}^{(1)}, t^{(1)}) v_j(I, \mathbf{x}^{(2)}, t^{(2)}) \rangle = \langle v_i(I, \mathbf{x}^{(1)}, t^{(1)} + b) v_j(I, \mathbf{x}^{(2)}, t^{(2)} + b) \rangle.$$

In the remainder of this paper, the terms homogeneous and stationary will be used to mean covariance homogeneous and covariance stationary, respectively.

3. STRUCTURE OF AUTOCORRELATION TENSOR

Assuming periodicity in ξ^1, ξ^2, and ξ^3, we can Fourier analyze the velocity fluctuation field according to

$$v_i(\boldsymbol{\xi}, t) = \sum_{\mathbf{p}} A_i(\mathbf{p}, t) \exp(i\mathbf{p}\cdot\boldsymbol{\xi}) \quad (3.1)$$

where $\mathbf{p}\cdot\boldsymbol{\xi} = p_i \xi^i$, summed over the repeated indices. Without loss of generality, it can be supposed that at $t = 0$, all members of the ensemble of velocity fields have a rectangular periodic lattice as observed in frame I. The periodic lattice in $\boldsymbol{\xi}$-space, which for convenience may be visualized with the three ξ^i axes drawn mutually perpendicular, can then be considered rectangular for all times. For simplicity, we suppose that a periodic element at $t = 0$ is a cube of volume L^3. Hence, the covariant components p_i of wave vector $\mathbf{p}$ are $(2\pi/L)n_i$, $i = 1, 2, 3$, where the n_i are integers, and the sum in (3.1) is over all integers.

The inverse of (3.1) is

$$A_i(\mathbf{p}, t) = L^{-3} \int d\boldsymbol{\xi} \exp(-i\mathbf{p}\cdot\boldsymbol{\xi}) v_i(\boldsymbol{\xi}, t),$$

where the region of integration is any periodic element in $\boldsymbol{\xi}$-space.

As a consequence of (2.4), the covariance function of A_i reduces to

$$\langle A_i(\mathbf{p}, t) A_j(\mathbf{p}', t') \rangle = \Delta(\mathbf{p} + \mathbf{p}') \alpha_{ij}'(\mathbf{p}, t, t'), \quad (3.2)$$

where

$$\Delta(\mathbf{p}) = 1, \quad \text{if } \mathbf{p} = 0,$$
$$= 0, \quad \text{otherwise},$$

and

$$\alpha_{ij}'(\mathbf{p}, t, t') = L^{-3} \int F_{ij}(\boldsymbol{\xi} - \boldsymbol{\xi}', t, t') \times \exp[-i\mathbf{p}\cdot(\boldsymbol{\xi} - \boldsymbol{\xi}')] d(\boldsymbol{\xi} - \boldsymbol{\xi}').$$

We have set

$$F_{ij}(\boldsymbol{\xi} - \boldsymbol{\xi}', t, t') = \langle v_i(\boldsymbol{\xi}, t) v_j(\boldsymbol{\xi}', t') \rangle,$$

which follows from (2.4). Conversely, it is the $\Delta(\mathbf{p} + \mathbf{p}')$ structure of the covariance function that assures and therefore is equivalent to homogeneity.

We shall now derive the condition on α_{ij}' that is equivalent to stationarity. Reverting to frame I coordinates and imposing stationarity, we have

$$\begin{aligned} F_{ij}(\boldsymbol{\xi} - \boldsymbol{\xi}', t, t') &= F_{ij}(\xi^1 - \xi'^1, \xi^2 - \xi'^2, \xi^3 - \xi'^3, t, t') \\ &= \langle v_i(I, \mathbf{x}, t) v_j(I, \mathbf{x}', t') \rangle \\ &= \langle v_i(I, \mathbf{x}, t + b) v_j(I, \mathbf{x}', t' + b) \rangle \\ &= \langle v_i(\xi^1 - \omega b \xi^2, \xi^2, \xi^3, t + b) \\ &\quad \times v_j(\xi'^1 - \omega b \xi'^2, \xi'^2, \xi'^3, t' + b) \rangle \\ &= F_{ij}(\xi^1 - \xi'^1 - \omega b(\xi^2 - \xi'^2), \xi^2 - \xi'^2, \\ &\quad \xi^3 - \xi'^3, t + b, t' + b). \end{aligned}$$

Hence, F_{ij} must be invariant under the substitutions

$$t \to t + b, \quad t' \to t' + b,$$
$$\xi^1 - \xi'^1 \to \xi^1 - \xi'^1 - \omega b(\xi^2 - \xi'^2), \quad (3.3)$$

for any b.

We have written F_{ij} as a function of the five independent space-time variables $\xi^1 - \xi'^1, \xi^2 - \xi'^2, \xi^3 - \xi'^3, t$, and t'. It is clear that we could have used, just as well, the five independent variables

$$\begin{aligned} &\xi^1 - \xi'^1 + \tfrac{1}{2}\omega(t + t')(\xi^2 - \xi'^2), \\ &\xi^2 - \xi'^2, \quad \xi^3 - \xi'^3, \\ &t - t', \quad t + t'. \end{aligned} \tag{3.4}$$

Of these, only $t + t'$ changes under the substitutions (3.3). It follows that F_{ij} must depend only on the other four variables in (3.4). Hence, stationarity is equivalent to the statement that

$$\begin{aligned} F_{ij}(\xi^1 - \xi'^1, \xi^2 - \xi'^2, \xi^3 - \xi'^3, t, t') \\ = G_{ij}(\xi^1 - \xi'^1 + \tfrac{1}{2}\omega(t + t')(\xi^2 - \xi'^2), \xi^2 \\ - \xi'^2, \xi^3 - \xi'^3, t - t'). \end{aligned}$$

Equation (3.2) then becomes

$$\alpha'_{ij}(\mathbf{p}, t, t') = L^{-3} \int G_{ij}(\xi''^1 + \tfrac{1}{2}\omega(t + t')\xi''^2, \xi''^2, \xi''^3, t - t') \times \exp(-i\mathbf{p}\cdot\xi'')d\xi''. \tag{3.5}$$

The region of integration in (3.5) can be a cube of volume L^3 with edges parallel to the ξ'' axes. For the moment, we choose a cube ($ABCD$ in Fig. 1) centered at $\xi'' = 0$, which incidentally lies in the neighborhood of the maximum contours of G_{ij}.

Since G_{ij} is defined as the covariance function of fields periodic in the components of ξ and ξ', it is clearly periodic in the components of $\xi'' = \xi - \xi'$. Periodicity in ξ''^1 allows the ξ''^1, ξ''^2 region of integration to be distored into a parallelogram ($A'B'C'D$ in Fig. 1) as long as $A'D' = B'C' = L$. If we let the independent variables in G_{ij} be

$$\begin{aligned} &\eta_1 = \xi''^1 + \omega T\xi''^2, \\ &\eta_2 = \xi''^2, \quad \eta_3 = \xi''^3, \\ &\tau = t - t', \end{aligned} \tag{3.6}$$

where $T = \frac{1}{2}(t + t')$, and take $A'A = \frac{1}{2}\omega TL$, then (3.5) becomes

$$\begin{aligned} \alpha'_{ij}(p, t, t') = L^{-3} \int_{-L/2}^{L/2} d\eta_1 \int_{-L/2}^{L/2} d\eta_2 \int_{-L/2}^{L/2} d\eta_3 \\ \times \exp[-i\eta_1 p_1 - i\eta_2(p_2 - \omega T p_1) - i\eta_3 p_3] \\ \times G_{ij}(\eta_1, \eta_2, \eta_3, \tau). \end{aligned} \tag{3.7}$$

The right-hand side of (3.7) is a function only of the four independent variables $p_1, p_2 - \omega T p_1, p_3$, and τ. An equivalent choice of independent variables is $p_2 - \omega t p_1$, $p_2 - \omega t' p_1, p_3$, and $t - t'$. Hence, a necessary condition for stationarity is that α'_{ij} be of the form

$$\alpha'_{ij}(\mathbf{p}, t, t') = \alpha_{ij}(p_2 - \omega t p_1, p_2 - \omega t' p_1, p_3, t - t'). \tag{3.8}$$

That (3.8) is not sufficient for stationarity will now be demonstrated. An additional, but almost trivial, condition is required on the covariance function. Given an α'_{ij} that satisfies (3.8), we can take the Fourier inverse of (3.5) to obtain the function G_{ij}. The latter will be periodic on the boundaries of $ABCD$. On the other hand, inverting Eqs. (3.6), we have

$$\begin{aligned} &\xi''^1 = \eta_1 - \omega T\eta_2 \\ &\xi''^2 = \eta_2, \quad \xi''^3 = \eta_3, \\ &t - t' = \tau. \end{aligned} \tag{3.9}$$

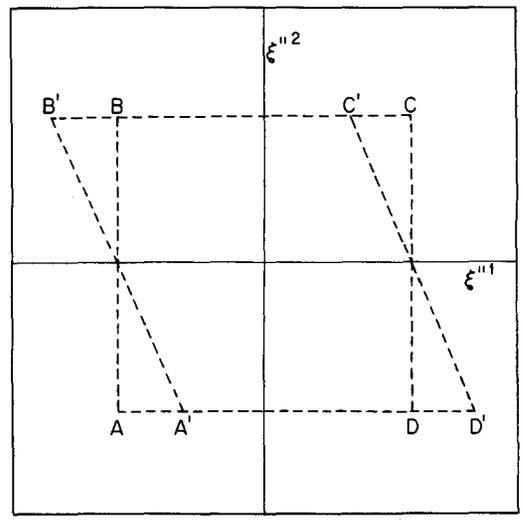


FIG. 1. Permissible regions of integration in Eq. (3.5) projected onto the ξ''^1, ξ''^2 plane.

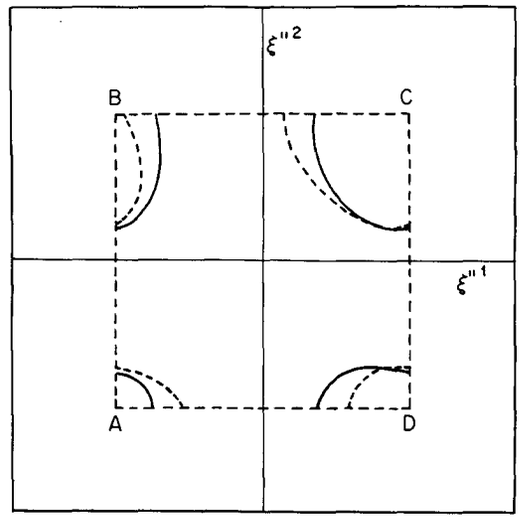


FIG. 2. Periodicity violating contours of G_{ij} for fixed ξ''^3.

Constant values of η_1, η_2, η_3, and τ imply constant values of G_{ij} for any T, if stationarity is to hold. Suppose that at $T = 0$, a contour of G_{ij} for fixed ξ''^3 is located as shown by the solid curves in Fig. 2. The way the contours evolve with T is determined by the condition of stationarity, embodied in Eqs. (3.9). In particular, for a slightly greater value of T, the contours will have shifted and distorted into the dotted curves indicated in Fig. 2, and periodicity is destroyed. It is clear that periodicity and stationarity can be rigorously compatible inside the square only if no contour of G_{ij} crosses the boundaries AD or BC.

A sensible extension of this condition is that G_{ij} is zero along the planes $\xi''^2 = \pm\frac{1}{2}L$. For if $\xi''^2 = \pm\frac{1}{2}L$, then $|x_2 - x'_2| = \frac{1}{2}L$. Hence a physically reasonable form for the additional condition is that L must be large enough so that the velocity covariance is zero (or nearly zero in practice) for spatial separations of $\frac{1}{2}L$ in the x_2 direction.

4. CONTRAVARIANT COMPONENTS OF v: INCOMPRESSIBILITY

Up to this point we have not dealt with the transformation of the velocity field components into the quasi-Lagrangian frame. For applications involving dynamical

equations of motion, turbulent advection of passive field, or simply the statement of incompressibility, it is analytically advantageous to introduce the transformed components. The contravariant components of $\mathbf{v}$ in the quasi-Lagrangian coordinate system are

$$\begin{aligned} u^1(\xi, t) &= v_1(\xi, t) - \omega t v_2(\xi, t), \\ u^2(\xi, t) &= v_2(\xi, t), \\ u^3(\xi, t) &= v_3(\xi, t). \end{aligned} \tag{4.1}$$

While we shall not discuss the transformed Navier-Stokes or passive field advection equations, it is well to point out that the field $\mathbf{u} = (u^1, u^2, u^3)$ has the following two essentially self-evident properties:

(a) The form of the substantial derivative D/Dt is preserved in quasi-Lagrangian coordinates. Thus,

$$\frac{\partial}{\partial t} + \sum_{i=1}^{3} V_i \frac{\partial}{\partial x_i} = \frac{\partial}{\partial t} + \sum_{i=1}^{3} u^i \frac{\partial}{\partial \xi^i},$$

where x_i and t are the independent variables on the left-hand side, while ξ^i and t are the independent variables on the right-hand side.

(b) The form of the incompressibility condition is preserved. If $\nabla \cdot \mathbf{V} = 0$ in frame I, then

$$\sum_{i=1}^{3} \frac{\partial}{\partial \xi^i} u^i = 0. \tag{4.2}$$

A simple physical interpretation of the contravariant velocity components can be made. Consider a wave in quasi-Lagrangian coordinates

$$u_0{}^j(\xi, t) = C^j(t) \exp(i\mathbf{p} \cdot \xi)$$

with $\mathbf{C} \cdot \mathbf{p} = 0$ so that the wave is formally transverse in quasi-Lagrangian space. Using (4.1), and reverting to frame I coordinates, we have

$$\mathbf{v}_0(\xi, t) = \mathbf{v}_0(I, x, t) = \mathbf{A}(t) \exp[i\mathbf{q}(t) \cdot \mathbf{x}], \tag{4.3}$$

where

$$\begin{aligned} A_1(t) &= C^1(t) + \omega t C^2(t), \\ A_2(t) &= C^2(t), \quad A_3(t) = C^3(t), \end{aligned} \tag{4.4}$$

and

$$\begin{aligned} q_1 &= p_1, \quad q_3 = p_3, \\ q_2 &= p_2 - \omega t p_1. \end{aligned} \tag{4.5}$$

Equation (4.4) gives the rule for transforming contravariant components in the quasi-Lagrangian frame back into frame I components. Similarly, (4.5) applies to covariant components. Then, since $\mathbf{A}(t) \cdot \mathbf{q}(t) = 0$, (4.3) is a transverse wave in real space as well. The wave vector $\mathbf{q}(t)$ is time-varying in such a way that, in addition to their normal propagation, planes of constant phase are advected or rotated by the shear mean flow. Conversely, if the mean flow has a linear profile, a transverse wave in real space is observed as a transverse wave in ξ-space.

Modes of this type were derived by Moffat[9] in a study of plane wave perturbations in an otherwise laminar uniform shear flow. Linearized Navier-Stokes dynamics led to explicit solutions for the $C^j(t)$, but such considerations are beyond the scope of this paper.

It should be clear that the definition of homogeneity in Sec. 1 applies equally well to the contravariant components of the velocity fluctuation. Thus, any v_i in (2.4) may be replaced by a u^i. Moments defined on one type of field are linear combinations of moments of the other type, with time-dependent coefficients. A velocity fluctuation field can then be defined by the contravariant components

$$u^j(\xi, t) = \sum_{\mathbf{p}} C^j(\mathbf{p}, t) \exp(i\mathbf{p} \cdot \xi)$$

and homogeneity is satisfied if

$$\langle C^i(\mathbf{p}, t) C^j(\mathbf{p}', t') \rangle \sim \Delta(\mathbf{p} + \mathbf{p}').$$

Incompressibility is assured if we let

$$C^i(\mathbf{p}, t) = \epsilon^{ijk} B_j'(\mathbf{p}, t) p_k, \tag{4.6}$$

where ϵ^{ijk} is the permutation symbol and $\mathbf{B}'$ is any complex-valued vector function of $\mathbf{p}$ and t. Without loss of generality, we may write

$$\begin{aligned} B_1' &= B_1(\mathbf{p}, t), \\ B_2' &= B_2(\mathbf{p}, t) + \omega t B_1(\mathbf{p}, t), \\ B_3' &= B_3(\mathbf{p}, t). \end{aligned} \tag{4.7}$$

The components of $\mathbf{B}'$ may be considered the covariant components of $\mathbf{B}$ transformed into the quasi-Lagrangian Fourier space. Then, using (4.1) and (3.1), we find

$$\mathbf{A}(\mathbf{p}, t) = \mathbf{B}(\mathbf{p}, t) \times \mathbf{q}(t), \tag{4.8}$$

where the components of $\mathbf{q}$ are given by (4.5). Equation (4.8) is the incompressibility condition for the Fourier modes in (3.1). The real-space field with Fourier mode coefficients $\mathbf{B}(\mathbf{p}, t)$ is the vector potential for the velocity fluctuation field.

The component of $\mathbf{B}$ parallel to $\mathbf{q}$ does not contribute to $\mathbf{A}$ (or to $\mathbf{C}$). It is, therefore, appropriate to decompose $\mathbf{B}$ into perpendicular and parallel parts,

$$\mathbf{B} = \mathbf{B}_\perp + \mathbf{B}_\parallel.$$

Equation (4.8) can now be solved for the perpendicular part.

We find

$$B_\perp(\mathbf{p}, t) = q^{-2} \mathbf{q}(t) \times \mathbf{A}(\mathbf{p}, t), \tag{4.9}$$

an immediate consequence of which is that the covariances of the components of $\mathbf{B}_\perp(\mathbf{p}, t)$ and $\mathbf{B}_\perp(\mathbf{p}', t')$ are all proportional to $\Delta(\mathbf{p} + \mathbf{p}')$. Since the statistical properties of $\mathbf{B}_\parallel$ are immaterial, we can then write

$$\langle B_i(\mathbf{p}, t) B_j(\mathbf{p}', t') \rangle \sim \Delta(\mathbf{p} + \mathbf{p}') \tag{4.10}$$

as the statement of homogeneity suitable for an incompressible velocity field.

Again, from (4.9), it is seen that the covariances of the components of $\mathbf{B}_\perp(\mathbf{p}, t)$ and $\mathbf{B}_\perp(-\mathbf{p}', t')$ are linear combinations of the covariances α_{ij} with coefficients [functions of $\mathbf{q}(t)$ and $\mathbf{q}(t')$] that depend only on the independent variables $p_1, p_2 - \omega t p_1, p_2 - \omega t' p_1$, and p_3. Since

$$p_1 \equiv [(p_2 - \omega t' p_1) - (p_2 - \omega t p_1)]/\omega(t - t'), \tag{4.11}$$

and stationarity requires the α_{ij} to depend only on $p_2 - \omega t p_1, p_2 - \omega t' p_1, p_3$, and $t - t'$, it follows that the covariances of the components of $\mathbf{B}_\perp$ depend only on

these same variables. It is obviously convenient to choose the statistical properties of $B_{\parallel}$ to be the same as those of $B_{\perp}$. Hence, stationarity for the incompressible field is equivalent to the statement that

$$\langle B_i(\mathbf{p}, t)B_j(-\mathbf{p}, t')\rangle = \beta_{ij}(p_2 - \omega t p_1, p_2 - \omega t' p_1, p_3, t - t'). \quad (4.12)$$

A field $\mathbf{B}(\mathbf{p}, t)$ satisfying (4.10) and (4.12) gives the Fourier modes for an incompressible, homogeneous, and stationary velocity fluctuation field according to (4.8), or the corresponding contravariant modes according to (4.7) and (4.6).

5. A TWO-DIMENSIONAL RANDOM VELOCITY FIELD

In this section we shall construct a two-dimensional example of an incompressible random velocity fluctuation field in a mean shear flow, satisfying homogeneity and stationarity. The example serves as a concrete demonstration of the methodology and may be a useful model in numerical simulations of random shear flows. The extension to three dimensions is not difficult.

Suppose g is a random complex function of $\mathbf{p}$ and t such that

$$g(\mathbf{p}, t) = -g^*(-\mathbf{p}, t) \quad (5.1)$$

and

$$\langle g(\mathbf{p}, t)g(\mathbf{p}', t')\rangle = -\Delta(\mathbf{p} + \mathbf{p}')\phi(t - t'). \quad (5.2)$$

Property (5.1) assures that the velocity is real. Two-dimensionality of $\mathbf{A}$ follows if we take

$$B_1(\mathbf{p}, t) = B_2(\mathbf{p}, t) = 0$$

but let

$$B_3(\mathbf{p}, t) = b(p_1, p_2 - \omega t p_1)g(\mathbf{p}, t)$$

where $b(\cdots)$ is a real nonrandom even function of its two arguments. Combining the above properties, we have

$$B_3(\mathbf{p}, t)B_3(\mathbf{p}', t') = -\Delta(\mathbf{p} + \mathbf{p}')b(p_1, p_2 - \omega t p_1)b(p_1, p_2 - \omega t' p_1)\phi(t - t').$$

Recalling (4.11), we see that (4.10) and (4.12) are satisfied by this $\mathbf{B}$. The corresponding components of $\mathbf{C}$ and $\mathbf{A}$ are, respectively,

$$C_1(\mathbf{p}, t) = -p_2 b(p_1, p_2 - \omega t p_1)g(\mathbf{p}, t),$$

$$C_2(\mathbf{p}, t) = p_1 b(p_1, p_2 - \omega t p_1)g(\mathbf{p}, t),$$

and

$$A_1(\mathbf{p}, t) = -(p_2 - \omega t p_1)b(p_1, p_2 - \omega t p_1)g(\mathbf{p}, t),$$

$$A_2(\mathbf{p}, t) = C_2(\mathbf{p}, t).$$

An example of a random function satisfying (5.1) and (5.2) is

$$g(\mathbf{p}, t) = \gamma(\mathbf{p})\exp[i\Omega(p)t + i\delta(p)] - \gamma(-\mathbf{p}) \times \exp[-i\Omega(-p)t - i\delta(-p)]$$

where $\gamma(p)$, $\Omega(p)$, and $\delta(p)$ are real random variables with the following properties:

(a) $\gamma(\mathbf{p})$, $\Omega(\mathbf{p}')$, and $\delta(\mathbf{p}'')$ are independent random variables for any $\mathbf{p}, \mathbf{p}', \mathbf{p}''$.

(b) $\delta(\mathbf{p})$ and $\delta(\mathbf{p}')$ are independent for any $\mathbf{p} \neq \mathbf{p}'$.

(c) $\delta(\mathbf{p})$ is uniformly distributed in the interval $[0, 2\pi]$ for all $\mathbf{p}$.

(d) $\langle \gamma^2(\mathbf{p})\rangle = \frac{1}{2}$ for all $\mathbf{p}$.

(e) The probability distribution of $\Omega(\mathbf{p})$ is the same for all $\mathbf{p}$.

The function ϕ is given by

$$\phi(\tau) = \langle \cos \Omega\tau \rangle. \quad (5.3)$$

We may note that the probability laws of $\gamma(\mathbf{p})$ and $\Omega(\mathbf{p})$ are essentially unspecified. However, two important special cases of (5.3) are (i) $\phi(\tau)$ is a symmetric Gaussian function of τ if Ω has a symmetric normal distribution, and (ii) $\phi(\tau)$ is an exponential function of $|\tau|$ if Ω has a symmetric Cauchy distribution.

To be specific, let us suppose that

$$b(p_1, p_2 - \omega t p_1) = (2\pi)^{1/2} v_0 L^{-1} Q^{-2} \times \exp\{-[p_1^2 + (p_2 - \omega t p_1)^2]/4Q^2\}, \quad (5.4)$$

where Q is a cutoff parameter in wave-vector space and v_0 is the root-mean-square of one component of the fluctuation velocity, and that Ω is Cauchy-distributed with the probability density function

$$f(\Omega) = (1/\pi)\tau_0/(1 + \tau_0^2\Omega^2), \quad -\infty < \Omega < \infty,$$

so that

$$\phi(\tau) = \exp(-|\tau|/\tau_0).$$

Then, we find

$$\langle C_i(\mathbf{p}, t)C_j(\mathbf{p}', t')\rangle = \Delta(\mathbf{p} + \mathbf{p}')(\delta_{ij}p^2 - p_i p_j)E,$$

$$\langle A_i(\mathbf{p}, t)A_j(\mathbf{p}', t')\rangle = \Delta(\mathbf{p} + \mathbf{p}')(\delta_{ij}\mathbf{q}'\cdot\mathbf{q} - q_i' q_j)E,$$

where

$$E \equiv 2\pi v_0^2 L^{-2} Q^{-4} \exp(-|t - t'|/\tau_0) \times \exp\{-\tfrac{1}{2}p_1^2 Q^{-2}[1 + \tfrac{1}{4}\omega^2(t - t')^2] - \tfrac{1}{2}Q^{-2}[p_2 - \tfrac{1}{2}\omega(t + t')p_1]^2\},$$

$$q_1 = q_1' = p_1,$$

$$q_2 = p_2 - \omega t p_1,$$

$$q_2' = p_2 - \omega t' p_1.$$

The velocity fluctuation moments are

$$\langle v_i(\xi, t)v_j(\xi', t')\rangle = \sum_{\mathbf{p}} \langle A_i(\mathbf{p}, t)A_j(-\mathbf{p}, t')\rangle \exp[i\mathbf{p}\cdot(\xi - \xi')].$$

Using the prescription

$$\sum_{\mathbf{p}} \to (L/2\pi)\int d\mathbf{p},$$

we find

$$\langle v_i(\xi, t)v_j(\xi', t')\rangle = v_0^2 S^{-5} \exp[-\tfrac{1}{2}Q^2(\eta_1^2 S^{-2} + \eta_2^2) - |t - t'|/\tau_0]\rho_{ij} \quad (5.5)$$

where

$$\rho_{11} = Q^2[\tfrac{1}{4}\omega^2(t - t')^2\eta_1^2 - S^4\eta_2^2] + S^2,$$

$$\rho_{12} = Q^2 S^2 \eta_1\eta_2 + \tfrac{1}{2}\omega(t - t')(S^2 - Q^2\eta_1^2),$$

$$\rho_{21} = Q^2 S^2 \eta_1\eta_2 - \tfrac{1}{2}\omega(t - t')(S - Q^2\eta_1^2),$$

$$\rho_{22} = S^2 - Q^2\eta_1{}^2,$$

$$S = [1 + \tfrac{1}{4}\omega^2(t - t')^2]^{1/2},$$

$$\eta_1 = \xi^1 - \xi'^1 + \tfrac{1}{2}\omega(t + t')(\xi^2 - \xi'^2),$$

$$\eta_2 = \xi^2 - \xi'^2.$$

Although the space-time structure of $\langle v_i v_j \rangle$ in (5.5) appears complicated and artificially contrived from *ad hoc* premises, the assumptions that led to this structure are actually very simple. With some redundance, the essential ingredients are summarized below:

(a) homogeneity, stationarity, and incompressibility.

(b) decomposition of the fluctuation field into plane waves of statistically independent amplitudes, frequencies, and phases. Only the probability law of the phases needs to be specified completely, and this is dictated by homogeneity.

(c) the form of the spectrum $b(\cdots)$ in (5.4). This model spectrum would appear to be the most artificial assumption, but it is also perhaps the simplest extension to a shear flow, consistent with the stationarity requirement (4.12), of a commonly used Gaussian spectrum.

*This work was performed while the author held a National Research Council Resident Research Associateship.

[1]W. G. Rose, J. Fluid Mech. 25, 97 (1966).
[2]F. H. Champagne, V. G. Harris, and S. Corrsin, J. Fluid Mech. 41, 81 (1970).
[3]J. L. Lumley, J. Fluid Mech. 41, 413 (1970).
[4]R. H. Kraichnan, Phys. Fluids 13, 22 (1970).
[5]F. A. Gifford, Proc. USAEC Meteorological Information Meeting (AECL-2787), 485 (1967).
[6]G. K. Batchelor, *The Theory of Homogeneous Turbulence* (Cambridge U. P., Cambridge, 1953).
[7]J. L. Lumley, Proc. of the International Colloquium on Atmospheric Turbulence and Radio Wave Propagation, Moscow (edited by A. M. Yaglom and V. I. Tatarsky), 166 (1967).
[8]M. Loève, *Probability Theory* (Van Nostrand, Princeton, N. J., 1963), p. 478.
[9]H. K. Moffat, in Ref. 7, 139 (1967).

Prequantization of charge*

Jędrzej Śniatycki

Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary, Alberta, Canada
(Received 15 October 1973)

It is shown that a necessary and sufficient condition for quantization of relativistic dynamics of a particle with charge e moving in an external electromagnetic field F is that $(e/2\pi)F$ should define an integral de Rham cohomology class on the space-time manifold.

1. INTRODUCTION

All known particles have charges which are integral multiples of the charge of electron. The first attempt to explain this fact theoretically was made by Dirac in his paper on existence of magnetic poles.[1] He has shown that the assumption that a particle with charge e moving in the field of a magnetic pole with strangth m has a single valued wavefunction leads to the condition $em = 2\pi n$, where n is an integer and the rationalized units with $\hbar = c = 1$ are used. Theoretical implications of existence of magnetic poles and the problems of quantization of charge have been subsequently studied by several authors.[2]

An alternative formulation of the Dirac condition $em = 2\pi n$ is possible if one removes the world lines of magnetic poles, allowing thus for a more complicated topology of the space—time X. Then, the electromagnetic field F becomes a closed 2-form on X but it is not exact. In this situation the Dirac condition is replaced by the requirement that $(e/2\pi)F$ should define an integral de Rham cohomology class: $[(e/2\pi)F] \in H^2(X, Z)$. Thus, the existence of a nonexact background electromagnetic field F in the space—time X would lead to the quantization of charge as follows: If there exists the smallest positive charge e_0 such that $[(e_0/2\pi)F] \in H^2(X, Z)$, then the dynamics of a particle with charge e moving in the electromagnetic field F can be quantized if and only if e is an integral multiple of e_0. The value of e_0 depends on the background electromagnetic field F; in general a positive e_0 satisfying the condition $[(e_0/2\pi)F] \in H^2(X, Z)$ need not exist.

The aim of this note is to show that the condition $[(e/2\pi)F] \in H^2(X, Z)$ is a necessary and sufficient condition for a geometric quantization of the phase space of relativistic particles with charge e moving in the electromagnetic field F. This result is a direct consequence of the Hamiltonian dynamics of relativistic charged particles which is reviewed in Sec. 2 and of the necessary and sufficient condition for prequantization of a symplectic manifold reviewed in Sec. 3.

Canonical quantization of relativistic charged particles has been studied by Torrence and Tulczyjew[3] under the assumption that F is exact, i.e., $[F] = 0$. In this case there are no restrictions on the value of the charge.

The mathematical technique used in this paper is that of symplectic geometry and topology of manifolds. The definitions and the notation used here can be found in any text on these subjects.[4]

2. HAMILTONIAN DYNAMICS OF RELATIVISTIC CHARGED PARTICLES

Let X denote a four-dimensional C^∞ manifold representing the space—time and let g be a Lorentzian metric on X representing the gravitational field. Further, let T^*X denote the cotangent bundle of X, ρ: $T^*X \to X$ the cotangent bundle projection, and θ the Liouville form on T^*X defined by $\theta(u) = p(T\rho(u))$, for each $p \in T^*X$ and each $u \in T_pT^*X$. The exterior derivative of θ is a symplectic form on T^*X.

The sympletic manifold $(T^*X, d\theta)$ is the phase space for relativistic particles. An element $p \in T_x^*X$ represents a canonical momentum at a point $x \in X$, and $d\theta$ is the Lagrange bracket. The function H on T^*X, defined by $H(p) = \frac{1}{2}g(p, p)$, has the physical interpretation of $\frac{1}{2}$ times the square of the mass of a particle with momentum p, and it determines the Hamiltonian dynamics of a relativistic particle as follows. Let u_H be the Hamiltonian vector field on $(T^*X, d\theta)$ associated with the Hamiltonian H, i.e., u_H satisfies the equation $u_H \lrcorner d\theta = -dH$. Then, the integral lines of u_H projected to X give the world lines of particles moving in the gravitational field g.[5] In the presence of an electromagnetic field, described by a closed 2-form F on X, the Hamiltonian dynamics of particles with charge e is given by the Hamiltonian vector field v_H on T^*X associated with the same Hamiltonian H but with respect to a modified symplectic form $\omega = d\theta - e\rho^*F$, where ρ^*F is the pull back of F to T^*X.[6] Thus, the phase space for relativistic particles with charge e moving in a gravitational field g and an electromagnetic field F is given by a symplectic manifold (T^*X, ω), where $\omega = d\theta - e\rho^*F$.

3. GEOMETRIC QUANTIZATION

Geometric quantization is a scheme of setting up a quantum theory starting from an arbitrary phase space, provided certain consistency conditions are satisfied.[7] It consists of two steps. The first step in quantizing a symplectic manifold (P, ω), called prequantization, consists of constructing a complex line bundle L over P with connection ∇ such that $(1/2\pi)\omega$ is the curvature form of ∇.[8] The main result of the theory of prequantization which will be needed here is the following:

Theorem: For a given symplectic manifold (P, ω) there exists a line bundle L over P with connection ∇ such that $(1/2\pi)\omega$ is the curvature form of ∇ if and only if $[(1/2\pi)\omega] \in H^2(P, Z)$. Proof of this theorem as well as a detailed exposition of prequantization can be found in the first of the papers by Kostant referred to above.[7] Given such a line bundle L over P, to each function on P there

corresponds a differential operator on the space of sections of L in such a way that the canonical commutation relations are satisfied. The second step in geometric quantization corresponds to the choice of position variables and the construction of the Hilbert space of wave functions. It will not be discussed here since it is not needed for our considerations.

From Sec. 2 it follows that the phase space for relativistic particles with charge e moving in a gravitational field g and an electromagnetic field F is given by a symplectic manifold (T^*X, ω), where $\omega = d\theta - e\rho^*F$. Applying the condition of the theorem above to this symplectic manifold we get the condition $[(1/2\pi)(d\theta - e\rho^*F)] \in H^2(T^*X, Z)$ which is equivalent to $[(e/2\pi)F] \in H^2(X, Z)$. Thus $[(e/2\pi)F] \in H^2(X, Z)$ is a necessary and sufficient condition for prequantization of the phase space $(T^*X, d\theta - e\rho^*F)$ for relativistic charged particles with charge e moving in an electromagnetic field F.

ACKNOWLEDGMENTS

The author is greatly indebted to R. Geroch, N. Hurt, B. Kostant, and D.J. Simms for stimulating discussions.

*Work done when the author visited Mathematisches Institut der Universität Bonn supported by Sonderforschungsbereich "Theoretische Mathematik."

[1]P.A.M. Dirac, Proc. Roy. Soc. A **133**, 60 (1931).

[2]See, e.g., E. Amaldi, "On the Dirac Magnetic Poles," in *Old and New Problems in Elementary Particles*, edited by G. Puppi (Academic, New York, 1968), and H. Jehle, "Flux Quantization, Charge Quantization and the Electromagnetic Interaction Constant" in *Lectures on Theoretical Physics*, edited by A.O. Barut and W.E. Britten (Gordon and Breach, New York, 1968), Vol. XB, and the references quoted there.

[3]R.J. Torrence and W.M. Tulczyjew, J. Math. Phys. **14**, 1727 (1973).

[4]See, e.g., R. Hermann, *Vector Bundles in Mathematical Physics* (Benjamin, New York, 1970).

[5]This form of the differential equation for geodesics is in S. Lang, *Introduction to Differentiable Manifolds* (Interscience, New York, 1962), Chap. 7, Sec. 6.

[6]J.-M. Sourian, *Structure des systèmes dynamiques* (Dunod, Paris, 1970), Chap. 3, Sec. 15.

[7]For literature on geometric quantization method see: R.J. Blattner, "Quantization and representation theory," to appear in *Proceedings of A.M.S. 1972 Summer Institute on Group Representations*; N.E. Hurt, Ann. Inst. Henri Poincaré, **14**, 153 (1971); B. Kostant, "Quantization and unitary representations," *Lectures in Modern Analysis and Applications* III, edited by C.T. Taam, Lecture Notes in Mathematics 170 (Springer, Berlin, 1970); B. Kostant, "Symplectic spinors," to appear in *Proceedings of Convengo di Geometria Simplettica e Fisica Matematica, INDAM Rome* (1973), Symposia Math. (Academic); J. Rawnsley, "Some applications of quantization," thesis, Oxford (1972); P. Renouard, "Variétés symplectiques et quantization," thèse, Orsay (1969); D.J. Simms, "Geometric quantization of energy levels in the Kepler problem," to appear in *Proceedings of Convengo di Geometria Siplettica e Fisica Matematica, INDAM Rome* (1973), Symposia Math. (Academic); D.J. Simms, Z. Naturfors. **28a**, 538 (1973); D.J. Simms, "Geometric quantization of the harmonic oscillator with diagonalized Hamiltonian," to appear in *Proceedings of 2nd International Colloquium on Group Theoretical Methods in Physics, Nijmegen*, (1973); J.-M. Souriau, *Structure des systèmes dynamiques*, (Dunod, Paris, 1970), Chap. 5.

[8]The difference in the coefficient $1/2\pi$ between this formulation and that of Kostant is due to the requirement that a constant function c should correspond to the operator of multiplication by c rather than $2\pi ic$.

Smoothing operators for field domains

Gerhard C. Hegerfeldt

Institut für Theoretische Physik, Universität Göttingen, Göttingen, Germany
(Received 21 August 1972)

We consider sharp-time fields and write down, in terms of the fields, simple and explicit expressions for operators with very strong smoothing properties. When applied to any vector in Hilbert space, the resulting vector is in the domain of any power of the space-smeared fields, and it even is entire for the fields. It is shown that in this way one obtains a common dense domain of definition on which the field operators are essentially self-adjoint. Attention is focused on the space $\mathscr{S}$ of rapidly decreasing C^∞ functions as smearing functions for the fields; here the smoothing operators are simply products of exponentials of the field smeared with Hermite functions.

I. INTRODUCTION

Bose fields become, after smearing with suitable test functions, unbounded operators in Hilbert space. As a consequence, the field operators cannot be applied to every vector; this is sometimes expressed by saying that the resulting vector has infinite norm. There are also vectors to which the smeared fields, but not their square etc. can be applied. It may even happen that the ground state is not in the domain of the field operators, as for instance in the ultralocal model[1] where the space-smeared fields are well-defined self-adjoint operators, yet in some cases cannot see the ground state. A further complication arises from the fact that the domain of a field operator depends on the particular smearing function. It is therefore conceivable that the intersection of the domains for all different smearing functions could turn out to be zero; then there would exist no common domain. Such a situation would make it practically impossible to calculate with the fields directly.

In a previous paper,[2] the author has made a general investigation of this question for fields which can be defined for sharp time, i.e., which need only space-smearing,

$$\varphi(f) \equiv \int \varphi(\mathbf{x},0) f(\mathbf{x}) d^3x, \tag{1.1}$$

f real. For fixed $t=0$ these operators have to commute. It is more convenient, in view of the domain question, to consider the unitary Weyl operators

$$U(f) = e^{i\varphi(f)}. \tag{1.2}$$

From the properties of $\varphi(f)$ one is then lead to the relation

$$U(f_1+f_2) = U(f_1)U(f_2). \tag{1.3}$$

To recover $\varphi(f)$ as generator from $U(f)$ one only needs ray continuity, i.e., $U(\lambda f)$ has to be (weakly or strongly) continuous in λ. In general, however, one deals with stronger continuity conditions; commonly the smearing functions are taken to be in the Schwarz space $\mathscr{S}(\mathbb{R}^3)$ of real rapidly decreasing, infinitely differentiable functions, and $U(f)$ is assumed to be continuous in f.

For such fields the author[2] has shown the existence of a dense set of vectors to which every $\varphi(f)$ can be applied arbitrarily often and on which one can even consider power series; in particular,[3]

$$\sum_{n=0}^{\infty} \frac{t^n}{n!} \left\| \varphi(f)^n |\psi\rangle \right\| < \infty, \tag{1.4}$$

for all t. The domain is invariant under the field and the unitary Weyl operators, and the fields are essentially self-adjoint on it ("Gårding domain").

A similar result was obtained if in addition one has a conjugate field Π fulfilling canonical commuation relations (CCR's) with φ,

$$[\varphi(\mathbf{x},t), \Pi(\mathbf{x}',t)] = i\delta^{(3)}(\mathbf{x}-\mathbf{x}').$$

For the smeared field and the Weyl operators,

$$\begin{aligned} &\Pi(g) \equiv \int \Pi(\mathbf{x},0) g(\mathbf{x}) d^3x \\ &V(g) \equiv e^{i\Pi(g)} \end{aligned} \tag{1.5}$$

one has, respectively.

$$\begin{aligned} &[\varphi(f), \Pi(g)] = i\int f(\mathbf{x}) g(\mathbf{x}) d^3x \equiv i(f,g), \\ &V(g)U(f) = e^{i(g,f)} U(f) V(g). \end{aligned} \tag{1.6}$$

In this case the existence of a common Gårding domain for both φ and Π was proved.

The fairly simple characterization of the domain was based on a direct integral realization of Hilbert space with a certain measure associated with the field operators. However, in order to know this domain explicitly for calculational purposes one would have to determine the measure, a task which is usually very difficult and which would have to be repeated for each particular model theory. For *practical* purposes it is therefore desirable to be able to write down a Gårding domain directly in terms of the fields, independent of any particular field realization.

This *explicit* specification of a domain for the fields will be carried out in the following, and the result turns out to be of great simplicity. We define simple smoothing operators which, when applied to any vector, yield a vector in the domain of the fields, indeed even an entire[3] vector. The set of vectors thus obtained is dense.

Theorem: (i) Let φ be a sharp-time field as in Eqs. (1.1–4) and let $U(f) = \exp\{i\varphi(f)\}$ be (weakly or strongly) continuous for $f \in \mathscr{S}(\mathbb{R}^3)$. Let $h_n(x)$ be the nth Hermite function, let $\nu = (\nu_1, \nu_2, \nu_3)$, with ν_i a nonnegative integer, and put

$$h_\nu(\mathbf{x}) = h_{\nu_1}(x_1) h_{\nu_2}(x_2) h_{\nu_3}(x_3). \tag{1.7}$$

Let $\{c_\nu\}$ be a triple sequence of positive numbers satisfying the growth restriction

$$c_\nu \leq k \prod_{i=1}^{3} (\nu_i + 1)^r \tag{1.8}$$

for some positive constants k and r, and define the smoothing operator $A_{\{c\}}$ by[4]

$$A_{\{c\}} = \exp\left[-\sum_\nu c_\nu^{-2}\varphi(h_\nu)^2\right]. \tag{1.9}$$

Then every vector of the form

$$|\hat{\psi}\rangle \equiv A_{\{c\}}|\psi\rangle, \quad |\psi\rangle \in \mathscr{H}, \tag{1.10}$$

is in the domain of $\varphi(f)^n$ for each n and $f \in \mathcal{S}(\mathbb{R}^3)$. Moreover, $|\hat{\psi}\rangle$ as well as its image under $\varphi(g)$ and $U(g)$, $g \in \mathcal{S}(\mathbb{R}^3)$, are entire[3] for each $\varphi(f)$.

One obtains a dense domain on which the smeared fields are essentially self-adjoint if one chooses

$$c_\nu = \prod_i (\nu_i + 1)^r, \quad \nu_i = 0,1,2,\ldots, \quad r = 1,2,\ldots, \tag{1.11}$$

varies $|\psi\rangle$ through a dense set of $\mathscr{H}$ and takes finite linear combinations of the resulting vectors $|\hat{\psi}\rangle$. Applying the field and Weyl operators repeatedly one obtains an invariant Gårding domain.

(ii) Let φ and Π be fields satisying CCR's and the continuity condition of part (i). In addition to $A_{\{c\}}$ we define the smoothing operator[5]

$$B_{\{\hat{c}\}} = \exp\left[-\sum_\nu \hat{c}_\nu^{-2}\Pi(h_\nu)^2\right], \tag{1.12}$$

where $\{\hat{c}\}$ satisfies the growth restriction of Eq. (1.8). Then every vector of the form

$$|\hat{\psi}\rangle = A_{\{c\}}B_{\{\hat{c}\}}|\psi\rangle, \quad |\psi\rangle \in \mathscr{H}, \tag{1.13}$$

is in the domain of both $\varphi(f)^n$ and $\Pi(f)^n$ for each n and $f \in \mathcal{S}(\mathbb{R}^3)$. Moreover, $|\hat{\psi}\rangle$ and its image under $\varphi(g)^n$, $\Pi(g)^n$, $U(g)$, $V(g)$, $g \in \mathcal{S}(\mathbb{R}^3)$, are entire vectors for $\varphi(f)$ and $\Pi(f)$. One already obtains a dense domain of essential self-adjointness for all $\varphi(f)$ and $\Pi(f)$, $f \in \mathcal{S}(\mathbb{R}^3)$, if one takes $\{\hat{c}\} = \{c\}$, varies $\{c\}$ through the sequences given by Eq. (1.11) and proceeds as in (i).

Remarks: (i) That the smoothed vectors $|\hat{\psi}\rangle$ in Eq. (1.10) are entire if they are nonzero, is not unexpected. The main statement of the theorem is the result that sufficiently many of the smoothing operators do not vanish identically[6] and that one obtains a dense set of nonzero vectors.

(ii) The special choice of the basis $\{h_\nu\}$ as products of Hermite functions is essential.

(iii) For multicomponent fields and configuration space $\mathbb{R}^n$ the result carries over with the obvious changes.

The theorem will be proved in Sec. 2. General test function spaces and arbitrary bases are considered in Sec. 3 and a result slightly weaker than the above theorem is derived. In Sec. 4 the results are discussed. It is pointed out that for *irreducible* representations of the CCR's a single smoothing operator suffices, and the smoothing operators are expressed as an integral over the Weyl operators.

II. THE ROLE OF $\mathcal{S}$ SMOOTHING OPERATORS IN Q-SPACE

The special role played by the Hermite function is connected to the fact that they form an absolute basis for $\mathcal{S}$.[7] For $f \in \mathcal{S}(\mathbb{R}^3)$ we put

$$\xi_\nu \equiv \int f(\mathbf{x})h_\nu(\mathbf{x})d^3x. \tag{2.1}$$

Then

$$f(\mathbf{x}) = \sum_\nu \xi_\nu h_\nu(\mathbf{x}), \tag{2.2}$$

where the convergence is that of $\mathcal{S}$. Thus one has a unique correspondence between f and a triple sequence.

$$f \longleftrightarrow \{\xi\} \equiv \{\xi_{000}, \cdots, \xi_\nu, \cdots\}. \tag{2.3}$$

A sequence $\{\xi\}$ belongs to an element of $\mathcal{S}(\mathbb{R}^3)$ if and only if all norms of the form

$$|\xi|_r \equiv \sup_\nu |(\nu+1)^r \xi_\nu| \tag{2.4}$$

are finite where $(\nu+1)^r \equiv \prod_i (\nu_i+1)^r$ with r a positive integer. Furthermore, $\lim_{i\to\infty} |\xi^{(i)}|_r = 0$ for all r if and only if the associated functions f_i converge to zero in $\mathcal{S}(\mathbb{R}^3)$. This means that $\mathcal{S}(\mathbb{R}^3)$ is isomorphic[7] to a sequence space $\mathbf{s}$ with norms (2.4).

It is convenient to replace the norms (2.4) by

$$\|\xi\|_r^2 \equiv \sum_\nu (\nu+1)^{2r}\xi_\nu^2. \tag{2.5}$$

Clearly $|\xi|_r \leq \|\xi\|_r$, and from

$$\sum_\nu (\nu+1)^{2r}\xi_\nu^2 = \sum_\nu (\nu+1)^{-2r}(\nu+1)^{4r}\xi_\nu^2 \leq |\xi|_{2r}^2 \sum_\nu (\nu+1)^{-2r}$$

it follows that $\|\xi\|_r \leq k|\xi|_{2r}$. Hence the two systems of norms are equivalent; the latter, however, are Hilbertian, they can be derived from the scalar product

$$\langle \xi, \eta\rangle_r \equiv \sum_\nu (\nu+1)^{2r}\xi_\nu\eta_\nu. \tag{2.6}$$

Since $\mathcal{S}$ and $\mathbf{s}$ are isomorphic, so are their duals $\mathcal{S}'$ and $\mathbf{s}'$. The dual $\mathbf{s}'$ consists of all triple sequences $\{\eta'\}$ such that

$$\|\eta'\|_{-r}^2 \equiv \sum_\nu (\nu+1)^{-2r}\eta_\nu'^2 < \infty \tag{2.7}$$

for *some* positive integer r, since then

$$(\eta', \xi) \equiv \sum_\nu \eta_\nu'\xi_\nu \leq \|\eta'\|_{-r} \cdot \|\xi\|_r. \tag{2.8}$$

Q-space and P-space

In Ref. 8 a realization of the Hilbert space has been given in which the field becomes multiplication by a scalar factor, just as in the Schrödinger representation for finitely many degrees of freedom ("Q-space"). The Hilbert space consists of vector-valued functions $\psi(\Lambda)$, where the variable Λ runs through all linear functionals, the dual of the test function space. Then $U(f) = \exp\{i\varphi(f)\}$ simply acts as

$$U(f)\psi(\Lambda) = e^{i(\Lambda, f)}\psi(\Lambda), \tag{2.9}$$

and

$$\varphi(f)\psi(\Lambda) = (\Lambda, f)\psi(\Lambda). \tag{2.10}$$

The scalar product in Q-space is expressed by means of a measure which is determined by the field. In a similar way one can diagonalize $\Pi(f)$ if one deals with CCR's ("P-space").

In Ref. 2 it was shown that, for a large class of test function spaces, including $\mathcal{S}$, one obtains an entire vector for φ if one multiplies by an exponentially decreasing function.

$$\hat{\psi}(\Lambda) = \exp(-\|\Lambda\|_c^2)\psi(\Lambda), \tag{2.11}$$

where $\|f\|_c$ is a continuous Hilbertian norm on the test function space and $\|\Lambda\|_{-c}$ the associated dual norm.[9] Furthermore, one obtains a total set of entire vectors if one takes the norms which determine the topology of the test function space.

Specialization to $\mathscr{S}$

If, in the above correspondance to sequences, f corresponds to $\{\xi\}$ and Λ to $\{\eta'\}$, then one has $(\Lambda, f) = (\eta', \xi)$, by definition. We can therefore write $\psi(\eta')$ instead of $\psi(\Lambda)$ and

$$U(f)\psi(\eta') = e^{i(\eta',\xi)}\psi(\eta'), \qquad \varphi(f)\psi(\eta') = (\eta', \xi)\psi(\eta'). \tag{2.12}$$

For the norms in Eq. (2.11) we can then take $\|\eta'\|_{-r}$ of Eq. (2.7) which is the dual norm to $\|\xi\|_r$ of Eq. (2.5).

Let $\{\delta^{(n)}\}$ denote the sequence corresponding to h_n; it consists of zeros except for a 1 at the nth place. By Eq. (2.12) one has, for any $k > 0$.

$$\exp[-k\varphi(h_\nu)^2]\psi(\eta') = \exp[-k(\eta', \delta^{(\nu)})^2]\psi(\eta') = \exp(-k\eta_\nu'^2)\psi(\eta'). \tag{2.13}$$

Hence

$$\lim_{|n|\to\infty} \prod_{|\nu|\le|n|} \exp[-(\nu+1)^{-2r}\mathcal{S}(h_\nu)^2]\psi(\eta') = \exp(-\|\eta'\|_{-r}^2)\psi(\eta'), \tag{2.14}$$

where the limit exists strongly since $\exp\{-\|\eta'\|_{-r}^2\} \le 1$ for all $\{\eta'\}$ so that one can apply Lebesque's bounded convergence theorem. The operator on the lhs of Eq. (2.14) is just $A_{\{c\}}$ of Eq. (1.9) with $c_\nu = (\nu+1)^r$.

In case of a general $\{c\}$ satisfying the growth restriction of Eq. (1.9) we note that

$$\|\xi\|_c^2 \equiv \sum_\nu c_\nu^2 \xi_\nu^2 \le k^2 \|\xi\|_r^2.$$

Hence $\|\xi\|_c$ is a continuous norm, with dual norm

$$\|\eta'\|_{-c}^2 \equiv \sum_\nu c_\nu^{-2}\eta_\nu'^2,$$

and Eqs. (2.12) and (2.13) apply again. Thus multiplication by $\exp\{-\|\eta'\|_{-c}^2\}$ in Q-space corresponds to applying the operator $A_{\{c\}}$ whose form is independent of any particular realization of the Hilbert space. By the quoted results of Ref. 2, this proves part (i) of the theorem.[10]

In the case of CCR's it is seen in exactly the same way that the smoothing operator $B_{\{c\}}$ corresponds to multiplication by a similar exponential function in P-space. Part (ii) of the theorem follows then immediately from Sec. 3 in Ref. 2.

III. GENERAL TEST FUNCTION SPACES AND ARBITRARY BASES

In this section we derive a slightly weaker result for an arbitrary test function space $\mathscr{V}$ and an arbitrary independent set of test functions $g_1, g_2, \ldots$. Only ray continuity will be assumed. In the case of CCR's we take the g_i's to be orthonormal, $(g_i, g_j) = \delta_{ij}$. Let

$$\mathscr{V}_0 \equiv \mathfrak{L}\{g_1, g_2, \ldots\} \tag{3.1}$$

be the set of *finite* linear combinations of the g_i's. In the following, smoothing operators and field domains will be derived *for test functions in* $\mathscr{V}_0$.

Theorem 3.1: Let $\{c\} \equiv \{c_1, c_2, \ldots\}$ be any sequence of positive numbers. Define the smoothing operator $A_{\{c\}}$ by[11]

$$A_{\{c\}} \equiv \exp\left[-\sum_i c_i^{-2}\varphi(g_i)^2\right] \tag{3.2}$$

and, in case of CCR's, also

$$B_{\{c\}} = \exp\left[-\sum_i c_i^{-2}\Pi(g_i)^2\right]. \tag{3.3}$$

Then, for any $|\psi\rangle \in \mathscr{H}$, the application of $A_{\{c\}}$ (or $A_{\{c\}}B_{\{\hat{c}\}}$) yields an entire vector for $\varphi(f), f \in \mathscr{V}_0$ [or, respectively, for $\varphi(f)$ and $\Pi(f), f \in \mathscr{V}_0$]. To obtain a Gårding domain, or rather a dense invariant domain of entire vectors it suffices to take $\{c\} = \{\hat{c}\}$, to vary $\{c\}$ through the positive integers and to proceed as in the theorem of the Introduction.

Proof: If $f \in \mathscr{V}_0$, then $f = \sum_{\nu=1}^n \xi_\nu g_\nu$ for some n and ξ_ν. Thus $\mathscr{V}_0$ can be identified with the space of all terminating sequences $\{\xi\} = \{\xi_\nu \ldots, \xi_n, 0, \ldots\}$. For every (infinite) sequence $\{c\} \equiv \{c_1, \cdots\}$ of positive numbers the Hilbertian norm

$$\|\xi\|_c^2 \equiv \sum_\nu c_\nu^2 \xi_\nu^2 \tag{3.4}$$

exists since the sum is finite. The topology induced by these norms is just that needed for ray continuity. One obtains an equivalent system of norms if one restricts c_ν to the positive integers. The dual $\mathscr{V}_0'$ now can be identified with the space of *all* sequences $\{\eta'\} \equiv \{\eta_1', \cdots\}$, with no growth restriction.

Again one has a Q-space realization (and a P-space in case of CCR's) as in Sec. 2, now $\{\eta'\}$ varying through $\mathscr{V}_0'$. As before one shows that the operator $A_{\{c\}}$ corresponds to multiplication by

$$\exp(-\|\eta'\|_{-c}^2)$$

in Q-space,[12] and similarly for $B_{\{c\}}$ in P-space. Theorem 3.1 then follows from the results of Sec. 3 of Ref. 2. QED

Remark: If $\hat{f}$ is in $\mathscr{V}$ but not in $\mathscr{V}_0$ and if one does not have additional continuity properties, in general nothing can be said whether or not $|\psi\rangle$ is in the domain of $\varphi(\hat{f})$. If it should be possible to construct smoothing operators of this kind for all test functions in $\mathscr{V}$, one certainly will have to impose growth restrictions on the c_ν, as in the case of $\mathscr{S}$. Otherwise c_ν^{-2} and the associated smoothing operator might not decrease sufficiently fast to control $\varphi(\hat{f})$.

IV. DISCUSSION

The main idea in Secs. 2 and 3 has been to express $\exp\{-\|\Lambda\|_{-c}^2\}$ in terms of field operators. From the procedure it is clear that one can obtain a similar construction for those nuclear test function spaces whose topology can be expressed by a set of (semi-) norms which, for a suitable basis, become "diagonal" as in Eq. (2.5).

To obtain a dense domain of definition for the fields

we needed, in the case of $\mathcal{S}$, a countable number of smoothing operators. It can also be shown for the case considered in Theorem 3.1 that a countable number is sufficient; however, it depends on the particular example which of the smoothing operators have to be taken. The question arises if one always needs an infinite number of smoothing operators. In general nothing can be said about this. But when one has an *irreducible representation of the* CCR's a single smoothing operator AB suffices. This is due to the following.

Corollary 4.1: If the representation of the CCR's is irreducible, then either the smoothing operator $A_{\{c\}}$ is identically zero or maps nonzero vectors onto nonzero vectors. If $A_{\{c\}} \neq 0$ and if $\mathcal{D}$ is a dense domain in $\mathcal{H}$, then $A_{\{c\}}\mathcal{D}$ is again dense in $\mathcal{H}$. The same holds for the smoothing operator $B_{\{c\}}$. Hence in the case of $\mathcal{S}$, there is a $\{c\}$ as given in Eq. (1.11) such that $A_{\{c\}}B_{\{c\}}$ is a dense Gårding domain for φ and Π. In the case of Theorem 3.1 a similar result holds.

This result follows in a straightforward way from Theorem 4.2 of Ref. 2. The particular sequence $\{c\}$ that will give the desired smoothing operator can, however, depend on the particular representation of the CCR's.

It is possible to express the smoothing operators in terms of the unitary Weyl operators. From the integral relation

$$e^{-(\Lambda,f)^2/k^2} = k(4\pi)^{-1/2}\int_{-\infty}^{\infty} d\lambda\ e^{-k^2\lambda^2/4}e^{i\lambda(\Lambda,f)},$$

one obtains

$$e^{-k^{-2}\varphi(f)^2} = k(4\pi)^{-1/2}\int_{-\infty}^{\infty} d\lambda\ e^{-k^2\lambda^2/4}e^{i\lambda\varphi(f)}, \tag{4.1}$$

and similarly for Π in the case of CCR's. Hence we find for the smoothing operator of Eq. (3.2)

$$A_{\{c\}} = \text{s-}\lim_{n\to\infty}\int \frac{d\lambda_1}{\sqrt{4\pi}}\cdots\frac{d\lambda_n}{\sqrt{4\pi}}\left(\prod_1^n c_\nu\right)\exp\left\{-\tfrac{1}{4}\sum_{\nu=1}^n c_\nu^2\lambda_\nu^2\right\}U\left(\sum_1^n \lambda_\nu g_\nu\right), \tag{4.2}$$

and a similar expression for that in Eq. (1.9). For CCR's, $A_{\{c\}}B_{\{\hat{c}\}}$ can be expressed as[13]

$$A_{\{c\}}A_{\{\hat{c}\}} = \text{s-}\lim_{n\to\infty}\int\frac{d\lambda_1 d\hat{\lambda}_1}{4\pi}\cdots\frac{d\lambda_n d\hat{\lambda}_n}{4\pi}$$
$$\times\left(\prod_1^n c_\nu\hat{c}_\nu\right)\exp\left(-\tfrac{1}{4}\sum_1^n(c_\nu^2\lambda_\nu^2+\hat{c}_\nu^2\hat{\lambda}_\nu^2)\right)U\left(\sum_1^n\lambda_\nu g_\nu\right)V\left(\sum_1^n\lambda_\nu g_\nu\right) \tag{4.3}$$

and an analogous formula with the Hermite functions for the test function space $\mathcal{S}$.

The expressions of Eqs. (4.2) and (4.3) are very similar to the ones given by Gårding[14] for representations of finite-dimensional Lie groups where the group operators are integrated with C_0^∞ functions. One of course can in the present case integrate the group operators also with other functions; the main difference to Gårding's procedure is that one first has to integrate over n-dimensional subgroups and then take the limit $n\to\infty$.[15]

[1] J. R. Klauder, Commun. Math. Phys. **18,** 307 (1970); Acta Phys. Austriaca Suppl. **8,** 227 (1971) (Schladming Lectures 1971).

[2] G. C. Hegerfeldt, J. Math. Phys. **13,** 821 (1972). For the special case of Gårding–Wightman representations of the CCR's see also M. C. Reed, Commun. Math. Phys. **14,** 336 (1969).

[3] A vector $|\psi>$ satisfying Eq. (1.4) for some $t>0$ (or for all t) is called *analytic* for $\mathcal{P}(f)$ (or *entire*, respectively).

[4] Note that Eq. (1.9) has to be regarded as a strong limit. $A_{\{c\}} = \text{s-lim}_{|n|\to\infty}\Pi_{|\nu|\le|n|}\exp[-\mathcal{P}(h_\nu)^2/c^2{}_\nu]$, where $|\nu| = \nu_1+\nu_2+\nu_3$. The strong limit exists and defines a bounded operator which may or may not be zero.

[5] Again $B_{\{\hat{c}\}}$ is a strong limit.

[6] $\Sigma c_\nu^{-2}\mathcal{P}(h_\nu)^2|\psi>$ may well diverge for a particular $\{c\}$ and every $|\psi>$ for which each term is defined. Then $A_{\{c\}}$ would be identically zero. But also when $A_{\{c\}}$ is not the zero operator it may annihilate some vectors. This is reflected by the fact that the above sum may not be densely defined.

[7] A. Pietsch, *Nukleare lokalkonvexe Räme* (Springer-Verlag, Berlin, 1969), 2nd ed., Sec. 10.3.4; L. Schwartz, *Théorie des distributions* (Hermann, Paris, 1966), p. 262. One uses $(\nu+1)^\nu$ rather than ν^ν in Eq. (2.4) in order to cover the case $\nu_i = 0$.

[8] G. C. Hegerfeldt and O. Melsheimer, Commun. Math. Phys. **12,** 304 (1969).

[9] $||\Lambda||_{-c}$ may become infinite for some Λ. Then we put $\hat{\psi}(\Lambda) = 0$. If the set of Λ's for which $||\Lambda||_{-c}$ is finite has measure zero, then $\hat{\psi}(\Lambda)$ corresponds to the zero vector.

[10] One can easily see that $||\mathcal{P}(h_\alpha)\Pi_\nu\exp\{-c_\nu^{-2}\mathcal{P}(h_\nu)^2\}||\le||\mathcal{P}(h_\alpha)\exp[-c_\alpha^{-2}\mathcal{P}(h_\alpha)^2]|| = \sup|\eta'_\alpha\exp\{-c_\alpha^{-2}\eta'^2_\alpha\}| = \max|x|\exp(-c_\alpha^{-2}\cdot x^2)$. In this way one can show that $|\hat{\psi}>$ is an entire vector.

[11] This is again a limit as in footnote 4.

[12] the same remark as in footnote 9 applies since $||\eta'||^2_{-c} = \Sigma_\nu c_\nu^{-2}\eta'^2_\nu$ may be infinite.

[13] If $A_n\to A$ and $B_n\to B$ strongly and if $\|A_n\|\le M$ and $\|B_n\|\le M$, then $A_nB_n\to AB$ strongly.

[14] L. Gårding, Proc. Natl. Acad. Sci. USA **33,** 331 (1947).

[15] A related type of group integral for the CCR's has been investigated by J. R. Klauder and J. McKenna, J. Math. Phys. **6,** 68 (1965); H. Araki and E. J. Woods, Publ. Res. Inst. Math. Sci. **A2,** 157 (1966); G. C. Hegerfeldt, J. Math. Phys. **10,** 1681 (1969); J. Math. Phys. **11,** 21 (1970).

Limit of the most degenerate representations of $SO(p,1)$

F. Ardalan*

Department of Physics, Arya-Mehr University of Technology, Tehran, Iran
(Received 4 September 1973)

The limit of the group $SO(p,1)$, for large Lorentz transformations, are studied. It is found that the basis functions of the most degenerate representations of $SO(p,1)$, constructed according to the subgroups $SO(p)$ and $SO(p-1,1)$, asymptotically tend to the basis functions, constructed according to the corresponding contracted subgroups. This result is applied to the matrix elements of $SO(p,1)$ to derive novel relations. The limit of the overlap functions are considered.

1. INTRODUCTION

Theory of groups and their representations has become an established part of modern physical theory and diverse aspects of this theory have been extensively studied in recent years. One such aspect is the concept of group limit and group contraction for Lie groups.[1,2] By group contraction it is generally understood that one has in mind a certain deformation of the Lie algebra which corresponds to a singular transformation of the algebra. Such is the case for the contraction of the rotation group $SO(3)$ with respect to its subgroup $SO(2)$, where the generator of the $SO(2)$, J_3, is left intact, while the other two generators J_1, and J_2 are multiplied by ϵ, which in turn tends to zero. Thus in the limit $\epsilon \to 0$, ϵJ_1, ϵJ_2, and J_3 form the Lie algebra of the group of rigid rotations of the plane, i.e., $E(2)$. In other words the $SO(3)$ group manifold is deformed into the $E(2)$ manifold. To see more clearly what is involved in this contraction, it is sufficient to consider the homogeneous spaces $SO(3)/SO(2)=S^2$ and $E(2)/SO(2)=R^2$. Under contraction, finite rotations on S^2 generated by J_1 and J_2 become infinitesimal rotations on S^2 and finite translations on R^2. R^2 is the tangent plane to S^2.

The concept of group contraction is applicable to a wide range of Lie groups and is essentially a local concept.[1] Another important related concept is that of the group limit for noncompact groups.[2] By group limit we mean the asymptotic form of the group manifold when one or several of the parameters of the group are allowed to become large. In other words one is interested in the behavior of the group manifold at infinity. A simple visual aid is again the homogeneous space of the noncompact group, for example $SO(2,1)/SO(2)=H^2$, the two-dimensional hyperboloid. A large boost will take a finite point of H^2 to its asymptotic region. It will be seen that in this region the group acts as the group of linear transformations of the line. Clearly this group limit is not the same as $E(1,1)$, the group of rigid motions of the Lorentzian plane, which is obtained by contracting $SO(2,1)$ in the conventional manner.

From the foregoing it is clear that it is the group limit which is most relevent to the physics of the infinite momentum, and generally to high energy phenomena.[3] But in the studies of group theory and their repesentations the group contraction has been more closely pursued than the group limit in the above sense. Even when the latter has been considered, the concern has been the group structure and the representation types.[2] The question of the behavior of the representation space under the limiting procedure has been largely untouched.[4] It is the purpose of this work to tackle this problem in a specific example in detail.

The concrete group we will address ourselves to is $SO(p,1)$, the group of Lorentz transformations in $p+1$ dimensions. This group plays an important role in physical theory, both for its significance in relativistic kimematics when $p=3$, and for its usefulness in algebraic descriptions of hadrons.[5] The representations of this group have been studied in some detail; in particular, the most degenerate continuous representations of $SO(p,1)$ have been constructed both according to the maximal compact subgroup $SO(p,1)\supset SO(p)$, and according to the $SO(p,1)\supset SO(p-1,1)$.[6] In this article, the limit of these representations are studied and various consequences of the limit relations are exhibited.

In Sec. 2, we will review the basic elements of the $SO(p,1)$ representation theory. In Sec. 3, we will construct the most degenerate representations of $SO(p,1)$ according to the $E(p-1)$ subgroup. In Sec. 4 the relevant aspects of the group limit are discussed and then in Sec. 5 applied to a study of the behavior of the basis elements for the representation space under a large boost. It is found that the vectors of the canonical basis, the basis constructed by diagonalizing the Casimir operator of the $SO(p)$ subgroup, are transformed into the basis vectors constructed by diagonalizing the Casimir operator of the Euclidean subgroup $E(p-1)$. This limit relation will allow us, then, to read off the limit relation between the matrix elements of a group element in a representation constructed according to the subgroup $SO(p)$ and a representation constructed according to its contraction $E(p-1)$. Similarly the relation between two sets of overlap functions are found.

Finally, matrix elements of the form $\langle A|\exp(-iaK)|B\rangle$, where $|A\rangle$ and $|B\rangle$ are two "rest" states and K is the generator of the boost, are found in the limit of large a. These matrix elements emerge in resonance model calculations of structure functions in the deep inelastic scattering region.[7]

2. THE MOST DEGENERATE REPRESENTATION OF $SO(p,1)$

The most degenerate representation of the group $SO(p,1)$ may be built on $\mathcal{L}^2(H^{p-1})$, the Hilbert space of square integrable functions on the two sheeted hyperboloid,[8]

$$\eta^\mu \in H^{p-1} \Rightarrow \eta^\mu \eta_\mu = 1, \quad \mu = 0,1,2,\ldots,p-1, \tag{2.1}$$

$$f,g \in \mathcal{L}^2(H^{p-1}), \quad (f,g)=\int_{H_{p-1}} d\eta \bar{f}(\eta) g(\eta). \tag{2.2}$$

The generators of the group $SO(p,1)$ for this representation are

$$S_{\mu\nu}=i(\eta_\mu\delta_\nu-\eta_\nu\delta_\mu),\quad \delta_\mu=\frac{\partial}{\partial\eta^\mu}-\eta_\mu\eta^\nu\frac{\partial}{\partial\eta^\nu}, \tag{2.3}$$

which form the Lie algebra of an $SO(p-1,1)$ subalgebra, and

$$\Gamma_\mu=i(\delta_\mu+\sigma\eta_\mu), \tag{2.4}$$

which together with $S_{\mu\nu}$ commute as

$$[S_{\mu\nu},S_{\sigma\lambda}]=i(g_{\nu\sigma}S_{\mu\lambda}-g_{\mu\sigma}S_{\nu\lambda}-g_{\nu\lambda}S_{\mu\sigma}+g_{\mu\lambda}S_{\nu\sigma}), \tag{2.5}$$

$$[S_{\mu\nu},\Gamma_\lambda]=i(g_{\nu\lambda}\Gamma_\mu-g_{\mu\lambda}\Gamma_\nu), \tag{2.6}$$

$$[\Gamma_\mu,\Gamma_\nu]=iS_{\mu\nu}, \tag{2.7}$$

to form the algebra of $SO(p,1)$.

It is known that $\{S_{\mu\nu},\Gamma_\mu\}$ form an irreducible representation of $SO(p,1)$ and for

$$\sigma=-\tfrac{1}{2}(p-1)+i\rho,\quad \rho \text{ real}. \tag{2.8}$$

The representation is unitary.[6] The canonical basis for these $SO(p,1)$ representations may be obtained by decomposition according to the $SO(p)$ subgroup generated by

$$\{S_{\alpha\beta},\Gamma_\alpha\},\quad \alpha,\beta=1,2,\ldots,p-1$$

and are

$$\Psi_{(n)}(\eta)=N_{(n)}\cosh^\sigma a\tanh^l a C_{n-l}^{l+(p-2)/2}\left(\mp\frac{1}{\cosh a}\right)Y_{(l)}(\Omega_{p-2}),$$

$$N_{(n)}=2^{l+(p-3)/2}\Gamma\left(l+\frac{p-2}{2}\right)\left(\frac{[n+(p-2)/2]\Gamma(n-l+1)}{\pi\Gamma(n+l+p-2)}\right)^{1/2} \tag{2.9}$$

Here $(n)=(n_1,n_2,\ldots,n_{p-1})$ symbolizes the set of eigenvalues of the decomposition chain

$$SO(p)\supset SO(p-1)\cdots\supset SO(2),\quad n=n_{p-1} \tag{2.10}$$

and $(l)=(n_1,n_2,\ldots,n_{p-2})$ symbolizes the eigenvalues of the decomposition chain

$$SO(p-1)\supset SO(p-2)\cdots\supset SO(2),\quad n_{p-2}=l. \tag{2.11}$$

Here the spherical coordinates of H^{p-1} are employed:

$$\begin{aligned}
&\eta^0=\mp\cosh a, && 0\le a<\infty\\
&\eta^1=\sinh a\sin\theta_{p-2}\cdots\sin\theta_1, && 0\le\theta_1<2\pi\\
&\eta^2=\sinh a\sin\theta_{p-2}\cdots\sin\theta_2\cos\theta_1, && 0\le\theta_i\le\pi\\
&\vdots\\
&\eta^{p-2}=\sinh a\sin\theta_{p-2}\cos\theta_{p-3},\\
&\eta^{p-1}=\sinh a\cos\theta_{p-2}.
\end{aligned} \tag{2.12}$$

$Y_{(l)}(\Omega)$ are the usual spherical harmonics, forming the canonical basis for the $SO(p-1)$ subgroup.

Similarly a pseudobasis may be constructed according to the decomposition

$$SO(p,1)\supset SO(p-1,1)\supset SO(p-1)\supset SO(p-2)\cdots\supset SO(2), \tag{2.13}$$

$$\Psi_{\nu,(l)}(\eta)=N_\nu\sinh a^{-(p-3)/2}P^{-[l+(p-3)/2]}_{-1/2+i\nu}(\cosh a)Y_{(l)}(\Omega_{p-1}),$$

$$N_\nu=\left(\frac{\nu\sinh\pi\nu}{\pi}\right)^{1/2}\left|\Gamma\left(i\nu+l+\frac{p-2}{2}\right)\right|. \tag{2.14}$$

3. THE DECOMPOSITION $SO(p,1)\supset E(p-1)$

We accomplish the decomposition of the most degenerate representation of the group $SO(p,1)$ with respect to its Euclidean subgroup $E(p-1)$ by introducing the parabolic coordinate system on H^{p-1} which parametrizes an element η^μ of H^{p-1} according to

$$\begin{aligned}
&\eta^0+\eta^{p-1}=\mp e^{a'},\quad \eta^0-\eta^{p-1}=\mp(e^{-a'}+e^{a'}r^2), && -\infty<a'<\infty\\
&\eta^1=e^{a'}r\sin\theta_{p-3}\cdots\sin\theta_1, && 0\le\theta_1<2\pi\\
&\eta^2=e^{a'}r\sin\theta_{p-3}\cdots\sin\theta_2\cos\theta_1, && 0\le\theta_i\le\pi\\
&\vdots\\
&\eta^{p-3}=e^{a'}r\sin\theta_{p-3}\cos\theta_{p-2},\\
&\eta^{p-2}=e^{a'}r\cos\theta_{p-3}.
\end{aligned} \tag{3.1}$$

The subgroup $E(p-1)$ is generated by

$$\begin{aligned}
&E_i=K_i-S_{i,p-1},\quad E_{p-1}=\Gamma_0-\Gamma_{p-1},\\
&L_{ij}=S_{ij},\quad L_{i,p-1}=\Gamma_i,\quad i=1,2,\ldots,p-2,
\end{aligned} \tag{3.2}$$

where $K_i=-S_{0i}$. It is easily verified that

$$[E_\alpha,E_\beta]=0 \tag{3.3}$$

$$[L_{\alpha\beta},E_\gamma]=i(g_{\beta\gamma}E_\alpha-g_{\alpha\gamma}E_\beta). \tag{3.4}$$

Note that E_i and L_{ij} generate an $E(p-2)$ subgroup of $E(p-1)$; and L_{ij} generate an $SO(p-2)$ subgroup of $E(p-2)$.

The representation of $SO(p,1)$ are decomposed according to the chain,

$$\begin{aligned}
SO(p,1)&\supset E(p-1)\supset E(p-2)\supset SO(p-2)\\
&\supset SO(p-3)\cdots\supset SO(2).
\end{aligned} \tag{3.5}$$

The Casimir operator of $E(p-1)$ is $E_\alpha E_\alpha$, where in terms of the differential operators on H^{p-1}, E_α are

$$E_i=-i\frac{\partial}{\partial x^i},\quad \eta^i=\frac{x^i}{z},\quad z=e^{-a'}, \tag{3.6}$$

$$E_{p-1}=-i\frac{\partial}{\partial t}-i\frac{\sigma}{z}. \tag{3.7}$$

We call the eigenvalues of this Casimir operator Λ and find the corresponding pseudobasis functions:

$$\Psi^{(\mp)}_{\Lambda,x,(m)}(\eta)=(x\Lambda)^{1/2}z^{-\sigma+1/2}J_{\mp1/2}(\mu z)J_{m+(p-4)/2}(xr)Y_{(m)}(\Omega_{p-3}). \tag{3.8}$$

$\Psi_{\Lambda,x,(m)}$ are eigenfunctions of the Casimir operators of the above subgroup chain,

$$(E_iE_i)\Psi^{(\mp)}_{\Lambda,x,(m)}=x^2\Psi^{(\mp)}_{\Lambda,x,(m)}, \tag{3.9}$$

$$\tfrac{1}{2}(L_{ig}L_{ig})\Psi^{(\mp)}_{\Lambda,x,(m)}=m(m+p-4)\Psi^{(\mp)}_{\Lambda,x,(m)},\ \text{etc.}, \tag{3.10}$$

and satisfy the following orthogonality and completeness relations

$$\int_{H^{p-1}}d\eta\,\overline{\Psi^{(\mp)}_{\Lambda,x,(m)}(\eta)}\Psi^{(\mp)}_{\Lambda',x',(m')}(\eta)=\delta(\Lambda-\Lambda')\delta(x-x')\delta_{(m),(m')}, \tag{3.11}$$

$$\sum_{(m)}\int d\Lambda\,dx\Psi_{\Lambda,x,(m)}(\eta)\overline{\Psi_{\Lambda,x,(m)}(\eta')}=\delta(\eta-\eta'). \tag{3.12}$$

For reference we will write down the pseudobasis func-

tions for the $SO(p,1)$ representation according to the decomposition chain,

$$SO(p,1)\supset SO(p-1,1)\supset E(p-2)\supset SO(p-2)\supset SO(p-3)\cdots\supset SO(2). \tag{3.13}$$

These functions are the generalized eigenfunctions of the Casimir operator of $SO(p-1,1)$, the Laplace–Beltrami operator on H^{p-1}:

$$\tfrac{1}{2}S_{\mu\nu}S^{\mu\nu}=-\delta^2=z^2\frac{\partial^2}{\partial z^2}-(p-3)z\frac{\partial}{\partial z}+z^2\nabla^2_{p-2}, \tag{3.14}$$

where ∇^2_{p-2} is the Laplacian in $(p-2)$-dimensional Euclidean space. The eigenfunctions of δ^2 are

$$\Psi_{\nu,x,(m)}(\eta)=\left(\frac{2x\nu\sinh\pi\nu}{\pi^2}\right)^{1/2}z^{(p-2)/2}K_{i\nu}(xz)r^{-(p-4)/2}\times J_{m+(p-4)/2}(xr)Y_{(m)}(\Omega_{p-3}), \tag{3.15}$$

where the $K_{i\nu}$ are the modified Bessel functions of order $i\nu$.[9]

4. GROUP LIMIT

In this section some elementary results on group limit theory as applied to the group $SO(p,1)$ will be discussed. Under the action of $\mathrm{Ad}\exp(-itK_{p-1})$, one can easily verify, by considering matrix representations, that the Lie algebra of $SO(p,1)$ approaches, the following limits:

$$\begin{aligned}
&\lim \mathrm{Ad}\exp(itK_{p-1})(S_{ij})=S_{ij},\\
&\lim \mathrm{Ad}\exp(itK_{p-1})(S_{i,p-1})=\tfrac{1}{2}e^tE_i,\\
&\lim \mathrm{Ad}\exp(itK_{p-1})(K_i)=\tfrac{1}{2}e^tE_i,\\
&\lim \mathrm{Ad}\exp(itK_{p-1})(K_{p-1})=K_{p-1}\\
&\lim \mathrm{Ad}\exp(itK_{p-1})(\Gamma_i)=\Gamma_i,\\
&\lim \mathrm{Ad}\exp(itK_{p-1})(\Gamma_{p-1})=\tfrac{1}{2}e^tE_{p-1},\\
&\lim \mathrm{Ad}\exp(itK_{p-1})(\Gamma_0)=\tfrac{1}{2}e^tE_{p-1},\quad i=1,2,\ldots,p-2.
\end{aligned} \tag{4.1}$$

Thus the Lie algebra of $SO(p,1)$ approaches that of $E(p-1)\otimes D$, where D is dilation on R^{p-1}. Here $D=K_{p-1}$,

$$[K_{p-1},E_\alpha]=iE_\alpha,\quad \alpha=1,2,\ldots,p-1, \tag{4.2}$$

$$[K_{p-1},\Gamma_i]=0, \tag{4.3}$$

$$[K_{p-1},S_{ij}]=0. \tag{4.4}$$

It must be noted that to be exact one should consider the limit of $\mathrm{Ad}\exp(itK_{p-1})(K_i/\frac{1}{2}e^t)$, etc., in order to obtain a well defined, finite Lie algebra instead of the asymptotic limits considered above; and it is in this sense that one Lie algebra is the limit of the other. But we prefer to keep our asymptotic notation, as it makes contact with its physical applications more readily.[3]

In a related but different procedure, contraction of $SO(p-1)$ with respect to its subgroup $SO(p-2)$, a similar connection between the two algebras may be obtained. In this case the elements $S_{i,p-1}$ are replaced by $\epsilon S_{i,p-1}$ and ϵ is let to go to zero. The resultant Lie algebra is that of $E(p-2)$. As a consequence it can be shown that the Legendre functions asymptotically approach the Bessel functions. The usual derivation of this result is as follows[10]: One considers a representation of $SO(p-1)$ with the weight l, and calculates the matrix elements of a group element, e.g., $\exp[(i\theta/l)S_{i,p-1}]$ in this representation, which are the Legendre functions, and then allows l to go to infinity, $(1/l)S_{i,p-1}$ tends to E_i, and as direct calculation shows, the matrix elements of $(1/l)S_{i,p-1}$ between the two $SO(p-1)$ states approach the matrix elements of E_i between two $E(p-2)$ states. Since the matrix elements of $\exp(i\theta E_i)$ are the Bessel functions, one has the desired relation.

In the next section we will prove the stronger result that the basis elements in the $SO(p,1)$ representations approach the basis elements of the representations of the limit group $E(p-1)\otimes D$. From this result, the connection between the matrix elements of two groups in addition to some novel relations will emerge immediately.

5. LIMIT OF REPRESENTATION BASIS

The elements of the canonical basis for the most degenerate representation of the group $SO(p,1)$ which are constructed according to the decomposition (2.10) are determined by the eigenvalue equation chain,

$$\begin{aligned}
&(\Gamma_\mu\Gamma^\mu+\tfrac{1}{2}S_{\mu\nu}S^{\mu\nu})\Psi_{(n)}=\{-[(p-1)/2]^2-\rho^2\}\Psi_{(n)},\\
&\qquad\qquad \mu=0,1,2,\ldots,p-1,\\
&(\Gamma_\alpha\Gamma_\alpha+\tfrac{1}{2}S_{\alpha\beta}S_{\alpha\beta})\Psi_{(n)}=n(n+p-2)\Psi_{(n)},\\
&\qquad\qquad \alpha,\beta=1,2,\ldots,p-1,\\
&\tfrac{1}{2}(S_{\alpha\beta}S_{\alpha\beta})\Psi_{(n)}=l(l+p-3)\Psi_{(n)},\\
&\tfrac{1}{2}(S_{ij}S_{ij})\Psi_{(n)}=m(m+p-4)\Psi_{(n)},\quad i=1,2,\ldots,p-2.
\end{aligned}$$

Under the action of a Lorentz boost $\exp(itK_{p-1})$, these eigenvalue equations are transformed, asymptotically, into the chain

$$(\Gamma_\mu\Gamma^\mu+\tfrac{1}{2}S_{\mu\nu}S^{\mu\nu})\exp(itK_{p-1})\Psi_{(n)}=\{-[(p-2)/2]^2-\rho^2\}\exp(itK_{p-1})\Psi_{(n)}, \tag{5.1}$$

$$(\tfrac{1}{4}e^{2t}E_\alpha E_\alpha)\exp(itK_{p-1})\Psi_{(n)}=n(n+p-2)\exp(itK_{p-1})\Psi_{(n)}, \tag{5.2}$$

$$(\tfrac{1}{4}e^{2t}E_iE_i)\exp(itK_{p-1})\Psi_{(n)}=l(l+p-3)\exp(itK_{p-1})\Psi_{(n)}, \tag{5.3}$$

$$\tfrac{1}{2}(S_{ij}S_{ij})\exp(itK_{p-1})\Psi_{(n)}=m(m+p-4)\exp(itK_{p-1})\Psi_{(n)}. \tag{5.4}$$

These equations suggest that the boosted states $\exp(itK_{p-1})\Psi_{(n)}$ form a basis for the representation of $SO(p,1)$ according to the decomposition (3.5), with the characteristic parameters,

$$\Lambda=\lim_{\substack{t\to\infty\\ n\to\infty}}2ne^{-t}, \tag{5.5}$$

$$x=\lim_{\substack{t\to\infty\\ n\to\infty}}2le^{-t}, \tag{5.6}$$

other parameters being identical to those of the canonical basis.

Similarly, the pseudobasis for the decomposition (2.13) is suggested to be transformed into the pseudobasis for the decomposition (3.13). In the following we will prove that these functions behave as suggested and in fact the limit relations are pointwise.

To begin with, we consider the effect of a large boost on one of our basis functions. By definition the effect of

$\exp(itK_{p-1})$ on such a function f is

$$[\exp(itK_{p-1})f](\eta)=f[\exp(-itK_{p-1})\eta]. \tag{5.7}$$

Therefore, it is sufficient to consider the effect of the boost $\exp(-itK_{p-1})$ on the vector η^μ as t tends to infinity, and find the limit of $f[\exp(-itK_{p-1})\eta]$ as a function of η^μ. We consider the spherical parametrization of the hyperboloid and seek the parameters of $\exp(-itK_{p-1})\eta$ in the parabolic coordinate representation (3.1). The result is

$$\cosh a_{t\ t\to\infty}\to \lambda/z,\quad \lambda=\tfrac{1}{2}e^t, \tag{5.8}$$

$$\coth a_{t\ t\to\infty}\to \cosh(z/\lambda), \tag{5.9}$$

$$\theta_{t\ t\to\infty}\to r/\lambda, \tag{5.10}$$

where $a_t, \theta_t=\theta^t_{p-2}$ are the relevant parameters of $\exp(-itK_{p-1})\eta$, and z and r are the relevant parabolic parameters of η^μ.

As a first case, consider the pseudobasis functions (2.14) of the decomposition, $SO(p,1)\supset SO(p-1,1)\supset SO(p-1)$. The asymptotic limit of the normalized spherical harmonics,

$$Y_{(l)}(\Omega_{p-2})=2^{m+(p-4)/2}\Gamma\left(m+\frac{p-3}{2}\right)\times\left(\frac{[l+(p-3)/2]\Gamma(l-m+1)}{\pi\Gamma(l+m+p-3)}\right)^{1/2}\times C^{m+(p-3)/2}_{l-m}(\cos\theta)\sin^m\theta Y_{(m)}(\Omega_{p-3}), \tag{5.11}$$

where $m=n_{p-3}$, are easily found to be,

$$(-1)^{(m+p)/2}\lambda^{(p-3)/2}[x^{1/2}\gamma^{-[(p-4)/2]}J_{m+(p-4)/2}(x\gamma)Y_{(m)}(\Omega_{p-2})]. \tag{5.12}$$

Here $x=\lim_{\lambda,\,l\to\infty}(l/\lambda)$, and the expression inside the bracket is the normalized eigenfunction of $E(p-2)$ Casimir operator, as expected from the fact the $E(p-2)$ is the contraction of $SO(p-1)$. The factor $(-1)^{(m+p)/2}$ is a phase factor related to the arbitrariness in the definition of the functions in the limiting expressions; and the $\lambda^{(p-3)/2}$ factor is the result of a change of normalization; the $Y_{(l)}$ are normalized to $\delta_{(l),(l')}$, while the generalized eigenfunctions of the $E(p-2)$ Casimir operator are normalized according to

$$\delta(x-x')=\lim_{\lambda,\,l,\,l'\to\infty}\lambda\delta(l-l').$$

Note that the limiting procedure invoked here, introduces additional powers of λ, coming from the change of measure,

$$d\Omega_{p-2}=\sin^{p-3}\theta\,d\theta\,d\Omega_{p-3}\,\lambda^{-(p-2)}\gamma^{p-3}\,dr\,d\Omega_{p-3}, \tag{5.13}$$

which explains the factor $\lambda^{(p-3)/2}$ appearing in the expression (5.12).

In the derivation of (5.12), we have used the connection between the Gegenbauer polynomials and Legendre functions, and then the limit relation.[11]

$$\lim_{\nu\to\infty}\nu^\mu P^{-\mu}_\nu(\cos x/\nu)=J_\mu(x). \tag{5.14}$$

Similarly the limit of the remaining factors in the expression (2.14) for the pseudobasis functions may be obtained by exploiting Whipple's formula,[11] connecting the Legendre functions of the first kind with those of the second kind, and using the limit relation[11]

$$\lim_{\nu\to\infty}\nu^{-\mu}\exp(-i\pi\mu)Q^\mu_\nu[\cosh(z/\nu)]=K_\mu(z). \tag{5.15}$$

The final result is

$$[\exp(itK_{p-1})\Psi_{\nu,(l)}](\eta)_{\,l,\,t\to\infty}\to(-1)^{(m+p)/2}\lambda^{-1/2}\Psi_{\nu,x,(m)}(\eta). \tag{5.16}$$

The coefficient $(-1)^{(m+p)/2}$ is a phase factor and the additional power of λ is again due to the change of normalization.

The limit of the canonical basis (2.9) can be similarly found. In this case, we obtain the pseudobasis (3.8) of the decomposition $SO(p,1)\supset E(p-1)$, only when we let n, l, and $n-l$, tend to infinity simultaneously with t. To find the limit behavior of (2.9), we observe that when $n-l=2N$, for nonnegative integer N,[11]

$$C^{l+(p-2)/2}_{n-l}\left(\mp\frac{1}{\cosh a}\right)=(-1)^N\frac{2^N N!\,\Gamma[N+l+(p-2)/2]}{(2N)!\,\Gamma[l+(p-2)/2]}\times P^{[-1/2,\,l+(p-3)/2]}_N\left(1-\frac{2}{\cosh^2 a}\right). \tag{5.17}$$

Then we use the expansion for the Jacobi polynomials,[11]

$$P^{[-1/2,\,l+(p-3)/2]}_N\left(1-\frac{2z^2}{\lambda^2}\right)=\sum_{k=0}^N b_k\left(\frac{z}{\lambda}\right)^{2k}, \tag{5.18}$$

$$b_k=\frac{(-)^k}{k!(N-k)!}\cdot\frac{\Gamma[N+l+k+(p-2)/2]}{\Gamma[N+l+(p-2)/2]}\cdot\frac{\Gamma(N+\frac{1}{2})}{\Gamma(k+\frac{1}{2})},$$

and find that

$$\lim_{\lambda,\,l,\,N\to\infty}P^{[-1/2,\,l+(p-3)/2]}_N\left(1-\frac{2z^2}{\lambda^2}\right)=\left(\frac{\mu z}{2N}\right)^{1/2}J_{-1/2}(\mu z) \tag{5.19}$$

where

$$\mu=(\Lambda^2-x^2)^{1/2},\quad \Lambda=\lim_{n,\,\lambda\to\infty}(n/\lambda),\quad x=\lim_{l,\,\lambda\to\infty}(l/\lambda). \tag{5.20}$$

The limit of the canonical basis is then obtained:

$$\Psi_{(n)}(\eta)_{\,l,n,\,\lambda\to\infty}\to(-)^{N+m}\lambda^{\sigma+(p-3)/2}\Psi^{(-)}_{\Lambda,x,(m)}(\eta), \tag{5.21}$$

where $\Psi^{(-)}$ is defined in (3.8). The factor $\lambda^{\sigma+(p-3)/2}$ is again due to a change of normalization.

In exactly the same manner we find that, for the case $n-l=2N+1$,

$$\Psi_{(n)}(\eta)_{\,l,n,\,\lambda\to\infty}\to\mp(-)^{N+m}\lambda^{\sigma+(p-3)/2}\Psi^{(+)}_{\Lambda,x,(m)}(\eta). \tag{5.22}$$

These limit relations can be used now to relate matrix elements of $SO(p-1)$ group elements with those of $E(p-2)$ group elements. To illustrate the point we choose $p=4$, and calculate the limit of,[12]

$$P^l_{m,m'}[\cos(\alpha/\lambda)]\equiv\int d\Omega\,Y^m_l(\Omega)\exp[i(\alpha/\lambda)J_2]Y^{m'}_l(\Omega),$$

when l and λ tend to infinity. Here $P^l_{m,m'}$ are related to the Jacobi polynomials. From (5.12), we have,

$$Y^m_l(\Omega)\to i^m\lambda^{1/2}\left(x^{1/2}J_m(x\gamma)\frac{1}{(2\pi)^{1/2}}\exp(im\phi)\right),$$

where the expression in the large parenthesis is the normalized generalized eigenstates of the Casimir

operator of $E(2)$. Also from (4.1) we have,

$$J_2 \to \lambda E_1.$$

Using these two relations, and paying attention to the change of normalization, we obtain the known result,[12]

$$P^l_{m,m'}[\cos(\alpha/\lambda)]_{l,\vec{\lambda}\to\infty} i^{m'-m} J_{m'-m}(x\alpha). \tag{5.23}$$

A further consequence of our limit relations between the basis elements is the simple calculation of limit of the overlap functions as expressed in Ref. 6. These functions, $K^\rho_l(\nu, n)$, are inner products of the elements of the canonical basis (2.9) and the corresponding elements of the pseudobasis (2.14),

$$K^\rho_l(\nu, n) = \int_{H^{p-1}} d\eta\, \overline{\Psi}_{n,(l)}(\eta)\Psi_{\nu,(l)}(\eta). \tag{5.24}$$

They have been expressed in terms of the Meijer's G functions; and a direct calculation of their limits for large l and n is not trivial. However, we can use our Eqs. (5.16), (5.21), and (5.22), to find the limit of $K^\rho_l(\nu, n)$ easily:

$$K^\rho_l(\nu, n)_{l,\vec{n}\to\infty} (\mp)^{(1+\epsilon)/2}(-1)^{N+(m+\rho)/2}\lambda^{-1/2-i\rho}\left(\frac{\Lambda\nu\sinh\pi\nu}{\pi^2}\right)^{1/2} I,$$

$$I = \int_0^\infty dz z^{i\rho} K_{i\nu}(xz) J_{\epsilon/2}(\mu z), \tag{5.25}$$

where $\epsilon = -1$, for the case $n - l = 2N$, and $\epsilon = +1$, for $n - l = 2N+1$. $(\mp)$ refers to either sheets of the hyperboloid. The integral can be calculated, and the result is

$$K^\rho_l(\nu, n)_{l,n,\vec{\lambda}\to\infty} (\mp)^{(1+\epsilon)/2}(-)^{N+(m+\rho)/2}\lambda^{-1/2-i\rho} C_{\rho,\nu} B_{\rho,\nu}(x, \Lambda),$$

$$C_{\rho,\nu} = 2^{i\rho-1}\left(\frac{\nu\sinh\pi\nu}{\pi^2}\right)^{1/2}\Gamma\left(\frac{1+\epsilon/2+i(\rho+\nu)}{2}\right)$$
$$\times\Gamma\left(\frac{1+\epsilon/2+i(\rho-\nu)}{2}\right)\Big/\Gamma(1+\epsilon/2),$$

$$B_{\rho,\nu}(x, \Lambda) = \mu^{\epsilon/2}\Lambda^{1/2}x^{-(1+\epsilon/2+i\rho)}$$
$$\times F\left(\frac{1+\epsilon/2+i(\rho+\nu)}{2}, \frac{1+\epsilon/2+i(\rho-\nu)}{2}; 1+\epsilon/2; -\frac{\mu^2}{x^2}\right). \tag{5.26}$$

Finally, we will evaluate the matrix elements of a large boost between two vectors of the pseudobasis (2.13), and then between two vectors of the canonical basis (2.19). These matrix elements are known for a finite boost,[14] from which the limits can be obtained. Here we use our limit relations between the basis elements and find the matrix elements immediately. The matrix elements are

$$\langle\rho, \nu, (0)|\exp(itK_{p-1})|\rho, \nu, l, (0)\rangle \tag{5.27}$$

where the state on the left is the ground state of the basis (2.13), while the state on the right is a general element of that basis. Now as t and l tend to infinity, (5.27) tends to

$$(\mp)^{(1+\epsilon)/2}(-)^{N+m}\lambda^{-1+i\rho}\langle\rho, \nu, (0)|\rho, \nu, x, (0)\rangle, \tag{5.28}$$

where the state on the right is an element (3.8) of the Euclidean decomposition (3.13). We have used the limit relation (5.16).

To calculate the overlap functions in (5.28), we consider

$$\langle\rho, \nu, (0)|E_iE_i|\rho, \nu, x, (0)\rangle = x^2\langle\rho, \nu, (0)|\rho, \nu, x, (0)\rangle$$
$$= \langle\rho, \nu, (0)|[(p-2)/2]^2 + \nu^2 - K^2_{p-1}$$
$$+ (p-2)iK_{p-1}|\rho, \nu, x, (0)\rangle, \tag{5.29}$$

and observe that

$$K_{p-1}|\rho, \nu, x, (0)\rangle = i\left(x\frac{d}{dx} + \frac{1}{2}\right)|\rho, \nu, x, (0)\rangle; \tag{5.30}$$

the Eq. (5.29) will lead to a differential equation for the overlap functions, whose solutions are

$$\langle\rho, \nu, (0)|\rho, \nu', x, (0)\rangle = cx^{(p-3)/2}K_{i\nu}(x)\delta(\nu - \nu') \tag{5.31}$$

with the coefficient c obtained from the completeness relation for the $|\rho, \nu, x, (m)\rangle$ and is

$$c = \frac{2}{\pi^{1/4}}\left(\frac{\Gamma[(p-1)/2]}{\Gamma[(p-2)/2]}\right)^{1/2}\left[\left|\Gamma\left(\frac{p-2}{2} + i\nu\right)\right|\right]^{-1}. \tag{5.32}$$

In the special case $P = 3$, this result agrees with that obtained by Chang and O'Raifeartaigh.[15]

The matrix elements of a large boost between two canonical basis elements can be obtained from the above result by observing that a (X_{p-1}, X_0) hyperbolic rotation may be obtained by a combination of a (X_p, X_0) hyperbolic rotation and (X_{p-1}, X_p) Euclidean rotation; alternatively one may calculate the overlap functions

$$\langle\rho, (0)|\rho, \Lambda, x, (0)\rangle$$

directly. The result is

$$\langle\rho, (0)|\exp(itK_{p-1})|\rho, n, l, (0)\rangle_{n,\vec{l},\lambda\to 0} 0, \quad n - l = 2N+1,$$
$$N = 0, 1, 2, \ldots,$$

$$\langle\rho, (0)|\exp(itK_{p-1})|\rho, n, l, (0)\rangle_{n,\vec{l},\lambda\to 0} (-)^{N+m}\lambda^{-1+i\rho}$$
$$\times\frac{4}{\Gamma(-\bar{\sigma})}\left(\frac{\Gamma(p/2)}{\pi\Gamma[(p-2)/2]}\right)^{1/2}\mu^{-1/2}\Lambda^{i\rho+1/2}x^{(p-3)/2}K_{i\rho}(\Lambda),$$
$$n - l = 2N. \tag{5.33}$$

ACKNOWLEDGMENTS

The author would like to acknowledge numerous helpful discussions with Dr. H. Partovi.

*Present address: Department of Physics, Yale University, New Haven, Connecticut 06520.

[1]E. Inonu, in *Group Theoretical Concepts and Methods in Elementary Particle Physics* edited by F. Gürsey (Gordon & Breach, New York, 1964).
[2]R. Hermann, *Lie Groups for Physicists* (Benjamin, New York, 1966); R. Hermann, *Fourier Analysis on Groups and Partial Wave Analysis* (Benjamin, New York, 1969).
[3]Tung-Mow Yan, Phys. Rev. D7 1780 (1973), and references therein; for a review of the subject in its earlier stages, see. M. B. Halpern, in *Lectures in Theoretical Physics* (Gordon and Breach, New York, 1968), Vol. XI-A.
[4]The limit of a spinor representation basis for $SO(4,1)$ has been considered in W. J. Hollman III, J. Math. Phys. 10, 1888 (1969).
[5]G. Mack and A. Salam, Ann. Phys. 53, 174 (1969); B. Hu, Nuovo Cimento 11A, 937 (1972), and references cited therein.

[6]C. P. Boyer and F. Ardalan, J. Math. Phys. 12, 2070 (1971); C. P. Boyer, J. Math. Phys. 14, 606 (1972).
[7]Raj Wilson and R. F. Peierls, "Proton Structure Functions in an Infinite Resonance $O(4,2)$ Model," BNL 17475 Preprint, Dec. 72, and references cited therein.
[8]Throughout this paper higher Greek indices will designate p-dimensional Lorentzian vectors with the metric $g_{00}=1$, $g_{11}=-1,\ldots, g_{p-1,p-1}=-1$; lower Greek indices will designate $(p-1)$-dimensional Euclidean vectors, and middle italic indices will designate $p-2$ Euclidean vectors. Unless otherwise specified Einstein summation rule will be followed.
[9]A. Erdelyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 2.
[10]W. Miller Jr., *Lie Theory and Special Functions* (Academic Press, New York, 1968).
[11]W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer-Verlag, New York, 1966).
[12]N. Ya Vilenkin, *Special Functions and the Theory of Group Representations* (American Mathematical Society, Providence, R. I., 1968).
[13]I. S. Gradshteyn and I. M Ryzhik, *Table of Integrals, Series, and Products* (Academic Press, New York, 1965).
[14]C. P. Boyer, J. Math. Phys. 12, 1599 (1971).
[15]Shau-Jin Chang and L. O'Raifeartaigh, J. Math. Phys. 10, 21 (1969).

High-frequency backscattering from an elliptic metal plate *

P. L. E. Uslenghi

Department of Information Engineering, University of Illinois at Chicago Circle, Chicago, Illinois 60680

S. W. Lee

Department of Electrical Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

(Received 11 August 1972)

The high-frequency backscattered field produced by a plane, linearly polarized electromagnetic wave obliquely incident on a perfectly conducting disk of elliptical shape, is considered. The leading term of the asymptotic expansion is obtained by means of the geometrical theory of diffraction, and it is matched to the physical optics result for normal incidence via Bessel functions. The formula thus obtained is uniformly valid for all directions of incidence and polarization; it reduces to the known result for a circular disk in the case of zero eccentricity. Numerical results are presented for direct- and cross-polarized monostatic cross sections. The difficulties encountered in obtaining higher-order terms of the asymptotic expansion are discussed in detail.

1. INTRODUCTION

The high-frequency scattering behavior of flat metallic plates has received considerable attention in recent years, mainly because such plates represent good approximations to portions of airplanes and missiles such as wings, fins, and tail structures. Two especially important studies are those on the rectangular plate by Ross[1] and on the circular disk by Knott *et al.*[2] The flat metallic plate of infinitesimal thickness and elliptical shape considered in this paper is important in its own right, because it is one of the simplest edge structures with variable radius of curvature on which Keller's theory can be tested and because it includes the circular disk and the strip as limiting cases. Also, any scattering results for an elliptic plate are at once applicable to the diffraction by an elliptic aperture in an infinite plane screen, via Babinet's principle. From a practical viewpoint, an elliptic plate is a reasonably good approximation for studying the scattering properties of the wings of certain aircraft at centimeter wavelengths.

Two general techniques are available in dealing with high-frequency scattering problems. The first is Keller's geometrical theory of diffraction (GTD), which represents by the inclusion of diffracted rays a systematic improvement of geometrical optics. The second is Ufimtsev's theory of fringe waves, which represents a systematic improvement of physical optics surface currents through addition of nonuniform components. The leading terms of the far-field asymptotic expansions are the same in both theories; however, discrepancies arise between the higher-order terms. Since both theories are heuristic in nature, a preferential choice can only be made on the basis of either experimental evidence or asymptotic expansions of exact solutions of canonical problems. Measurement data are inconclusive because the difference between the two theories is often numerically small and thus rests within limits of experimental errors. On the other hand, it has been theoretically proven that Ufimtsev's second-order term is definitely incorrect for the strip[3] and most probably also for the circular disk.[2] For these reasons, we chose GTD as the method of analysis in this paper.

In Sec. 2, the first term in the high-frequency expansion of the backscattered field is obtained for oblique incidence and arbitrary polarization of the incident wave. It arises from two optical rays which are singly diffracted at the two flash points on the edge of the elliptic plate. These are the two points where the tangent to the edge is perpendicular to the incident Poynting vector. The expression thus derived for oblique incidence fails at normal incidence, due to the caustic behavior of the diffracted rays. However, the leading term of the backscattered field for normal incidence is trivially obtained by physical optics, and a matching between the on-axis and off-axis results can be achieved by employing Bessel functions (Sec. 3). The resulting formula is uniformly valid for all directions of incidence, and reduces to a previously known result in the particular case of the circular disk.[2] Numerical values of the first-order monostatic cross section for vertical, horizontal, and cross polarization are plotted in Sec. 4. Finally, a detailed discussion of the difficulties encountered in extending the analysis to higher-order terms of the asymptotic expansion is presented in Sec. 5. In particular, the second-order backscattered field is derived for certain directions of incidence.

Criticism of GTD is usually twofold: First, the theory suffers from inherent difficulties at caustics; second, it leads to unacceptable results at or near geometric optics boundaries. Both objections can be overcome, at the price of cumbersome calculations, by boundary-layer techniques.[4] In particular, uniform asymptotic expansions which are valid at caustics have been obtained by Kravtsov[5-7] and Ludwig,[8] while an asymptotic solution which is uniformly valid near edges and shadow boundaries has been provided by Lewis and Boersma[9] and by Ahluwalia.[10] Despite these advances, it is still sometimes claimed that Keller's theory leads to an infinite value of the diffracted field at geometric optics boundaries (see, e.g., Ref. 11). Since this confusion arises from an incorrect use of the diffraction coefficients, in Appendix A we restate their correct values at and near geometric optics boundaries. These values are needed in the discussion of second-order contributions performed in Sec. 5.

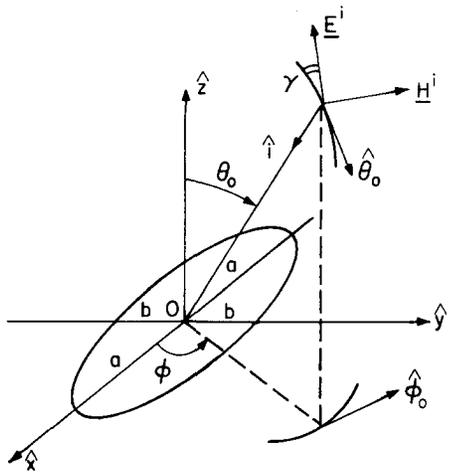


FIG. 1. Geometry for the scattering problem.

2. THE BACKSCATTERED FIELD

The geometry of the scattering problem is illustrated in Fig. 1. The elliptic plate is centered at the origin of coordinates in the (x, y) plane, has major axis $2a$ oriented in the x direction and minor axis $2b$. The incident plane electromagnetic wave with wavenumber $k=2\pi/\lambda$ and angular frequency ω propagates in the direction of the unit vector

$$\hat{i}=-\hat{x}\sin\theta_0\cos\phi_0-\hat{y}\sin\theta_0\sin\phi_0-\hat{z}\cos\theta_0 \tag{1}$$

with an incident electric field

$$\mathbf{E}^i=\hat{e}^i\exp[-ik(x\sin\theta_0\cos\phi_0+y\sin\theta_0\sin\phi_0+z\cos\theta_0)], \tag{2}$$

where

$$\begin{aligned}\hat{e}^i&=\hat{\phi}_0\sin\gamma-\hat{\theta}_0\cos\gamma\\&=-\hat{x}(\sin\gamma\sin\phi_0+\cos\gamma\cos\theta_0\cos\phi_0)+\hat{y}(\sin\gamma\cos\phi_0\\&\quad-\cos\gamma\cos\theta_0\sin\phi_0)+\hat{z}\cos\gamma\sin\theta_0\end{aligned} \tag{3}$$

and the time-dependence factor $\exp(-i\omega t)$ is omitted. Without loss of generality, the angles θ_0 and ϕ_0, which define the direction of incidence, and the polarization angle γ are restricted to the intervals

$$0\le\theta_0\le\pi/2,\quad -\pi/2\le\phi_0\le\pi/2,\quad 0\le\gamma\le\pi/2. \tag{4}$$

The far backscattered electric field may be written as

$$\mathbf{E}^{\text{b.s.}}=(e^{ikr}/kr)(\hat{\phi}_0 S_E\sin\gamma+\hat{\theta}_0 S_H\cos\gamma), \tag{5}$$

where r is the distance of the observation point from the center 0 of the plate. If the minimum radius of curvature of the edge is large compared to the wavelength, i.e.,

$$kb^2/a\ll 1, \tag{6}$$

then the far-field coefficients S_E and S_H for vertical $(\gamma=\pi/2)$ and horizontal $(\gamma=0)$ polarizations can be asymptotically expanded in series of powers of $k^{-1/2}$. The dominant terms of these series arise from optical rays which are backscattered from the flash points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ on the edge of the plate, as shown in Fig. 2. The coordinates of these points and the radius R of curvature of the edge at either point are given by

$$\begin{aligned}&x_1=-x_2=(a/A)\cos\phi_0,\quad y_1=-y_2=(b^2/aA)\sin\phi_0,\\&R=b^2/aA^3,\end{aligned} \tag{7}$$

where

$$A=A(\delta,\phi_0)=\cos\phi_0(1+\delta^2\tan^2\phi_0)^{1/2} \tag{8}$$

with

$$\delta=b/a. \tag{9}$$

By following the prescriptions of GTD,[12] the divergence factors Γ_1 and Γ_2 associated with the rays backscattered from P_1 and P_2 are easily determined

$$\begin{Bmatrix}\Gamma_1\\ \Gamma_2\end{Bmatrix}=\begin{Bmatrix}1\\ -i\end{Bmatrix}\frac{b}{r}A^{-3/2}(2a\sin\theta_0)^{-1/2}; \tag{10}$$

these expressions are valid if θ_0 is bounded away from zero, i.e., for oblique incidence only.

The first-order backscattered field is

$$\begin{aligned}\mathbf{E}^{\text{b.s.}}\sim[\exp(ikr+i\pi/4)/2\sqrt{2\pi k}]\,[&\Gamma_1\Delta_1\{\hat{e}^i\}_{P_1}\exp(-i2kA\sin\theta_0)\\&+\Gamma_2\Delta_2\{\hat{e}^i\}_{P_2}\exp(i2kA\sin\theta_0)],\end{aligned} \tag{11}$$

where the column vectors $\{\hat{e}^i\}_{P_1}$ and $\{\hat{e}^i\}_{P_2}$ are expressed in terms of the local base vectors $\hat{T}_1$, $\hat{N}_1$, $\hat{B}_1$ at P_1 and $\hat{T}_2$, $\hat{N}_2$, $\hat{B}_2$ at P_2, respectively (see the Appendix):

$$\begin{aligned}&\hat{T}_1=-\hat{T}_2=-\hat{x}\sin\phi_0+\hat{y}\cos\phi_0,\\&\hat{N}_1=-\hat{N}_2=\hat{x}\cos\phi_0+\hat{y}\sin\phi_0,\\&\hat{B}_1=\hat{B}_2=-\hat{z};\end{aligned} \tag{12}$$

hence,

$$\{\hat{e}^i\}_{P_1}=\begin{pmatrix}\sin\gamma\\ -\cos\gamma\cos\theta_0\\ -\cos\gamma\sin\theta_0\end{pmatrix},\quad \{\hat{e}^i\}_{P_2}=\begin{pmatrix}-\sin\gamma\\ \cos\delta\cos\theta_0\\ -\cos\gamma\sin\theta_0\end{pmatrix}. \tag{13}$$

The matrices Δ_1 and Δ_2 are given by Eqs. (A4), (A7), and (A9) in the Appendix with $\beta=\pi/2$ and

$$\alpha=\theta=\theta_0\quad\text{for }\Delta_1,\qquad \alpha=\theta=-\theta_0\quad\text{for }\Delta_2; \tag{14}$$

thus,

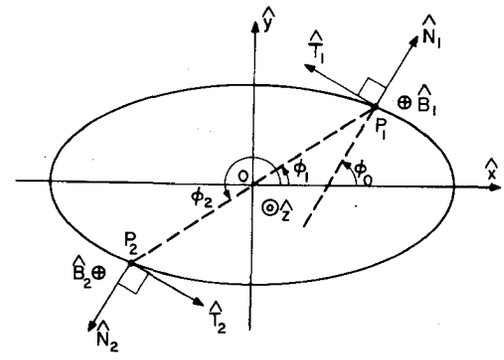


FIG. 2. First-order scattering centers.

$$\Delta_1 = \begin{pmatrix} -1-\frac{1}{\sin\theta_0} & 0 & 0 \\ 0 & \left(1-\frac{1}{\sin\theta_0}\right)\cos^2\theta_0 & \left(1-\frac{1}{\sin\theta_0}\right) \times \sin\theta_0\cos\theta_0 \\ 0 & \left(1-\frac{1}{\sin\theta_0}\right)\sin\theta_0\cos\theta_0 & \left(1-\frac{1}{\sin\theta_0}\right) \times \sin^2\theta_0 \end{pmatrix}, \tag{15}$$

$$\Delta_2 = \begin{pmatrix} -1+\frac{1}{\sin\theta_0} & 0 & 0 \\ 0 & \left(1+\frac{1}{\sin\theta_0}\right)\cos^2\theta_0 & -\left(1+\frac{1}{\sin\theta_0}\right) \times \sin\theta_0\cos\theta_0 \\ 0 & -\left(1+\frac{1}{\sin\theta_0}\right)\sin\theta_0\cos\theta_0 & \left(1+\frac{1}{\sin\theta_0}\right) \times \sin^2\theta_0 \end{pmatrix}. \tag{16}$$

It follows from Eqs. (5), (10)–(13), (15), and (16) that

$$S_{\substack{E\\H}} \sim -\frac{\delta}{4}\left(\frac{ka}{\pi\sin\theta_0}\right)^{1/2} A^{-3/2}\left[\left(1\mp\frac{1}{\sin\theta_0}\right)\exp\left(i2kaA\sin\theta_0 - i\frac{\pi}{4}\right) + \left(1\pm\frac{1}{\sin\theta_0}\right)\exp\left(-i2kaA\sin\theta_0 + i\frac{\pi}{4}\right)\right], \quad \theta_0 \neq 0. \tag{17}$$

In the particular case of the circular disk, $\delta = A = 1$ and Eq. (17) reduces to a previously known result.[2] Observe that (17) remains finite at edge-on incidence ($\theta_0 = \pi/2$). (An infinity for edge-on incidence occurs in the second term of the asymptotic expansion of S_H.)

For axial incidence ($\theta_0 = 0$), the far-backscattered field is easily obtained by physical optics

$$S_{\substack{E\\H}} \sim \pm\tfrac{1}{2} i\,\delta k^2 a^2, \quad \theta_0 = 0. \tag{18}$$

The first-order solution is given by Eqs. (5), (17), (18).

3. CAUSTIC MATCHING

Since the first-order field obtained in the previous section is nonuniform in θ_0, we now introduce interpolating functions which yield results (17) and (18) as particular cases. A correct interpolation could be obtained by asymptotically expanding the exact solution of the scattering problem, if such a solution were available. A discussion similar to that performed by Keller[12] for the circular edge would yield an axial correction factor containing a combination of Mathieu functions. For numerical purposes, it is preferable to introduce matching functions which are more easily calculated, such as Bessel functions. We do not pretend that the formula thus produced is the correct interpolation, but only that it is sufficiently accurate for practical purposes.

A comparison between Eqs. (17), (18), and the corresponding results for a circular disk, as well as an inspection of the interpolation formula used for the circular disk,[2,13,14] leads us to suggest the following expression:

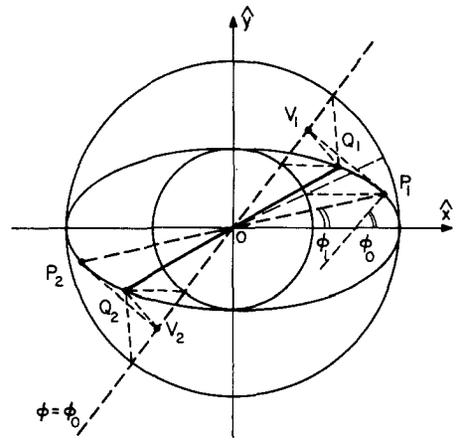


FIG. 3. Geometric meaning of effective diameter.

$$S_{\substack{E\\H}} \sim \frac{i\delta ka}{2A}\left(\pm\frac{J_1(2kaA\sin\theta_0)}{\sin\theta_0} - iJ_2(2kaA\sin\theta_0)\right). \tag{19}$$

The above equation reduces to the result for the circular disk when $\delta = 1$, to the physical optics result (18) when $\theta_0 = 0$ and to formula (17) plus terms $O[(ka)^{-1/2}]$ when the argument of the Bessel functions is large with respect to unity.

A comparison of (19) with the corresponding result for a circular disk shows that the elliptic plate behaves like a circular plate of effective diameter $2Aa$. A geometrical interpretation is given in Fig. 3: P_1 and P_2 are the first-order scattering centers, Q_1 and Q_2 are points of the edge with coordinates $x_{Q_1} = -x_{Q_2} = a\cos\phi_0$, $y_{Q_1} = -y_{Q_2} = b\sin\phi_0$, and V_1 and V_2 are the projections of P_1 and P_2 on the straight line $y = x\tan\phi_0$; then,

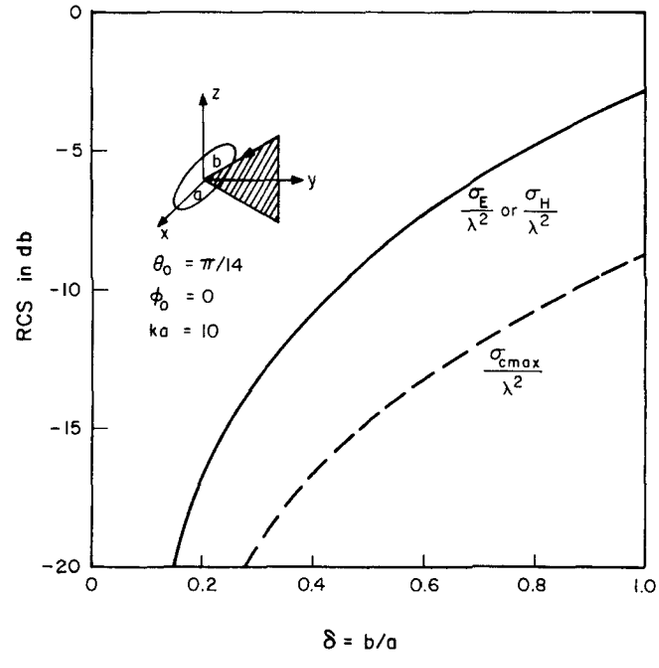


FIG. 4. Monostatic RCS of an elliptical plate as a function of axial ratio computed by the first-order theory of GDT.

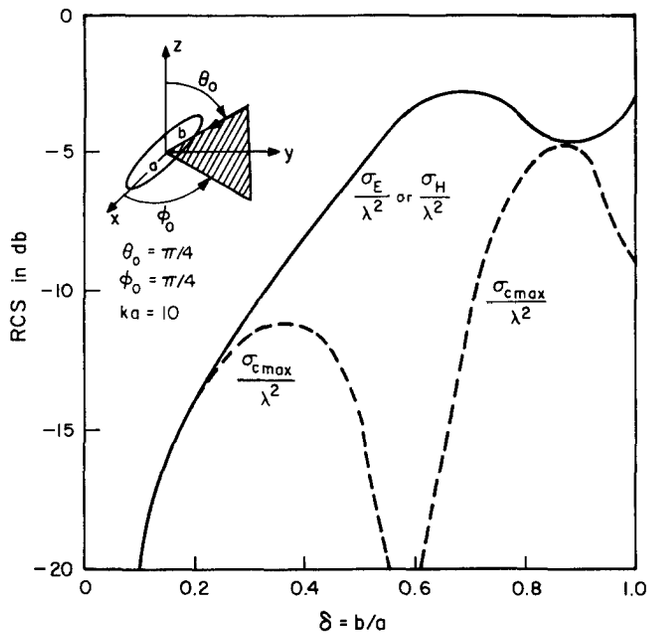


FIG. 5. Monostatic RCS of an elliptical plate as a function of axial ratio computed by the first-order theory of GDT.

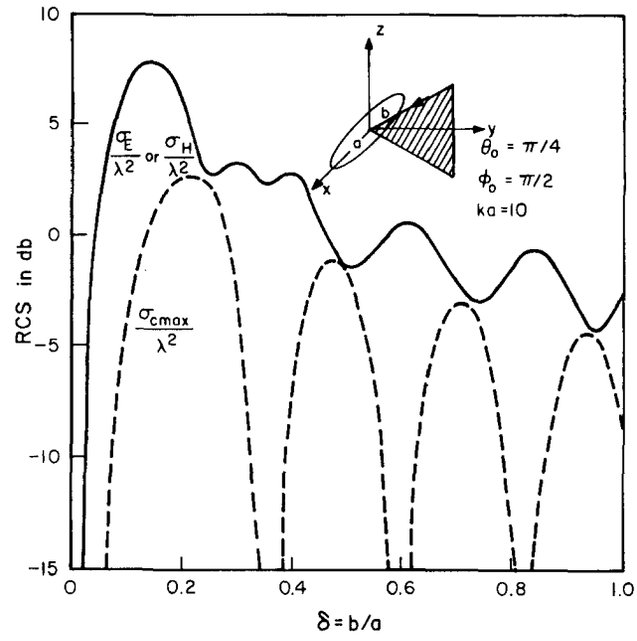


FIG. 6. Monostatic RCS of an elliptical plate as a function of axial ratio computed by the first-order theory of GDT.

$$2Aa = V_1V_2 = Q_1Q_2 = 20P_1\cos(\phi_0 - \phi_1). \tag{20}$$

The argument of the Bessel functions in (19) is therefore equal to the product of the wavenumber k times the projection of the distance P_1P_2 on the direction of incidence.

In general, the polarizations of the incident and backscattered fields differ. The cross-polarized return may be defined as

$$\mathbf{E}^{\text{b.s.}} \cdot \hat{i}X\,\hat{e}^i = (e^{ikr}/kr)\tfrac{1}{2}(S_E + S_H)\sin 2\gamma. \tag{21}$$

Therefore, we introduce the far-field coefficient S_{cross} of the cross-polarized component of the backscattered field:

$$S_{\text{cross}} = \tfrac{1}{2}(S_E + S_H)\sin 2\gamma. \tag{22}$$

From formula (19) it follows that

$$S_{\text{cross}} = (kb/2A)J_2(2kaA\sin\theta_0)\sin 2\gamma; \tag{23}$$

hence, $S_{\text{cross}} = 0$ if $\theta_0 = 0$, or $\gamma = 0$, or $\gamma = \pi/2$, whereas, for given θ_0 and ϕ_0, $|S_{\text{cross}}|$ is maximum when $\gamma = \pi/4$.

The monostatic cross sections for vertical and horizontal polarizations are

$$\sigma_{E,H} = (\lambda^2/\pi)|S_{E,H}|^2, \tag{24}$$

while the maximum cross-polarized cross section as a function of polarization angle is

$$\sigma_{\text{cross}}\big|_{\gamma=\pi/4} = (\lambda^2/\pi)\,|(kb/2A)J_2(2kaA\sin\theta_0)|^2. \tag{25}$$

4. NUMERICAL RESULTS

The final results for the monostatic radar cross sections (RCS) are given in (19), (24), and (25). We will now present some numerical computations based on these three formulas.

In Figs. 4, 5, and 6, the RCS for different polarizations is presented as a function of the axial ϕ ratio $\delta = b/a$. [Because of the restriction on (6), the numerical results in Figs. 5, 6, and 7 may not be accurate, say, for $\delta < 0.4$.] Since only the first-order field is included in the calculation, we have $\sigma_E = \sigma_H$ and, hence, there are only two distinct curves in each figure. For an incidence along the narrow direction of the elliptical plate ($\phi_0 = 0$), it may be shown that

$$\sigma(\text{elliptical}) = (b/a)^2\,\sigma(\text{circular with radius } a), \quad \text{for } \phi_0 = 0, \tag{26}$$

for corresponding polarizations. Thus, for a fixed ka, the RCS in Fig. 4 increases with the ratio of (b/a). For an incidence along the broadside direction ($\phi_0 = \pi/2$), we have the relation

$$\sigma(\text{elliptical}) = (a/b)^2\,\sigma(\text{circular with radius } b),$$

which, as expected, is symmetrical with respect to (26). The presence of oscillatory behavior for the RCS in Fig. 6 is due to the fact that Fig. 6 is computed for fixed ka, not for fixed kb. Thus, the oscillatory behavior in Fig. 6 may be explained in terms of the variation of the RCS with a radius for a circular plate.

The second group of computations presented in Figs. 7 and 8 is designed to study the azimuthal variation of RCS. The elliptical plate has the axial ratio 2:1. Generally $\sigma_E\,(=\sigma_H)$ increases as the incident angle moves toward the broadside direction of the plate. The maximum cross-polarization, cross-section $\sigma_{c\,\max}$ is oscillatory, and is identically zero when ϕ satisfies the following equation:

$$2ka\sin\theta_0[\cos^2\phi_0 + (b/a)^2\sin^2\phi_0]^{1/2} = j_{2\nu}, \quad \text{for } \nu = 1, 2, 3, \ldots$$

where $\{j_{2\nu}\}$ are the zeros of the Bessel function $J_2(x)$. The first few $\{j_{2\nu}\}$ have the following numerical values:

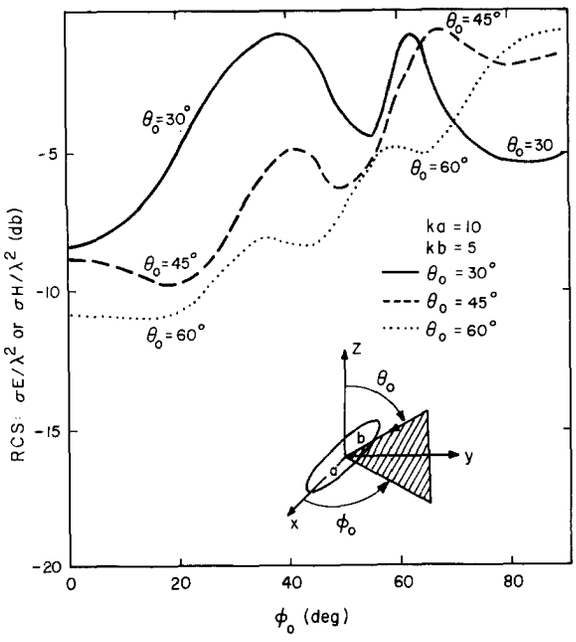


FIG. 7. Monostatic RCS of an elliptical plate as a function of the incident azimuthal angle computed by the first-order theory of GDT.

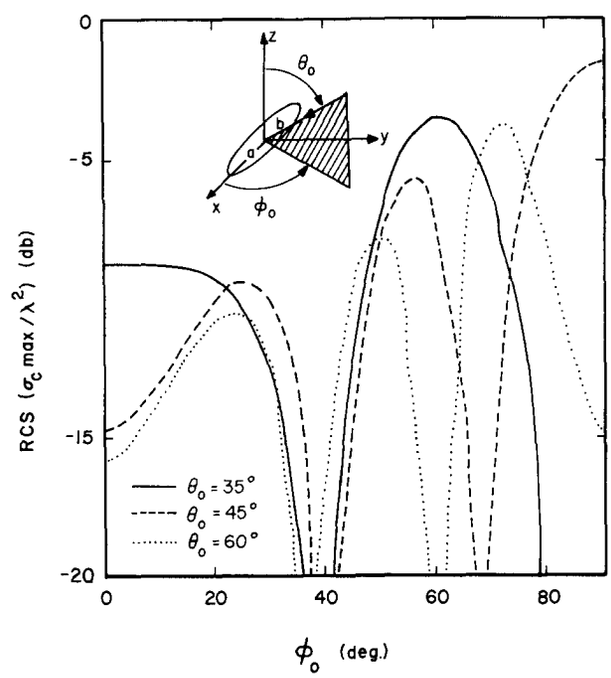


FIG. 8. Monostatic RCS of an elliptical plate as a function of axial ratio computed by the first-order theory of GDT.

$$j_{2\nu} = 5.14,\ 8.41,\ 11.62,\ 14.80,\ 17.96,\ \ldots$$

Between zeros, the peak values of $\sigma_{c\,max}$ increase again as the incident azimuthal angle ϕ_0 moves toward the broadside direction.

5. DISCUSSION OF HIGHER-ORDER CONTRIBUTIONS

The asymptotic expansions obtained in Sec. 3 can be extended beyond the dominant term only at the price of cumbersome calculations, and only for oblique incidence ($\theta_0 \neq 0$). Rather than deriving the complicated expressions which apply to the case of arbitrary incidence (θ_0, ϕ_0), we consider the simpler case of a primary wave propagating in a direction parallel to the (y, z) plane (i.e., $\phi_0 = \pi/2$). A similar derivation is at once applicable to incidence in a direction parallel to the (x, z) plane (i.e., $\phi_0 = 0$). The second-order term in the expansion of the backscattered field is due to optical rays which strike the edge of the plate, are diffracted along its surface, and, after striking the edge once more, are finally diffracted in the backscattering direction. It can be shown that two independent paths exist, each corresponding to two rays; these paths are shown in Fig. 9 for the particular case considered here. The first path of Fig. 9 lies in the plane $x = 0$: The incident ray is diffracted at P_3 (or P_4) toward P_4 (or P_3), whence it proceeds in the backscattering direction. The second path lies in a plane parallel to the x axis; the incident ray is diffracted at P_5 (or P_6) in the direction $-\hat{x}$ (or $+\hat{x}$), and it is then backscattered from P_6 (or P_5). The positions of P_5 and P_6 along the edge vary with the angle θ_0 of incidence. Thus, this second path corresponds to two "migrating" rays, which were first introduced by Knott *et al.*[2] for the circular disk and by Senior and Uslenghi[15] for the flat-based cone. For all four doubly diffracted rays, propagation occurs on both illuminated and shadowed faces of the plate. However, the backscattered field has the same form for rays propagating on either face[2] and, therefore, we consider rays on the illuminated face of the plate only.

It is easily seen that the rays doubly diffracted at P_3 and P_4 do not contribute to vertical polarization ($\gamma = \pi/2$),

$$S_E^{II}\big|_{3,4} = 0.$$

For horizontal polarization ($\gamma = 0$), the incident field

$$\mathbf{E}^i = -\hat{\theta}_0 \exp[-ik(y\sin\theta_0 + z\cos\theta_0)] \tag{27}$$

generates the second-order contribution

$$\mathbf{E}^{b.s.}\big|_{3,4}^{II} = \hat{\theta}_0 (e^{ikr}/kr) S_H^{II}\big|_{3,4}, \tag{28}$$

where

$$S_H^{II}\big|_{3,4} = kr\, e^{i2kb}(D_{i34}^H D_{34r}^H \Gamma_{i34}\Gamma_{34r} + D_{i43}^H D_{43r}^H \Gamma_{i43}\Gamma_{43r}). \tag{29}$$

D_{hjl}^H and Γ_{hjl} are, respectively, the Neumann diffraction coefficient and the divergence factor associated with the

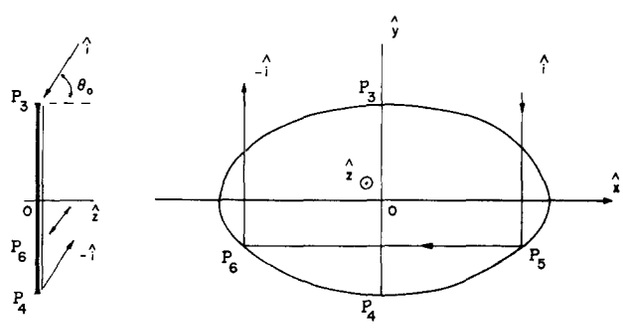


FIG. 9. Optical ray paths for second-order contributions, when $\phi_0 = \pi/2$.

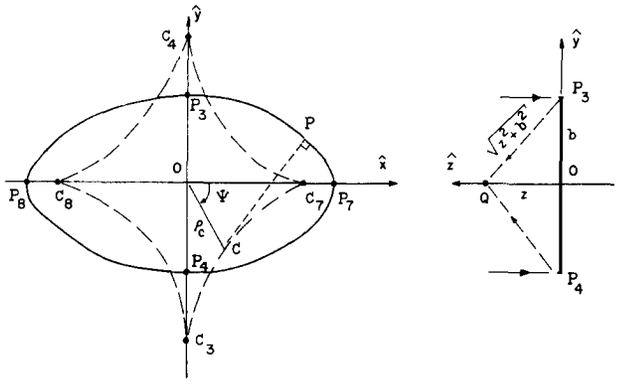


FIG. 10. Astigmatic cylinder for axial backscattering.

ray coming from P_h, hitting the edge at P_j, and diffracted from P_j toward P_1. The subscripts "i" and "r" represent the incident and backscattered directions. It can be shown that

$$D^H_{i34}D^H_{34r}=D^H_{i43}D^H_{43r}=i/(\pi k\cos\theta_0) \tag{30}$$

and that

$$\Gamma_{i34}\Gamma_{34r}=\Gamma_{i43}\Gamma_{43r}=-i/[2r\delta(1-\delta^2\cos^2\theta_0)^{1/2}]. \tag{31}$$

This last equation is valid if

$$r\gg[a/2\delta(1-\delta^2\cos^2\theta_0)]\,|2\delta^2(1+\sin\theta_0)-1|. \tag{32}$$

Substitution of (30) and (31) into (29) yields

$$S^{II}_H\big|_{3,4}=e^{i2kb}/[\pi\delta\cos\theta_0(1-\delta^2\cos^2\theta_0)^{1/2}]. \tag{33}$$

The coordinates of the migrating points P_5 and P_6 are

$$x_5=-x_6=a/(1+\delta^2\sin^2\theta_0)^{1/2},\quad y_5=y_6=-x_5\delta^2\sin\theta_0, \tag{34}$$

whereas the angle β between the direction of propagation of the incident ray and the tangent to the edge at the point of incidence P_5 or P_6 is given by

$$\sin\beta=(1+\sin^2\theta_0)^{-1/2}. \tag{35}$$

The second-order contribution to the backscattered field due to the two migrating rays is

$$\mathbf{E}^{b.s.}\big|^{II}_{5,6}=(i/8\pi k)(1+\sin^2\theta_0)e^{ikr}\exp[i2ka(1+\delta^2\sin^2\theta_0)^{1/2}]$$
$$\times(\Gamma_{i56}\Gamma_{56r}\Delta_{56r}\{\Delta_{i56}\hat{e}^i_5\}+\Gamma_{i65}\Gamma_{65r}\Delta_{65r}\{\Delta_{i65}\hat{e}^i_6\}); \tag{36}$$

here $\hat{e}^i_5$ and $\hat{e}^i_6$ are column vectors with components along the local base vectors $\hat{T}_5$, $\hat{N}_5$, $\hat{B}_5$ at P_5 and $\hat{T}_6$, $\hat{N}_6$, $\hat{B}_6$ at P_6 (see the Appendix), where

$$\hat{T}_{\substack5\\6}=\frac{\hat{x}\sin\theta_0\pm\hat{y}}{(1+\sin^2\theta_0)^{1/2}},\quad \hat{N}_{\substack5\\6}=\frac{\pm\hat{x}-\hat{y}\sin\theta_0}{(1+\sin^2\theta_0)^{1/2}},\quad \hat{B}_5=\hat{B}_6=-\hat{z}. \tag{37}$$

The matrix Δ_{hjl} is given by Eq. (A4), for a ray coming from P_h and edge-diffracted at P_j toward P_1. The column vectors $\{\Delta_{i56}\hat{e}^i_5\}$ and $\{\Delta_{i65}\hat{e}^i_6\}$ are expressed in terms of the local base vectors at P_6 and P_5, respectively. Thus we find that

$$\left.\begin{matrix}\Delta_{56r}\{\Delta_{i56}\hat{e}^i_5\}\\ \Delta_{65r}\{\Delta_{i65}\hat{e}^i_6\}\end{matrix}\right\}=\frac{-8\sin(\gamma\mp\theta_0)}{1+\sin^2\theta_0}$$
$$\times(-\hat{x}\cos\theta_0\mp\hat{y}\sin\theta_0\cos\theta_0\pm\hat{z}\sin^2\theta_0). \tag{38}$$

The determination of the divergence factors is accomplished with the aid of Frenét's formulas of differential geometry. The final result is

$$\Gamma_{i56}\Gamma_{56r}=\Gamma_{i56}\Gamma_{65r}$$
$$=-(\delta^2/2r)[(1+\delta^2\sin^2\theta_0)(1-\delta^2\cos^2\theta_0)]^{-1/2} \tag{39}$$

and is valid provided that

$$r\gg\frac{a\delta^2(2-\delta^2\cos2\theta_0)}{2(1-\delta^2\cos^2\theta_0)}(1+\delta^2\sin^2\theta_0)^{-3/2}. \tag{40}$$

Substitution of (38) and (39) into (36) yields

$$S^{II}_{\substack E\\H}\Big|_{5,6}=\frac{i\delta^2}{\pi}\begin{Bmatrix}\cos^2\theta_0\\ \sin^2\theta_0\end{Bmatrix}[(1+\delta^2\sin^2\theta_0)$$
$$\times(1-\delta^2\cos^2\theta_0)]^{-1/2}\exp[i2ka(1+\delta^2\sin^2\theta_0)^{1/2}]. \tag{41}$$

The total second-order contribution

$$S^{II}_{\substack E\\H}=S^{II}_{\substack E\\H}\Big|_{3,4}+S^{II}_{\substack E\\H}\Big|_{5,6} \tag{42}$$

is therefore given by Eqs. (26), (33), and (41); it is valid under the limitations (32) and (40), which are satisfied for all θ_0 if $\delta<1$. For the circular plate ($\delta=1$), both (32) and (40) fail at $\theta_0=0$, due to the fact that the z axis is now a caustic of the diffracted rays. However, this does not mean that the backscattering behavior of a circular plate at and near axial incidence is more difficult to obtain than the corresponding behavior of an elliptic plate. On the contrary, in the latter case an astigmatic cylinder of diffracted rays effectively precludes the determination of the second-order terms of the backscattered field at and near $\theta_0=0$. A detailed study of this phenomenon is given below.

Consider the case of axial backscattering. To the first order, the only contributions arise from the four rays diffracted at the edge points P_3, P_4, P_7, and P_8 of Fig. 10, provided that $a\neq b$. In the degenerate case of a circular plate, rays diffracted at all edge points must be taken into account because the z axis is now a caustic When $a\neq b$, the center of curvature C at the edge point P is no longer on the z axis; its position in polar co-

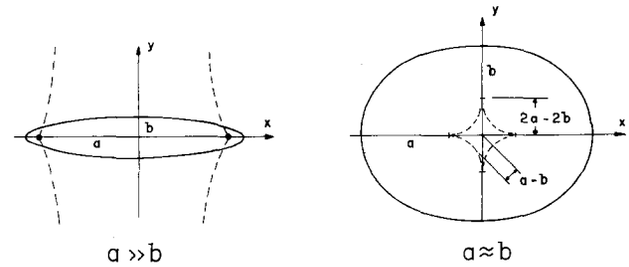


FIG. 11. Two limiting cases for the astigmatic cylinder.

ordinates (ρ_c, Ψ) is given by the equation

$$\rho_c(\Psi) = [(a^2 - b^2)/(a\cos\Psi)]\,[1 + (\delta\tan\Psi)^{2/3}]^{-3/2}, \tag{43}$$

from which it follows that the point C moves from $C_7(x_{c_7} = a - b^2/a)$ when $P \equiv P_7$ to $C_3(y_{c_3} = b - a^2/b)$ when $P \equiv P_3$. The other three portions of the curve $\rho_c(\Psi)$ are obtained from C_7C_3 by specular symmetries with respect to the x and y axes, and are indicated by broken lines in Fig. 10. Since

$$\left.\frac{d\rho_c}{d\Psi}\right|_{\Psi=0+} = -\infty, \quad \left.\frac{d\rho_c}{d\Psi}\right|_{\Psi=\pi/2-} = +\infty, \tag{44}$$

the points C_3, C_4, C_7, and C_8 are cusps. Thus, the curve $C_7C_3C_8C_4$ is the intersection with the $z=0$ plane of an astigmatic cylinder with four cusps, whose generators are parallel to the z axis. The minimum distance of the cylinder surface from the z axis occurs at the four points

$$\rho_c|_{\min} = a - b, \quad \Psi_{\min} = \pm\arctan\sqrt{\delta} + \begin{cases} 0 \\ \pi \end{cases}. \tag{45}$$

Two limiting cases are shown in Fig. 11. If $a \gg b$ (almost a strip), the cusps C_3 and C_4 are very far from the ellipse, while, if $a \approx b$ (almost a circle), all four cusps are at a distance $2(a-b)$ from the origin 0, and the astigmatic cylinder collapses on the z axis in the limit $a = b$.

An asymptotic expansion of the scattered field that is uniformly valid near a single cusped caustic has been obtained by Ludwig[8] for the scalar case; the derivations have subsequently been simplified by Kravtsov.[7] Ludwig's results cannot be easily extended to our case because of the vector nature of the field and because we are here considering the four cusps instead of one. However, some insights into the difficulties of the problem can be gained by examining the form that the scattered field would take at a point Q on the z axis (see Fig. 10) if no caustic correction is introduced. To this end, let us consider the incident field

$$\mathbf{E}^i = (\hat{x}\cos\xi + \hat{y}\sin\xi)\,e^{-ikz} \tag{46}$$

which produces the far-backscattered field

$$\mathbf{E}^{\text{b.s.}} = (\hat{x}E_x^{\text{b.s.}}\cos\xi + \hat{y}E_y^{\text{b.s.}}\sin\xi)\,e^{ikz}, \tag{47}$$

where, according to first-order GTD,

$$E_x^{\text{b.s.}} = (D^E_{i3Q}\,\Gamma_{i3Q} + D^E_{i4Q}\,\Gamma_{i4Q})\exp[ik(z^2+b^2)^{1/2}] - (D^H_{i7Q}\,\Gamma_{i7Q} + D^H_{i8Q}\,\Gamma_{i8Q})\exp[ik(z^2+a^2)^{1/2}], \tag{48}$$

$$E_y^{\text{b.s.}} = (D^E_{i7Q}\,\Gamma_{i7Q} + D^H_{i8Q}\,\Gamma_{i8Q})\exp[ik(z^2+a^2)^{1/2}] - (D^H_{i3Q}\,\Gamma_{i3Q} + D^H_{i4Q}\,\Gamma_{i4Q})\exp[ik(z^2+b^2)^{1/2}]. \tag{49}$$

The superscripts E and H respectively indicate the Dirichlet and Neumann diffraction coefficients of the scalar theory. Since Q is near the boundary of the reflected field, Eq. (A10) of the Appendix must be used:

$$D^E_{i7Q} = D^E_{i8Q} = -\frac{\exp[i(\pi/4)]}{2\sqrt{\pi k}}\left[\left(1 + \frac{z}{(z^2+a^2)^{1/2}}\right)^{-1/2} \pm i2k^{1/2}(z^2+a^2)^{1/4}F(k^{1/2}[(z^2+a^2)^{1/2} - z]^{1/2}) \times\exp\{ik[z - (z^2+a^2)^{1/2}]\}\right], \tag{50}$$

$$D^H_{i3Q} = D^H_{i4Q} = -\frac{\exp[i(\pi/4)]}{2\sqrt{\pi k}}\left[\left(1 + \frac{z}{(z^2+b^2)^{1/2}}\right)^{-1/2} \pm i2k^{1/2}(z^2+b^2)^{1/4}F(k^{1/2}[(z^2+b^2)^{1/2} - z]^{1/2}) \times\exp\{ik[z - (z^2+b^2)^{1/2}]\}\right]. \tag{51}$$

If Q is very far from the plate

$$F(k^{1/2}[(z^2+a^2)^{1/2} - z]^{1/2}) \sim \tfrac{1}{2}\sqrt{\pi}\,\exp[i(\pi/4)] - (ka^2/2z)^{1/2} + O(z^{-3/2}), \quad z \gg a, \tag{52}$$

with a similar approximation when a is replaced by b.

The divergence factors are evaluated by disregarding the proximity of Q to the caustic cylinder

$$\Gamma_{i7Q} = \Gamma_{i8Q} = -ib(a^2-b^2)^{-1/2}(z^2+a^2)^{-1/4}, \tag{53}$$

$$\Gamma_{i3Q} = \Gamma_{i4Q} = a(a^2-b^2)^{-1/2}(z^2+b^2)^{-1/4}. \tag{54}$$

The final result is

$$E_x^{\text{b.s.}} \sim (a - ib)\,(a^2-b^2)^{-1/2}\,e^{ikz} + [\pi kz(a^2-b^2)]^{-1/2}\{i2kab \mp \tfrac{1}{2}[(a-b) + i(a+b)]\}\,e^{ikz} + O(z^{-3/2}). \tag{55}$$

Formula (55) is obviously incorrect. The first term does not cancel the reflected field, and the second term does not agree with the physical optics result (18). In fact, it is not even proportional to z^{-1}. These errors are attributable to the divergence factors (53)—(54), and can be corrected only by an extension of Ludwig's theory that is not presently available. However, let us consider the case $a \gg b$ of Fig. 11, in which the caustic cylinder is as far as possible from the z axis. Formula (55) becomes

$$E_x^{\text{b.s.}} \sim e^{ikz} + (e^{ikz}/\sqrt{\pi kz})\,(i2kb \mp 2^{-1/2}\exp[i(\pi/4)]) + O(z^{-3/2}). \tag{56}$$

The first term of Eq. (56) cancels the reflected field. The second term resembles the dominant term in the expansion of the field backscattered by a strip. For broadside incidence, it exactly coincides with that term if the contributions from P_7 and P_8 are neglected.

In conclusion, it is presently possible to obtain second-order results in GTD only for oblique incidence; even in that case, the calculations are tolerable only for particular directions of incidence. The extension of GTD to terms higher than the second would be very complicated even if the caustic difficulties were overcome. Specifically, one should then consider the following additional contributions: (i) The variations in the radius of curvature along the edge would require higher-order corrections for the divergence factors; (ii) the diffraction coefficients considered in this paper are only the leading terms of asymptotic expansions whose next term, needed to calculate third-order GTD contributions, can be found by a lengthy analysis involving boundary-layer techniques; (iii) the number of optical ray paths increases and varies with the direction (θ_0, ϕ_0) of incidence.

6. CONCLUSION

Even though the geometrical theory of diffraction was developed about fifteen years ago, its application to truly three-dimensional problems was carried out only recently. The elliptical plate considered in this paper is

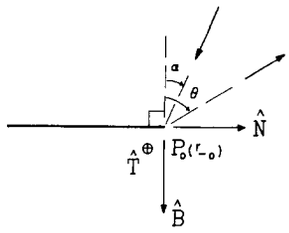


FIG. 12. Geometry for diffraction matrix.

one of the simplest edge structures with variable radius of curvature; thus, it provides a good testing ground for the geometrical theory of diffraction. Following Keller's theory, the dominant term with a caustic correction of the backscattering for an incident plane wave with arbitrary polarization is readily obtained. The final expression is given in (19) in terms of Bessel functions. To improve its accuracy, one may attempt to include the next order or even higher-order ray contributions. However, we have illustrated that such an attempt usually leads to not only cumbersome results but also unsolved problems associated with caustic corrections.

APPENDIX: THE DIFFRACTION MATRIX

Consider the perfectly conducting half-plane shown in Fig. 12. The incident electric field is

$$\mathbf{E}^i = \hat{e}^i \exp(ik\hat{i}\cdot\mathbf{r}), \tag{A1}$$

where $\hat{i}$ forms the angle β with the edge $0<\beta\leq\pi/2$. The diffracted portion of the total field is

$$\mathbf{E}^d = \hat{e}^d \frac{\exp[ik\hat{i}\cdot\mathbf{r}_0 + iks + i(\pi/4)]}{2\sqrt{2\pi ks}\,\sin\beta}, \tag{A2}$$

where s is the distance from the point $P_0(\mathbf{r}_0)$ of the edge to the observation point P and the diffracted ray P_0P forms the angle β with the edge. The unit vectors $\hat{e}^i$ and $\hat{e}^d$ give the polarizations of the incident and diffracted fields, and are related by

$$\hat{e}^d = \Delta\hat{e}^i, \tag{A3}$$

where

$$\Delta = \begin{pmatrix} -(B+C) & 0 & 0 \\ (B+C)\cot\beta\,\sin\theta & (B-C)\cos\theta\cos\alpha & (B-C)\cos\theta\sin\alpha \\ -(B+C)\cot\beta\,\cos\theta & (B-C)\sin\theta\cos\alpha & (B-C)\sin\theta\sin\alpha \end{pmatrix}. \tag{A4}$$

Both $\hat{e}^i$ and $\hat{e}^d$ are column vectors with components along the base vectors $\hat{T}$, $\hat{N}$, $\hat{B}$; $\hat{N}$ is normal to the edge and points from metal into free space, $\hat{B}$ is normal to the half-plane and points into the shadowed half-space, and $\hat{T}=\hat{N}\times\hat{B}$. The angle β is the angle that $\hat{T}$ forms with the direction $\hat{i}$. The incident and diffracted rays are on opposite sides of the plane normal to $\hat{T}$ and containing P_0; their projections on this plane form angles α and θ with $-\hat{B}$ (see Fig. 12), where

$$-\pi/2\leq\alpha\leq\pi/2, \quad -\pi/2<\theta<3\pi/2. \tag{A5}$$

If the observation point P is not near the boundary $\theta=\alpha+\pi$ of the shadow region, i.e., if

$$|\cos\tfrac{1}{2}(\theta-\alpha)|\sqrt{2ks\sin\beta}\gg 1, \tag{A6}$$

then

$$B\sim 1/\cos\tfrac{1}{2}(\theta-\alpha), \tag{A7}$$

whereas if P is not near the boundary $\theta=-\alpha$ of the reflected field, i.e., if

$$|\sin\tfrac{1}{2}(\theta+\alpha)|\sqrt{2ks\sin\beta}\gg 1, \tag{A8}$$

then

$$C\sim 1/\sin\tfrac{1}{2}(\theta+\alpha). \tag{A9}$$

The result embodied in (A4), (A7), and (A9) is a particular case of the wedge diffraction matrix of Senior and Uslenghi.[15] An equivalent matrix, expressed in terms of ray coordinates rather than edge coordinates, has been subsequently derived by Kauyaumjian and Pathak.[16] It represents an extension to the vector case of the scalar results of Wolfe,[17] Ahluwalia *et al.*,[18] Lewis and Boersma,[9] and Ahluwalia,[10] and has the advantages of being simpler than (A4) and of remaining valid at the optical boundaries.

In this paper, consistent use has been made of (A7) and (A9) except for the backscattered field at axial incidence discussed in Sec. 5. In that case inequality (A8) is not satisfied, but a simple inspection of Sommerfeld's half-plane solution[19,20] shows that, for $\beta=\pi/2$,

$$C = \pm i\,2\sqrt{2ks}\,F[\mp\sqrt{2ks}\,\sin\tfrac{1}{2}(\theta+\alpha)] \times \exp\{-iks[1-\cos(\theta+\alpha)]\} \tag{A10}$$

at and near the boundary of the reflected field, where

$$F(u) = \int_u^\infty e^{iv^2}\,dv \tag{A11}$$

is a Fresnel integral and the upper (lower) signs apply when $\theta+\alpha<0$ ($\theta+\alpha>0$). Similarly, if (A6) is not satisfied and $\beta=\pi/2$,

$$B = \pm i2\sqrt{2ks}\,F[\mp\sqrt{2ks}\cos\tfrac{1}{2}(\theta-\alpha)] \times \exp\{-iks[1+\cos(\theta-\alpha)]\}, \tag{A12}$$

where the upper (lower) signs apply when $\theta-\alpha>\pi$ ($\theta-\alpha<\pi$).

*This work was supported by U.S. Army Missile Command under Contract DAAH01-72-C-0329 and National Science Foundation Grant GK-37839.

[1]R. A. Ross, IEEE Trans. Antennas Propag. **14,** 329 (1966).
[2]E. F. Knott, T. B. A. Senior, and P. L. E. Uslenghi, **118,** 736 (1971).
[3]T. B. A. Senior and P. L. E. Uslenghi, IEEE Trans. Antennas Propag. **19,** 557 (1971).
[4]R. N. Buchal and J. B. Keller, Commun. Pure Appl. Math. **13,** 85 (1960).
[5]Yu. A. Kravtsov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **7,** 664 (1964).
[6]Yu. A. Kravtsov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **7,** 1049 (1964).
[7]Yu. A. Kravtsov, Akust. Zh. **14,** 1 (1968) [Sov. Phys.-Acoust. **14,** 1 (1968)].
[8]D. Ludwig, Commun. Pure Appl. Math. **19,** 215 (1966).
[9]R.M. Lewis and J. Boersma, J. Math. + Phys. **10,** 2291 (1969).
[10]D. S. Ahluwalia, SIAM J. Appl. Math. (Soc. Ind. Appl. Math.) **18,** 287 (1970).
[11]F. K. Oshiro, K. M. Mitzner, and S. S. Locus, "Calculation of Radar Cross-Section," Technical Report AFAL-TR-71-50, Northrop Corporation, Aircraft Division, Hawthorne, California, 1971.
[12]J. B. Keller, J. Opt. Soc. Am. **52,** 116 (1962).
[13]P. Ia. Ufimtsev, Zh. Tekh. Fiz. **28,** 2604 (1958) [Sov. Phys.-Tech. Phys. **3,** 2386 (1958)].
[14]P. Ia. Ufimtsev, Zh. Tekh. Fiz. **28,** 583 (1958) [Sov. Phys.-Tech. Phys. **3,** 549 (1958)].
[15]T. B. A. Senior and P. L. E. Uslenghi, Radio Sci. **6,** 393 (1971).
[16]R. G. Kauyaumjian and P. H. Pathak, "The Dyadic Diffraction Coefficient for a Perfectly Conducting Edge Structure," presented at URSI Fall Meeting, Los Angeles, California, 1971.
[17]P. Wolfe, thesis (New York University, 1965).
[18]D. S. Ahluwalia, R. M. Lewis, and J. Boersma, SIAM J. Appl. Math. (Soc. Ind. Appl. Math.) **16,** 783 (1968).
[19]J. J. Bowman, T. B. A. Senior, and P. L. E. Uslenghi, Eds., *Electromagnetic and Acoustic Scattering by Simple Shapes* (North-Holland, Amsterdam, 1969).
[20]R. Mittra, and S. W. Lee, *Analytical Techniques in the Theory of Guided Waves* (Macmillan, New York, 1971).

Quantum conservation laws for charged particle systems in magnetic fields

Zygmunt K. Petru

Institute of Theoretical Physics, University of Wrocław, Wrocław, Poland
(Received 21 July 1972)

The conservation laws in the exact operator form are derived for the two-component system of interacting charged particles in magnetic and electric fields. These laws represent the equations of motion for the mass-, current-, and energy-density operators, where both the pressure tensor and energy-current operators are calculated exactly.

1. INTRODUCTION

The purpose of the present paper is to obtain equations of motion for the density operators for the interacting charged Fermi (or Bose) particle system in magnetic and electric fields. These equations are the differential conservation laws for the charge-, current-, and energy-density operators, where the stress tensor and energy-current operators are calculated exactly, it means, without any approximations concerned with a range of the interparticle forces.

The quantum hydrodynamic (QHD) equations were first introduced to one-particle quantum mechanics by Madelung.[1] In the many-body case, under the influence of Landau's phenomenological theory of helium, the quantum-statistical hydrodynamic approach was successfully used to calculate Green's functions (or correlation functions) for the normal and superfluid bosons by Bogolubov,[2] Kadanoff and Martin,[3] Hohenberg and Martin,[4] and for the fermions by Galasiewicz.[5] The derivations of QHD equations in operator form from the many-body Hamiltonian were given by Puff and Gillis,[6] Kugler,[7] and recently by Robertson.[8] In the last five years the QHD method has been studied extensively, especially by the Fröhlich School[9] (structure of the equations of motion and its application to superfluids), and by the group of authors: Bierter, Garrison, Morrison, and Wong[10] (mathematical structure of the algebra of density operators and its physical consequences).

However the QHD description of a many-charged-particles system in a magnetic field has not been considered, although, e.g., an electron gas and an ion lattice seem to be an interesting subject for application of such a theory. This is done in the present paper. Starting from the Hamiltonian for the interacting charged particles in the magnetic (and electric) field the exact equations and conservations laws are derived in Sec. 2, and in the Appendix.

2. CONSERVATION LAWS IN OPERATOR FORM

Switching on the magnetic field generated by vector potential $\mathbf{A}(\mathbf{r},t)$ leads to the change (we put $\hbar=1=c$)

$$\partial_\alpha \to \partial_\alpha - ieA_\alpha$$

so we introduce the operation denoted by semicolon as follows:

$$\psi_{;\alpha} \equiv (\psi_{,\alpha} - ieA_\alpha\psi) \equiv (\partial_\alpha - ieA_\alpha)\psi, \tag{1}$$

for destruction operators, and

$$\psi^+_{;\alpha} \equiv \psi^+_{,\alpha} + ieA_\alpha\psi^+, \tag{2}$$

for creation operators of the fermion (or boson) field.

In this notation the Hamiltonian of the two-component system to be considered here is

$$H = \sum_{i=1,2.} H^i,$$

$$H^i = (2m^i)^{-1}\int d\mathbf{r}\,\hat\psi^i_{;\alpha}(\mathbf{r}t)\psi^i_{;\alpha}(\mathbf{r}t) + \int d\mathbf{r}\,e^i n^i(\mathbf{r}t)\,U(\mathbf{r}t)$$
$$+\tfrac{1}{2}\sum_{k=1,2}\int d\mathbf{r}\,d\mathbf{r}'V^{ik}(|\mathbf{r}-\mathbf{r}'|)\hat\psi^i(\mathbf{r}t)\,n^k(\mathbf{r}'t)\,\psi(\mathbf{r}t), \tag{3}$$

where we assume a summation over repeated indexes α. $U(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$ are the time-and space-dependent external scalar and vector potentials. The charge-density operators satisfy the neutrality condition

$$\sum_{i=1,2}\int d\mathbf{r}\,e^i n^i(\mathbf{r},t) = 0,$$

which insures the stability of the system. [In (3) we omit the spin.]

The charge- (or mass-), current-, and energy-density operators are defined in the symmetric form

$$\rho^e \equiv \sum_i^2 e^i n^i, \quad \rho^m \equiv \sum_i^2 m^i n^i,$$
$$n^i(\mathbf{r}t) \equiv \hat\psi^i(\mathbf{r}t)\,\psi^i(\mathbf{r}t), \tag{4}$$

$$j^e_\alpha \equiv \sum_i j^i_\alpha e^i, \quad j^m_\alpha \equiv \sum_i j^i_\alpha,$$
$$j^k_\alpha(\mathbf{r}t) \equiv \tfrac{1}{2}i(\hat\psi^k_{;\alpha}\psi^k - \hat\psi^k\psi^k_{;\alpha})(\mathbf{r}t), \tag{5}$$

and

$$(\rho\epsilon)(\mathbf{r}t) \equiv -\sum_i (4m^i)^{-1}(\hat\psi^i_{;\alpha\alpha}\psi^i + \hat\psi^i\psi^i_{;\alpha\alpha})(\mathbf{r}t)$$
$$+\sum_i e^i n^i(\mathbf{r}t)\,U(\mathbf{r}t) \tag{6}$$
$$+\tfrac{1}{2}\sum_{i,k}\hat\psi^i(\mathbf{r}t)\int V^{ik}|\mathbf{r}-\mathbf{r}'|\,n^k(\mathbf{r}'t)\,d\mathbf{r}'\psi^i(\mathbf{r}t),$$

respectively.

From the commutation relations with the Hamiltonian (3) it follows that

$$i\partial_t\psi^k(\mathbf{r}t) = [\psi^k, H] = -(2m^k)^{-1}\psi^k_{;\alpha\alpha}(\mathbf{r}t) + e^k\psi^k(\mathbf{r}t)\,U(\mathbf{r}t)$$
$$+\sum_i\int V^{ki}|\mathbf{r}-\mathbf{r}'|\,n^i(\mathbf{r}'t)\,d\mathbf{r}'\,\psi^k(\mathbf{r}t). \tag{7}$$

The equations of motion for the density operators (4)–(6) can be calculated with the aid of (7) and have the following form (see Appendix):

$$\partial_t\rho^m = -j^m_{\alpha,\alpha}, \quad \partial_t\rho^e = -\sum_i (e^i/m^i)\,j^i_{\alpha,\alpha}, \tag{8}$$

$$\partial_t j^e_\alpha = -T_{\alpha\beta,\beta} + \sum_i [(e^i)^2/m^i][\mathbf{j}^i\times\mathrm{rot}\mathbf{A}]_\alpha$$
$$-\sum_i (e^i)^2 n^i(U_{,\alpha} + \partial_t A_\alpha), \tag{9}$$

$$\partial_t(\rho\epsilon) = -I_{\alpha,\alpha} - \sum_i (e^i/m^i) j^i_\alpha (\partial_\alpha U + \partial_t A_\alpha) + \partial_t \sum_i e^i n^i U. \tag{10}$$

The stress tensor operator, $T_{\alpha\beta}$, is symmetric and gauge-invariant, and is given by

$$T_{\alpha\beta} \equiv T'_{\alpha\beta} + T''_{\alpha\beta},$$

$$T'_{\alpha\beta} = \sum_{i=1}^{2} \frac{e^i}{2m^i} (\psi^i_{;\alpha}\psi^i_{;\beta} + \psi^i_{;\beta}\psi^i_{;\alpha}) - \sum_i \frac{e^i}{4m^i} n^i_{,\alpha\beta}, \tag{11a}$$

$$T''_{\alpha\beta} = -\frac{1}{4}\int d\mathbf{R}\,\frac{R_\alpha R_\beta}{R}\int_{-1}^{+1} d\lambda \sum_{ik} \frac{\partial V^{ik}(R)}{\partial R}\,\psi^i(x_+) e^k n^k(x_-)\,\psi^i(x_+), \tag{11b}$$

where

$$x_\pm = [\mathbf{r} + \tfrac{1}{2}(\lambda \pm 1)\mathbf{R}, t].$$

The energy-current operator is given by

$$I_\beta \equiv I'_\beta + I''_\beta,$$

$$I'_\beta = -(i/4)\sum_k (1/m_k^2)(\psi^k_{;\beta}\psi^k_{;\alpha\alpha} - \psi^k_{;\alpha\alpha}\psi^k_{;\beta}) - \sum_k (4m_k^2)^{-1} j^k_{\alpha,\alpha\beta} + \tfrac{1}{2}\sum_{i,k} (m_k)^{-1}\int V^{ik}(\mathbf{r}-\mathbf{r}')\,\psi^i(\mathbf{r}'t)\, j^k_\beta(\mathbf{r}t)\,\psi^i(\mathbf{r}'t)\,d\mathbf{r}', \tag{12a}$$

$$I''_\beta = -\frac{1}{4}\int d\mathbf{R}\,\frac{R_\alpha R_\beta}{R}\int_{-1}^{+1} d\lambda \sum_{ik}\frac{1}{m_k}\frac{\partial V^{ik}(R)}{\partial R}\,\psi^i(x_+) j^k_\alpha(x_-)\,\psi^i(x_+). \tag{12b}$$

Note, that the "semicolon notation" (1)—(2) underlines the external similarity of Eqs. (8)—(12) to those found for the case $A \equiv 0$ (see, e.g., Ref. 6). The Lorentz term $(e/m)[\mathbf{j}\times\mathcal{H}] + \rho\mathbf{E}$ has a form identical with the classical one, but in (9) ρ and $\mathbf{j}$ are operators.

3. DISCUSSION

We want to emphasize that the operators (11) and (12) are calculated exactly (see Appendix), and then the formulas (8)—(10) are valid for the various types of two-particle interactions. Since these conservation laws are written in operator form and, consequently, the method of averaging can be arbitrary, they can be applied also to systems far away from equilibrium.

The exact methods of closure of the set of QHD equations were given by Zubarev[11] and Robertson[8] for neutral particles. These methods can be adapted to the case of charged particles in a magnetic field and then apply to an electron gas and/or an ion lattice. This is the object of our current investigation. The approximate approach has been made[12] with the aid of the Bogolubov method[6,9] (hydrodynamic and acoustic approximation) and the Gorkov quasiclassical ansatz.[13]

ACKNOWLEDGMENTS

The author is grateful to Professor Z. Galasiewicz for suggesting the direction of research and to Dr. J. Czerwonko for constructive criticism.

APPENDIX

In this appendix the formula (11) for $T''_{\alpha\beta}$ is derived. Putting (7) into $\partial_t j_\alpha$, where j_α is given by (5), we obtain by simple but tedious calculations the right-hand side of (9) with the following expression for $T''_{\alpha\beta}$:

$$\sum_\alpha T''_{\alpha\beta,\beta} = -\int d\mathbf{r} \sum_{i,k}^{1,2} \frac{\partial V^{ik}(|\mathbf{r}-\mathbf{r}'|)}{\partial r_\alpha}\,\psi^i(\mathbf{r}t)\, n^k(\mathbf{r}'t)\,\psi^i(\mathbf{r}t)\, e^k. \tag{A1}$$

It is convenient to pass to a new integration variable $\mathbf{R} = \mathbf{r} - \mathbf{r}'$:

$$\sum_\beta T''_{\alpha\beta,\beta} = \frac{1}{2}\int d\mathbf{R}\,\frac{R_\alpha}{R}\sum_{i,k}\frac{\partial V^{ik}(R)}{\partial R}\,\psi^i(\mathbf{r}t)[n^k(\mathbf{r}+\mathbf{R},t) - n^k(\mathbf{r}-\mathbf{R},t)]\,\psi^i(\mathbf{r}t)\,e^k. \tag{A2}$$

We use, in analogy to Ref. 6, the relations

$$i[\psi^k(\mathbf{r},0), \mathbf{P}(t)] = [\nabla_r - (ie^k/c)\mathbf{A}(\mathbf{r}t)]\,\psi^k(\mathbf{r},0),$$

$$\exp[-i\mathbf{R}\mathbf{P}(t)]\, n^k(\mathbf{r},0)\exp[i\mathbf{R}\mathbf{P}(t)] = n^k(\mathbf{r}+\mathbf{R},0), \tag{A3}$$

where $\mathbf{P}$ is the total momentum:

$$P_\alpha(t) \equiv \int d\mathbf{r}\sum_k \{(i/2)[\psi^k_{,\alpha}(\mathbf{r},0)\,\psi^k(\mathbf{r},0) - \psi^k(\mathbf{r},0)\,\psi^k_{,\alpha}(\mathbf{r},0)] - c^{-1}e^k n^k(\mathbf{r},0)A_\alpha(\mathbf{r},t)\}.$$

With the aid of (A3) we have

$$iU^+(t)[\psi^k(\mathbf{r})\, n^l(\mathbf{r}-\mathbf{R})\,\psi^k(\mathbf{r}),\ P_\alpha(t)]U(t) = \frac{\partial}{\partial r_\alpha}[\psi^k(\mathbf{r}t)\, n^l(\mathbf{r}-\mathbf{R},t)\,\psi^k(\mathbf{r}t)]. \tag{A4}$$

[$U(t)$ is the time evolution operator.]

Consider now the integral

$$J(t\mathbf{r}) \equiv U^+(t)\int_{-1}^{+1} d\lambda \frac{d}{d\lambda}\exp[-i(1+\lambda)\mathbf{R}\mathbf{P}/2]\,\psi^i(\mathbf{r})\, n^k(\mathbf{r}-\mathbf{R})\,\psi^i(\mathbf{r}) \times\exp[i(1+\lambda)\mathbf{R}\mathbf{P}/2]\,U(t)$$

$$= -\{\psi^i(\mathbf{r}t)\, n^k(\mathbf{r}-\mathbf{R},t)\,\psi^i(\mathbf{r}t) - \psi^i(\mathbf{r}+\mathbf{R},t)\, n^k(\mathbf{r}t)\,\psi^i(\mathbf{r}+\mathbf{R},t)\}. \tag{A5}$$

On the other hand, by differentiation of the integrand, this integral takes the form

$$J(t\mathbf{r}) = (i/2)\mathbf{R}\int_{-1}^{1} d\lambda\, U^+(t)\exp[-i(\lambda+1)\mathbf{R}\mathbf{P}/2] \times[\psi^i(\mathbf{r})\, n^k(\mathbf{r}-\mathbf{R})\,\psi^i(\mathbf{r}), \mathbf{P}]\exp[i(\lambda+1)\mathbf{R}\mathbf{P}/2]\,U(t)$$

$$= \nabla_r \mathbf{R}\tfrac{1}{2}\int_{-1}^{1} d\lambda\,\psi^i[\mathbf{r}+\tfrac{1}{2}(\lambda+1)\mathbf{R},t]\, n^k[\mathbf{r}+\tfrac{1}{2}(\lambda-1)\mathbf{R},t] \times\psi^i[\mathbf{r}+\tfrac{1}{2}(\lambda+1)\mathbf{R},t]. \tag{A6}$$

By substituting (A5) and (A6) into (A2) we find the final result given by Eq. (11b).

In an analogous way one can prove the validity of Eq. (12b) for I''_β.

[1] E. Madelung, Z. Physik, 40, 322 (1927).
[2] N. N. Bogolubov, Dubna preprint No. R-1395 (1963), see also N. N. Bogolubov, *Collected Papers* (Naukova Dumka, Kiev, 1971, in Russian), Vol. 3.
[3] L. P. Kadanoff and P. C. Martin, Ann. Phys. (N. Y.) 24, 419 (1963).

[4]P. C. Hohenberg and P. C. Martin, Ann. Phys. (N. Y.) **34**, 291 (1965).
[5]Z. M. Galasiewicz, Nucl. Phys. **44**, 107 (1966), see also Z. M. Galasiewicz, *Superconductivity and Quantum Fluids* (Pergamon, Oxford, 1970).
[6]R. D. Puff and N. S. Gillis, Ann. Phys. (N. Y.) **46**, 364 (1968).
[7]A. Kugler, Z. Phys. **198**, 236 (1967).
[8]B. Robertson, J. Math. Phys. **11**, 2482 (1970).
[9]H. Frohlich, Physica **37**, 215 (1967); J. Phys. Soc. Japan **26**, Suppl. 189 (1969); Phys. kondens. Materie **9**, 350 (1969); G. J. Hyland and G. Rowlands, Physica **54**, 542 (1971), Phys. Lett. **33A**, 197 (1970); **36A**, 233 (1971); J. Low-Temp. Phys. (to be published). H. Haug and K. Weiss, Phys. Rev. A **3**, 717 (1971); Nuovo. Cimento Lett. **2**, 887 (1971); Physica **59**, 29 (1972).
[10]W. Bierter and H. L. Morrison, J. Low-Temp. Phys. **1**, 65 (1969); Nuove Cimento Lett. **1**, 701 (1969); J. C. Garrison and H. L. Morrison, and J. Wong, J. Math. Phys. **11**, 630 (1970); W. Bierter, Preprint of Max Planck Institute (1971), submitted to J. Low Temp. Phys.
[11]D. N. Zubarev, Dokl. Akad. Nauk. SSSR **161**, 537 (1965) [Sov. Phys. Dokl. **10**, 850 (1966)]; see also D. N. Zubarev, *Non-Equilibrium Statistical Thermodynamics* (Nauka, Moscow, 1971, in Russian).
[12]Z. K. Petru, Bull. Acad. Polon. Sci., Serie Sci. Math. Astr. Phys. **19**, 159 (1971).
[13]L. P. Gorkov, Zh. Eksp. Teor. Fiz. **34**, 735 (1958) (in Russian).

Momentum distribution in the Tomonaga model at finite temperatures

M. Aizenman* and H. Gutfreund

The Racah Institute of Physics, The Hebrew University of Jerusalem, Jerusalem, Israel
(Received 16 August 1972)

The momentum distribution function in the one-dimensional Tomonaga model is evaluated and its properties are discussed. The effect of finite temperatures on the applicability of the model is investigated.

I. INTRODUCTION

Tomonaga has shown[1] that a one-dimensional system of electrons can be described in terms of collective degrees of freedom which approximately satisfy boson commutation relations. This model was used to investigate the possibility of flux quantization in a one-dimensional ring.[2] Recently it has also been applied to the Kondo problem[3] and to the threshold behavior of x-ray absorption.[4]

The operators of the Tomonaga model which create and destroy eigenstates of the total Hamiltonian obey boson commutation relations only when evaluated within a subspace of states, all of which have a filled core in k space, i.e., have no holes in a certain interval $[-k^*, k^*]$, $k^* < k_F$, where k_F is the Fermi momentum. The consistency of the model depends on whether the calculated eigenstates share this property. Tomonaga argued that for sufficiently weak and long-range interactions and for sufficiently low temperatures all the relevant states belong to the subspace described above. This assumption was checked in Ref. 5 (hereafter referred to as I) by calculating the momentum distribution in the interacting ground state at $T=0$ and showing explicitly to what extent are the single particle states within the core $[-k^*, k^*]$ actually filled. The analysis in I leads to quantitative conditions on the strength and range of the interaction necessary for the applicability of the Tomonaga model. The purpose of this paper is to perform a similar analysis for finite temperatures.

The Tomonaga model is reviewed briefly in Sec. 2. In Sec. 3, the momentum distribution function at finite temperatures is calculated. The properties of the momentum distribution and the effect of temperature on the applicability conditions are discussed in Sec. 4.

In view of the equivalence between the Tomonaga and the Luttinger[6] models exposed in I, the present derivation of the momentum distribution applies with minor changes also to the Luttinger model. However, there is no problem of consistency in the latter case, because in the Luttinger model the boson commutation relations are exact.

II. THE TOMONAGA MODEL

Consider a system of fermions on a line of length L. The second quantized particle field is written in terms of plane waves as

$$\psi(x) = L^{-1/2} \sum a_n \exp(ik_n x), \qquad k_n = (2\pi/L)n, \quad n = 0, 1, \cdots, \tag{2.1}$$

where a_n, a_n^+ are the fermion destruction and creation operators. We omit spin indices as they play no role in the present discussion. The particle density may be expanded as

$$\rho(x) = \psi^*(x)\psi(x) = L^{-1} \sum_n \rho_n \exp(ik_n x), \tag{2.2}$$

where

$$\rho_n = \sum_l a_l^+ a_{l+n}.$$

Tomonaga decomposes the operators ρ_n into two parts

$$\rho_n = \rho_n^+ + \rho_n^-,$$
$$\rho_n^+ = \sum\nolimits_{l \geq -n/2} a_l^+ a_{l+n}, \tag{2.3}$$
$$\rho_n^- = \sum\nolimits_{l < -n/2} a_l^+ a_{l+n}.$$

The operators $\rho_n^\pm$ satisfy complicated commutation relations, which simplify considerably if the commutators are evaluated acting on a subspace S of functions in the neighborhood of the noninteracting ground state. These functions are specified by the absence of holes in a certain interval $[-n^*, n^*]$, $n^* < n_F$, where $n_F = (L/2\pi)k_F$. Within this subspace the operators $\rho_n^\pm$ satisfy, for $|n|, |n'| < (2/3)n^*$, the following bosonlike commutation relations

$$[\rho_n^+, \rho_{n'}^+] = n\delta_{n,-n'},$$
$$[\rho_n^-, \rho_{n'}^-] = -n\delta_{n,-n'}, \tag{2.4}$$
$$[\rho_n^+, \rho_{n'}^-] = 0.$$

Let us define operators

$$\sqrt{|n|}\, b_n^+ = \rho_{-n}^+ \quad \text{for } n > 0$$
$$= \rho_{-n}^- \quad \text{for } n < 0 \tag{2.5}$$
$$\sqrt{|n|}\, b_n = \rho_n^+ \quad \text{for } n > 0$$
$$= \rho_n^- \quad \text{for } n < 0.$$

These operators satisfy the commutation relations

$$[b_n, b_{n'}^+] = \delta_{nn'}, \quad [b_n, b_{n'}] = [b_n^+, b_{n'}^+] = 0. \tag{2.6}$$

Since our treatment is restricted to states with particles and holes near the Fermi momentum, we can approximate the free particle energy by a linear form

$$k_n^2/2m = (2m)^{-1}(k_F + k_n - k_F)^2 \approx (2m)^{-1}(2k_F k_n - k_F^2). \tag{2.7}$$

Tomonaga showed that the expression

$$H_{K.E.} = (\hbar^2 k_F/m) k_n b_n^+ b_n + E_0 \tag{2.8}$$

is equivalent to the Hamiltonian of the noninteracting

system. The constant E_0 is equal to the kinetic energy of the filled Fermi sea.[7]

Any two-body interaction of the form

$$H_I=\tfrac{1}{2}\int\int\rho(x)\rho(x')J(|x-x'|)\,dxdx'-\tfrac{1}{2}\int\rho(x)J(0)\,dx \tag{2.9}$$

can be expressed in terms of the density fluctuation operators as

$$H_I=\tfrac{1}{2}\sum_{n\neq0}J_n\rho_n\rho_{-n}+\tfrac{1}{2}\rho_0^2J_0-\tfrac{1}{2}\rho_0J(0) \tag{2.10}$$

or in terms of the boson operators b_n, b_n^+:

$$H_I=\tfrac{1}{2}\sum_{n\neq0}|n|\,J_n(b_n^++b_{-n})(b_{-n}^++b_n)+\tfrac{1}{2}\rho_0^2J_0-\tfrac{1}{2}\rho_0J(0), \tag{2.11}$$

where

$$J_n=J_{-n}=L^{-1}\int J(x)\exp(ik_nx)\,dx.$$

The full Hamiltonian, $H_{\text{K.E.}}+H_I$, can be diagonalized by a canonical transformation generated by the operator

$$G=\tfrac{1}{2}i\sum_{n\neq0}\theta_n(b_n^+b_{-n}^+-b_nb_{-n}), \tag{2.12}$$

where

$$\theta_n=\tfrac{1}{4}\ln[1+(NJ_n/2\epsilon_F)]$$

and N is the total number of particles. The Hamiltonian may be expressed in terms of the "dressed" boson operators as

$$H=\sum_{n\neq0}\Omega_n\tilde{b}_n\tilde{b}_n+\text{const}, \tag{2.13}$$

where

$$\begin{aligned}\tilde{b}_n&=\exp(iG)\,b_n\exp(-iG)=b_n\cosh\theta_n+b_{-n}^+\sinh\theta_n,\\ \tilde{b}_n^+&=\exp(iG)\,b_n^+\exp(-iG)=b_n^+\cosh\theta_n+b_{-n}\sinh\theta_n,\end{aligned} \tag{2.14}$$

and

$$\Omega_n=(\hbar^2k_F|k_n|/m)\exp(2\theta_n). \tag{2.15}$$

The interacting ground state is obtained by

$$|\tilde{0}>=\exp(iG)|0>. \tag{2.16}$$

The excited states may be expressed in terms of the "bare" or the "dressed" boson operators

$$|\{m_i(n_i)\}>=\prod_i(\tilde{b}_{n_i}^+)^{m_i}|\tilde{0}>=\exp(iG)\prod_i(b_{n_i}^+)^{m_i}|0>. \tag{2.17}$$

It was shown in Ref. 2 that the states of the form $\prod_i(b_{n_i}^+)^{m_i}|0>$ which belong to subspace S form a complete basis in this subspace.

III. THE MOMENTUM DISTRIBUTION

As in I, the field operators are decomposed into positive and negative momentum parts:

$$\begin{aligned}\psi_+(x)&=(\sqrt{L})^{-1}\sum_{n\geq0}a_n\exp(ik_nx),\\ \psi_-(x)&=(\sqrt{L})^{-1}\sum_{n<0}a_n\exp(ik_nx).\end{aligned} \tag{3.1}$$

We then start with the following expression for the number of particles of momentum k_n:

$$\begin{aligned}n(k_n)&=<a_n^+a_n>\\ &=1-\frac{1}{L}\int_{-L/2}^{L/2}\int<\psi_\pm(x)\,\psi_\pm^+(y)>\exp[ik_n(x-y)]\,dxdy.\end{aligned} \tag{3.2}$$

The upper sign refers to positive and the lower to negative momenta. The symbol $<\ldots>$ denotes the usual canonical ensemble averaging. In the absence of external fields the momentum distribution is symmetric about the origin and it suffices to calculate $n(k_n)$ for positive n. To do this we shall derive an operator which is equivalent to $\psi_+(x)\,\psi_+^+(y)$ in the subspace S and is expressed in terms of the boson density fluctuation operators. We shall need the commutation relations of the boson operators with the field operators ψ_+, ψ_+^+. These are easily found to be (for $n>0$):

$$[b_n,\psi_+^+(x)]=(\sqrt{n})^{-1}\exp(-ik_nx)\,\psi_+^+(x)+(\sqrt{n})^{-1}\exp(ik_nx)\,(\sqrt{L})^{-1}\times\sum_{-n/2<l<0}a_l^+\exp(-ik_lx), \tag{3.3}$$

$$[b_n^+,\psi_+^+(x)]=(\sqrt{n})^{-1}\exp(ik_nx)\,\psi_+^+(x)-(\sqrt{n})^{-1}\exp(ik_nx)\,(\sqrt{L})^{-1}\times\sum_{0<l<n}a_l^+\exp(-ik_lx).$$

The other two relevant commutation relations may be obtained from these two by

$$\begin{aligned}[b_n,\psi_+(x)]&=-[b_n^+,\psi_+^+(x)]^+,\\ [b_n^+,\psi_+(x)]&=-[b_n,\psi_+^+(x)]^+.\end{aligned} \tag{3.4}$$

These commutation relations are rather complicated, but they simplify greatly when evaluated between two states of subspace S. The second terms in Eqs. (3.3) have zero matrix elements in subspace S and we are left with the expressions ($n>0$)

$$\begin{aligned}[b_n,\psi_+^+(x)]&=(\sqrt{n})^{-1}\exp(-ik_nx)\,\psi_+^+(x),\\ [b_n^+,\psi_+^+(x)]&=(\sqrt{n})^{-1}\exp(ik_nx)\,\psi_+^+(x),\end{aligned} \tag{3.5}$$

and Eqs. (3.4). The commutation relations for negative n with ψ_+, ψ_+^+ are either identically zero, or they give terms which vanish between two states of subspace S.

Using these commutation relations we get ($n>0$)

$$\begin{aligned}[b_n,\psi_+(x)\,\psi_+^+(y)]&=i\alpha_n^*(x,y)\,\psi_+(x)\,\psi_+^+(y),\\ [b_n^+,\psi_+(x)\,\psi_+^+(y)]&=-i\alpha_n(x,y)\,\psi_+(x)\,\psi_+^+(y),\\ [b_{-n},\psi_+(x)\,\psi_+^+(y)]&=[b_{-n}^+,\psi_+(x)\,\psi_+^+(y)]=0,\end{aligned} \tag{3.6}$$

where

$$\alpha_n(x,y)=(i/\sqrt{n})[\exp(ik_ny)-\exp(ik_nx)]. \tag{3.7}$$

We are looking for an operator U, which is a function of b_n, b_n^+, satisfying the same commutation relations with b_n, b_n^+ as $\psi_+(x)\,\psi_+^+(y)$, and in addition to that we require

$$<0|U|0>=<0|\psi_+(x)\,\psi_+^+(y)|0>=L^{-1}\sum_{n>n_F}\exp[ik_n(y-x)]. \tag{3.8}$$

Such an operator U has the same matrix elements as

$\psi_+(x)\,\psi_+^*(y)$ between any two states of the type $\Pi_n(b_n^*)^{m_n}|0>$. It is easily seen that the desired commutation relations are satisfied by the operator

$$U=D(x,y)\exp[i\sum_{n>0}(\alpha_n b_n+\alpha_n^* b_n^*)]. \tag{3.9}$$

The coefficient $D(x,y)$ is determined by the condition (3.8). Using the well-known identity

$$e^{A+B}=e^A e^B e^{-[A,B]/2}, \tag{3.10}$$

we get

$$<0|U|0>=D(x,y)\exp(-\tfrac{1}{2}\sum_{n>0}|\alpha_n|^2)<0|\exp[i\sum\alpha_n^* b_n^*]\times\exp[i\sum\alpha_n b_n]|0> \tag{3.11}$$

and hence we find from Eq. (3.8)

$$D(x,y)=\exp[\tfrac{1}{2}\sum_{n>0}|\alpha_n|^2](1/L)\sum_{n>n_F}\exp[ik_n(y-x)]. \tag{3.12}$$

We shall now express the operator U by means of the "dressed" density fluctuation operators [the inverse of Eqs. (2.14)]:

$$\begin{aligned}U&=D(x,y)\exp(i\sum_{n>0}[\alpha_n(\tilde{b}_n\cosh\theta_n-\tilde{b}_{-n}^+\sinh\theta_n)\\&\quad+\alpha_n^*(\tilde{b}_n^+\cosh\theta_n-\tilde{b}_{-n}\sinh\theta_n)])\\&=D(x,y)\exp(i\sum_{n>0}\cosh\theta_n(\alpha_n\tilde{b}_n+\alpha_n^*\tilde{b}_n^+))\\&\quad\times\exp(-i\sum_{n>0}\sinh\theta_n(\alpha_n\tilde{b}_{-n}^+\alpha_n^*\tilde{b}_{-n})).\end{aligned} \tag{3.13}$$

The thermal averaging is easily performed with the help of the known identity for the creation and destruction operators of a harmonic oscillator[8]

$$\langle\exp[i(\alpha b+\alpha^* b^+)]\rangle=\exp[-|\alpha|^2\langle\hat{N}+\tfrac{1}{2}\rangle], \tag{3.14}$$

where the operator $\hat{N}$ measures the number of excitations of the oscillator. Applying this identity to our case of a system of uncoupled harmonic oscillators, we get

$$\begin{aligned}\langle\psi_+(x)\,\psi_+^*(y)\rangle&=D(x,y)\exp[-\sum_{n>0}|\alpha_n|^2(\cosh^2\theta_n+\sinh^2\theta_n)\langle\hat{N}+\tfrac{1}{2}\rangle]\\&=D(x,y)\exp[-\tfrac{1}{2}\sum|\alpha_n|^2]\exp[-\sum|\alpha_n|^2(\sinh^2\theta_n\\&\quad+\cosh2\theta_n\langle\hat{N}\rangle)],\end{aligned}$$

with

$$\langle\hat{N}\rangle=[\exp(\beta\Omega_n)-1]^{-1},$$

where Ω_n is given by Eq. (2.15), and $\beta=1/kT$. In view of Eqs. (3.7) and (3.12), we find from (3.15)

$$\langle\psi_+(x)\,\psi_+^*(y)\rangle=\exp[-Q(x-y),T]\;L^{-1}\sum_{n>n_F}\exp[ik_n(y-x)], \tag{3.16}$$

where

$$\begin{aligned}Q(x-y,T)&=\sum_{n>0}(2/n)\{\sinh^2\theta_n+[\cosh2\theta_n/(\exp(\beta\Omega_n)-1)]\}\\&\quad\times[1-\cos k_n(x-y)].\end{aligned} \tag{3.17}$$

For $T=0$, the second term in the parantheses vanishes and one gets Eq. (3.21) of I. Inserting Eq. (3.17) into Eq. (3.2) and converting sums into integrals, we get for the momentum distribution

$$\begin{aligned}n(k)&=1-L^{-1}\int_{-L/2}^{L/2}\int dxdy\,\exp[-Q(x-y,T)]\\&\quad\times\{1/2\pi\int_{k_F}^{\infty}\exp[i(k'-k)(x-y)]\,dk'.\end{aligned} \tag{3.18}$$

We define $x-y=r$, $k'-k=q$, integrate over one of the coordinates and pass to the macroscopic limit $N,L\to\infty$. Since $Q(r,T)$ is symmetric in r, we obtain

$$n(k)=1-1/2\pi\int_{-\infty}^{\infty}\{\exp[-Q(r,T)]\}(\int_{k_F-k}^{\infty}\cos qr\,dq)\,dr. \tag{3.19}$$

The range of integration over q may be split into two parts: $[k_F-k,0]$ and $[0,\infty]$. The integral from zero to infinity gives a δ function and hence

$$1/2\pi\int_{-\infty}^{\infty}[\exp(-Q(r,T)][\int_0^{\infty}\cos qr\,dq]\,dr=\tfrac{1}{2}, \tag{3.20}$$

since $Q(0,T)=0$. Finally,

$$\begin{aligned}n(k)&=\frac{1}{2}+\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp[-Q(r,T)]\left[\int_0^{k_F-k}\cos qr\,dq\right]dr\\&=\frac{1}{2}+\frac{1}{\pi}\int_0^{\infty}\left[\frac{\exp[-Q(r,T)]}{r}\right]\sin(k_F-k)\,r\,dr.\end{aligned} \tag{3.21}$$

IV. PROPERTIES OF $n(k)$

A. Free particles

Let us first prove that the last expression reduces to the ordinary Fermi–Dirac distribution in the case of free fermions. To this end we evaluate the function $Q(r,T)$, defined in Eq. (3.17), for zero interaction (we replace the sum by an integral)

$$Q(r,T)=2\int_0^{\infty}\frac{1-\cos kr}{q[\exp(\omega_k\beta)-1]}\,dk, \tag{4.1}$$

where

$$\omega_k=(\hbar^2k_F/m)k.$$

Denoting

$$\alpha=(\hbar k_F\beta/m)^{-1},$$

and substituting $x=\omega_k\beta$, one gets[9]

$$Q(r,T)=4\int_0^{\infty}\frac{\sin^2(\frac{1}{2}\alpha rx)\,dx}{x(\exp x-1)}=\ln\left[\frac{\sinh(\pi\alpha r)}{\pi\alpha r}\right]. \tag{4.2}$$

Inserting this expression into Eq. (3.21), we obtain[10]

$$\begin{aligned}n(k)&=\frac{1}{2}+\alpha\int_0^{\infty}\frac{\sin(k_F-k)r}{\sinh(\pi\alpha r)}\,dr=\frac{1}{2}+\frac{1}{2}\tanh\left(\frac{k_F-k}{2\alpha}\right)\\&=\left[\exp\left(\frac{\hbar^2k_F\beta}{m}(k-k_F)\right)+1\right]^{-1}.\end{aligned} \tag{4.3}$$

This is the Fermi–Dirac distribution with the linear relation between free particle energies and momenta, assumed in the Tomonaga model.

B. Behavior near k_F

The behavior of $n(k)$ near k_F, namely for small (k_F-k), is determined by the behavior of $\exp[-Q(r,T)]/r$ for large r. To obtain this behavior we split $Q(r,T)$ into two parts:

$$Q(r,T)=Q(r,0)+Q_1(r,T), \tag{4.4}$$

$$Q(r,0)=2\int_0^\infty \sinh^2\theta(k)\left(\frac{1-\cos kr}{k}\right)dk, \tag{4.5}$$

$$Q_1(r,T)=2\int_0^\infty \frac{\cosh 2\theta(k)}{\exp[\beta\tilde{\Omega}(k)]-1}\left(\frac{1-\cos kr}{k}\right)dk. \tag{4.6}$$

At $T=0$, one is left with $Q(r,0)$. The asymptotic behavior of $Q(r,0)$ was shown in I to be

$$Q(r,0)\to\lambda(0)\ln r + C + O(1/r^2), \tag{4.7}$$

where

$$\lambda(k)=2\sinh^2\theta(k),$$
$$C=2\int_0^\infty 1/k[\lambda(k)-\lambda(0)\cos k]dk. \tag{4.8}$$

It was also shown there that the behavior of $n(k)$ in the neighborhood of k_F is determined by the value of $\lambda(0)$. If $\lambda(0)=0$, $n(k)$ has a finite jump at k_F, if $\lambda(0)<1$, $n(k)$ is continuous at k_F but has an infinite derivative there, and finally, if $\lambda(0)>1$, $n(k)$ is continuous and has a finite derivative at k_F. It remains to derive the asymptotic behavior of $Q_1(r,T)$. To this end we rewrite Eq. (4.6) in the form

$$Q_1(r,T)=\int_0^\infty f(k)\left(\frac{1-\cos kr}{k^2}\right)dk, \tag{4.9}$$

where

$$f(k)=2k\cosh 2\theta(k)/\{\exp[\beta\Omega(k)]-1\}.$$

Note that $f(0)$ is finite. It is shown in the appendix that Eq. (4.9) is asymptotically equal to

$$Q_1(r,T)\to\tfrac{1}{2}\pi f(0)r + f'(0)\ln r + \text{const} + O(1/r^2). \tag{4.10}$$

The leading term in this expansion assures that the integral in Eq. (3.21) is absolutely integrable, and hence $n(k)$ is continuous at k_F and varies linearly in its immediate neighborhood for any interaction.

C. Consistency conditions

Let us now evaluate $n(k)$ for k significantly below k_F, say $k=\frac{1}{2}k_F$. It is evident that the main contribution to the integral in Eq. (3.21) comes in that case from the region $r<2k_F^{-1}$. We expand $Q(r,T)$, an even function of r, for small r:

$$Q(r,T)\simeq r^2/4b^2, \tag{4.11}$$

where

$$b^{-2}=4\int_0^\infty\left(\sinh^2\theta(k)+\frac{\cosh 2\theta(k)}{\exp[\beta\Omega(k)]-1}\right)kdk. \tag{4.12}$$

Inserting Eq. (4.11) into Eq. (3.21), we get

$$n(\tfrac{1}{2}k_F)=\tfrac{1}{2}+\tfrac{1}{2}\mathrm{erf}(\tfrac{1}{2}k_Fb). \tag{4.13}$$

For the Tomonaga model to apply, we want the error function to be close to one or k_Fb to be large. Actually, a value of $k_Fb\simeq 2.8$ is sufficient to give $n(\frac{1}{2}k_F)\simeq 0.99$. We write the consistency condition in the form

$$1\gg(k_Fb)^{-2}. \tag{4.14}$$

To obtain a qualitative understanding of this condition we choose a slowly varying potential of strength V and range R. The Fourier components of such a potential may be approximated by

$$J(k)=VR/L, \quad |k|\le k_R=2\pi/R,$$
$$=0, \quad |k|>k_R. \tag{4.15}$$

This gives

$$\theta(k)=\tfrac{1}{4}\ln[1+(V/2\epsilon_F)(R/r_0)], \quad |k|<k_R$$
$$=0, \quad k>k_R, \tag{4.16}$$

where r_0 is the average interparticle distance L/N. Note that a necessary condition for the Tomonaga model is $k_R\ll k_F$, which implies $R\gg r_0$. Let us now estimate the right-hand side of Eq. (4.15). Using Eqs. (4.12), (4.16), and (2.15), we obtain after some trivial changes of variables

$$(k_Fb)^{-2}=2(k_R/k_F)^2\sinh^2\theta + \frac{\cosh(2\theta)}{(\beta\epsilon_F)^2\exp(4\theta)}\int_0^{(k_R/k_F)2\beta\epsilon_F\exp(2\theta)}\frac{xdx}{\exp(x)-1} + \frac{1}{(\beta\epsilon_F)^2}\int_{(k_R/k_F)\beta\epsilon_F}^\infty\frac{xdx}{\exp(x)-1}, \tag{4.17}$$

where θ is the rhs of Eq. (4.16). We assume that $(\beta\epsilon_F)\gg 1$, otherwise states with $k\ll k_F$ are depleted because of thermal excitations regardless of the interaction and the Tomonaga model does not apply in that case. Thus, the third term hardly effects the consistency condition and we shall therefore omit it. In the second term we extend the integration limit to ∞. Thereby we overestimate the rhs of Eq. (4.14) and we shall therefore obtain only sufficient conditions for the applicability of the model. Next, we distinguish between repulsive and attractive interactions. In the first case the positive exponential in the hyperbolic functions is dominant, while in the latter case it is negative exponential which may become very large. Keeping this in mind, using Eq. (4.16) and putting $k_F=\pi/r_0$, we get for repulsive interactions

$$1\gg 8\left(\frac{r_0}{R}\right)^2\left[1+\left(\frac{V}{2\epsilon_F}\right)\left(\frac{R}{r_0}\right)\right]^{1/2}+\frac{\pi^2}{6(\beta\epsilon_F)^2}\left[1+\left(\frac{V}{2\epsilon_F}\right)\left(\frac{R}{r_0}\right)\right]^{-1/2}. \tag{4.18}$$

The second term is small and the applicability of the model depends on whether the first term satisfies the inequality or not. Thus, in the case of repulsion one finds the same conditions on the strength and range of the interaction as in I for $T=0$. For attractive interactions we get in the same way

$$1\gg 8\left(\frac{r_0}{R}\right)^2\left[1-\left(\frac{V}{2\epsilon_F}\right)\left(\frac{R}{r_0}\right)\right]^{-1/2}+\frac{\pi^2}{6(\beta\epsilon_F)^2}\left[1-\left(\frac{V}{2\epsilon_F}\right)\left(\frac{R}{r_0}\right)\right]^{-3/2}. \tag{4.19}$$

Note that the model breaks down unless the expression in the brackets is positive. Now, the second term may be as important as the first one. If the values of V and R are such that the model is still applicable at $T=0$, then it will remain applicable at temperatures for which

$$(\beta\epsilon_F)\gg\frac{\pi}{2\sqrt{12}}\left(\frac{R}{r_0}\right)\left[1-\left(\frac{V}{2\epsilon_F}\right)\left(\frac{R}{r_0}\right)\right]^{1/2}. \tag{4.20}$$

It follows that the range of temperatures for which the Tomonaga model still holds for attractive interactions decreases with increasing range and strength of the interaction and with increasing density of the system.

V. SUMMARY

The present paper is an extension of I to finite temperatures. The momentum distribution is derived by expressing products of fermion fields in terms of the boson operators which destroy and create eigenstates of the Hamiltonian. The thermal averaging may then be performed very simply. The $T=0$ results of I follow immediately. The interaction dependent singular behavior of the momentum distribution at the Fermi energy is washed out for $T>0$. The applicability condition in the case of repulsive interactions is almost unaffected by finite temperatures. On the other hand, one finds an interaction-dependent restriction on the range of T in the case of attraction.

APPENDIX

We want to prove Eq. (4.10) of the text. To this end we decompose Eq. (4.9) into three parts in the following manner:

$$\int_0^\infty f(k)\frac{1-\cos kr}{k^2}\,dk = I_1 + I_2 + I_3, \tag{A1}$$

where

$$I_1 = \int_0^\infty f(0)\frac{\cos k - \cos kr}{k^2}\,dk, \tag{A2}$$

$$I_2 = \int_0^\infty \left(\frac{f(k)-f(0)\cos k}{k}\right)\left(\frac{1-\cos kr}{k}\right)dk, \tag{A3}$$

$$I_3 = \int_0^\infty f(0)\frac{1-\cos k}{k^2}\cos kr\,dk. \tag{A4}$$

The first integral is equal to $[\pi f(0)/2]r$. Integrating Eq. (A4) twice by parts one finds that the asymptotic behavior of I_3 is $O(1/r^2)$. To evaluate the asymptotic behavior of (A3), we again decompose the integral I_2 into three parts:

$$I_2 = \int_0^\infty g(0)\frac{\cos k - \cos kr}{k}\,dk + \int_0^\infty \frac{g(k)-g(0)\cos k}{k}\,dk + \int_0^\infty \frac{g(k)-g(0)}{k}\cos kr\,dk, \tag{A5}$$

where we have defined

$$g(k) = \frac{f(k)-f(0)\cos k}{k} \tag{A6}$$

The first integral is equal to $g(0)\ln r$, the second is a constant, and the third behaves asymptotically as $1/r^2$. Note that $g(0)=f'(0)$. This completes the proof of Eq. (4.10).

*Present address: Belfer Graduate School of Science, Yeshiva University, New York, N. Y.

[1]S. Tomonaga, Prog. Theor. Phys. **5**, 544 (1950).
[2]M. Schick, Phys. Rev. **166**, 404 (1968).
[3]K. D. Schotte, Z. Phys. **230**, 99 (1970).
[4]K. D. Schotte and U. Schotte, Phys. Rev. **182**, 479 (1969).
[5]H. Gutfreund and M. Schick, Phys. Rev. **168**, 418 (1968).
[6]J. M. Luttinger, J. Math. Phys. **4**, 1154 (1963); D. Mattis and E. Lieb, J. Math. Phys. **6**, 304 (1963).
[7]It was shown in Ref. 2. that if one takes the true quadratic energy spectrum, it is still possible to express the kinetic energy in terms of the operators b_n, but the expression also contains trilinear terms.
[8]A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1961), p. 450.
[9]I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products* (Academic, New York, 1965), No. 3.951,21.
[10]Reference 9, No. 3.981,1.

Thermal conductivity of an isotopically disordered harmonic crystal

B. S. Semwal* and P. K. Sharma

Department of Physics, Universsity of Allahabad, Allahabad, India
(Received 27 March 1972; final manuscript received 15 September 1972).

Using the Kubo formalism, an expression is obtained for the lattice thermal conductivity of a three-dimensional harmonic Bravais crystal, containing a certain number of randomly distributed isotopic impurities, by the method of double-time thermal Green's functions. It is shown that the total thermal conductivity can be separated into two contributions, namely, diagonal and nondiagonal contributions; the former in the case of small half-width of the phonons reduces to the expression obtained from the Boltzmann transport equation. An approximate expression for the dominant nondiagonal contribution to the conductivity is obtained.

1. INTRODUCTION

There have been many attempts to calculate theoretically the thermal conductivity of the solids containing various defects using different techniques. Comprehensive review of the work is available in the literature.[1–4] These studies involve assumptions of one type or another and mostly use the phonon Boltzmann equation, first derived by Peierls,[5] in the relaxation time approximation. Although plausible from physical considerations the assumptions are ad hoc and difficult to incorporate consistently into a mathematical framework; this approach therefore suffers from the usual shortcomings of the kinetic theories as enumerated by Hardy,[6] and Allen and Ford.[7] The solution of the usual Boltzmann transport equation for the phonon occupation number $N_{\mathbf{ks}}$ in the relaxation time approximation gives

$$N_{\mathbf{ks}} = N^0_{\mathbf{ks}} - \tau_{\mathbf{ks}} \left(\frac{dN_{\mathbf{ks}}}{dt}\right) \mathbf{v}_{\mathbf{ks}} \cdot \nabla T, \tag{1}$$

where $N^0_{\mathbf{ks}}$, $\tau_{\mathbf{ks}}$, $\mathbf{v}_{\mathbf{ks}}$, and ∇T are the equilibrium occupation number, relaxation time, group velocity, and the temperature gradient, respectively. It is evident that as k goes to zero, $\tau_{\mathbf{ks}}$ tends to infinity which implies that for pure isotopic scattering, at some point, Eq. (1) predicts negative occupation numbers. It seems somewhat unphysical. Horie and Krumhansl,[8] and Kwok and Martin[9] have considered the generalization of the phonon Boltzmann equation using the equation of motion method for nonequilibrium Green's functions. This approach has recently been used by a number of authors[10–12] to derive expression for the transport coefficients.

Recent theories on phonon transport in solids express the thermal conductivity in terms of the correlation functions of the energy flux. This approach allows one to develop the theory of heat conduction in lattices on a more rigorous theoretical basis. Schieve and Peterson[13] have studied the elastic scattering of phonons by randomly arranged point defects in an anisotropic crystal using the correlation function method. Recently Allen and Ford[7] have investigated the energy transport properties of a one-dimensional isotopically disordered chain by using the Kubo formalism. In a later paper, they[14] extended the treatment to an infinite three-dimensional harmonic lattice of atoms connected by nearest-neighbor harmonic springs of equal strength. Woll[15] has also used the Kubo formalism to calculate the lattice thermal conductivity of a one-dimensional linear chain containing randomly distributed impurities using matrix propagator method. Maradudin[16] has made a similar calculation for a three-dimensional harmonic crystal through perturbation theory using diagram techniques and Chester and Thallung[17] have used van Hove's many-body technique. Krumhansi and Matthew[18] have considered the problem of one-dimensional lattice in a rather complete way. Elliott and Taylor[19] using double-time Green's function technique have evaluated the relaxation time of a crystal containing one changed mass at the origin. Yussouff and Mahanty,[20] and Klein[21] have given general theories of phonon scattering in cubic metals using the T matrix of generalized scattering theory.

We rely on the Kubo formalism[22] which is concise and rigorous and provides a systematic method to include the low-frequency phonons. This approach has recently been used by the authors[23] to obtain an expression for the thermal conductivity of an anharmonic crystal considering nondiagonal terms in the energy flux operator.[24] In this paper we have evaluated an expression for the thermal conductivity of a harmonic cubic Bravais crystal containing certain number of randomly distributed isotopic impurities using the method of double-time thermal Green's functions,[25] which has been widely used in a variety of many-body problems. In Sec. 2 we give a general formalism of Kubo formula for the thermal conductivity. In Sec. 3 the double-time Green's function is evaluated by the equation of motion method and the results are used to evaluate the correlation functions. In Sec. 4 we obtain an expression for the thermal conductivity of the system.

2. GENERAL FORMALISM

The Kubo formula for the thermal conductivity of a crystal is given by[26]

$$K = \lim_{\epsilon \to 0} \frac{k_B \beta}{3\Omega} \int_0^\infty dt e^{-\epsilon t} \int_0^\beta d\lambda \langle \mathbf{Q}(0) \cdot \mathbf{Q}(t + i\hbar\lambda) \rangle . \tag{2}$$

Here k_B is the Boltzmann constant, Ω is the volume of the crystal, t is the parameter with units of time, λ is the parameter with units (energy)$^{-1}$, $\hbar$ is the Planck's constant divided by 2π, $\beta = (k_B T)^{-1}$, T being the absolute temperature, $\mathbf{Q}(t)$ is the energy flux operator in the Heisenberg representation, and the angular bracket represents the thermal average, namely, for any operator O

$$\langle O \rangle = \mathrm{Tr}(e^{-\beta H} O)/\mathrm{Tr}(e^{-\beta H}), \tag{3}$$

where H is the crystal Hamiltonian. Hardy[24] has given a general expression for the energy flux operator for a three-dimensional crystal with anharmonic forces and has shown that even in the harmonic approximation the total energy flux operator contains some nondiagonal contribution besides the usual diagonal term. The total energy flux operator introduced by Hardy[24] can be written as

$$\mathbf{Q}(t)=\mathbf{Q}_d^0(t)+\mathbf{Q}_{nd}^0(t), \tag{4}$$

where $\mathbf{Q}_d^0(t)$ and $\mathbf{Q}_{nd}^0(t)$ represent the diagonal and nondiagonal parts of the energy flux operator, respectively, and are given by

$$\mathbf{Q}_d^0(t)=\sum_{\mathbf{ks}}\hbar\omega_{\mathbf{ks}}N_{\mathbf{ks}}(t)\mathbf{v}_{\mathbf{ks}}, \tag{5}$$

and

$$\mathbf{Q}_{nd}^0(t)=\frac{1}{2}{\sum_{\mathbf{kss'}}}'\hbar\omega_{\mathbf{ks}}\mathbf{v}_{\mathbf{kss'}}A_{\mathbf{ks}}B^{\dagger}_{\mathbf{ks'}}. \tag{6}$$

Here, $\omega_{\mathbf{ks}}$, $\mathbf{v}_{\mathbf{ks}}$, and $N_{\mathbf{ks}}=a^{\dagger}_{\mathbf{ks}}a_{\mathbf{ks}}$ are the frequency, group velocity, and the occupation number of the phonon in the mode (ks), $A_{\mathbf{ks}}$ and $B_{\mathbf{ks}}$ are phonon operators which are defined in terms of usual phonon creation and annihilation operators by $A_{\mathbf{ks}}=a_{\mathbf{ks}}+a^{\dagger}_{-\mathbf{ks}}$, $B_{\mathbf{ks}}=a_{\mathbf{ks}}-a^{\dagger}_{-\mathbf{ks}}$, $a^{\dagger}_{\mathbf{ks}}$ and $a_{\mathbf{ks}}$ are the creation and annihilation operators of the phonon of wave vector k and the polarization index s and $\mathbf{v}_{\mathbf{kss'}}$ $(s\neq s')$ has the dimension of the velocity. The prime over the summation in Eq. (6) denotes that $s\neq s'$. Using the expression (4) for $\mathbf{Q}(t)$ into Eq. (2), the expression for the thermal conductivity can be written as

$$K=K_0+K_1, \tag{7}$$

where

$$K_0=\lim_{\epsilon\to 0}\frac{\hbar^2k_B\beta}{3\Omega}\int_0^{\infty}dte^{-\epsilon t}\int_0^{\beta}d\lambda$$

$$\times\sum_{\mathbf{ks}}\sum_{\mathbf{k's'}}\omega_{\mathbf{ks}}\omega_{\mathbf{k's'}}\mathbf{v}_{\mathbf{ks}}\cdot\mathbf{v}_{\mathbf{k's'}}R_{\mathbf{ks},\mathbf{k's'}}(t+i\hbar\lambda). \tag{8}$$

and

$$K_1=\lim_{\epsilon\to 0}\frac{\hbar^2k_B\beta}{12\Omega}\int_0^{\infty}dte^{-\epsilon t}\int_0^{\beta}d\lambda$$

$$\times{\sum_{\mathbf{kss'}}}'\sum_{\mathbf{k}_1s_1s_1'}{}'\omega_{\mathbf{ks}}\omega_{\mathbf{k}_1s_1'}\mathbf{v}_{\mathbf{kss'}}\cdot\mathbf{v}_{\mathbf{k}_1s_1s_1'}L_{\mathbf{kss'},\mathbf{k}_1s_1s_1'}(t+i\hbar\lambda). \tag{9}$$

The correlation functions R and L are given by

$$R_{\mathbf{ks},\mathbf{k's'}}(t)=\langle a^{\dagger}_{\mathbf{ks}}(0)a_{\mathbf{ks}}(0)a^{\dagger}_{\mathbf{k's'}}(t)a_{\mathbf{k's'}}(t)\rangle, \tag{10}$$

and

$$L_{\mathbf{kss'},\mathbf{k}_1s_1s_1'}(t)=\langle A_{\mathbf{ks}}(0)B^{\dagger}_{\mathbf{ks'}}(0)A_{\mathbf{k}_1s_1}(t)B^{\dagger}_{\mathbf{k}_1s_1'}(t)\rangle. \tag{11}$$

The evaluation of the thermal conductivity thus involves the calculation of the two-time correlation functions given by Eqs. (10) and (11). It is difficult to evaluate them exactly. The evaluation is considerably simplified if the correlation function $\langle abcd\rangle$ is unlinked in the following manner.[27–29]

$$\langle abcd\rangle=\langle ab\rangle\langle cd\rangle+\langle ac\rangle\langle bd\rangle+\langle ad\rangle\langle bc\rangle, \tag{12}$$

where a,b,c, and d are operators. With this approximation, the correlation functions (10) and (11) can be written as

$$R_{\mathbf{ks},\mathbf{k's'}}(t)=\langle a^{\dagger}_{\mathbf{ks}}(0)a_{\mathbf{k's'}}(t)\rangle\langle a_{\mathbf{ks}}(0)a^{\dagger}_{\mathbf{k's'}}(t)\rangle$$
$$+\langle a^{\dagger}_{\mathbf{ks}}(0)a^{\dagger}_{\mathbf{k's'}}(t)\rangle\langle a_{\mathbf{ks}}(0)a_{\mathbf{k's'}}(t)\rangle, \tag{13}$$

$$L_{\mathbf{kss'},\mathbf{k}_1s_1s_1'}(t)=\langle A_{\mathbf{ks}}(0)A_{\mathbf{k}_1s_1}(t)\rangle\langle B^{\dagger}_{\mathbf{ks'}}(0)B^{\dagger}_{\mathbf{k}_1s_1'}(t)\rangle$$
$$+\langle A_{\mathbf{ks}}(0)B^{\dagger}_{\mathbf{k}_1s_1'}(t)\rangle\langle B^{\dagger}_{\mathbf{ks'}}(0)A_{\mathbf{k}_1s_1}(t)\rangle. \tag{14}$$

The approximation (12) is sufficiently simple and reflects the number of interesting properties of the system.[30,31] As will be seen, in the present problem the perturbation Hamiltonian is a quadratic function of the phonon operators. Since only two-phonon vertices occur in the problem, we are led to the result that the correlation function can be written in the form (13) and (14). It is to be noted that we have neglected the correlation functions of the operators with the same time argument because they will not contribute to the thermal conductivity. For the sake of convenience, in what follows we use index k as short form for ks and k' for $k's'$.

3. GREEN'S FUNCTIONS AND THE HAMILTONIAN

We evaluate the correlation functions occurring in Eqs. (13) and (14) with the help of one-particle retarded Green's function defined by[25]

$$G_{kk'}(t-t')=\langle\langle a_k(t);\ a^{\dagger}_{k'}(t')\rangle\rangle$$
$$=-i\theta(t-t')\langle[a_k(t),\ a^{\dagger}_{k'}(t')]\rangle, \tag{15}$$

where the step function $\theta(t-t')=1$ for $t>t'$, $=0$ for $t<t'$. The correlation function is given by

$$f_{kk'}(t)=\langle a^{\dagger}_{k'}(0)a_k(t)\rangle=\int_{-\infty}^{\infty}d\omega J_{kk'}(\omega)e^{-i\omega t}, \tag{16}$$

where $J_{kk'}(\omega)$ is the spectral density function and is related to the Green's function through the relation

$$J_{kk'}(\omega)=[i/(e^{\beta\hbar\omega}-1)][G_{kk'}(\omega+i\epsilon)-G_{kk'}(\omega-i\epsilon)]. \tag{17}$$

$G(\omega)$ is the Fourier transform of the double-time Green's function

$$G_{kk'}(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}dt\,G_{kk'}(t)e^{i\omega t}. \tag{18}$$

We consider a three-dimensional cubic crystal with volume Ω of whose $N-n$ lattice sites are occupied by atoms of mass M, while n sites are occupied by isotopic impurities of equal mass M'. It is assumed that the impurities are distributed randomly and n is quite small compared with N so that the distance between any two impurities is so large that their mutual interactions can be neglected. The total Hamiltonian for the disordered system may be written as

$$H=\sum_{l\alpha}\frac{p_{\alpha}^2(l)}{2M}+\sum_{i\alpha}{}'\frac{p_{\alpha}^2(i)}{2M'}\left(1-\frac{M'}{M}\right)$$
$$+\frac{1}{2}\sum_{l\alpha}\sum_{l'\beta}\phi_{\alpha\beta}(l,l')u_{\alpha}(l)u_{\beta}(l'), \tag{19}$$

where $p_\alpha(l)$ and $u_\alpha(l)$ are the α-Cartesian components of momentum and displacement from the equilibrium position of the lth atom, M and M' are the masses of the host and impurity atoms, $\phi_{\alpha\beta}(l,l')$ is the component of the harmonic force constant between the atoms l and l'. In writing Eq. (19) we have neglected any change in force constant resulting from the introduction of defect atoms. This will be true when the concentration of impurity atoms is small compared to N and the impurities are distributed randomly.

We now define the weighted harmonic mean M_0 of the masses of all the atoms by the relation

$$1/M_0=(f/M')+[(1-f)/M], \tag{20}$$

where $f=n/N$. One might call M_0 the effective atomic mass as seen by the phonons in the crystal. As seen from the above equation, it differs from atom mass by an amount depending upon the concentration of impurity. Using relation (20) in Eq. (19) and expressing the momentum and displacement vectors in terms of phonon creation and annihilation operators in the usual manner, the Hamiltonian H in the second quantized form can be written as

$$H=\sum_k \hbar\omega_k(a_k^\dagger a_k+\tfrac{1}{2})+H', \tag{21}$$

where H' is the perturbation part of the Hamiltonian which is responsible for the scattering of the phonons due to the isotopic impurities and is given by

$$H'=-\sum_{k_1,k_2}\hbar C(k_1,k_2)B_{k_1}B_{k_2}, \tag{22}$$

with

$$C(k_1,k_2)=(2\mu)^{-1}(M_0/2N)(\omega_{k_1}\omega_{k_2})^{1/2}\,\mathbf{e}_{k_1}\cdot\mathbf{e}_{k_2}\times\left[\sum_l^N f e^{i(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{R}_l}-\sum_l^n e^{i(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{R}_l}\right]. \tag{23}$$

Here ω_k is the frequency of the normal mode of the crystal for wave vector k, $\mathbf{e}_k$ is the polarization vector, $\mathbf{R}_l$ is the equilibrium position vector of the lth atom, and $\mu=MM'/(M'-M)$. $C(k_1,k_2)$ vanishes when n is either zero or N.

In order to evaluate the correlation functions occurring in Eqs. (13) and (14), we introduce the following one-particle Green's functions:

$$G_{kk'}(t-t')=\langle\langle a_k(t);a_{k'}^\dagger(t')\rangle\rangle, \tag{24a}$$
$$H_{kk'}(t-t')=\langle\langle a_k^\dagger(t);a_{k'}^\dagger(t')\rangle\rangle, \tag{24b}$$
$$h_{kk'}(t-t')=\langle\langle a_k(t);a_{k'}(t')\rangle\rangle, \tag{24c}$$
$$g_{kk'}(t-t')=\langle\langle A_k(t);A_{k'}(t')\rangle\rangle, \tag{24d}$$
$$D_{kk'}(t-t')=\langle\langle B_k^\dagger(t);B_{k'}^\dagger(t')\rangle\rangle, \tag{24e}$$
$$E_{kk'}(t-t')=\langle\langle A_k(t);B_{k'}^\dagger(t')\rangle\rangle. \tag{24f}$$

A convenient starting point for the calculation of these Green's functions is the equation of motion method. Differentiating Eq. (24a) with respect to time t and taking a Fourier transform like Eq. (18), we obtain

$$(\omega-\omega_k)G_{kk'}(\omega)=(\delta_{kk'}/2\pi)+2\sum_{k_1}C(-k,k_1)[G_{k_1k'}(\omega)-F_{k_1k'}(\omega)], \tag{25}$$

where $F_{kk'}(\omega)$ is the Fourier transform of the Green's function

$$F_{kk'}(t-t')=\langle\langle a_{-k}^\dagger(t);a_{k'}^\dagger(t')\rangle\rangle. \tag{26}$$

Similarly the equation for $F_{kk'}(\omega)$ is

$$(\omega+\omega_k)F_{kk'}(\omega)=2\sum_{k_1}C(-k,k_1)[G_{k_1k'}(\omega)-F_{k_1k'}(\omega)]. \tag{27}$$

Combination of Eqs. (25) and (27) leads to a simple relation between G and F as

$$(\omega-\omega_k)G_{kk'}(\omega)=(\delta_{kk'}/2\pi)+(\omega+\omega_k)F_{kk'}(\omega). \tag{28}$$

Inserting the value of $F_{kk'}(\omega)$ from Eq. (28) into Eq. (25) we get

$$(\omega-\omega_k)G_{kk'}(\omega)=(\delta_{kk'}/2\pi)+[C(-k,k')/\pi(\omega+\omega_{k'})]+4\sum_{k_1}[\omega_{k_1}C(-k,k_1)/(\omega+\omega_{k_1})]G_{k_1k'}(\omega). \tag{29}$$

An exact solution of Eq. (29) presents great difficulties. In order to get some physical results we use the iteration process and to the second-order approximation in the perturbation we write

$$(\omega-\omega_k)G_{kk'}(\omega)=\frac{\delta_{kk'}}{2\pi}+\frac{1}{\pi}\frac{C(-k,k')}{(\omega-\omega_{k'})}+\frac{4}{\pi}\sum_{k_1}\frac{C(-k,k_1)C(-k_1,k')\omega_{k_1}}{(\omega+\omega_{k'})(\omega^2-\omega_{k_1}^2)}+16\sum_{k_1,k_2}\frac{C(-k,k_1)C(-k_1,k_2)\omega_{k_1}\omega_{k_2}}{(\omega+\omega_{k_2})(\omega^2-\omega_{k_1}^2)}G_{k_2k'}(\omega). \tag{30}$$

The best approximation to cut-off the long chain of the Green's function can be obtained by taking $k_2=k$ in the third term which means the higher-order scattering processes are neglected. This reduces Eq. (30) to the form

$$[\omega-\omega_k-f_k(\omega)M_k(\omega)]G_{kk'}(\omega)=\frac{\delta_{kk'}}{2\pi}+\frac{1}{\pi}\frac{C(-k,k')}{(\omega-\omega_{k'})}+\frac{4}{\pi}\sum_{k_1}\frac{C(-k,k_1)C(-k_1,k')\omega_{k_1}}{(\omega^2-\omega_{k_1}^2)(\omega+\omega_{k'})}, \tag{31}$$

where

$$f_k(\omega)=(\omega+\omega_k)^{-1} \tag{32a}$$

and

$$M_k(\omega)=16\omega_k\sum_{k_1}[C(-k,k_1)C(-k_1,k)\omega_{k1}/(\omega^2-\omega_{k_1}^2)]. \tag{32b}$$

The solution of Eq. (31) can conveniently be written as

$$G_{kk'}(\omega)=G_{kk'}^0(\omega)+G_{kk'}'(\omega)+G_{kk'}''(\omega), \tag{33}$$

where

$$G_{kk'}^0(\omega)=\delta_{kk'}/2\pi[\omega-\omega_k-f_k(\omega)M_k(\omega)], \tag{34}$$

$$G_{kk'}'(\omega)=\frac{C(-k,k')}{\pi(\omega-\omega_{k'})[\omega-\omega_k-f_k(\omega)M_k(\omega)]}, \tag{35}$$

and

$$G''_{kk'}(\omega)=\frac{4}{\pi}\frac{\Sigma_{k_1}[C(-k,k_1)\omega_{k_1}C(-k_1,k')/(\omega+\omega_{k'})(\omega^2-\omega_{k_1}^2)]}{[\omega-\omega_k-f_k(\omega)M_k(\omega)]}. \tag{36}$$

For a small quantity ϵ tending to zero, Eqs. (32a) and (32b) can be written as

$$f_k(\omega\pm i\epsilon)=\alpha_k(\omega)\mp i\gamma_k(\omega), \tag{37a}$$

$$M_k(\omega\pm i\epsilon)=\xi_k(\omega)\mp i\eta_k(\omega), \tag{37b}$$

where $\alpha_k(\omega)$ and $\xi_k(\omega)$ represent the shift in the frequency, and $\gamma_k(\omega)$ and $\eta_k(\omega)$ are the half-width of the phonons of wave vector k. The explicit expressions for them are given by

$$\alpha_k(\omega)=P(\omega+\omega_k)^{-1}, \tag{38}$$

$$\xi_k(\omega)=16\omega_k\sum_{k_1}C(-k,k_1)C(-k_1,k)\omega_{k_1}P(\omega^2-\omega_{k_1}^2)^{-1}, \tag{39}$$

$$\gamma_k(\omega)=\pi\delta(\omega+\omega_k), \tag{40}$$

and

$$\eta_k(\omega)=8\pi\omega_k\sum_{k_1}C(-k,k_1)C(-k_1,k)\{\delta(\omega-\omega_{k_1})-\delta(\omega+\omega_{k_1})\}, \tag{41}$$

where P denotes the principal value. From Eqs. (37a) and (37b) the total shift in frequency $\Delta_k(\omega)$ and the half-width $\Gamma_k(\omega)$ are given by

$$\begin{aligned}\Delta_k(\omega)=&16\omega_kP(\omega+\omega_k)^{-1}\sum_{k_1}C(-k,k_1)C(-k_1,k)\\&\times\omega_{k_1}P(\omega^2-\omega_{k_1}^2)^{-1}\\&-8\pi^2\omega_k\delta(\omega+\omega_k)\sum_{k_1}C(-k,k_1)C(-k_1,k)\\&\times\{\delta(\omega-\omega_{k_1})-\delta(\omega+\omega_{k_1})\},\end{aligned} \tag{42}$$

and

$$\begin{aligned}\Gamma_k(\omega)=8\pi\omega_k\sum_{k_1}&C(-k,k_1)C(-k_1,k)\{\delta(\omega+\omega_k)P[2\omega_{k_1}/(\omega^2-\omega_{k_1}^2)]\\&+P(\omega+\omega_k)^{-1}[\delta(\omega-\omega_{k_1})-\delta(\omega+\omega_{k_1})]\}.\end{aligned} \tag{43}$$

Similarly if we proceed with the equation of motion for the Green's functions (24b), (24c), (24d), (24e), and (24f) and follow the procedure as used above, we obtain

$$H_{kk'}(\omega)=\frac{C(k,k')}{\pi(\omega-\omega_{k'})[\omega+\omega_k-f'_k(\omega)M_k(\omega)]}, \tag{44a}$$

$$h_{kk'}(\omega)=\frac{C(-k,-k')}{\pi(\omega+\omega_{k'})[\omega-\omega_k-f_k(\omega)M_k(\omega)]} \tag{44b}$$

$$g_{kk'}(\omega)=\frac{\omega_k\delta_{k-k'}}{\pi[\omega^2-\omega_k^2-M_k(\omega)]}, \tag{44c}$$

$$D_{kk'}(\omega)=-\frac{\omega_k\delta_{k-k'}}{\pi[\omega^2-\omega_k^2-M_k(\omega)]}, \tag{44d}$$

and

$$E_{kk'}(\omega)=\frac{\omega\delta_{kk'}}{\pi[\omega^2-\omega_k^2-M_k(\omega)]}, \tag{44e}$$

where $f'_k(\omega)=(\omega-\omega_k)^{-1}$. In the above equations for simplicity we have retained only the dominating term in the expansion of the type (33) and higher-order terms have been neglected.

4. THERMAL CONDUCTIVITY

With the use of Eqs. (13), (16), and (17) in Eq. (8), the contribution K_0 to the thermal conductivity can be written as

$$K_0=K'_0+K''_0, \tag{45}$$

where

$$\begin{aligned}K'_0=&-\lim_{\epsilon\to0}\frac{\hbar^2k_B\beta}{3\Omega}\sum_{k,k'}\omega_k\omega_{k'}\mathbf{v}_k\cdot\mathbf{v}_{k'}\int_0^\infty dte^{-\epsilon t}\\&\times\int_0^\beta d\lambda\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}d\omega_1d\omega_2\frac{e^{-i(\omega_2-\omega_1)t}e^{(\omega_2-\omega_1)\hbar\lambda}}{(e^{\beta\hbar\omega_1}-1)(e^{\beta\hbar\omega_2}-1)}\\&\times e^{\beta\hbar\omega_1}[G_{kk'}(\omega_1+i\epsilon)-G_{kk'}(\omega_1-i\epsilon)]\\&\times[G_{k'k}(\omega_2+i\epsilon)-G_{k'k}(\omega_2-i\epsilon)],\end{aligned} \tag{46}$$

and

$$\begin{aligned}K''_0=&-\lim_{\epsilon\to0}\frac{\hbar^2k_B\beta}{3\Omega}\sum_{k,k'}\omega_k\omega_{k'}\mathbf{v}_k\cdot\mathbf{v}_{k'}\int_0^\infty dte^{-\epsilon t}\\&\times\int_0^\beta d\lambda\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}d\omega_1d\omega_2\frac{e^{-i(\omega_2+\omega_1)t}e^{(\omega_2+\omega_1)\hbar\lambda}}{(e^{\beta\hbar\omega_1}-1)(e^{\beta\hbar\omega_2}-1)}\\&\times[H_{k'k}(\omega_1+i\epsilon)-H_{k'k}(\omega_1-i\epsilon)]\\&\times[h_{k'k}(\omega_2+i\epsilon)-h_{k'k}(\omega_2-i\epsilon)].\end{aligned} \tag{47}$$

Interchanging ω_1 and ω_2 and noting that

$$(\omega_1-\omega_2-i\epsilon)^{-1}-(\omega_1-\omega_2+i\epsilon)^{-1}=2\pi i\delta(\omega_1-\omega_2), \tag{48}$$

Eq. (46) reduces to

$$\begin{aligned}K'_0=&-\lim_{\epsilon\to0}\frac{\hbar^2k_B\beta^2\pi}{3\Omega}\sum_{k,k'}\omega_k\omega_{k'}\mathbf{v}_k\cdot\mathbf{v}_{k'}\\&\times\int_{-\infty}^{\infty}d\omega\frac{e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega}-1)^2}[G_{kk'}(\omega+i\epsilon)-G_{kk'}(\omega-i\epsilon)]\\&\times[G_{k'k}(\omega+i\epsilon)-G_{k'k}(\omega-i\epsilon)].\end{aligned} \tag{49}$$

If we substitute the value of $G(\omega)$ from Eqs. (33), (34), (35), and (36) in the above expression we obtain an explicit expression for K'_0. But as may be seen, the major contribution to the thermal conductivity comes from the Green's function $G^0(\omega)$. The other terms give nondiagonal contribution, i.e., the contribution from the waves of different wave vectors but in the same state of polarization. Their net contributions will however be small as compared to the diagonal one. We evaluate here the diagonal and nondiagonal contributions to the thermal conductivity K'_0 separately. Substituting the value of $G^0(\omega)$ from Eq. (34) into Eq. (49) we obtain for the diagonal part of the conductivity K'_0 as

$$K'_{0d}=\frac{\hbar^2k_B\beta^2}{3\Omega\pi}\sum_k\omega_k^2v_k^2\int_{-\infty}^{\infty}d\omega\frac{e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega}-1)^2}$$

$$\times \frac{\Gamma_k^2(\omega)}{\{[\omega-\epsilon_k(\omega)]^2+\Gamma_k^2(\omega)\}^2}, \tag{50}$$

where $\epsilon_k(\omega)=\omega_k+\Delta_k(\omega)$. In the limiting case when $\Gamma_k(\omega)$ is small the integrand in Eq. (50) has sharp maximum around $\omega\approx\epsilon_k$ and the integral can be evaluated analytically. We then obtain the following expression for the thermal conductivity:

$$K'_{0d}=\frac{\hbar^2 k_B\beta^2}{3\Omega}\sum_k \omega_k^2 v_k^2 \frac{e^{\beta\hbar\epsilon_k}}{(e^{\beta\hbar\epsilon_k}-1)^2}\tau_k, \tag{51}$$

where the relaxation time τ_k is given by

$$\tau_k=1/2\Gamma_k. \tag{52}$$

Equation (51) has the form similar to that obtained from kinetic theory Boltzmann equation[1,2] for the isotopic mass scattering and is obtained here as a direct consequence of correlation function formula and the choice of the Hamiltonian. Many ad hoc assumptions (i.e., phenomenological phonon lifetime) employed in the kinetic theory are avoided here. The lifetime in our result arises in a natural way.

In Eq. (43) only positive values of the frequency are meaningful. The density of states in k space is sufficiently great, we may replace the summation over k_1 by integration so that Γ_k becomes

$$\Gamma_k=8\pi\omega_k(\Omega/8\pi^3)\int \sin\theta d\theta d\phi k_1^2 dk_1 |C(-k,k_1)|^2 \times P(\omega+\omega_k)^{-1}\delta(\omega-\omega_{k_1}), \tag{53}$$

where (θ,ϕ) measures the direction of k_1 with respect to k. Because of the delta function the integration is to be carried out over a constant energy surface characterized by the energy $\hbar\omega$ in wave vector space. In Debye approximation, $\omega=kc$, where c is the sound velocity, we get

$$\Gamma_k=\frac{\omega_k^2\Omega}{2\pi^2c^3}\int \left| C(k,k_1)\right|^2_{|\mathbf{k}_1|=|\mathbf{k}|} \sin\theta d\theta d\phi. \tag{54}$$

Expression (54) gives Γ_k and hence the thermal conductivity for a particular configuration of the two kinds of atoms over the lattice sites. In the case of random distribution of impurities, averaging the terms in bracket in Eq. (23) over all possible atomic configuration, and taking the directional average we obtain

$$\Gamma_k=(\omega_k^4\Omega/24N\pi c^3)(M_0/\mu)^2 f(1-f), \tag{55}$$

which shows that Γ_k and inverse relaxation time τ_k^{-1} of the scattering of phonons by mass defect is proportional to fourth power[32] of the phonon frequency ω_k and the square of the mass change $M'-M$.[33]

With Eq. (55), for small ω_k, the expression (51) for the thermal conductivity diverges. The divergence, as is evident, arises because we consider an infinite harmonic crystal and elastic processes do not lead to the energy dissipation. To avoid this difficulty, one should consider scattering processes in a pure crystal using the Mathiessen rule[34] which states that different scattering mechanisms (boundary, Umklapp scattering, etc.) give additive contribution to τ_k^{-1}. τ_k^{-1} is therefore replaced by $\tau_{\rm tot}^{-1}=\tau_0^{-1}+\tau_k^{-1}$, where τ_0^{-1} is the reciprocal relaxation time for pure crystal due to Umklapp processes and scattering by boundaries and uncontrolled impurities.

The nondiagonal part of the thermal conductivity K'_0 can be obtained similarly by substituting the values of $G'(\omega)$ and $G''(\omega)$ from Eqs. (35) and (36), respectively, into Eq. (49). The result is

$$\begin{aligned}K'_{0nd}=&(4\hbar^2k_B\beta^2/3\Omega\pi)\sum_{k,k'}\omega_k\omega_{k'}\mathbf{v}_k\cdot\mathbf{v}_{k'}|C(-k,k')|^2\\ &\times\int_{-\infty}^{\infty}d\omega[e^{\beta\hbar\omega}/(e^{\beta\hbar\omega}-1)^2](P[(\omega-\omega_k)(\omega-\omega_{k'})]^{-1}\Gamma_k(\omega)\Gamma_{k'}(\omega)\\ &+\pi\{P(\omega-\omega_k)^{-1}[\omega-\epsilon_k(\omega)]\delta(\omega-\omega_{k'})\Gamma_{k'}(\omega)\\ &+P(\omega-\omega_{k'})^{-1}[\omega-\epsilon_{k'}(\omega)]\delta(\omega-\omega_k)\Gamma_k(\omega)\\ &+\pi[\omega-\epsilon_k(\omega)][\omega-\epsilon_{k'}(\omega)]\delta(\omega-\omega_k)\delta(\omega-\omega_{k'})\})\\ &/\{[(\omega-\epsilon_k(\omega))^2+\Gamma_k^2(\omega)][(\omega-\epsilon_{k'}(\omega))^2+\Gamma_{k'}^2(\omega)]\}\\ &+\ldots,\end{aligned} \tag{56}$$

where the dots represent the contributions arising from higher-order Green's functions in the expansion (33). For $\epsilon_k\approx\epsilon_{k'}$, $\Gamma_k\approx\Gamma_{k'}$, $\omega_k\approx\omega_{k'}$ $(k\neq k')$, the above expression is simplified to

$$\begin{aligned}K'_{0nd}=&\frac{4\hbar^2k_B\beta^2}{3\Omega}\sum_{k,k'}\omega_k\omega_{k'}\mathbf{v}_k\cdot\mathbf{v}_{k'}|C(-k,k')|^2\\ &\times\frac{e^{\beta\hbar\epsilon_k}}{(e^{\beta\hbar\epsilon_k}-1)^2}\left[\frac{1}{2\Delta_k^2\Gamma_k}+\frac{\Gamma_k^2-\Delta_k^2}{(\Delta_k^2+\Gamma_k^2)^3}\right.\\ &\left.+\pi\delta(\omega_k-\omega_{k'})\frac{\Delta_k^2}{(\Delta_k^2+\Gamma_k^2)^2}\right].\end{aligned} \tag{57}$$

The nondiagonal contribution K''_0 to the thermal conductivity arising from two-creation and two-annihilation correlation functions is given by Eq. (47). Substituting the values of the Green's functions $H_{kk'}$ and $h_{kk'}$ from Eqs. (44a) and (44b), this contribution turns out to be equal to zero. In the higher-order approximation of the Green's functions, this is negligibly small. This is no longer small in the case of superconductors because of the law of conservation of momentum. Thus for normal matter the contribution of the diagonal part of the energy flux operator to the thermal conductivity is given by Eqs. (50) and (57).

We come finally to the nondiagonal contribution K_1 to the thermal conductivity due to the nondiagonal term in the energy flux operator. If we make use of Eqs. (14), (16), (17), (44c), (44d), and (44e), we see that the expression for K_1 can be written in the form

$$K_1=K'_1+K''_1, \tag{58}$$

where

$$\begin{aligned}K'_1=&-(\hbar^2k_B\beta^2/3\Omega\pi)\sum_{\mathbf{k}ss'}{}'\sum_{\mathbf{k}_1s_1s'_1}{}'\omega_{\mathbf{k}s}\omega_{\mathbf{k}_1s_1}\mathbf{v}_{\mathbf{k}ss'}\cdot\mathbf{v}_{\mathbf{k}_1s_1s'_1}\\ &\times\int_{-\infty}^{\infty}d\omega[e^{\beta\hbar\omega}/(e^{\beta\hbar\omega}-1)^2]\omega_{\mathbf{k}s}\omega_{\mathbf{k}s'}\eta_{\mathbf{k}s}(\omega)\eta_{\mathbf{k}s'}(\omega)\\ &\times\frac{\delta_{k-k_1}\delta_{ss_1}\delta_{s's'_1}}{[(\omega^2-\varphi_{\mathbf{k}s}^2(\omega))^2+\eta_{\mathbf{k}s}^2(\omega)][(\omega^2-\varphi_{\mathbf{k}s'}^2(\omega))^2+\eta_{\mathbf{k}s'}^2(\omega)]},\end{aligned} \tag{59}$$

where $\eta_{\mathbf{k}s}(\omega)$ is given by Eq. (41) and

$$\varphi_{\mathbf{k}s}^2(\omega)=\omega_{\mathbf{k}s}^2+\xi_{\mathbf{k}s}(\omega).$$

Using the symmetry relations $v_{kss'} = -v_{-k\,ss'}$ and $\omega_{ks} = \omega_{-ks}$, the above equation becomes

$$K_1' = \frac{\hbar^2 k_B \beta^2}{3\Omega\pi} \sum_{kss'}{}' \omega_{ks}^2 \omega_{ks'}^2 v_{kss'}^2 \int_{-\infty}^{\infty} d\omega \frac{e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega}-1)^2} \times \frac{\eta_{ks}(\omega)\eta_{ks'}(\omega)}{[(\omega^2 - \varphi_{ks}^2(\omega))^2 + \eta_{ks}^2(\omega)][(\omega^2 - \varphi_{ks'}^2(\omega))^2 + \eta_{ks'}^2(\omega)]}. \quad (60)$$

Similarly, the expression for K_1'' is given by

$$K_1'' = \frac{\hbar^2 k_B \beta^2}{3\Omega\pi} \sum_{kss'}{}' \omega_{ks}\omega_{ks'} v_{kss'} \cdot v_{ks's} \int_{-\infty}^{\infty} d\omega\ \omega^2 \frac{e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega}-1)^2} \times \frac{\eta_{ks}(\omega)\eta_{ks'}(\omega)}{[(\omega^2 - \varphi_{ks}^2(\omega))^2 + \eta_{ks}^2(\omega)][(\omega^2 - \varphi_{ks'}^2(\omega))^2 + \eta_{ks'}^2(\omega)]}. \quad (61)$$

Equations (50), (57), (60), and (61) give the expression for the thermal conductivity of a cubic Bravais crystal containing isotopic impurities. Even in the diagonal approximation for the energy flux operator, the total thermal conductivity can be expressed as the sum of two contributions, namely, diagonal and nondiagonal ones. The relaxation time of the former in the case of small half-width of phonons is approximately proportional to the fourth power of the phonon frequency and is similar to what Klemens[32] has obtained using the Boltzmann equation. Under the approximation mentioned earlier, the nondiagonal contribution K_{ond}' [Eq. (57)] depends upon Δ_k^{-2} and $(2\Gamma_k)^{-1}$ and will thus be much smaller as compared to the diagonal one. Expressions (60) and (61) show that the nondiagonal contribution K_1 from the nondiagonal part of the energy flux operator comes from modes with different polarizations. Further the contribution K_1 depends on higher order concentration of impurities than K_0. At temperatures which are so small that only those phonons which suffer negligible impurity scattering are excited, the diagonal contribution (50) gives a reasonable approximation to conductivity of crystals which contain impurities. At high temperatures, the full form should be used. Hardy,[24] using a classical treatment, has argued that the nondiagonal energy flux operator is made up of oscillating terms whose frequencies are the sum and difference of the frequencies of different branches and will give negligible contribution to thermal conductivity compared to diagonal one when averaged over a long period of time. The present work and the earlier study[23] show that their contribution is finite, though expected to be much smaller than diagonal contribution.

The present study shows that from the Kubo formula using the double-time Green's function technique, the thermal conductivity due to isotopic mass scattering can easily be obtained. Further the relaxation time of the phonon scattering by impurities can be estimated much more accurately than from the simple Boltzmann transport equation. It may be emphasized that the results obtained in this paper differ from that of the conventional theory based on the Peierls transport equation since the latter makes no provision for the nondiagonal contribution to energy flux operator.

ACKNOWLEDGMENTS

The authors are thankful to Professor K.S. Singwi and Professor Vachaspati for encouragement. Thanks are also due to the Council of Scientific and Industrial Research, New Delhi for financial assistance.

*Present address: Department of Physics, Birla Govt. Post-Graduate College, Srinagar (Garhwal) (U.P.), India.

[1]P. G. Klemens, Handb. Phys. 14, 198 (1956); Solid State Phys. 7, 1 (1958).
[2]G. Leibfried, Handb. Phys. 7, 293 (1955).
[3]P. Carruthers, Rev. Mod. Phys. 33, 92 (1961).
[4]P. C. Kwok, Solid State Phys. 20, 214 (1967).
[5]R. E. Peierls, Ann. Physik (Leipzig) 3, 1055 (1929).
[6]R. J. Hardy, J. Math. Phys. 6, 1749 (1965); J. Math. Phys. 7, 1435 (1966).
[7]K. R. Allen and J. Ford, Phys. Rev. 176, 1046 (1968).
[8]C. Horie and J. A. Krumhansl, Phys. Rev. 136, A1397 (1964).
[9]P. C. Kwok and P. C. Martin, Phys. Rev. 142, 495 (1966).
[10]L. J. Sham, Phys. Rev. 156, 494 (1967); Phys. Rev. 163, 401 (1967).
[11]R. Klein and R. K. Wehner, Phys. Kondens. Mater. 8, 141 (1968).
[12]W. Götze and K. H. Michel, Z. Phys. 223, 199 (1969).
[13]W. C. Schieve and R. L. Peterson, Phys. Rev. 126, 1458 (1962).
[14]K. R. Allen and J. Ford, Phys. Rev. 187, 1132 (1968).
[15]E. J. Woll, Jr., Phys. Rev. 137, A95 (1965).
[16]A. A. Maradudin, J. Am. Chem. Soc. 86, 3405 (1964).
[17]G. V. Chester and A. Thellung, Proc. Phys. Soc. Lond. 73, 745 (1959); Proc. Phys. Soc. Lond. 77, 1005 (1961).
[18]J. A. Krumhansl and J. A. D. Matthew, Phys. Rev. 140, A1812 (1965).
[19]R. J. Elliott and D. W. Taylor, Proc. Phys. Soc. Lond. 83, 189 (1964).
[20]M. Yussouff and J. Mahanty, Proc. Phys. Soc. Lond. 85, 1223 (1965); Proc. Phys. Soc. Lond. 87, 689 (1966).
[21]M. V. Klein, Phys. Rev. 131, 1500 (1963); Phys. Rev. 141, 716 (1966).
[22]R. Kubo, M. Yokota, and S. Nakajima, J. Phys. Soc. Jap. 12, 1203 (1957).
[23]B. S. Semwal and P. K. Sharma, Phys. Rev. B 5, 3909 (1972).
[24]R. J. Hardy, Phys. Rev. 132, 168 (1963).
[25]D. N. Zubarev, Usp. Fiz. Nauk 71, 71 (1960) [Sov. Phys.-Usp. 3, 320 (1960)].
[26]H. Mori, I. Oppenheim, and J. Ross, in *Studies in Statistical Mechanics*, edited by J. de Boer and G. E. Uhlenbeck (Interscience, New York, 1962), Vol. 1, pp. 271-298.
[27]L. P. Gorkov, Zh. Eksp. Teor. Fiz. 34, 735 (1958) [Sov. Phys.-JETP 7, 505 (1958)].
[28]C. Bloch and C. de Dominicis, Nucl. Phys. 7, 459 (1958).
[29]D. N. Zubarev, Dokl. Akad. Nauk SSSR 132, 1055 (1960) [Sov. Phys.-Dokl. 5, 570 (1960)].
[30]N. N. Bogolubov and S. V. Tyablikov, Dokl. Akad. Nauk SSSR 126, 53 (1959) [Sov. Phys.-Dokl. 4, 589 (1959)].
[31]K. N. Pathak, Phys. Rev. 139, A1569 (1965).
[32]P. G. Klemens, Proc. R. Soc. Lond. 208, 108 (1951).
[33]P. G. Klemens, Proc. Phys. Soc. Lond. A68, 1113 (1955).
[34]J. M. Ziman, *Electrons and Phonons* (Oxford U. P., New York, 1960).

Deformation of the $so(2, c)$ subgroup of the Lorentz group

M. Huszár

Central Research Institute for Physics of the Hungarian Academy of Sciences, Budapest, Hungary
(Received 29 August 1972)

The two-dimensional complex sphere $S_1^2+S_2^2+S_3^2=S^2$ forms a homogeneous space under the $SL(2,C)$ group. The little group of a point in this space is the $SO(2,C)$ group or the horospheric group $T(2)$, according to whether $S\neq 0$ or $S=0$. Deformation of the $SO(2,C)$ group into $T(2)$ is investigated and is demonstrated on unitary representations. This deformation is a counterpart to that of the little groups $SO(3)$ or $SO(2,1)$ into $E(2)$. We conclude with a formula relating the matrix elements of unitary representations of the $SL(2,C)$ group in $SO(2)\times SO(1,1)=SO(2,C)$ basis to those in horospheric basis.

It is well known that horospheric subgroups, as well as horospheres themselves, play an important role in the representation theory of the $SL(2,C)$ group. A horosphere in a homogeneous space is the orbit of a point under the subgroup of the form $\left(\begin{smallmatrix}1&\beta\\0&1\end{smallmatrix}\right)$ (or any conjugate subgroup).[1] The mathematical significance of horospheres lies, among other things, in the existence of horospheric integral transformations, since by means of these we can associate with each function defined over a space Σ, homogeneous under the $SL(2,C)$ group, another function defined over the manifold of horospheres. This reversible integral transformation provides an opportunity to investigate functions defined over the manifold of horospheres Σ_0 instead of the functions defined over the homogeneous space Σ. This considerably simplifies the problem of, say, finding functions on Σ that transform according to unitary irreducible representations, since the functions in question, generally speaking, assume a much simpler form on Σ_0. The difficulties are in this way shifted to the subsequent integration accomplishing the inverse horospheric transformation.

The two-dimensional complex sphere Σ investigated in the present paper furnishes a good instance of this. As it will be shown, vectors on the complex sphere of zero radius Σ_0 can be associated with the horospheres situated on Σ. The space Σ_0 is equivalent to the space of spinors z_1, z_2, in the sense that they are both quotient spaces $SL(2,C)/H$ with respect to the same, so-called horospheric subgroup H. It is well known that in the space of spinors, functions transforming according to irreducible unitary representations in any basis assume a rather simple form. Actually, the facts mentioned above render it possible for the representation theory of the $SL(2,C)$ group to be built up on spinors in a relatively simple way.[1] All more specific problems, such as the search for spherical functions in certain homogeneous spaces, the evaluation of matrix elements of representations in different bases, etc., reduce to performing an inverse horospheric transformation. (In many cases this is done implicitly, when the matrix elements of representations are expressed in terms of integrations over the spinor space.) Even if, in a particular case, this does not prove to be a simple task, and even if an integration is not the simplest way of evaluation, at any rate the method concentrates the difficulties where they are in fact present.

The subject of the present paper is the deformation of the $SO(2,C)$ subgroup of the $SL(2,C)$ group into $T(2)$, the horospheric subgroup. As is known, the little group of a point on the two-sheeted (one-sheeted) real hyperboloid is the $SO(3)$ $[SO(2,1)]$ group, which can be continuously deformed into $E(2)$, the little group of the light cone.[3-5] The subgroups $SO(2,C)$ and $T(2)$ arise when studying the two-dimensional complex sphere $S_1^2+S_2^2+S_3^2=S^2$, homogeneous under the $SL(2,C)$ group. It will be shown that if $S\neq 0$, the little group of a certain fixed point on the complex sphere is the $SO(2,C)$ group, consisting of matrices of the form $\left(\begin{smallmatrix}\alpha&0\\0&\alpha^{-1}\end{smallmatrix}\right)$. On the other hand, the little group of a point on the complex sphere of zero radius, $S_1^2+S_2^2+S_3^2=0$, is the horospheric subgroup $\left(\begin{smallmatrix}1&\beta\\0&1\end{smallmatrix}\right)$ isomorphic to $T(2)$, the group of Euclidean translations in two dimensions. Deformation of the $SO(2,C)$ group into the $T(2)$ group will be accomplished by considering the little group of a point on a complex sphere with a radius depending on a real parameter τ, which is introduced in such a way that the radius of the complex sphere should tend to zero in the limit $\tau\to\infty$.

Orbits arising under the little group obtained for arbitrary values of the deformation parameter τ can be investigated in any space homogeneous under the $SL(2,C)$ group. Here we are concerned only with the orbits produced by the above little group on the two-dimensional complex sphere of nonzero radius. It is found that these orbits are determined by the intersections of the sphere with complex planes having a normal vector of length tending to zero in the limit of the horospheric subgroup. Thus, the vectors of the complex sphere of zero radius can be associated with the horospheres, and vice versa.

To be strict, the above planes become undefined when the horospheres degenerate into complex straight lines. These are the horospheres of the second kind, consisting of straight generators of the complex sphere. The

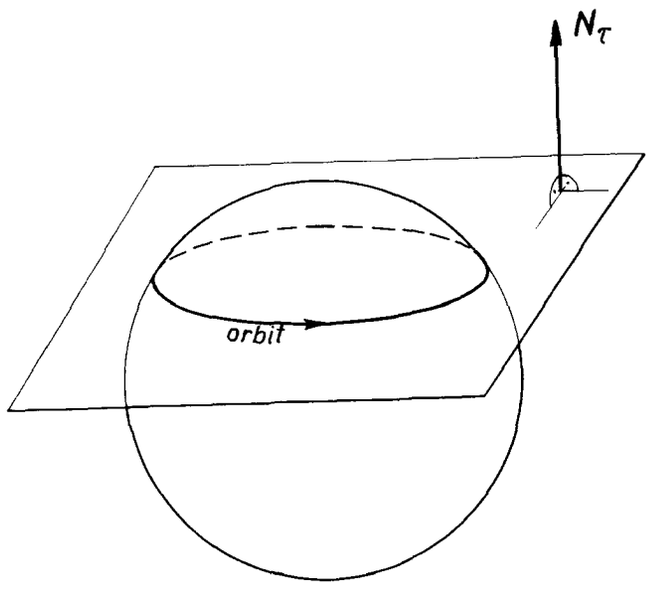


FIG. 1. Illustration of orbits on the complex sphere Σ. The orbit on the complex sphere Σ under the little group $H_\tau'(\varphi)$ is situated on the intersection of the complex sphere Σ and a complex plane with a normal vector $\mathbf{N}_\tau$. In the limit $\tau \to \infty$ the orbit approaches the horosphere, which cannot be visualized so simply, since in this case the normal becomes a complex vector of zero length (i.e., $\mathbf{N}_\infty^2 = 0$, though $\mathbf{N}_\infty \neq 0$).

existence of horospheres of the second kind is well known for the one-sheeted real hyperboloid,[1] and their appearance in the present case is not surprising, since the one-sheeted hyperboloid forms a subspace of the complex sphere.

In the last section the process of deformation is demonstrated on unitary representations. It is shown how the spherical functions defined on the familiar $\varphi(z, z^*)$ space, given in $SO(2,C)$ basis, can be deformed continuously into the spherical functions, in $T(2)$ basis, by tending the deformation parameter τ to infinity. The deformation formula obtained is a counterpart to the Wigner–Inonu contraction of the $SO(3)$ $[SO(2,1)]$ group into the $E(2)$ group.[3] We conclude with a formula relating the matrix elements of unitary representations of the $SL(2,C)$ group in $SO(2,C)$ basis to those in horospheric basis by means of a deformation process.

I. LITTLE GROUPS ON THE COMPLEX SPHERE OF ZERO AND NONZERO RADIUS

Homogeneous spaces are useful tools in representation theory, because by embedding a homogeneous space into the group G one obtains, on one hand, a subgroup H as a little group of a point of the homogeneous space and, on the other, an arrangement of the group into cosets with respect to the subgroup H. The most frequently used homogeneous spaces are those of the hyperboloid family yielding the subgroups $SO(3)$, $E(2)$, $SO(2,1)$ for $SL(2,C)$. Here we shall investigate the deformation of the two-dimensional complex rotation group isomorphic to the two-dimensional real rotation group with dilatations into the two-dimensional real rotation group with dilatations into the horospheric group isomorphic to the two-dimensional translation group $T(2)$. To this end a family of homogeneous spaces should be used the little groups of which are apt for demonstration of the deformation process.

Since the proper Lorentz group is isomorphic to the connected part of the three-dimensional complex rotation group,[6] the two-dimensional complex sphere

$$S_1^2 + S_2^2 + S_3^2 = S^2 \tag{1.1}$$

(hereafter Σ_s) forms a homogeneous space under the proper Lorentz group, as well as under $SL(2,C)$. The three-dimensional complex vector $\mathbf{S} = (S_1, S_2, S_3)$, which is the self-dual part of a Lorentz covariant antisymmetric tensor $S_{\alpha\beta}(\alpha, \beta = 0, 1, 2, 3)$ under $g \in SL(2,C)$, transforms as

$$\hat{S}' = g\hat{S}g^{-1}, \tag{1.2}$$

where $\hat{S} = \boldsymbol{\sigma}\mathbf{S} = \sigma_1 S_1 + \sigma_2 S_2 + \sigma_3 S_3$ and the σ_i stand for the Pauli matrices.[7]

We choose a standard vector on Σ_s

$$\mathbf{S}_0 = (0, 0, S). \tag{1.3}$$

Here S is supposed to be nonzero. The little group of this vector, that is, the subgroup of $SL(2,C)$ satisfying the condition

$$\hat{S}_0 = H_0 \hat{S}_0 H_0^{-1} \tag{1.4}$$

is clearly of the form

$$H_0(\varphi) = \exp(-i\varphi\sigma_3/2) = \begin{pmatrix} e^{-i\varphi/2} & 0 \\ 0 & e^{i\varphi/2} \end{pmatrix}, \tag{1.5}$$

where $\varphi = \varphi_1 + i\varphi_2$ is a complex angle, with a real part describing a rotation about the z axis and varying in the range $-2\pi \leq \varphi_1 < 2\pi$, and with an imaginary part describing a boost along the z axis and varying in the range $-\infty < \varphi_2 < \infty$. It follows that this group is $SO(2,C) = SO(2) \times SO(1,1)$.

In a similar way, by choosing the standard vector

$$\mathbf{S}_\infty = (S, iS, 0) \quad (S \neq 0) \tag{1.6}$$

on the complex sphere of zero radius Σ_0, we arrive at the horospheric little group isomorphic to the two-dimensional real Euclidian translation group $T(2)$ (cf. Ref. 8):

$$H_\infty(\varphi) = \exp(-i\varphi\sigma_+/2) = \begin{pmatrix} 1 & -i\varphi \\ 0 & 1 \end{pmatrix}, \tag{1.7}$$

where $\sigma_+ = \sigma_1 + i\sigma_2$ and $\varphi = \varphi_1 + i\varphi_2$. In the present case both φ_1 and φ_2 vary from $-\infty$ to ∞. It is easy to see the validity of the inverse statement; namely, if the little group of a three-dimensional complex vector is $T(2)$ $[SO(2,C)]$, then it is situated on the complex sphere of zero (nonzero) radius. Here and throughout this paper it is supposed that the point $\mathbf{S} = 0$ is excluded from Σ_0,

since this point is itself invariant under $SL(2,C)$ and thus the homogeneity of Σ_0 would be spoiled by including it.

Consider now the vector

$$\mathbf{S}_\tau = (\,[\tau/(1+\tau)]S,\ [i\tau/(1+\tau)]S,\ [1/(1+\tau)]S) \qquad (1.8)$$

interpolating between $\mathbf{S}_0$ and $\mathbf{S}_\infty$. Here τ is a real parameter describing the deformation varying in the range

$$0 \leq \tau < \infty.$$

The limits of the vector (1.8) as $\tau \to 0$ and $\tau \to \infty$ are $\mathbf{S}_0$ and $\mathbf{S}_\infty$, as given by Eqs. (1.3) and (1.6). Since the length of the vector $\mathbf{S}_\tau$

$$[(S_{\tau 1})^2 + (S_{\tau 2})^2 + (S_{\tau 3})^2]^{1/2} = S/(1+\tau) \qquad (1.9)$$

is nonzero for $\tau < \infty$, the little group of $\mathbf{S}_\tau$ is a $SO(2,C)$ group isomorphic to H_0. For $\tau \to \infty$ the little group H_∞, as given by Eq. (1.7), is obtained. By making use of Eqs. (1.2) and (1.8) we get an explicit form of the little group of $\mathbf{S}_\tau$ for an arbitrary value of τ:

$$H_\tau(\varphi) = \exp\left(-i\frac{\varphi}{2}(\sigma_3 + \tau\sigma_+)\frac{1}{1+\tau}\right) = \begin{pmatrix} \exp\left(-i\frac{1}{1+\tau}\frac{\varphi}{2}\right) & -2i\tau\sin\left(\frac{1}{1+\tau}\frac{\varphi}{2}\right) \\ 0 & \exp\left(i\frac{1}{1+\tau}\frac{\varphi}{2}\right) \end{pmatrix} \qquad (1.10)$$

In other words, this is the subgroup satisfying the equation

$$H_\tau(\varphi)\hat{S}_\tau H_\tau^{-1}(\varphi) = \hat{S}_\tau, \qquad (1.11)$$

with $\mathbf{S}_\tau$ given by (1.8). The range of $\varphi = \varphi_1 + i\varphi_2$ in this case is given by the inequalities

$$-2\pi(1+\tau) \leq \varphi_1 < 2\pi(1+\tau), \quad -\infty < \varphi_2 < \infty.$$

In the next section we proceed to an investigation of orbits generated by the above subgroup in the space of complex vectors. In particular, we are interested in the orbits as $\tau \to \infty$.

II. ORBITS ON THE COMPLEX SPHERE

According to Eq. (1.9) the final point of the vector $\mathbf{S}_\tau$ is situated on the complex sphere $\Sigma_{S/(1+\tau)}$ of radius $S/(1+\tau)$, which is nonzero for finite τ but tends to zero as $\tau \to \infty$. In either case, $\mathbf{S}_\tau$ has the little group as given by Eq. (1.10).

Let us fix the value of τ for the time being and see what little group $H'_\tau(\varphi)$ is obtained if another standard vector of the same length is chosen instead of $\mathbf{S}_\tau$. The answer is trivial, since, as a consequence of the homogeneity of $\Sigma_{S/(1+\tau)}$, there exists a $g \in SL(2,C)$ which translates $\mathbf{S}_\tau$ into $\mathbf{S}'_\tau$:

$$\hat{S}'_\tau = g\hat{S}_\tau g^{-1}. \qquad (2.1)$$

It follows then from Eq. (1.11) that

$$H'_\tau(\varphi)\,\hat{S}'_\tau H'_\tau(\varphi)^{-1} = \hat{S}'_\tau, \qquad (2.2)$$

where

$$H'_\tau(\varphi) = gH_\tau(\varphi)\,g^{-1}. \qquad (2.3)$$

It is obvious that $H'_\tau(\varphi)$ is isomorphic to $SO(2,C)$ when $\tau < \infty$, and isomorphic to $T(2)$ as $\tau \to \infty$.

Now, we are interested in orbits of a complex vector $\mathbf{S}$ under the group (2.3). It is supposed that the final point of $\mathbf{S}$ is situated on a complex sphere Σ of nonzero radius [not to be confused with the sphere $\Sigma_{S/(1+\tau)}$]. Under the group $H'_\tau(\varphi)$ the vector $\mathbf{S}$ describes the orbit

$$\hat{S}_\tau(\varphi) = H'_\tau(\varphi)\hat{S}H'_\tau(\varphi)^{-1}. \qquad (2.4)$$

Since $H'_\tau(\varphi) \subset SL(2,C)$ the orbit is obviously situated on the sphere Σ. Moreover, it will be verified that the orbit lies in a complex plane $\mathbf{S}_\tau(\varphi)\mathbf{N}_\tau = C_\tau = \text{const.}$, where $\mathbf{N}_\tau$ proves to be identical with the $\mathbf{S}'_\tau$ obtained previously in Eq. (2.1). Indeed, it follows from Eqs. (2.3) and (2.4) that

$$\mathbf{S}_\tau(\varphi)\mathbf{S}'_\tau = \tfrac{1}{2}\mathrm{Tr}[\hat{S}_\tau(\varphi)\hat{S}'_\tau] = \tfrac{1}{2}\mathrm{Tr}[H'_\tau(\varphi)\hat{S}H'_\tau(\varphi)^{-1}\hat{S}'_\tau]$$

$$= \tfrac{1}{2}\mathrm{Tr}[\hat{S}H'_\tau(\varphi)^{-1}\hat{S}'_\tau H'_\tau(\varphi)] = \tfrac{1}{2}\mathrm{Tr}(\hat{S}\hat{S}'_\tau) = \mathbf{S}\mathbf{S}'_\tau = C_\tau = \text{const.} \qquad (2.5)$$

According to this equation one can associate with each orbit generated by the little group $H'_\tau(\varphi)$ a normal vector $\hat{S}'_\tau$. The orbit can be given by the homogeneous coordinates $(\mathbf{S}'_\tau, C_\tau)$; nevertheless, apart from the singular case $C_\tau = 0$, one can normalize C_τ to 1 by an appropriate dilatation of $\mathbf{S}'_\tau$.

As the above statements are independent of the value of τ, we can take the limit $\tau \to \infty$, which produces horospheres. So, according to (2.3) and (2.4), horospheres on Σ are orbits described by the horospheric subgroup

$$H'_\infty(\varphi) = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\begin{pmatrix} 1 & -i\varphi \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}^{-1} = gH_\infty g^{-1} \quad (\alpha\delta - \gamma\beta = 1)$$

for fixed g. Taking into account Eq. (1.9) and the fact that transformation (2.1) leaves the length of $\mathbf{S}_\tau$ unchanged, we get for $\tau \to \infty$

$$(S'_{\infty 1})^2 + (S'_{\infty 2})^2 + (S'_{\infty 3})^2 = 0.$$

Thus, when the $SO(2,C)$ group deforms into the horospheric group as $\tau \to \infty$, the normal vector $\mathbf{N}_\tau = \mathbf{S}'_\tau$ characteristic for the orbits arrives at the complex sphere of zero radius; that is, $\mathbf{N}_\tau^2 = \mathbf{S}_\infty'^2 = 0$. Horospheres of Σ are therefore determined by the equation $\mathbf{S}\mathbf{N}_\infty = 1$, where $\mathbf{N}_\infty^2 = 0$.[9]

At this point a remark is in order. We did not investigate the question whether, in Eq. (2.5), the normal vector of the plane of horospheres is unique up to a factor. From a more detailed investigation, which for the sake of brevity is left to the reader, the following can be shown. A single fixed point $\mathbf{S}$ of the space Σ is crossed by a one-parametric manifold of horospheres. These are second-order curves, which, generally speaking, unambiguously define a plane with a normal of zero length, as indicated above. However, in the manifold of horospheres crossing a fixed point there are two positions where the horosphere degenerates into a complex straight line. These lines can be given in the form

$$\mathbf{S}_\infty(\varphi) = \mathbf{A}\varphi + \mathbf{S}, \quad \mathbf{S}_\infty(\varphi) = \mathbf{B}\varphi + \mathbf{S},$$

with

$$\mathbf{A} = (A_1 + iA_2,\ A_1 - iA_2,\ A_3) = (S[S_+^2/(S+S_3)^2],\ -S,\ S[S_+/(S+S_3)]),$$

$$\mathbf{B}=(B_1+iB_2,\ B_1-iB_2,\ B_3)=(-S[S_+^2/(S-S_3)^2],\ \times S,\ S[S_+/(S-S_3)]),\quad (S\pm S_3\neq 0),$$

where the usual notation $S_\pm=S_1\pm iS_2$ is used. Each point **S** is thus crossed by two straight horospheres, that are determined by the position of **S** alone. These horospheres can be called horospheres of the second kind, as distinguished from those of the first kind, which are in one-to-one correspondence with the vectors of the sphere of zero radius. Having obtained the horospheres a horosphere transformation between functions on the complex sphere of nonzero and zero radius could be defined but the investigation of this problem goes beyond the scope of present paper.

III. DEFORMATION OF UNITARY REPRESENTATIONS

To demonstrate the deformation on unitary representations, let us consider the linear fractional mapping of the z plane which is a factor space $SL(2,C)/\left(\begin{smallmatrix}\alpha & \beta\\ 0 & \alpha^{-1}\end{smallmatrix}\right)$:

$$z'=(\alpha z+\beta)/(\gamma z+\delta).$$

In the case of the $SO(2,C)$ subgroup given by (1.5) this reduces to

$$z'=e^{-i\varphi}z=e^{\varphi_2}e^{-i\varphi_1}z \tag{3.1}$$

which is a rotation followed by a dilatation. In a similar way, the horospheric transformation on the z plane takes the form of an Euclidian displacement

$$z'=z-i\varphi=z+(\varphi_2-i\varphi_1).$$

The interpolating subgroup given by Eq. (1.10) accomplishes a transformation similar to that of (3.1) on the displaced z plane, i.e.,

$$z'+\tau=\exp\{-i[\varphi/(1+\tau)]\}(z+\tau).$$

For realizing unitary representations the representation on the familiar $\varphi(z,z^*)$ functions will be used.[1] Action of an element $g=\left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)\in SL(2,C)$ on these functions is defined as

$$T_g\varphi(z)=(-\gamma z+\alpha)^{2j}(-\gamma^*z^*+\alpha^*)^{2k}\times\varphi\left(\frac{\delta z-\beta}{-\gamma z+\alpha},\ \frac{\delta^*z^*-\beta^*}{-\gamma^*z^*+\alpha^*}\right), \tag{3.2}$$

where

$$j=\tfrac{1}{2}(j_0-1+i\sigma),\quad k=\tfrac{1}{2}(-j_0-1+i\sigma). \tag{3.3}$$

Here j_0 takes integer and half-integer values, while σ is an arbitrary complex number. In what follows we restrict ourselves to the principal series of unitary representations for which σ is real. In Eq. (3.2) representations are defined by displacement from the left, which results in the following form of infinitesimal generators:

$$J_+=J_1+iJ_2=-\frac{\partial}{\partial z},\quad K_+=K_1+iK_2=2kz^*-z^{*2}\frac{\partial}{\partial z^*},$$

$$J_-=J_1-iJ_2=-2jz+z^2\frac{\partial}{\partial z},\quad K_-=K_1-iK_2=\frac{\partial}{\partial z^*},$$

$$J_3=j-z\frac{\partial}{\partial z},\quad K_3=-k+z^*\frac{\partial}{\partial z^*}. \tag{3.4}$$

These generators are related to the generators of spatial rotations about kth axis, M_k $(k=1,\ 2,\ 3)$, and to the generators of boosts along kth axis, N_k, as $J_k=\frac{1}{2}(M_k+iN_k)$, $K_k=\frac{1}{2}(M_k-iN_k)$.

Spherical functions in $SO(2,C)$ basis satisfy the eigenvalue equations

$$J_3\varphi^0_{mm^*}(z,z^*)=m\varphi^0_{mm^*}(z,z^*),\qquad K_3\varphi^0_{mm^*}(z,z^*)=m^*\varphi^0_{mm^*}(z,z^*), \tag{3.5}$$

where

$$m=\tfrac{1}{2}(\mu+i\nu),\quad m^*=\tfrac{1}{2}(\mu-i\nu), \tag{3.6}$$

with $\mu=0,\ \pm\frac{1}{2},\ \pm1,\ \ldots$ and ν continuous. The above basis is a generalization of finite dimensional spinors to the unitary case, where m and m^* correspond to undotted and dotted indices of spinors. Unitary spinors can be successfully applied to the evaluation of matrix elements of unitary representations of the Lorentz group, simplifying to a considerable extent the results obtained in angular momentum basis.[10–13]

Using the explicit form of the infinitesimal generators given by Eq. (3.4), the solution of (3.5) can be written as

$$\varphi^0_{mm^*}(z,z^*)=(2\pi)^{-1}z^{j-m}z^{*k+m^*}. \tag{3.7}$$

The requirement of single-valuedness on the complex z plane yields the condition $2\mu=$ integer, or to be strict, μ takes integer and half-integer values along with j_0. The functions (3.7) are normalized as

$$(\varphi^0_{m'm'^*},\ \varphi^0_{mm^*})=\tfrac{1}{2}i\int dz dz^*\varphi^0_{m'm'^*}(z,z^*)^*\,\varphi^0_{mm^*}(z,z^*)=\delta_{\mu'\mu}\delta(\nu'-\nu).$$

The horospheric group as given by Eq. (1.7) is generated by the Hermitian generators M_1-N_2 and M_2+N_1 or—equivalently—by the non-Hermitian generators J_+ and K_-. In this basis spherical functions are solutions of the eigenvalue equations

$$J_+\varphi^\infty_{mm^*}(z,z^*)=m\varphi^\infty_{mm^*}(z,z^*),\qquad K_-\varphi^\infty_{mm^*}(z,z^*)=m^*\varphi^\infty_{mm^*}(z,z^*), \tag{3.8}$$

where m, m^* are again written in the form (3.6) but now, due to the infinite range of the parameters φ_1 and φ_2, both μ and ν take continuous values.

The solution of (3.8) assumes the form of two-dimensional "plane waves"

$$\varphi^\infty_{mm^*}(z,z^*)=(2\pi)^{-1}\exp(-mz+m^*z^*). \tag{3.9}$$

In the present case the normalization assumes the form

$$(\varphi^\infty_{m'm'^*},\varphi^\infty_{mm^*})=\delta(\mu'-\mu)\,\delta(\nu'-\nu).$$

In case of the interpolating subgroup (1.10) the eigenvalue equation reads

$$\{(1+\tau)^{-1}J_3+[\tau/(1+\tau)]J_+\}\varphi^\tau_{mm^*}(z,z^*)=m\varphi^\tau_{mm^*}(z,z^*),\qquad \{(1+\tau)^{-1}K_3+[\tau/(1+\tau)]K_-\}\varphi^\tau_{mm^*}(z,z^*)=m^*\varphi^\tau_{mm^*}(z,z^*). \tag{3.10}$$

Solutions are as follows:

$$\varphi^\tau_{mm^*}(z,z^*)=(2\pi)^{-1}\eta_\tau(m)(z+\tau)^{j-m(1+\tau)}(z^*+\tau)^{k+m^*(1+\tau)}, \tag{3.11}$$

where

$$\eta_\tau(m)=\exp\{i[\nu\tau\ln\tau+(\nu-\sigma)\ln(1+\tau)]\}(1+\tau). \quad (3.12)$$

The parameters σ, ν were defined in Eqs. (3.3) and (3.6).

Functions (3.11) have a branch point at $z=-\tau$. The requirement of single-valuedness yields the quantization $\mu(1+\tau)=$ integer or half-integer along with j_0. As τ grows this quantization imposes a weaker and weaker condition on μ, which is *supposed to be fixed*, until in the limit $\tau\to\infty$, μ becomes continuous. Functions (3.11) are normalized as

$$(\varphi^\tau_{m'm'*},\ \varphi^\tau_{mm*})=\Delta_\tau(\mu',\mu)\,\delta(\nu'-\nu),$$

with

$$\Delta_\tau(\mu';\mu)=\frac{\sin[\pi(1+\tau)(\mu'-\mu)]}{\pi(\mu'-\mu)}=\begin{cases}\delta_{\mu'\mu} & \text{if } \tau=0,\\ (1+\tau)\delta_{\mu'\mu} & \text{if } 0<\tau<\infty,\\ \delta(\mu'-\mu) & \text{if } \tau\to\infty.\end{cases}$$

Here the quantization of μ has been taken into account. It is easy to verify that spherical functions φ^τ_{mm*} approach those given in $SO(2,C)$ $[T(2)]$ basis as $\tau\to 0$ $(\tau\to\infty)$; that is

$$\lim_{\tau\to 0}\varphi^\tau_{mm*}(z,z^*)=(2\pi)^{-1}z^{j-m}z^{*k+m*}$$

and

$$\lim_{\tau\to\infty}\varphi^\tau_{mm*}(z,z^*)=(2\pi^{-1})\exp(-mz+m^*z^*).$$

The phase factor in Eq. (3.12) was chosen to fit the phases of (3.7) and (3.9) at $\tau\to 0$ and $\tau\to\infty$.

It is deduced from (3.11) that "plane waves" of horospheric basis (3.9) can be obtained from spherical functions given in unitary spinor basis by means of the following deformation procedure. Labels m and m^* are to be substituted by $m_\tau\equiv m(1+\tau)$ and $m^*_\tau\equiv m^*(1+\tau)$, respectively. Afterwards, a horospheric displacement is to be performed by the special element $h_\tau=\left(\begin{smallmatrix}1 & -\tau\\ 0 & 1\end{smallmatrix}\right)$ as indicated by (3.2). Then, following a multiplication by the phase $\eta_\tau(m)$, spherical functions $\varphi^\infty_{mm*}(z,z^*)$ are obtained by means of the limit $\tau\to\infty$:

$$\varphi^\infty_{mm*}(z,z^*)=\lim_{\tau\to\infty}\eta_\tau(m)T_{h_\tau}\varphi^0_{m_\tau m^*_\tau}(z,z^*), \quad (3.13)$$

i.e.,

$$(2\pi)^{-1}\exp(-mz+m^*z^*)=\lim_{\tau\to\infty}\eta_\tau(m)T_{h_\tau}(2\pi)^{-1}z^{j-m_\tau}z^{*k+m^*_\tau}$$

This limit is a counterpart to the Wigner—Inonu contraction of Legendre polynomials into Bessel functions.[2] It permits us to relate matrix elements of unitary representations of the $SL(2,C)$ group in horospheric basis to those given in $SO(2,C)$ basis by means of a deformation. Owing to the unitarity of the operator T_{h_τ}, we obtain from Eq. (3.13)

$$(\varphi^\infty_{m'm'*},T_g\varphi^\infty_{mm*})=\lim_{\tau\to\infty}\eta_\tau(m')^*\eta_\tau(m)(T_{h_\tau}\varphi^0_{m'_\tau},T_gT_{h_\tau}\varphi^0_{m_\tau})$$

$$=\lim_{\tau\to\infty}\eta_\tau(m')^*\eta_\tau(m)(\varphi^0_{m'_\tau},T^{-1}_{h_\tau}T_gT_{h_\tau}\varphi^0_{m_\tau})$$

$$=\lim_{\tau\to\infty}\eta_\tau(m')^*\eta_\tau(m)(\varphi^0_{m'_\tau},T_{g_\tau}\varphi^0_{m_\tau}),$$

where $g\in SL(2,C)$ and $g_\tau=h_\tau^{-1}gh_\tau$. Action of T_g, the representation of the $SL(2,C)$ group, on functions $\varphi(z,z^*)$ is given by Eq. (3.2). The principal series of unitary representations of the $SL(2,C)$ group is known in unitary spinor basis,[11] i.e., the matrix elements $(\varphi^0_m,T_g\varphi^0_m)$ are expressed in terms of special functions. With the aid of the above formula the matrix elements of unitary representations of the $SL(2,C)$ group can be obtained in the basis defined by the horosheric subgroup

$$\begin{matrix}1 & \beta\\ 0 & 1\end{matrix}.$$

ACKNOWLEDGMENTS

The author expresses his gratitude to A. Sebestyén and K. Szegö for their critical remarks.

[1] I. M. Gelfand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5.
[2] J. D. Bjorken, Phys. Rev. **179**, 1547 (1969); R. P. Feynman, Phys. Rev. Lett. **23**, 1415 (1969).
[3] E. Inonu and E. P. Wigner, Proc. Natl. Acad. Sci. USA **39**, 510 (1953).
[4] R. Hermann, *Fourier Analysis on Groups and Partial Wave Analysis* (Benjamin, New York, 1969), p. 35–47.
[5] K. Szegö and K. Tóth, J. Math. Phys. **12**, 486 (1971); J. Math. Phys. **12**, 853 (1971).
[6] A. J. Macfarlane, J. Math. Phys. **3**, 1116 (1962).
[7] Under proper Lorentz transformations, a complex vector **S** transforms like **B** + *i***E**, where **B** and **E** are the magnetic and electric field strengths. The invariance of the square $(\mathbf{B}+i\mathbf{E})^2=\mathbf{B}^2-\mathbf{E}^2+2i\mathbf{BE}$ is also well known from electrodynamics.
[8] M. Huszár, Commun. Math. Phys. **23**, 132 (1971).
[9] The real and imaginary parts of complex vector on the complex vector on the complex sphere of zero radius are quantities analogous to the field strengths **B** and **E** of an electromagnetic plane wave, where **B** and **E** are of the same modulus and are perpendicular to each other $\mathbf{B}^2-\mathbf{E}^2=0$, $\mathbf{BE}=0$.
[10] H. Joos and R. Schrader, DESY Preprint 68/40, Hamburg (1968).
[11] M. Huszár and J. Smorodinsky, JINR Preprint E2—4225 (1968). Theor. Math. Phys. 4, 328 (1970).
[12] M. Carmeli, J. Math. Phys. **11**, 1917 (1970).
[13] M. Carmeli and S. Malin, J. Math. Phys. **12**, 225 (1971).

Lattice Green's function for the diced lattice

T. Horiguchi and C. C. Chen

Department of Physics, Ohio University, Athens, Ohio 45701
(Received 14 January 1974)

It is shown that the lattice Green's function for the diced lattice is expressed in terms of the one for the triangular lattice. The distribution function of the squared frequency is shown by a graph and compared with the one for the honeycomb lattice.

The frequency distribution function for the honeycomb lattice was obtained by Hobson and Nierenberg[1] and by Huckaby.[2] The frequency distribution function is related to the imaginary part of the lattice Green's function at the origin. The lattice Green's function[3] for the honeycomb lattice was shown to be expressed in terms of the one for the triangular lattice which was exactly solved in terms of the complete elliptic integrals of the first and second kind at an arbitrary lattice site.[4] The arrangement of lattice sites in the diced lattice, the network of which is shown in Fig. 1, is the same as that of the triangular lattice, but the path of the interaction among the lattice sites is rather similar to the one for the honeycomb lattice. In this short note, the lattice Green's function for the diced lattice is shown to be expressed in terms of the one for the triangular lattice.

As shown in Fig. 1, the three kinds of the sublattices are introduced. Each of them constitutes a triangular lattice. When the initial lattice site belongs to the sublattice α (α takes one of A, B, or C sublattice) the lattice Green's function is the solution of the following set of three difference equations:

$$\begin{aligned} t\,G^{A\alpha}(la,mb) &- \tfrac{1}{2}[G^{B\alpha}(la+a,mb+b)+G^{B\alpha}(la-a,mb+b) \\ &+ G^{B\alpha}(la,mb-2b)] \\ &- (\gamma/2)[G^{C\alpha}(la+a,mb-b) \\ &+ G^{C\alpha}(la-a,mb-b)+G^{C\alpha}(la,mb+2b)] \\ &= \delta_{l,0}\delta_{m,0}\delta_{A,\alpha}, \end{aligned} \tag{1}$$

$$\begin{aligned} t\,G^{B\alpha}(la,mb) &- \tfrac{1}{2}[G^{A\alpha}(la+a,mb-b)+G^{A\alpha}(la-a,mb-b) \\ &+ G^{A\alpha}(la,mb+2b)] \\ &= \delta_{l,0}\,\delta_{m,0}\delta_{B,\alpha}, \end{aligned} \tag{2}$$

$$\begin{aligned} t\,G^{C\alpha}(la,mb) &- (\gamma/2)[G^{A\alpha}(la+a,mb+b)+G^{A\alpha}(la-a,mb+b) \\ &+ G^{A\alpha}(la,mb-2b)] \\ &= \delta_{l,0}\delta_{m,0}\delta_{C,\alpha}, \end{aligned} \tag{3}$$

where $la=(\mathbf{r}_f-\mathbf{r}_i)_x$ and $mb=(\mathbf{r}_f-\mathbf{r}_i)_y$, and $\mathbf{r}_i$ and $\mathbf{r}_f$ are the initial and final lattice sites, respectively. $a=\sqrt{3}/2$ and $b=1/2$ if the nearest neighbor distance is chosen to be equal to 1. γ is the interaction parameter between A and C sublattices when the one between A and B sublattices is put equal to 1. t is the parameter corresponding to the squared frequency in the lattice vibration or to the energy of the electron in the tight binding approximation. t is assumed to be a complex variable in this note. The solution of these difference equations under the boundary condition that $G^{\alpha\beta}(la,mb)$ is equal to zero as l^2+m^2 tends to infinity, is satisfied by a certain complicated integral. From that integral, one can see that the parameters a and b may be deleted without loss of generality; we use the notation $G^{\alpha\beta}(l,m)$ instead of $G^{\alpha\beta}(la,mb)$.

After manipulating the integral obtained from Eqs. (1)–(3), we arrive at the following expressions:

$$G^{AA}(l,m)=\frac{2t}{1+\gamma^2}\,G_t\left(\tau;\,l,\,\frac{m}{3}\right), \tag{4}$$

$$G^{BB}(l,m)=\frac{\gamma^2}{t(1+\gamma^2)}\delta_{l,0}\,\delta_{m,0}+\frac{2t}{(1+\gamma^2)^2}G_t\left(\tau;\,l,\,\frac{m}{3}\right), \tag{5}$$

$$G^{CC}(l,m)=\frac{1}{t(1+\gamma^2)}\,\delta_{l,0}\,\delta_{m,0}+\frac{2t\gamma^2}{(1+\gamma^2)^2}\,G_t\left(\tau;\,l,\,\frac{m}{3}\right), \tag{6}$$

$$\begin{aligned} G^{AB}(l,m)=\frac{1}{1+\gamma^2}\,[&G_t(\tau;l+1,\tfrac{1}{3}(m+1)) \\ &+G_t(\tau;l-1,\tfrac{1}{3}(m+1))+G_t(\tau;l,\tfrac{1}{3}(m-2))], \end{aligned} \tag{7}$$

$$G^{BA}(l,m)=G^{AB}(-l,-m), \tag{8}$$

$$G^{AC}(l,m)=\gamma G^{BA}(l,m), \tag{9}$$

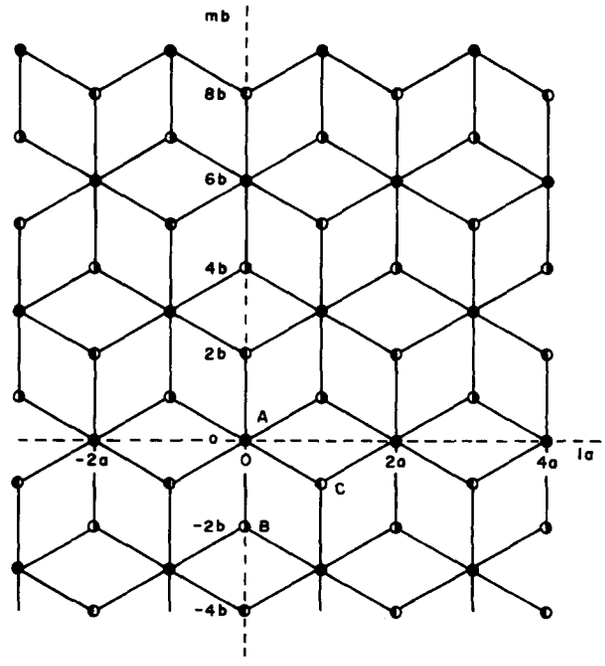


FIG. 1. The network of the diced lattice.

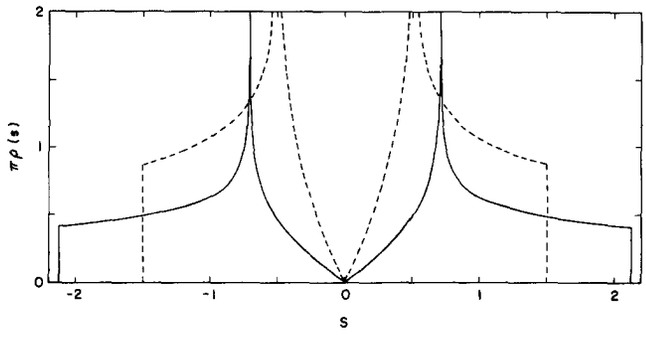


FIG. 2. The distribution functions of the squared frequency for the diced lattice (solid line) and for the honeycomb lattice (dotted line).

$$G^{CA}(l,m)=G^{AC}(-l,-m), \tag{10}$$

$$\begin{aligned}G^{BC}(l,m)=\frac{\gamma}{2t(1+\gamma^2)}\,[&G_t(\tau;l+2,\ \tfrac{1}{3}(m-2))\\&+2G_t(\tau;l+1,\ \tfrac{1}{3}(m+1))\\&+G_t(\tau;l,\ \tfrac{1}{3}(m+4))+2G_t(\tau;l,\ \tfrac{1}{3}(m-2))\\&+2G_t(\tau;l-1,\ \tfrac{1}{3}(m+1))\\&+G_t(\tau;l-2,\ \tfrac{1}{3}(m-2))],\end{aligned} \tag{11}$$

$$G^{CB}(l,m)=G^{BC}(-l,-m), \tag{12}$$

where $\tau=[4t^2-3(1+\gamma^2)]/2(1+\gamma^2)$. $G_t(\tau;l,m)$ is the lattice Green's function for the triangular lattice which is given by

$$G_t(\tau;l,m)=\frac{1}{\pi^2}\int_0^{\pi}dx\int_0^{\pi}dy\,\frac{\cos lx\cos my}{\tau-\cos 2x-2\cos x\cos y}. \tag{13}$$

$G_t(\tau;l,m)$ is known to be expressed in terms of the complete elliptic integrals of the first and second kinds at an arbitrary lattice site.[4] Then the lattice Green's function for the diced lattice at an arbitrary lattice site can be obtained from the knowledge of that for the triangular lattice. These expressions (4)—(12) can be reduced to the ones for the honeycomb lattice by putting $\gamma=0$.

When $t=s-i\epsilon$ where ϵ is an infinitesimal positive number, the distribution function of the squared frequency is obtained by

$$\rho(s)=\frac{1}{3\pi}\mathrm{Im}[G^{AA}(0,0)+G^{BB}(0,0)+G^{CC}(0,0)] \tag{14}$$

$$=\frac{4s}{3\pi(1+\gamma^2)}\mathrm{Im}\,G_t(\tau;0,0). \tag{15}$$

For $\gamma=1$, $\rho(s)$ is simply expressed in terms of the complete elliptic integral of the first kind as follows:

$$\rho(s)=\begin{cases}\dfrac{s^{1/2}2^{3/4}}{3\pi^2}K'\left(\dfrac{1}{k}\right) & \text{for } 1/\sqrt{2}\leqslant|s|\leqslant 3/\sqrt{2},\\[2ex] \dfrac{sg_I}{3\pi^2(1+k_I^2)^{1/2}}K\left(\dfrac{k_I}{(1+k_I^2)^{1/2}}\right) & \text{for } 0\leqslant|s|\leqslant 1/\sqrt{2},\end{cases} \tag{16, 17}$$

where

$$k=\frac{4\cdot 2^{1/4}\cdot s^{1/2}}{(2^{1/2}s-1)^{3/2}(2^{1/2}s+3)^{1/2}}, \tag{18}$$

$$k_I=\frac{4\cdot 2^{1/4}\cdot s^{1/2}}{(1-2^{1/2}s)^{3/2}(2^{1/2}s+3)^{1/2}}, \tag{19}$$

$$g_I=\frac{8}{(1-2^{1/2}s)^{3/2}(2^{1/2}s+3)^{1/2}}. \tag{20}$$

The curve for the distribution function of the squared frequency for the diced lattice with $\gamma=1$ is shown in Fig. 2 together with the one for the honeycomb lattice. When the sublattice C is incorporated into the sublattice B, the diced lattice is then composed of the two sublattices A and B, and the interactions exist only between the lattice sites belonging to different sublattices. Thus the diced lattice is one of the "alternating lattice."[5] The symmetry properties of the lattice Green's function for the alternating lattice have been studied in Ref. 5 and $\rho(s)=\rho(-s)$ is given as their special case.

[1]J. P. Hobson and W. A. Nierenberg, Phys. Rev. **89**, 662 (1953).
[2]D. A. Huckaby, J. Chem. Phys. **56**, 198 (1972).
[3]S. Katsura *et al.*, J. Math. Phys. **12**, 892 (1971).
[4]T. Horiguchi, J. Math. Phys. **13**, 1411 (1972).
[5]T. Morita and T. Horiguchi, J. Math. Phys. **13**, 1243 (1972).

Errata: Elastic general relativistic systems [J. Math. Phys. 13, 1934 (1972)]

E. N. Glass

Princeton University, Princeton, New Jersey 08540

J. Winicour

University of Pittsburgh, Pittsburgh, Pennsylvania 15260
(Received 19 February 1974)

Add the term $-P_{\alpha\beta}S$ to the right-hand sides of Eqs. (2.11), (3.25), (4.25), and (4.26).

On the left-hand side of Eq. (A14), add the term $\rho a^{\mu}S$ and delete $-P^{\alpha\beta}S_{;\alpha}\delta^{\mu}_{\beta}$ from the last term.

These corrections do not alter the fundamental theory or any major results, but they do modify the equations of motion.

Errata: A geometric generalization of Hooke's law [J. Math. Phys. 14, 1285 (1973)]

E. N. Glass

Princeton University, Princeton, New Jersey 08540

J. Winicour

University of Pittsburgh, Pittsburgh, Pennsylvania 15260
(Received 19 February 1974)

Equation (3.20) should read

$$\bar{T}^{ij} = DT^{ij} + g^{1/2}C^{ijkl}\eta_{kl} \tag{3.20}$$

where the Cauchy stress tensor density T^{ij} is defined in Eq. (3.18). The change in Hook's law from tensors to tensor densities adds the term $-P^{ij}\xi^{k}{}_{;k}$ to the right hand sides of Eqs. (3.23) and (3.24). As a check on these equations, note that $(\bar{P}^{ij} - P^{ij})\eta_{ij}$ is an exact differential in accord with the first law of thermodynamics.

The relationship between the perturbed stress tensor and tensor density below Eq. (3.25) should read

$$\bar{T}^{ij} = g^{1/2}\bar{P}^{ij},$$

since the deformation leaves the underlying geometry unchanged.

On the right-hand side of Eq. (4.6), the second term should read $P^{jk}\xi^{i}{}_{;kj}$ and the term $B^{i}\xi^{k}{}_{;k}$ should be added.

Delete the second term on the right-hand side of Eq. (4.7). Equation (4.9) should read

$$\rho\ddot{\xi}^{i} = [C^{ijkl}\xi_{(k;l)}]_{;j} - P\xi^{i;j}{}_{;j}. \tag{4.9}$$

On the right-hand side of Eq. (4.13), delete $\Theta^{;i}$ from the third term and delete the fourth term.

Equation (4.14) should read

$$\rho\ddot{\Theta} = (\lambda + 2\mu - P)\nabla^{2}\Theta - P_{;i}(\Theta^{;i} + \nabla^{2}\xi^{i}). \tag{4.14}$$

Erratum and Addendum: Unidirectional energy transfer in nonlinear wave–wave interactions [J. Math. Phys. 14, 911 (1973)]

P. K. C. Wang

School of Engineering and Applied Science, University of California, Los Angeles, California

(Received 27 November 1973)

It was brought to my attention by Professor L. Stenflo that the coupling coefficient μ_0^* [Eq. (34)] as given in Ref. 9 is in error. Consequently, the implications of condition (9) for unidirectional energy transfer in the magnetized electron plasma under consideration are altered. The correct expression for μ_0^* is given by

$$\mu_0^* = (1 + S_0S_1)\,\frac{jk_2u^2\omega_p^2\omega_0}{8N_0\omega_2^2}\,\frac{[k_2 - \omega_c(S_0k_0/\omega_0 - S_1k_1/\omega_1)]}{[(\omega_1 - S_1\omega_c)^2 + S_1\omega_c\omega_p^2/2\omega_1]}.$$

Straightforward calculations show that for $S_0 = S_1 = 1$, $\mathrm{Re}\{\mu_i^*\mu_k\} \leq 0$ for all $i \neq k$ except $\mathrm{Re}\{\mu_1^*\mu_2\} = \mathrm{Re}\{\mu_2^*\mu_1\} \geq 0$. Therefore, for right-hand polarized transverse waves condition (9) is not satisfied. For the case where $S_0 = S_1 = -1$, it can be readily deduced that

$$\text{(i)}\quad \mathrm{Re}\{\mu_0^*\mu_j\} > 0 \quad \text{for } j = 1, 2 \text{ if}$$
$$\mathrm{sgn}[(\omega_1 + \omega_c)^2 - \omega_c\omega_p^2/2\omega_1] = -\,\mathrm{sgn}[(\omega_0 + \omega_c)^2 - \omega_c\omega_p^2/2\omega_0]; \tag{45}$$

$$\text{(ii)}\quad \mathrm{Re}\{\mu_i^*\mu_j\} > 0 \quad \text{for } i = 1, 2, \quad j = 0, 1, 2 \text{ if}$$
$$\omega_0(\omega_0 + \omega_c)^2 < \omega_c\omega_p^2/2 < \omega_1(\omega_1 + \omega_c)^2. \tag{46}$$

Thus, condition (9) for unidirectional energy transfer is satisfied under condition (45) or (46). However, in view of the propagation condition for left-hand polarized transverse electromagnetic waves in a magnetized electron plasma given by

$$\omega_j(\omega_j + \omega_c) > \omega_p^2, \quad j = 0, 1, \tag{47}$$

it is evident that (47) is not compatible with (45) or (46) for positive ω_0 and ω_1.

In the sequel, we shall give examples of plasmas for which condition (9) has a simple interpretation. First, consider the nonlinear interaction of three longitudinal electromagnetic waves (ω_i, k_i), $i = 0, 1, 2$, in a magnetized collisionless plasma. Using the linearized Vlasov equation, Stenflo[12] showed that under the matching conditions $k_2 = k_0 - k_1$ and $\omega_2 = \omega_0 - \omega_1$, the equations governing the complex wave amplitudes have the form (1) with coupling coefficients μ_i given by

$$\mu_0 = C_L\left(\frac{\partial\epsilon_x}{\partial\omega}\Big|_{(\omega_0,k_0)}\right)^{-1}, \quad \mu_1 = -\,C_L\left(\frac{\partial\epsilon_x}{\partial\omega}\Big|_{(\omega_1,k_1)}\right)^{-1},$$
$$\mu_2 = -\,C_L\left(\frac{\partial\epsilon_x}{\partial\omega}\Big|_{(\omega_2,k_2)}\right)^{-1}, \tag{48}$$

where

$$\epsilon_x(\omega, k) = 1 + \frac{q^2}{\epsilon_0 mk}\int \frac{\partial F_0/\partial v_x}{(\omega - kv_x)}\,d\mathbf{v},$$
$$C_L = -\,\frac{q^3}{\epsilon_0 m^3}\int \frac{\partial F_0}{\partial v_x}\left(\prod_{i=0}^{2}(\omega_i - k_iv_x)\right)^{-1} d\mathbf{v}, \tag{49}$$

where $F_0 = F_0(v_x, v_y^2 + v_z^2)$ is the equilibrium velocity distribution function.

For the zeroth wave, condition (9) requires

$$\mathrm{Re}\{\mu_0^*\mu_1\} = -\,C_L^2\left[\left(\frac{\partial\epsilon_x}{\partial\omega}\Big|_{(\omega_0,k_0)}\right)\left(\frac{\partial\epsilon_x}{\partial\omega}\Big|_{(\omega_1,k_1)}\right)\right]^{-1} \geq 0,$$
$$\mathrm{Re}\{\mu_0^*\mu_2\} = C_L^2\left[\left(\frac{\partial\epsilon_x}{\partial\omega}\Big|_{(\omega_0,k_0)}\right)\left(\frac{\partial\epsilon_x}{\partial\omega}\Big|_{(\omega_2,k_2)}\right)\right]^{-1} \geq 0. \tag{50}$$

Clearly, (50) is satisfied with strict inequalities if and only if

$$\mathrm{sgn}\left(\frac{\partial\epsilon_x}{\partial\omega}\Big|_{(\omega_0,k_0)}\right) = \mathrm{sgn}\left(\frac{\partial\epsilon_x}{\partial\omega}\Big|_{(\omega_2,k_2)}\right) = -\,\mathrm{sgn}\left(\frac{\partial\epsilon_x}{\partial\omega}\Big|_{(\omega_1,k_1)}\right). \tag{51}$$

Moreover, (50) cannot be satisfied if all the waves have positive or negative energy. But if the zeroth and second waves have positive energy and the first wave has negative energy, then condition (9) or (50) is satisfied. In this case, if $\eta_0(\mathbf{a}_0) = \mathrm{Re}\{\mu_0^*a_0^*(0)a_1(0)a_2(0)\} \geq 0$, then $\eta_0(\mathbf{a}(t)) \geq 0$ for all $t \geq 0$ implying that the zeroth wave gains energy at all time. On the other hand, if the first wave has positive energy and the remaining waves have negative energy, then the zeroth wave loses energy at all time provided that $\eta_0(\mathbf{a}_0) \geq 0$. Analogous results can be also established for the first and second waves.

In a similar way, we can derive conditions for the unidirectional wave-energy transfer in the nonlinear interaction of two transverse waves (ω_0, k_0), (ω_1, k_1) (with same polarization) and a longitudinal wave (ω_2, k_2) in a magnetized collisionless electron plasma. For this case, the coupling coefficients μ_i are given by[12]

$$\mu_0 = \omega_0^2C_{21}^{\pm}/g^{\pm}(\omega_0, k_0), \quad \mu_1 = \omega_1^2C_{21}^{\pm}/g^{\pm}(\omega_1, k_1),$$
$$\mu_2 = -\,\tfrac{1}{2}C_{21}^{\pm}\left(\frac{\partial\epsilon_x}{\partial\omega}\Big|_{(\omega_2,k_2)}\right)^{-1}, \tag{52}$$

where

$$g^{\pm}(\omega_i, k_i) = \frac{\partial(\omega^2\epsilon^{\pm})}{\partial\omega}\Big|_{(\omega_i,k_i)}, \quad i = 0, 1. \tag{53}$$

The explicit expressions for $\epsilon^{\pm}$ and $C_{21}^{\pm}$ are given in Ref. 12. By direct computation, it can be deduced that for the zeroth wave, condition (9) is satisfied with strict inequalities if and only if

$$\mathrm{sgn}[g^{\pm}(\omega_0, k_0)] = \mathrm{sgn}[g^{\pm}(\omega_1, k_1)] = -\,\mathrm{sgn}\left(\frac{\partial\epsilon_x}{\partial\omega}\Big|_{(\omega_0,k_0)}\right). \tag{54}$$

Thus, condition (9) can be interpreted again in terms of the signs of the wave energy as in the case of three longitudinal waves.

ACKNOWLEDGMENTS

The author wishes to thank Professor L. Stenflo of Umeå University, Sweden for bringing to my attention the error in the coupling coefficient and also Dr. K. Jungwirth of the Institute of Plasma Physics, Czechoslovak Academy of Sciences for noting the inconsistency between the wave amplitude growth and the signs of the wave energies. This work was supported by AFOSR Grant 72-2303.

[12] L. Stenflo, J. Plasma Phys. 4, 585 (1970).