JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 13, NUMBER 9 SEPTEMBER 1972

The Dielectric Ring in a Uniform, Axial, Electrostatic Field

J. D. Love
Department of Physics, University of Toronto, Toronto 181, Canada
(Received 31 March 1972)

The exact solutions for the fields inside and outside a dielectric ring in a uniform, axial, electrostatic field are derived in the system of toroidal coordinates. Comparison is made with the perturbation solution generated by inverse aspect ratio expansion, and it is shown how the exact solution may be truncated to any desired accuracy. The cylindrical limits of the exact and truncated solutions are obtained.

1. INTRODUCTION

The determination of the electrostatic potentials for the regions inside and outside a uniform dielectric ring in an externally imposed electric field poses an inhomogeneous Neumann problem. The potentials are solutions of the Laplace equation, and their derivatives satisfy nonzero continuity conditions on the surface of the ring. A similar, though simpler, problem of this type was first considered by Hicks.[1] He obtained a solution for the potential external to a ring when the total normal derivative was prescribed on the surface of the ring. The conditions of our problem are significantly different from those of Hicks, so we must seek an alternative method of solution.

In Sec. 2 we derive an exact solution to the problem in the system of toroidal coordinates. This system is such that the solutions of the Laplace equation are separable, and such that the ring surface is simply described by a single coordinate. However, the toroidal harmonics we use are not orthogonal, which in turn generate three-term recurrence relations between their amplitudes when we satisfy the boundary conditions. We solve these inhomogeneous recurrence relations by using Green's functions for difference equations in a manner entirely analogous to the solution of second-order inhomogeneous differential equations by Green's functions. Convergence of the exact solutions is proved in Sec. 3. The exact solution is compared, in Sec. 4, with a solution of the same problem using a perturbation expansion in inverse aspect ratio of the ring up to second order. The latter solution is shown to break down at distances from the ring of order the aspect ratio of the ring multiplied by its minor radius.

For practical applications, the form of the exact solution is rather unwieldy. Thus, in Sec. 5, we demonstrate how for a given aspect ratio of the ring the exact solution can be truncated and still represent the fields to any desired accuracy at all points of space. Finally in Sec. 6 we impose the cylindrical limit of toroidal coordinates to demonstrate that both the exact and truncated solutions converge to the solution of the corresponding problem in cylindrical geometry.

2. EXACT SOLUTION

The toroidal coordinate system is described in detail by Hobson.[2] Unfortunately, however, Hobson's account of the properties of the associated Legendre functions which occur in the solution of the Laplace equation in toroidal coordinates is not without error. We shall, therefore, cite Bateman[3] for the properties of these functions, which are either given by or are readily derivable from this reference.

In terms of toroidal coordinates (η, τ, ϕ), the surface of the ring is described by $\eta = \eta_0$, where η_0 is a constant. The angle τ is a measure of displacement around the minor circumference of the ring, and the angle ϕ is the azimuthal angle about the axis of symmetry of the ring. The aspect ratio of the ring—the ratio of major to minor radii—is given by $\cosh\eta_0$, and the only length dimension d is the radius of the limiting torus when $\eta \to \infty$.

The solutions of the Laplace equation which are both axisymmetric and form a complete set in (η, τ) space are

$$(\cosh\eta - \cos\tau)^{1/2} P_{n-1/2}(\cosh\eta) e^{in\tau},$$

$$(\cosh\eta - \cos\tau)^{1/2} Q_{n-1/2}(\cosh\eta) e^{in\tau},$$

where n is an integer or zero. The $P_{n-1/2}(\cosh\eta)$ and $Q_{n-1/2}(\cosh\eta)$ are half odd-integral order associated Legendre functions of the first and second kinds, respectively. These functions are such that $P_{n-1/2}(\cosh\eta) \to 1$ as $\eta \to 0$ and $Q_{n-1/2}(\cosh\eta) \to 0$ as $\eta \to \infty$. Thus the P-type harmonics are suitable for describing the region external to the ring, where $\eta_0 \geq \eta \geq 0$,

and the Q-type harmonics are suitable for describing the region internal to the ring where $\eta_0 \leq \eta < \infty$.

We take the axis of symmetry of the ring to be the z axis, whence the uniform electrostatic field is represented by

$$\mathbf{E}_0 = E_0 \nabla z,$$

where E_0 is the constant field amplitude. The function z has the following expansion in terms of the Q-type harmonics:

$$z = (\cosh\eta - \cos\tau)^{1/2} \frac{2\sqrt{2}d}{i\pi} \sum_{n=-\infty}^{\infty} n Q_{n-1/2}(\cosh\eta) e^{in\tau}.$$

Thus appropriate expressions for the electrostatic potentials Φ^i and Φ^0 for the regions inside and outside $\eta = \eta_0$, respectively, are

$$\Phi^i = (\cosh\eta - \cos\tau)^{1/2} \sum_{n=-\infty}^{\infty} A_n Q_{n-1/2}(\cosh\eta) e^{in\tau}, \qquad \eta \geq \eta_0 \quad (2.1)$$

$$\Phi^0 = (\cosh\eta - \cos\tau)^{1/2} \sum_{n=-\infty}^{\infty} \{B_n P_{n-1/2}(\cosh\eta) + [2\sqrt{2}dE_0/i\pi] n Q_{n-1/2}(\cosh\eta)\} e^{in\tau}, \quad 0 \leq \eta \leq \eta_0, \quad (2.2)$$

where the A_n and B_n are constants to be determined.

The boundary conditions which Φ^i and Φ^0 must satisfy on $\eta = \eta_0$ reduce to

$$\epsilon \frac{\partial \Phi^i}{\partial \eta} = \frac{\partial \Phi^0}{\partial \eta}, \qquad \eta = \eta_0, \quad (2.3)$$

$$\frac{\partial \Phi^i}{\partial \tau} = \frac{\partial \Phi^0}{\partial \tau}, \qquad \eta = \eta_0, \quad (2.4)$$

where ϵ is the dielectric constant of the ring. To facilitate satisfying these conditions, we rearrange the expression (2.2) for Φ^0 into the form

$$\Phi^0 = (\cosh\eta - \cos\tau)^{1/2} \sum_{n=-\infty}^{\infty} \Big(C_n P_{n-1/2}(\cosh\eta) + (2\sqrt{2}dE_0/i\pi) n [P_{n-1/2}(\cosh\eta_0) Q_{n-1/2}(\cosh\eta) - Q_{n-1/2}(\cosh\eta_0) P_{n-1/2} \times (\cosh\eta)] / P_{n-1/2}(\cosh\eta_0) \Big) e^{in\tau}, \quad (2.5)$$

where the constants C_n are related to the B_n by

$$C_n = B_n + \frac{2\sqrt{2}dE_0}{i\pi} n \frac{Q_{n-1/2}(\cosh\eta_0)}{P_{n-1/2}(\cosh\eta_0)}. \quad (2.6)$$

The tangential boundary condition (2.4) is then easily satisfied if for all n,

$$A_n Q_{n-1/2}(\cosh\eta_0) = C_n P_{n-1/2}(\cosh\eta_0). \quad (2.7)$$

Substituting Eqs. (2.1) and (2.5) in the radial boundary condition (2.3) and eliminating the C_n using Eq. (2.7), we obtain

$$(\cosh\eta_0 - \cos\tau)^{1/2} \sum_{n=-\infty}^{\infty} A_n Q_n^0 \left(\epsilon \frac{Q_n^{0\prime}}{Q_n^0} - \frac{P_n^{0\prime}}{P_n^0} \right) e^{in\tau} + \frac{(\epsilon - 1)}{2(\cosh\eta_0 - \cos\tau)^{1/2}} \sum_{n=-\infty}^{\infty} A_n Q_n^0 e^{in\tau}$$

$$= (\cosh\eta_0 - \cos\tau)^{1/2} \frac{2\sqrt{2}dE_0}{i\pi} \sum_{n=-\infty}^{\infty} n \frac{W_n^0}{P_n^0} e^{in\tau}, \quad (2.8)$$

where $P_n^0 \equiv P_{n-1/2}(\cosh\eta_0)$, $Q_n^0 \equiv Q_{n-1/2}(\cosh\eta_0)$ and dash denotes differentation with respect to $\cosh\eta_0$. The Wronskian W_n^0 is defined by

$$W_n^0 \equiv P_n^0 Q_n^{0\prime} - P_n^{0\prime} Q_n^0 = -1/\sinh^2\eta_0 .$$

Upon introducing a simplification in notation,

$$D_n \equiv A_n Q_n^0, \qquad \Pi_n^0 \equiv \epsilon(Q_n^{0\prime}/Q_n^0) - (P_n^{0\prime}/P_n^0), \quad (2.9)$$

and rearranging Eq. (2.8), we obtain

$$(\epsilon - 1) \sum_{n=-\infty}^{\infty} D_n e^{in\tau} + 2(\cosh\eta_0 - \cos\tau) \sum_{n=-\infty}^{\infty} D_n \Pi_n^0 e^{in\tau} = 2(\cosh\eta_0 - \cos\tau) \sum_{n=-\infty}^{\infty} \lambda_n e^{in\tau}, \quad (2.10)$$

where

$$\lambda_n = [(i2\sqrt{2}E_0 d)/(\pi \sinh^2\eta_0)] (n/P_n^0). \quad (2.11)$$

We make a second change in notation to

$$E_n \equiv D_n \Pi_n^0, \qquad q_n \equiv 2\cosh\eta_0 + \frac{(\epsilon - 1)}{\Pi_n^0}, \quad (2.12)$$

and Fourier analyze Eq. (2.10) in τ, yielding

$$E_{n+1} - q_n E_n + E_{n-1} = \lambda_{n+1} - 2\cosh\eta_0 \lambda_n + \lambda_{n-1}, \qquad n = 0, \pm 1, \cdots. \quad (2.13)$$

Thus Eq. (2.13) is a second-order inhomogeneous difference equation for the E_n in terms of the given λ_n.

The solution of Eq. (2.13) is constructed by using Green's function techniques for difference equations. Let $G_{n,N}$ be a solution of the following equation, where $\delta_{i,j}$ is the Kronecker delta:

$$G_{n+1,N} - q_n G_{n,N} + G_{n-1,N} = \delta_{n,N+1} - 2\cosh\eta_0 \delta_{n,N} + \delta_{n,N-1}, \qquad n = 0, \pm 1, \ldots, \quad (2.14)$$

for each value of integer N, $-\infty < N < \infty$. The sense of Eq. (2.14) is such that for each fixed value of N, a solution $G_{n,N}$ is to be constructed valid for all n. A solution to Eq. (2.13) is then

$$E_n = \sum_{N=-\infty}^{\infty} G_{n,N} \lambda_N, \qquad n = 0, \pm 1, \ldots, \quad (2.15)$$

as can be verified by direct substitution.

To construct the functions $G_{n,N}$, we utilize the two independent solutions of the complementary difference equation of Eq. (2.14), namely

$$G_{n+1,N} - q_n G_{n,N} + G_{n-1,N} = 0, \qquad n = 0, \pm 1, \ldots, \quad (2.16)$$

for each fixed value of N. The complementary solutions of Eq. (2.16) are in turn constructed from the solutions of the characteristic equation, which is generated from Eq. (2.16) by taking the limit $|n| \to \infty$.

To determine this limit, we require the asymptotic forms of the associated Legendre functions for large $|n|$:

$$\lim_{|n|\to\infty} P_{n-1/2}(\cosh\eta_0) \simeq \frac{e^{|n|\eta_0}}{[2\pi(|n| + \tfrac{1}{2})\sinh\eta_0]^{1/2}} \tag{2.17}$$

$$\lim_{|n|\to\infty} Q_{n-1/2}(\cosh\eta_0) \simeq \frac{\pi^{1/2} e^{-|n|\eta_0}}{[2(|n| - \tfrac{1}{2})\sinh\eta_0]^{1/2}}. \tag{2.18}$$

Hence, from Eq. (2.9) and (2.12)

$$\Pi_n^0 \to 0 \quad \text{and} \quad q_n \to 2\cosh\eta_0 \quad \text{as } |n| \to \infty.$$

Thus the characteristic equation is

$$G_{n+1,N} - 2\cosh\eta_0 G_{n,N} + G_{n-1,N} = 0, \tag{2.19}$$

with roots $e^{\pm\eta_0}$ for the ratio $G_{n+1,N}/G_{n,N}$.

The roots of the characteristic equation are therefore distinct for $\eta_0 \neq 0$, and, consequently, we can utilize a theorem due to Perron.[4] This theorem states that there are two fundamental solutions to Eq. (2.16) for $G_{n+1,N}/G_{n,N}$, which tend to e^{η_0} and $e^{-\eta_0}$, respectively, as $|n| \to \infty$. From this result it is possible to construct two independent solutions of Eq. (2.16)[5]:

$$\frac{G_{n+1,N}}{G_{n,N}} = \cfrac{1}{q_{n+1} - \cfrac{1}{q_{n+2} - \cfrac{1}{q_{n+3} - \cdots}}} \equiv \alpha_{n+1}, \tag{2.20}$$

$$\frac{G_{n,N}}{G_{n-1,N}} = \cfrac{1}{q_{n-1} - \cfrac{1}{q_{n-2} - \cfrac{1}{q_{n-3} - \cdots}}} \equiv \beta_{n-1}, \tag{2.21}$$

such that both α_n and $\beta_n \to e^{-\eta_0}$ as $|n| \to \infty$. The values of each α_n and β_n are calculated from the respective continued fraction. It also follows from Eqs. (2.20) and (2.21), and the symmetry properties

$$Q_{-n}^0 \equiv Q_n^0, \qquad P_{-n}^0 \equiv P_n^0,$$

that

$$\alpha_n = \frac{1}{q_n - \alpha_{n+1}}, \qquad \beta_n = \frac{1}{q_n - \beta_{n-1}}, \qquad \alpha_n = \beta_{-n}.$$

The solution of Eq. (2.14) is now constructed from Eqs. (2.20) and (2.21) by setting

$$G_{n+1,N} = \alpha_{n+1} G_{n,N}, \qquad n \geq N+1,$$
$$G_{n-1,N} = \beta_{n-1} G_{n,N}, \qquad n \leq N-1.$$

The $G_{n,N}$ determined this way will be unique if $G_{N+1,N}$ and $G_{N-1,N}$ are specified. These two quantities, together with $G_{N,N}$, are given by Eq. (2.14) for $n = N-1, N$, and $N+1$, respectively

$$\begin{aligned} &(\beta_{N-2} - q_{N-1})G_{N-1,N} + G_{N,N} = 1, \\ &G_{N-1,N} - q_N G_{N,N} + G_{N+1,N} = -2\cosh\eta_0, \\ &G_{N,N} + (\alpha_{N+2} - q_{N+1})G_{N+1,N} = 1, \end{aligned} \tag{2.22}$$

with solution

$$\begin{aligned} G_{N-1,N} &= \frac{\beta_{N-1}(2\cosh\eta_0 - q_N)}{(q_N - \alpha_{N+1} - \beta_{N-1})}, \\ G_{N,N} &= \frac{(2\cosh\eta_0 - \alpha_{N+1} - \beta_{N-1})}{(q_N - \alpha_{N+1} - \beta_{N-1})}, \\ G_{N+1,N} &= \frac{\alpha_{N+1}(2\cosh\eta_0 - q_N)}{(q_N - \alpha_{N+1} - \beta_{N-1})}. \end{aligned} \tag{2.23}$$

Upon combining Eqs. (2.23), (2.15), (2.12), (2.9), we obtain the following expressions for the potentials:

$$\begin{aligned} \Phi^i &= \frac{2\sqrt{2}iE_0 d}{\pi\sinh^2\eta_0}(\cosh\eta - \cos\tau)^{1/2} \\ &\times \sum_{n=-\infty}^{\infty}\sum_{N=-\infty}^{\infty} a_{n,N}\frac{Q_{n-1/2}(\cosh\eta)}{Q_n^0} e^{in\tau}, \\ &\eta \geq \eta_0, \end{aligned} \tag{2.24}$$

$$\begin{aligned} \Phi^0 = \frac{2\sqrt{2}iE_0 d}{\pi}(\cosh\eta - \cos\tau)^{1/2} &\left(\sum_{n=-\infty}^{\infty}\frac{n}{P_n^0}\right. \\ &\times [Q_n^0 P_{n-1/2}(\cosh\eta) - P_n^0 Q_{n-1/2}(\cosh\eta)]e^{in\tau} \\ &+ \left.\frac{1}{\sinh^2\eta_0}\sum_{n=-\infty}^{\infty}\sum_{N=-\infty}^{\infty} a_{n,N}\frac{P_{n-1/2}(\cosh\eta)}{P_n^0} e^{in\tau}\right), \\ 0 \leq \eta \leq \eta_0, \end{aligned} \tag{2.25}$$

where

$$\begin{aligned} a_{n,N} = (N\{&\delta_{n,N}(2\cosh\eta_0 - \alpha_{N+1} - \beta_{N-1}) \\ &+ (2\cosh\eta_0 - q_N)[\theta(n-N)\prod_{m=N+1}^{n}\alpha_m + \theta(N-n) \\ &\times \prod_{m=n}^{N-1}\beta_m]\}/\Pi_n^0 P_N^0(q_N - \alpha_{N+1} - \beta_{N-1}) \end{aligned} \tag{2.26}$$

and $\theta(x) = 0$ if $x \leq 0$, $= 1$ if $x > 0$. Note that there are no terms in $n = 0$ in either Φ^i or Φ^0. The summation over N for $n = 0$ vanishes upon using the symmetry properties of the α_n and β_n.

The electic field components are given by

$$\mathbf{E} = \nabla\Phi \equiv \frac{(\cosh\eta - \cos\tau)}{d}\left(\frac{\partial}{\partial\eta}, \frac{\partial}{\partial\tau}, 0\right)\Phi. \tag{2.27}$$

In their respective domains, the electric field components satisfy all boundary conditions and, as will be proved in the next section, they are absolutely convergent. Thus the fields are unique, and, consequently, the corresponding potentials are only unique to within an arbitrary constant. We can utilize this arbitrariness to alter the expressions for the potentials, if we wish, by employing the expansion of unity as a series of toroidal harmonics:

$$1 \equiv \frac{\sqrt{2}}{\pi}(\cosh\eta - \cos\tau)^{1/2}\sum_{n=-\infty}^{\infty} Q_{n-1/2}(\cosh\eta)e^{in\tau},$$

which is valid for all values of η and τ.

3. CONVERGENCE

For the exact solutions obtained in the previous sections to be physically meaningful, we must show that the potentials and field components converge at all points of their respective domains. We shall show that the expressions for Φ^i and Φ^0 given by Eqs. (2.24) and (2.25) are absolutely convergent. The proof is easily adapted to showing absolute convergence of the corresponding field components.

We can ignore the series of Q-type harmonics in Eq.

(2.25) since they represent the uniform applied field. The single summation over P-type harmonics in Φ^0 will be absolutely convergent if the series

$$\sum_{n=-\infty}^{\infty} \left| n \frac{Q_n^0}{P_n^0} P_{n-1/2}(\cosh\eta) \right| \tag{3.1}$$

is convergent. Using the bounding properties of the associated Legendre functions

$$|P_{n-1/2}(\cosh\eta)| \leq |P_n^0|, \quad 0 \leq \eta \leq \eta_0, \quad n \neq 0, \tag{3.2}$$
$$|Q_{n-1/2}(\cosh\eta)| \leq |Q_n^0|, \quad \eta \geq \eta_0,$$

the sum (3.1) is bounded by

$$\sum_{n=-\infty}^{\infty} |nQ_n^0|. \tag{3.3}$$

For large values of $|n|$, we can use the asymptotic form for $Q_{n-1/2}(\cosh\eta)$ as given by Eq. (2.18) to show

$$|nQ_n^0| \sim (\sqrt{n} e^{-|n|\eta_0})/\sinh^{1/2}\eta_0 .$$

Hence, since $\sum \sqrt{n} e^{-|n|\eta_0}$ is convergent for $\eta_0 \neq 0$, the series (3.3) and, therefore, (3.1) are absolutely convergent for $\eta_0 \neq 0$. The case $\eta_0 = 0$ is of no physical interest, since it corresponds to a dielectric ring completely filling the whole of space.

If we apply the inequalities (3.2) to the remaining double sum in Φ^0 and to Φ^i, we find that both expressions will be absolutely convergent if the double sum

$$\sum_{n=-\infty}^{\infty} \sum_{N=-\infty}^{\infty} |a_{n,N}|, \tag{3.4}$$

is convergent. We first demonstrate convergence of the n summation for each value of N. Substituting for $a_{n,N}$ from Eq. (2.26) and rearranging, the convergence of Eq. (3.4) is determined by the convergence of

$$\sum_{N=-\infty}^{\infty} \frac{|N|\,|2\cosh\eta_0 - \alpha_{N+1} - \beta_{N-1}|}{|\Pi_N^0 P_N^0 (q_N - \alpha_{N+1} - \beta_{N-1})|}$$
$$+ \sum_{N=-\infty}^{\infty} \frac{|(2\cosh\eta_0 + q_N)|}{P_N^0 (q_N - \alpha_{N+1} - \beta_{N-1})}$$
$$\times \sum_{n=-\infty}^{\infty} \frac{\left| \theta(n-N) \prod_{m=N+1}^{n} \alpha_m + \theta(N-n) \prod_{m=n}^{N-1} \beta_m \right|}{|\Pi_n^0|}. \tag{3.5}$$

For sufficiently large values of n, such that $|n| \gg |N|$, $|n| \gg 1$, we recall from Eqs. (2.20) and (2.21) that

$$\alpha_n \simeq e^{-\eta_0}, \qquad \beta_n \simeq e^{-\eta_0}.$$

Further, from Eqs. (2.9), (2.17), and (2.18),

$$\Pi_n^0 \simeq [-(\epsilon+1)|n|]/\sinh\eta_0, \qquad |n| \to \infty .$$

For all but a finite number of values of n, the convergence of the summation over n in Eq. (3.5) is determined by the convergence of a series whose terms are of the form $|n| e^{-|n|\eta_0}$ for $|n| \gg |N|$, $|n| \gg 1$. Hence the n summation is convergent for $\eta_0 \neq 0$.

The convergence of Eq. (3.4) is now dominated by the convergence of

$$\sum_{N=-\infty}^{\infty} \frac{|N|}{|P_N^0|\,|\Pi_N^0|} \left(\frac{|2\cosh\eta_0 - \alpha_{N+1} - \beta_{N-1}|}{|q_N - \alpha_{N+1} - \beta_{N-1}|} + \frac{|(2\cosh\eta_0 - q_N)|\,|\Pi_N^0| M}{|q_N - \alpha_{N+1} - \beta_{N-1}|} \right), \tag{3.6}$$

where

$$M = \sup \sum_{n=-\infty}^{\infty} |\theta(n-N) \prod_{m=N+1}^{n} \alpha_m + \theta(N-n) \prod_{m=n}^{N-1} \beta_m| / |\Pi_n^0|$$

over all values of N, and M is, therefore, independent of N. As $|N| \to \infty$, each of the two expressions within brackets in Eq. (3.6) approaches an indeterminate limit. To avoid this difficulty, we use Eq. (2.23) to recast Eq. (3.6) as

$$\sum_{N=-\infty}^{\infty} \frac{|N|}{|P_N^0|\,|\Pi_N^0|} \left(|G_{N,N}| + \frac{|\Pi_N^0| M |G_{N+1,N}|}{|\alpha_{N+1}|} \right). \tag{3.7}$$

By letting $|N| \to \infty$ in the equations for $G_{N,N}$, (2.22), we find

$$|G_{N,N}| \to 1, \qquad |G_{N+1,N}| \to 0 \text{ as } |N| \to \infty.$$

As $|\alpha_{N+1}| \to e^{-\eta_0}$ in this limit, the asymptotic variation of terms in Eq. (3.7) is of order $|n|^{1/2} e^{-|n|\eta_0}$. Hence the series (3.7) and (3.4) are absolutely convergent for $\eta_0 \neq 0$.

The absolute convergence of the field components is proved using the same techniques as above. The only essential difference between the expressions for corresponding fields and potentials is that $P_{n-1/2}(\cosh\eta)$ is replaced by either $P'_{n-1/2}(\cosh\eta)$ or $n P_{n-1/2}(\cosh\eta)$ and similarly for the $Q_{n-1/2}(\cosh\eta)$. This means that wherever convergence is governed by series of terms varying as $|n|^P e^{-|n|\eta_0}$, the terms are replaced by $|n|^{P+1} e^{-|n|\eta_0}$, and convergence is again satisfied.

4. PERTURBATION SOLUTION

The exact solution obtained in Sec. 2 is valid at all points of space and for all values of the aspect ratio of the ring $\cosh\eta_0$. However, the expressions for the potentials as given by Eqs. (2.24) and (2.25) are rather elaborate, and it is, therefore, pertinent to ask whether simpler expressions could be formulated which would represent the solution to any desired accuracy. Since by definition $\cosh\eta_0 > 1$, it is clearly worth seeking a perturbation solution which proceeds in increasing powers of the inverse aspect ratio $(\cosh\eta_0)^{-1}$. This expansion is also physically attractive, since the lowest order solution ($\cosh\eta_0 = \infty$) will correspond to the solution for an infinite straight dielectric cylinder transverse to the applied field. Successive higher order solutions will then be expected to generate corrections to the lowest order solution to account for the effects of toroidicity.

One procedure for generating the inverse aspect ratio expansion would be to expand the exact solution in powers of $(\cosh\eta)^{-1}$ for $\eta \geq \eta_0$ and in powers of $(\cosh\eta/\cosh\eta_0)$ for $0 \leq \eta \leq \eta_0$. The disadvantages of such are twofold. Firstly, even to lowest order in such an expansion, the solution is not recognizable in terms of a straight cylinder without further geometric

expansions. Secondly, higher-order terms entail calculating power series expansions for the associated Legendre functions to greater order and then extracting the necessary coefficients of each order from the doubly infinite series of the exact solutions. While there is no formal difficulty in either of these two points, we have found that it is both physically more illuminating and algebraically more tractable to generate the inverse aspect ratio solution using another toroidal coordinate system.

The system of toroidal polar coordinates (r, θ, ϕ) is based on the center of the minor cross section of the ring as origin and shares the azimuthal angle about the axis of symmetry ϕ with toroidal coordinates. In the minor cross-section, (r, θ) are polar coordinates with the direction $\theta = 0$ corresponding to $\tau = 0$, radially outwards in the $z = 0$ plane. The major geometric difference between the two coordinate systems is that constant $-r$ surfaces form a set of nested concentric tori, while constant $-\eta$ surfaces form a set of nested coaxial tori.

In the toroidal polar coordinate system, the axisymmetric form of the Laplace equation is

$$\frac{\partial}{\partial r}\left(r(R + r\cos\theta)\frac{\partial \phi}{\partial r}\right) + \frac{\partial}{\partial \theta}\left(\frac{(R + r\cos\theta)}{r}\frac{\partial \phi}{\partial \theta}\right) = 0, \quad (4.1)$$

where R is the major radius of the ring. We define the inverse aspect ratio $\lambda = r/R$, so that the dimensionless form of Eq. (4.1) is

$$\lambda^2(1 + \lambda\cos\theta)\frac{\partial^2\Phi}{\partial\lambda^2} + \lambda(1 + 2\lambda\cos\theta)\frac{\partial\Phi}{\partial\lambda} + (1 + \lambda\cos\theta)\frac{\partial^2\Phi}{\partial\theta^2} - \lambda\sin\theta\frac{\partial\Phi}{\partial\theta} = 0. \quad (4.2)$$

It is not possible to solve Eq. (4.2) by simple separation of variables. For our inverse aspect ratio expansion we write the equation in the form

$$\mathcal{L}(\Phi) \equiv \lambda^2\frac{\partial^2\Phi}{\partial\lambda^2} + \lambda\frac{\partial\Phi}{\partial\lambda} + \frac{\partial^2\Phi}{\partial\theta^2} = -\lambda^3\cos\theta\frac{\partial^2\Phi}{\partial\lambda^2} - 2\lambda^2\cos\theta\frac{\partial\Phi}{\partial\lambda} - \lambda\cos\theta\frac{\partial^2\theta}{\partial\theta^2} + \lambda\sin\theta\frac{\partial\Phi}{\partial\theta}, \quad (4.3)$$

so that the right-hand side is one order higher in λ than the left-hand side.

The equation for the lowest order solution is

$$\mathcal{L}(\Phi_0) = 0, \quad (4.4)$$

where the subscript denotes the order of the solution. The applied field potential is $E_0 r\sin\theta$, so that solutions of Eq. (3.4) appropriate to the spatial dependence of the applied field are $\lambda\sin\theta$, $(1/\lambda)\sin\theta$, which together with the boundary conditions

$$\frac{\partial\Phi^i}{\partial\theta} = \frac{\partial\Phi^0}{\partial\theta}, \quad \epsilon\frac{\partial\Phi^i}{\partial\lambda} = \frac{\partial\Phi^0}{\partial\lambda}, \quad \lambda = \frac{a}{R} \quad (4.5)$$

yields the solutions

$$(\Phi_0^i/R) = [2E_0/(\epsilon + 1)]\lambda\sin\theta, \quad (4.6)$$

$$(\Phi_0^0/R) = E_0\lambda\sin\theta - [(\epsilon - 1)/(\epsilon + 1)]E_0(a/R)^2(\sin\theta/\lambda). \quad (4.7)$$

The zeroth order solution is just the solution for an infinite straight dielectric cylinder transverse to the applied field and corresponds, therefore, to the cylindrical limit of the ring when $R \to \infty$ for fixed a and r.

The equation for the first-order potentials is

$$\mathcal{L}(\Phi_1) = -\lambda^2\cos\theta\frac{\partial\Phi_0}{\partial\lambda} + \lambda\sin\theta\frac{\partial\Phi_0}{\partial\theta}. \quad (4.8)$$

From Eqs. (4.6) and (4.7) there are two possibilities: $\Phi_0 = \lambda^{-1}\sin\theta$ or $\Phi_0 = \lambda\sin\theta$. Substituting in Eq. (4.8) yields, respectively,

$$\mathcal{L}(\Phi_1) = \sin 2\theta \quad \text{or} \quad \mathcal{L}(\Phi_1) = 0,$$

which have particular solutions:

$$\Phi_1 = -\tfrac{1}{4}\sin 2\theta \quad \text{or} \quad \Phi_1 = 0. \quad (4.9)$$

The form of the particular solutions indicates that the only complementary solutions to be included are $\lambda^2\sin 2\theta$ and $(\sin 2\theta)/\lambda^2$.

Upon satisfying the boundary conditions (4.5), we obtain the first order potentials:

$$\frac{\Phi_1^i}{R} = \frac{E_0}{4}\frac{(\epsilon - 1)}{(\epsilon + 1)^2}\lambda^2\sin 2\theta,$$
$$\frac{\Phi_1^0}{R} = \frac{E_0}{4}\frac{(\epsilon - 1)}{(\epsilon + 1)}\left(\frac{a}{R}\right)^2 \times\left[1 - \frac{\epsilon}{(\epsilon + 1)}\left(\frac{a}{R}\right)^2\frac{1}{\lambda^2}\right]\sin 2\theta. \quad (4.10)$$

These expression contain the lowest order effects of toroidicity on the zeroth order solution, for if we allow R to increase indefinitely in (4.10) and hold a and r fixed, the potentials vanish. The field components corresponding to Φ_1^i and Φ_1^0 vary as λ inside the ring and as λ^{-1} outside the ring, so that the first-order solution satisfies all the physical properties we would expect of it. However, if we compare the first-order electric field with the zeroth order field, we find

$$\frac{|\mathbf{E}_1^i|}{|\mathbf{E}_0^i|} = \frac{(\epsilon - 1)}{(\epsilon + 1)}\frac{\lambda}{4}, \quad \lambda \lesssim \frac{a}{R}, \quad \frac{|\mathbf{E}_1^0|}{|\mathbf{E}_0^0|} \simeq \frac{\lambda}{2}, \quad \lambda \gg \frac{a}{R}.$$

Thus the correction for toroidicity on the internal field is at most of order the inverse aspect ratio of the ring when compared with the cylindrical solution. The same correction for the external field is only small near the surface of the ring, and is of order of the cylindrical solution or greater for $r \gtrsim 0(a)$. Thus the first-order solution can only be considered as a perturbation of the lowest-order solution inside or close to the surface of the ring. We shall, therefore, examine the second-order solutions to see if they improve the convergence of the inverse aspect ratio expansion.

The second-order potentials satisfy the equation

$$\mathcal{L}(\Phi_2) = -\lambda\cos\theta\mathcal{L}(\Phi_1) - \lambda^2\cos\theta\frac{\partial\Phi_1}{\partial\lambda} + \lambda\sin\theta\frac{\partial\Phi_1}{\Phi\theta}. \quad (4.11)$$

We see from Eq. (4.10) that there are three possibilities for Φ_1:

$$\sin 2\theta, \quad \lambda^2 \sin 2\theta, \quad \sin 2\theta/\lambda^2,$$

which give particular solutions of Eq. (4.11), respectively,

$$-\tfrac{3}{8}\lambda \sin 3\theta + \tfrac{1}{2}\lambda(\log\lambda)\sin\theta, \quad -\tfrac{1}{4}\lambda^3 \sin\theta, \quad -\sin 3\theta/4\lambda,$$

together with the appropriate complementary functions $\lambda \sin\theta$, $\sin\theta/\lambda$, $\lambda^3 \sin 3\theta$, $\sin 3\theta/\lambda^3$.

The boundary conditions (4.5) and continuity of the θ dependence, give the second-order potentials

$$\Phi_2^i = \frac{-E_0(\epsilon-1)}{16(\epsilon+1)}\left\{\lambda^3\sin\theta - 2\left(\frac{a}{R}\right)^2\left[\frac{(5\epsilon+3)}{2(\epsilon+1)} + \log\left(\frac{a}{R}\right)^2\right]\lambda\sin\theta + \frac{2}{3}\frac{(2\epsilon+3)}{(\epsilon+1)}\lambda^3\sin 3\theta\right\},$$

$$\Phi_2^0 = \frac{-E_0}{16}\frac{(\epsilon-1)}{(\epsilon+1)}\left(\frac{a}{R}\right)^2\left\{\frac{3}{2}\lambda\sin 3\theta - 2\lambda(\log\lambda)\times\sin\theta - \frac{\epsilon}{(\epsilon+1)}\left(\frac{a}{R}\right)^2\frac{\sin 3\theta}{\lambda} - 2\left(\frac{a}{R}\right)^2\left[\frac{2(\epsilon+1)}{(\epsilon+1)^2} - \frac{(\epsilon-1)}{(\epsilon+1)}\log\left(\frac{a}{R}\right)\right]\frac{\sin\theta}{\lambda} + \left(\frac{a}{R}\right)^4\frac{(3-4\epsilon-3\epsilon^2)}{6(\epsilon+1)^2}\frac{\sin 3\theta}{\lambda^3}\right\}. \tag{4.12}$$

The second-order electric field components derived from Eq. (4.12) vanish as R increases indefinitely for fixed a and r. However, at large distances from the ring the external field now has a logarithmic singularity in r.

If we now combine the terms in the perturbation expansion (4.6) (4.7), (4.10), and (4.12), we see that higher orders in the expansion produce higher order singularities in the far field, so that the domain of validity of our solution is restricted to $\lambda \lesssim 1$, or $r \gtrsim R = a(R/a)$. Hence as the aspect ratio of the ring decreases, for fixed a, the domain becomes increasingly smaller. We, therefore, conclude that the toroidicity of the ring cannot be described approximately at all points of space by an inverse ratio expansion that uses a straight cylinder as its lowest approximation. Nevertheless, it is possible to obtain an approximation to the exact solution that has any desired accuracy at all points of space. As we shall show in the next section, it is sufficient to retain the first few τ modes in the exact solution, as the dominant effects of toroidicity are contained in the η dependence of these modes.

5. APPROXIMATE SOLUTION

As we saw in the previous section, the inverse aspect ratio expansion technique was inadequate for describing an approximate solution at all points of space. We shall now show that such an approximation can be obtained by simply truncating the exact solutions, apart from the applied field terms, beyond the first few τ modes. These low order modes contain all the dominant effects of toroidicity in their η dependence. For example, as will be shown in the next section, the cylindrical limit of the exact solution is contained in the two terms $n = 1, N = 1$ and $n = -1, N = -1$.

We now demonstrate how the truncation procedure is effected for a given value of the ring aspect ratio $\cosh\eta_0$, and, further, show that the number of terms to be retained in the exact solutions for a given accuracy increases as the aspect ratio of the ring decreases.

The associated Legendre functions possess the following monotonic properties:

$$\begin{aligned} &Q_{n-1/2}(\cosh\eta),\ Q'_{n-1/2}(\cosh\eta) \text{ decrease as } \eta \text{ increases,} \\ &P_{n-1/2}(\cosh\eta),\ P'_{n-1/2}(\cosh\eta) \text{ decrease as } \eta \text{ decreases, } n \neq 0. \end{aligned} \tag{5.1}$$

In any truncation we shall always retain the terms $n = -1$ and 1. It then follows that the maximum truncation error in the potentials or the fields for both the regions inside and outside the ring occurs at the ring surface $\eta = \eta_0$, as can be seen by applying Eq. (5.1) to Eqs. (2.24)–(2.27). Thus we need only consider the exact solutions for $\eta = \eta_0$.

TABLE I: $a_{n,N}$ and b_n for aspect ratios of 10, 5, and 3. For values of $n < 0$ use the relation $a_{-n,-N} = a_{n,N}$.

	$\cosh\eta_0$	$n=1$	$n=2$	$n=3$	$n=4$
	10				
$a_{n,-1}$	5				
	3	0.003			
	10	-1.163	-0.049×10^{-1}	-0.17×10^{-2}	
$a_{n,1}$	5	-0.808	-0.007	-0.005×10^{-1}	
	3	-0.594	-0.009	-0.011×10^{-1}	
	10		-0.875×10^{-1}	-0.24×10^{-2}	
$a_{n,2}$	5		-0.123	-0.007×10^{-1}	
	3	-0.004	-0.155	-0.015×10^{-1}	
	10			-0.546×10^{-2}	
$a_{n,3}$	5			-0.155×10^{-1}	-0.016×10^{-2}
	3			-0.343×10^{-1}	-0.099×10^{-2}
	10				
$a_{n,4}$	5				-0.183×10^{-2}
	3			-0.006×10^{-1}	-0.665×10^{-2}
	10				
$a_{n,5}$	5				
	3				-0.002×10^{-2}
	10	-1.163	-0.924×10^{-1}	-0.956×10^{-2}	
b_n	5	-0.808	-0.130	-0.167×10^{-1}	-0.199×10^{-2}
	3	-0.595	-0.164	-0.375×10^{-1}	-0.766×10^{-2}

For each mode number n, we have to evaluate

$$b_n \equiv \sum_{N=-\infty}^{\infty} a_{n,N} \tag{5.2}$$

as a function of $\cosh\eta_0$. The expression for $a_{n,N}$, Eq. (2.26) falls off as $\exp(-|N-n|\eta_0)$ for $|N-n| \gg 1$, and the dominant contribution to (5.2) is always from the $n = N$ term. Thus in our calculated examples, as are shown in Table I,[6] for $\cosh\eta_0 = 10$, 5, and 3, it is not necessary to include more than five terms for each b_n to achieve an accuracy of 1%. For this accuracy the number of terms increases as $\cosh\eta_0$ decreases.

The expressions for the potentials and field components contain summation over n of terms of the form

$$b_n e^{in\tau}, \quad |n|b_n e^{in\tau}, \quad \frac{Q_n^{0\prime}}{Q_n^0} b_n e^{in\tau}, \quad \frac{P_n^{0\prime}}{P_n^0} b_n e^{in},$$
$$|n|Q_n^0 e^{in\tau}, \quad |n^2|Q_n^0 e^{in\tau}, \quad |n|Q_n^0 \frac{P_n^{0\prime}}{P_n^0} e^{in\tau}.$$

With the exception of the second term, the coefficients of $e^{in\tau}$ in all these expressions decrease as n increases, and therefore for a 1% accuracy in the n summation it is only necessary to retain the first few terms in n. The decrease in the value of b_n for increasing $|n|$ is also sufficiently rapid so that the summation over $|n|b_n$ converges to the given accuracy almost as quickly as the other summations. We also observe that for a given accuracy, the number of terms in the n summation to be retained increases as $\cosh\eta_0$ increases.

6. CYLINDRICAL LIMIT

The cylindrical limit of toroidal coordinates for this problem is affected by holding the value of the minor radius of the ring constant and allowing the major radius to increase indefinitely. This transformation expands the ring into an infinite straight dielectric cylinder of radius a, transverse to the uniform electric field. In terms of (r, θ) coordinates in the cross-section of the cylinder, the cylindrical potentials are given by Eqs. (4.5) and (4.6). We shall show that both the exact and approximate solutions in toroidal coordinates reduce to the cylindrical solution in this limit.

The following relationships exist between R, d, and r:

$$R = r\cosh\eta, \quad d = r\sinh\eta. \tag{6.1}$$

Further, the surface $\eta = \eta_0$ will correspond to the cylindrical surface $r = a$ and, therefore,

$$R = a\cosh\eta_0. \tag{6.2}$$

Thus, as we follow a particular point under the transformation R, d, n, and η_0 increases indefinitely, while a and r remain fixed. From the definitions of τ and θ it follows that $\tau \to \theta$ as $R \to \infty$.

For large values of η, the asymptotic forms of the associated Legendre functions are

$$\lim_{\eta\to\infty} P_{n-1/2}(\cosh\eta) \simeq \frac{2^{|n|-1/2}}{\pi^{1/2}} \frac{(|n|-1)!}{\Gamma(|n|+\frac{1}{2})} \times \cosh^{|n|-1/2}\eta, \quad |n| > 0, \tag{6.3}$$

$$\lim_{\eta\to\infty} Q_{n-1/2}(\cosh\eta) \simeq \frac{\pi^{1/2}}{2^{|n|+1/2}} \frac{\Gamma(|n|+\frac{1}{2})}{|n|!} \times \frac{1}{\cosh^{|n|+1/2}\eta}, \tag{6.4}$$

$$\lim_{\eta\to\infty} P_{-1/2}(\cosh\eta) \simeq \frac{1}{\pi\cosh\eta/2} \tag{6.5}$$

Applying the cylindrical limit to the series of Q-type harmonics in the expression for Φ^0, (2.25) and substituting for d from Eq. (6.1), we obtain, using Eq. (6.4),

$$\frac{-iE_0 r}{\pi^{1/2}} \lim_{\eta\to\infty} \sum_{n=-\infty}^{\infty} \frac{n\Gamma(|n|+\frac{1}{2})e^{in\theta}}{|n|!\,2^{|n|-1}\cosh^{|n|-1}\eta} = E_0 r\sin\theta,$$

which is the representation of the uniform field in Eq. (3.6).

To evaluate the cylindrical limit of the remaining terms in Eqs. (2.24) and (2.25) we need the asymptotic forms of Π_n^0, q_n, α_n, and β_n for large η_0. Substituting Eqs. (6.3)–(6.5) in Eq. (2.9) yields

$$\lim_{\eta_0\to\infty} \Pi_n^0 \simeq \frac{-[(\epsilon-1)+2|n|(\epsilon+1)]}{2\cosh\eta_0}, \quad n = 0, \pm1, \ldots, \tag{6.6}$$

and from Eq. (2.12),

$$\lim_{\eta_0\to\infty} q_n \simeq \frac{4|n|(\epsilon+1)\cosh\eta_0}{[(\epsilon-1)+2|n|(\epsilon+1)]}, \quad n = \pm1, \ldots,$$
$$\simeq 2\cosh\eta_0, \quad n = 0. \tag{6.7}$$

It therefore follows from Eqs. (2.25) and (2.26) that for all n

$$\lim_{\eta_0\to\infty} \alpha_n = \lim_{\eta_0\to\infty} \beta_n = 0. \tag{6.8}$$

Hence the only terms that remain in the summation over N in the expressions for Φ^i and Φ^0 are those for which $N = n$. This holds for both the exact and approximate expressions.

The cylindrical limit of Φ^i, consequently, is

$$\frac{4\sqrt{2}iE_0}{\pi} \lim_{\eta\to\infty} \frac{d(\cosh\eta-\cos\tau)^{1/2}}{\sinh^2\eta_0} \times \cosh\eta_0 \sum_{n=-\infty}^{\infty} \frac{nQ_{n-1/2}(\cosh\eta)e^{in\tau}}{\Pi_n^0 q_n^0 Q_n^0 P_n^0}$$
$$= \frac{-2\sqrt{2}iE_0}{\pi(\epsilon+1)} \lim_{\eta\to\infty} \frac{d\cosh^{1/2}\eta}{\cosh\eta_0} \times \sum_{n=\infty}^{\infty} \frac{nQ_{n-1/2}(\cosh\eta)e^{in\theta}}{|n|Q_n^0 P_n^0},$$

using Eqs. (6.6) and (6.7). Finally, using Eqs. (6.1) and (6.4), we have

$$\lim_{\eta\to\infty} \Phi^i \simeq \frac{-2iE_0 r}{\pi^{1/2}(\epsilon+1)} \times \lim_{\eta\to\infty} \sum_{n=-\infty}^{\infty} \frac{n\Gamma(|n|+\frac{1}{2})e^{in\theta}}{|n|2^{|n|-1}(|n|-1)!\cosh^{|n|-1}\eta}$$
$$= \frac{2E_0 r}{(\epsilon+1)}\sin\theta, \tag{6.9}$$

which is identical with Eq. (4.5).

Similarly, the cylindrical limit of the field perturbation in Φ^0 is

$$\frac{2\sqrt{2}iE_0}{\pi}\lim_{\eta\to\infty} d(\cosh\eta-\cos\tau)^{1/2}\sum_{n=-\infty}^{\infty}\frac{nP_{n-1/2}(\cosh\eta)}{P_n^0}\times\left(Q_n^0+\frac{2\cosh\eta_0}{\Pi_n^0 q_n\sinh^2\eta_0}\right)e^{in\tau}=\frac{iE_0(\epsilon-1)}{\pi^{1/2}(\epsilon+1)}r\times\lim_{\eta\to\infty}\sum_{n=-\infty}^{\infty}\frac{n\Gamma(|n|+\frac{1}{2})\cosh^{|n|+1}\eta}{|n|!2^{|n|-1}\cosh^{2|n|}\eta_0}e^{in\theta}, \quad (6.10)$$

by Eqs. (6.1), (6.3), (6.6), and (6.7). Hence

$$\lim_{\eta\to\infty}\Phi^0=E_0r\sin\theta-[E_0(\epsilon-1)/(\epsilon+1)]\times(a^2/r)\sin\theta,$$

which is exactly Eq. (4.6).

It is clear from Eqs. (6.9) and (6.10) that the only terms in the n summation that contribute to the cylindrical limit are those for which $n=-1$ and $n=1$. Since the truncation procedure of the previous section always retains these two terms, it follows that the cylindrical limit of the truncated potentials will approach the same limit as the exact solutions.

[1] W. M. Hicks, Phil. Trans. **176**, 161 (1884).
[2] E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics* (Chelsea, New York, 1955), p. 433.
[3] *Higher Transcendental Functions* edited by A. Erdelyi (McGraw-Hill, New York, 1953), Vol. I.
[4] L. M. Milne Thomsom, *The Calculus of Finite Differences* (MacMillan, London 1960), p. 531.
[5] A proof of this result is given in Ref. 4, pp. 532-34.
[6] The values of P_n^0 and Q_n^0 are taken from *Tables of Associated Legendre Functions*, National Bureau of Standards (Columbia U.P., New York, 1945).

Unitary Representations of the Homogeneous Lorentz Group in an $O(1,1)\otimes O(2)$ Basis and Some Applications to Relativistic Equations

E. G. Kalnins*

Department of Applied Mathematics, The University of Western Ontario, London, Ontario, Canada

(Received 6 March 1972)

Unitary irreducible representations of the homogeneous Lorentz group $O(3,1)$ belonging to the principal series are reduced with respect to the subgroup $O(1,1)\otimes O(2)$. As an application we determine the mixed basis matrix elements between $O(3)$ and $O(1,1)\otimes O(2)$ bases and derive recurrence relations for them. This set of functions is then used to obtain invariant expansions of solutions of the Dirac and Proca free field equations. These expansions are shown to have the correct nonrelativistic limit.

INTRODUCTION

In recent years there has been considerable interest in the unitary irreducible representations (UIR's) of the homogeneous Lorentz group in various bases.[1,2] Harmonic analysis of a scalar function in terms of the four subgroup bases [i.e., $O(3)$, $O(2,1)$, $E(2)$, and $O(1,1)\otimes O(2)$] has first been given by Smorodinski and Vilenkin.[2] Since this work most of the attention has been paid to the little group bases as these also play a role in the usual Poincaré invariant partial wave analysis[3,4] of scalar functions and helicity amplitudes. The properties of the reduction of $O(3,1)$ with respect to $O(1,1)\otimes O(2)$ are, however, not so well known. It is the purpose of this paper to develop these properties and indicate some possible uses. The content of the paper is arranged as follows. In Sec. 1 we collect the pertinent facts concerning $SL(2,C)$ [the covering group of $O(3,1)$], its Lie algebra and UIR's. In Sec. 2 we carry out the reduction of the principal series of $SL(2,C)$ with respect to $D(1,1)\otimes D(2)$ (see Sec. 2) the universal covering group of $O(1,1)\otimes O(2)$. The action of the infinitesimal generators of the Lie algebra in such a basis is also determined. In Sec. 3 we develop the expansion of a single particle helicity state in terms of mixed basis matrix elements. An explicit expression for these matrix elements is obtained for the first time. In Sec. 4 we derive recurrence relations for these mixed basis matrix elements, which are used in Sec. 5 to develop invariant expansions of solutions of the free field Proca and Dirac equations. Finally in Sec. 6 the nonrelativistic limit of these solutions is obtained.

1. RESUMÉ OF $SL(2,C)$ AND ITS UIR'S

The group $SL(2,C)$[5] is the universal covering group of the homogeneous Lorentz group $O(3,1)$. The elements of $SL(2,C)$ are the unimodular complex matrices in two dimensions

$$g=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix},\qquad \alpha\delta-\beta\gamma=1. \quad (1.1)$$

The subgroup $SU(2)$ consists of all unitary unimodular matrices of the form

$$\begin{pmatrix}\alpha & \beta\\ -\bar{\beta} & \bar{\alpha}\end{pmatrix},\qquad |\alpha|^2+|\beta|^2=1. \quad (1.2)$$

$SU(2)$ is of course the covering group of $O(3)$ the real orthogonal group in three dimensions. The covering group of $O(1,1)\otimes O(2)$ is denoted by $D(1,1)\otimes D(2)$ and consists of all diagonal unimodular matrices:

$$\begin{pmatrix}\alpha & 0\\ 0 & \beta\end{pmatrix},\qquad \alpha\beta=1. \quad (1.3)$$

[Note: $D(2)$ is the set of all diagonal matrices of the form

$$R(\psi)=\begin{pmatrix}e^{i\psi/2} & 0\\ 0 & e^{-i\psi/2}\end{pmatrix},\qquad 0\le\psi\le 2\pi,$$

such that to each rotation in the plane of the group $O(2)$ there corresponds the matrices $\pm R(\psi)$. This is just the usual two to one homomorphism between an orthogonal group and its spinor group. Similar remarks apply to $D(1,1)$ the set of matrices

which is identical with Eq. (4.5).

Similarly, the cylindrical limit of the field perturbation in Φ^0 is

$$\frac{2\sqrt{2}iE_0}{\pi}\lim_{\eta\to\infty} d(\cosh\eta-\cos\tau)^{1/2}\sum_{n=-\infty}^{\infty}\frac{nP_{n-1/2}(\cosh\eta)}{P_n^0}\times\left(Q_n^0+\frac{2\cosh\eta_0}{\Pi_n^0 q_n\sinh^2\eta_0}\right)e^{in\tau}=\frac{iE_0(\epsilon-1)}{\pi^{1/2}(\epsilon+1)}r\times\lim_{\eta\to\infty}\sum_{n=-\infty}^{\infty}\frac{n\Gamma(|n|+\frac{1}{2})\cosh^{|n|+1}\eta}{|n|!2^{|n|-1}\cosh^{2|n|}\eta_0}e^{in\theta}, \quad (6.10)$$

by Eqs. (6.1), (6.3), (6.6), and (6.7). Hence

$$\lim_{\eta\to\infty}\Phi^0=E_0r\sin\theta-[E_0(\epsilon-1)/(\epsilon+1)]\times(a^2/r)\sin\theta,$$

which is exactly Eq. (4.6).

It is clear from Eqs. (6.9) and (6.10) that the only terms in the n summation that contribute to the cylindrical limit are those for which $n=-1$ and $n=1$. Since the truncation procedure of the previous section always retains these two terms, it follows that the cylindrical limit of the truncated potentials will approach the same limit as the exact solutions.

[1] W. M. Hicks, Phil. Trans. **176**, 161 (1884).
[2] E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics* (Chelsea, New York, 1955), p. 433.
[3] *Higher Transcendental Functions* edited by A. Erdelyi (McGraw-Hill, New York, 1953), Vol. I.
[4] L. M. Milne Thomsom, *The Calculus of Finite Differences* (MacMillan, London 1960), p. 531.
[5] A proof of this result is given in Ref. 4, pp. 532-34.
[6] The values of P_n^0 and Q_n^0 are taken from *Tables of Associated Legendre Functions*, National Bureau of Standards (Columbia U.P., New York, 1945).

Unitary Representations of the Homogeneous Lorentz Group in an $O(1,1)\otimes O(2)$ Basis and Some Applications to Relativistic Equations

E. G. Kalnins*

Department of Applied Mathematics, The University of Western Ontario, London, Ontario, Canada

(Received 6 March 1972)

Unitary irreducible representations of the homogeneous Lorentz group $O(3,1)$ belonging to the principal series are reduced with respect to the subgroup $O(1,1)\otimes O(2)$. As an application we determine the mixed basis matrix elements between $O(3)$ and $O(1,1)\otimes O(2)$ bases and derive recurrence relations for them. This set of functions is then used to obtain invariant expansions of solutions of the Dirac and Proca free field equations. These expansions are shown to have the correct nonrelativistic limit.

INTRODUCTION

In recent years there has been considerable interest in the unitary irreducible representations (UIR's) of the homogeneous Lorentz group in various bases.[1,2] Harmonic analysis of a scalar function in terms of the four subgroup bases [i.e., $O(3)$, $O(2,1)$, $E(2)$, and $O(1,1)\otimes O(2)$] has first been given by Smorodinski and Vilenkin.[2] Since this work most of the attention has been paid to the little group bases as these also play a role in the usual Poincaré invariant partial wave analysis[3,4] of scalar functions and helicity amplitudes. The properties of the reduction of $O(3,1)$ with respect to $O(1,1)\otimes O(2)$ are, however, not so well known. It is the purpose of this paper to develop these properties and indicate some possible uses. The content of the paper is arranged as follows. In Sec. 1 we collect the pertinent facts concerning $SL(2,C)$ [the covering group of $O(3,1)$], its Lie algebra and UIR's. In Sec. 2 we carry out the reduction of the principal series of $SL(2,C)$ with respect to $D(1,1)\otimes D(2)$ (see Sec. 2) the universal covering group of $O(1,1)\otimes O(2)$. The action of the infinitesimal generators of the Lie algebra in such a basis is also determined. In Sec. 3 we develop the expansion of a single particle helicity state in terms of mixed basis matrix elements. An explicit expression for these matrix elements is obtained for the first time. In Sec. 4 we derive recurrence relations for these mixed basis matrix elements, which are used in Sec. 5 to develop invariant expansions of solutions of the free field Proca and Dirac equations. Finally in Sec. 6 the nonrelativistic limit of these solutions is obtained.

1. RESUMÉ OF $SL(2,C)$ AND ITS UIR'S

The group $SL(2,C)$[5] is the universal covering group of the homogeneous Lorentz group $O(3,1)$. The elements of $SL(2,C)$ are the unimodular complex matrices in two dimensions

$$g=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix},\qquad \alpha\delta-\beta\gamma=1. \quad (1.1)$$

The subgroup $SU(2)$ consists of all unitary unimodular matrices of the form

$$\begin{pmatrix}\alpha & \beta\\ -\bar{\beta} & \bar{\alpha}\end{pmatrix},\qquad |\alpha|^2+|\beta|^2=1. \quad (1.2)$$

$SU(2)$ is of course the covering group of $O(3)$ the real orthogonal group in three dimensions. The covering group of $O(1,1)\otimes O(2)$ is denoted by $D(1,1)\otimes D(2)$ and consists of all diagonal unimodular matrices:

$$\begin{pmatrix}\alpha & 0\\ 0 & \beta\end{pmatrix},\qquad \alpha\beta=1. \quad (1.3)$$

[Note: $D(2)$ is the set of all diagonal matrices of the form

$$R(\psi)=\begin{pmatrix}e^{i\psi/2} & 0\\ 0 & e^{-i\psi/2}\end{pmatrix},\qquad 0\le\psi\le 2\pi,$$

such that to each rotation in the plane of the group $O(2)$ there corresponds the matrices $\pm R(\psi)$. This is just the usual two to one homomorphism between an orthogonal group and its spinor group. Similar remarks apply to $D(1,1)$ the set of matrices

$$\pm\begin{pmatrix} e^{a/2} & 0 \\ 0 & e^{-a/2} \end{pmatrix}, \quad -\infty < a < +\infty\Big].$$

The Lie algebra of $SL(2, C)$ is six dimensional, being spanned by the generators $M_i, N_i (i = 1, 2, 3)$ which satisfy the commutation relations

$$[M_i, M_j] = \epsilon_{ijk} M_k, \quad [M_i, N_j] = \epsilon_{ijk} N_k, \quad [N_i, N_j] = -\epsilon_{ijk} M_k. \tag{1.4}$$

There are two independent Casimir invariants of $SL(2, C)$ which label each irreducible representation. They are

$$K_1 = \mathbf{M}^2 - \mathbf{N}^2, \quad K_2 = \mathbf{M}\cdot\mathbf{N}. \tag{1.5}$$

The Casimir invariant of $SU(2)$ is well known to be $\mathbf{M}^2$. Each inequivalent UIR of $SU(2)$ is labeled by the eigenvalue j, where

$$\mathbf{M}^2 = -j(j+1), \quad j = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \cdots. \tag{1.6}$$

Each UIR for given j is $(2j+1)$-dimensional and the spectrum of M_3 in it is

$$M_3 = -j, -j+1, \ldots, j-1, j. \tag{1.7}$$

A UIR of $D(1,1) \otimes D(2)$ is labeled by the two eigenvalues of M_3 and N_3, $\{m, \tau\}$ where

$$-\infty < \tau < +\infty, \quad m = 0, \pm\tfrac{1}{2}, \pm 1, \pm\tfrac{3}{2}, \cdots. \tag{1.8}$$

It is easy to see that each such UIR is one-dimensional.

We now give the spectrum of the Casimir operators K_1, K_2 corresponding to the principal series $\{j_0, \rho\}$ of $SL(2, C)$ together with the spectrum of j values of the UIR's of $SU(2)$ that appear in each such UIR of $SL(2, C)$. For the principal series

$$K_1 = 1 + \rho^2 - j_0^2, \quad K_2 = -\rho j_0, \quad j_0 = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \ldots, \quad -\infty < \rho < +\infty, \tag{1.9}$$

and the spectrum of j values is

$$j = j_0, j_0 + 1 \cdots.$$

The other set of UIR's of $SL(2, C)$ belong to the complementary series which we write as $\{0, i\rho\}$, where

$$K_1 = 1 - \rho^2, \quad K_2 = 0, \quad 0 < \rho < 1, \quad j_0 = 0, 1, 2, \cdots.$$

This set of UIR's does not figure in the completeness relation[5] of $SL(2, C)$ and so will not be considered subsequently.

Finally in this section we give the formulas for the action of the generators M_i, N_i on an $SU(2)$ basis of the principal series

$$\begin{aligned}
M_3 |j,m\rangle &= m |j,m\rangle, \\
M_+ |j,m\rangle &= -i\alpha^j_{\lambda+1} |j, m+1\rangle, \\
M_- |j,m\rangle &= -i\alpha^j_{\lambda} |j, m-1\rangle, \\
N_3 |j,m\rangle &= -i\sqrt{[j^2 - m^2]}\, C_j |j-1, m\rangle \\
&\quad + i A_j m |j,m\rangle \\
&\quad + i C_{j+1}\sqrt{[(j+1)^2 - m^2]} |j+1, m\rangle, \qquad (1.10) \\
N_+ |j,m\rangle &= -i C_j \sqrt{[(j-m)(j-m-1)]} |j-1, m+1\rangle \\
&\quad + i A_j \sqrt{[(j-m)(j+m+1)]} |j, m+1\rangle \\
&\quad - i C_{j+1}\sqrt{[(j+m+1)(j+m+2)]} |j+1, m+1\rangle, \\
N_- |j,m\rangle &= i C_j \sqrt{[(j+m)(j+m-1)]} |j-1, m-1\rangle \\
&\quad + i A_j \sqrt{[(j+m)(j-m+1)]} |j, m-1\rangle \\
&\quad + i C_{j+1}\sqrt{[(j-m+1)(j-m+2)]} |j+1, m-1\rangle,
\end{aligned}$$

where

$$A_j = \frac{-j_0\rho}{j(j+1)}, \quad C_j = \frac{i}{j}\left(\frac{(j^2 - j_0^2)(j^2 + \rho^2)}{4j^2 - 1}\right)^{1/2},$$

$$m = -j, -j+1, \ldots, j, \quad j = j_0, j_0 + 1, \ldots,$$

and $|j,m\rangle$ is an abbreviation for $|\rho j_0; jm\rangle$:

$$\alpha^j_\lambda = \sqrt{[j(j+1) - \lambda(\lambda - 1)]}.$$

2. REDUCTION OF THE PRINCIPAL SERIES OF $SL(2, C)$ UNDER $O(1,1) \otimes O(2)$

As is well known[5] the principal series of $SL(2, C)$ is realized via unitary transformations in a Hilbert space H of square integrable functions in a certain domain. The elements of H are specified by functions $f(z)$ of a single complex variable z varying over the entire complex plane. (This specification is only possible up to sets of measure zero.) The scalar product and norm are given by

$$(f,h) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, \overline{f(z)} h(z), \quad z = x + iy, \quad \|f\| = (f,f)^{1/2} < \infty. \tag{2.1}$$

In the UIR $\{j_0, \rho\}$ of the principal series, the unitary operator $U(g)$ representing the group element g acts on $f(z)$ in the following way:

$$[U(g)f](z) = (\delta + \beta z)^{j_0 - 1 + i\rho}(\bar{\delta} + \bar{\beta}\bar{z})^{-j_0 - 1 + i\rho} \times f[(\alpha z + \gamma)/(\beta z + \delta)] \tag{2.2}$$

This realization is not the most convenient one for our purposes. In order to realize the principal series in a $D(1,1) \otimes D(2)$ basis, we make the following transformation:

$$e^a = (x^2 + y^2)^{1/2}, \quad \tan\phi = y/x, \quad -\infty \le a \le +\infty, \quad 0 \le \phi \le 2\pi. \tag{2.3}$$

Instead of specifying an element of H by $f(z)$ we specify it by the new function

$$\tilde{f}(a, \phi) = e^{-ij_0\phi} e^{a(1-i\rho)} f(z). \tag{2.4}$$

With this indentification the scalar product can be written

$$(f,h) = \int_0^{2\pi} d\phi \int_{-\infty}^{\infty} da\, \overline{\tilde{f}(a,\phi)} \tilde{h}(a,\phi). \tag{2.5}$$

The generators M_i, N_i acting on the $\tilde{f}(a, \phi)$ functions can be expressed as differential operators acting on a and ϕ as

$$M_1 = j_0 \cosh a \cos\phi - (\rho + i)\sinh a \sin\phi + i\left(\sinh a \cos\phi \frac{\partial}{\partial\phi} - \cosh a \sin\phi \frac{\partial}{\partial a}\right),$$
$$M_3 = -i\frac{\partial}{\partial\phi},$$
$$N_1 = j_0 \sinh a \sin\phi + (\rho + i)\cosh a \cos\phi + i\left(\cosh a \sin\phi \frac{\partial}{\partial\phi} + \sinh a \cos\phi \frac{\partial}{\partial a}\right),$$
$$N_3 = -i\frac{\partial}{\partial a}. \tag{2.6}$$

The operators M_2, N_2 can be obtained from the expressions for M_1 and N_1, respectively, via the substitution $\phi \to -\frac{1}{2}\pi + \phi$. The principal series of $SL(2, C)$ is now realized as the set of functions $\tilde{f}(a, \phi)$ on the domain $(-\infty, +\infty) \otimes [0, 2\pi]$ which satisfy

$$(f, f) = \int_0^{2\pi} d\phi \int_{-\infty}^{\infty} da\, |\tilde{f}(a, \phi)|^2 < \infty. \tag{2.7}$$

The two Casimir invariants of $O(1, 1) \otimes O(2)$ are N_3 and M_3, so that the simultaneous eigenfunctions of N_3 and M_3 in this realization are

$$\Psi_{\tau m} = [1/(2\pi)]e^{i\tau a}e^{im\phi}, \tag{2.8}$$

where

$$N_3\Psi_{\tau m} = \tau\Psi_{\tau m}, \quad M_3\Psi_{\tau m} = m\Psi_{\tau m}, \quad (\Psi_{\tau' m'}, \Psi_{\tau m}) = \delta_{m' m}\delta(\tau' - \tau); \tag{2.9}$$

so together with the completeness relations[6]

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i(a'-a)\tau}\, d\tau = \delta(a' - a) \tag{2.10a}$$

$$\frac{1}{2\pi}\sum_{p=-\infty}^{\infty} e^{ip(\phi-\varphi')} = \sum_{n=-\infty}^{\infty}\delta(\phi - \phi' - 2\pi n), \tag{2.10b}$$

we get the following result.

Each UIR $\{j_0, \rho\}$ of the principal series of $SL(2, C)$ contains each UIR $\{m, \tau\}$ of $D(1, 1) \otimes D(2)$ exactly once, provided

$$m = j_0, j_0 \pm 1, j_0 \pm 2, \cdots. \tag{2.11}$$

Thus each $\tilde{f} \in H$ can be expanded in terms of the eigenfunctions $\Psi_{\tau m}$ according to

$$\tilde{f} = \sum_{m=-\infty}^{\infty}\int_{-\infty}^{\infty} d\tau\, f_m(\tau)\Psi_{\tau m}, \qquad f_m(\tau) = \int_0^{2\pi} d\phi \int_{-\infty}^{\infty} da\, \tilde{f}\overline{\Psi_{\tau m}}. \tag{2.12}$$

Finally in this section we calculate the action of the generators $M_\pm, N_\pm$ on the $\Psi_{\tau m}$ basis

$$M_\pm\Psi_{\tau m} = \tfrac{1}{2}(j_0 \pm i\rho \mp i\tau \mp 1 - m)\Psi_{\tau - i, m\pm 1} + \tfrac{1}{2}(j_0 \mp i\rho \mp i\tau \pm 1 + m)\Psi_{\tau + i, m\pm 1},$$
$$N_\pm\Psi_{\tau m} = \tfrac{1}{2}(\mp ij_0 + \rho - \tau + i \pm im)\Psi_{\tau - i, m\pm 1} + \tfrac{1}{2}(\pm ij_0 + \rho + \tau + i \pm im)\Psi_{\tau + i, m\pm 1},$$
$$N_\pm = N_1 \pm iN_2, M_\pm = M_1 \pm iM_2; \tag{2.13}$$

The action of M_3 and N_3 already having been given in Eq. (2.9).

3. CALCULATION OF THE MIXED BASIS MATRIX ELEMENTS AND SINGLE PARTICLE HELICITY STATES

In this section we construct relativistic functions with helicity for nonvanishing mass which are at the same time basis functions of a UIR $\{j_0, \rho\}$ of the homogeneous Lorentz group $O(3, 1)$ realized on the upper sheet of a double sheeted hyperboloid. In order to do this we use the method of Integral geometry.[2,7] In this method[7,8] a one-particle state of spin s, helicity λ, and four velocity u, denoted by $|u, s, \lambda\rangle$, is expressed in terms of a function on the light cone $\Phi_{j_0\rho}(\xi)$ via the relation

$$|u, s, \lambda\rangle = \frac{1}{2(2\pi)^3}\sum_{j_0=-s}^{s}\int_{-\infty}^{\infty} d\rho(\rho^2 + j_0^2) \times \int_\Gamma [u, \xi]^{-1-i\rho} D^s_{\lambda j_0}(R)\Phi_{j_0\rho}(\xi)d^2\xi, \tag{3.1}$$

where Γ is the integration path on the light cone, $d^2\xi$ the invariant measure on the cone, and $[u, \xi]$ the usual Lorentz scalar product

$$[u, \xi] = u_0\xi_0 - \mathbf{u}\cdot\boldsymbol{\xi}. \tag{3.2}$$

The rotation specified by $D_{\lambda j_0}(R)$ is the rotation necessary to account for the requantization of the helicity component from the direction ξ to that of $\mathbf{u}$. The parametrization of the four velocity u in the coordinate system of interest (the C system or cylindrical system[2]) is

$$u = (\cosh a \cosh b, \sinh a \cos\psi, \sinh a \sin\psi, \cosh a \sinh b), \tag{3.3}$$

and the 4-vector ξ is parametrized by

$$\xi = e^c(\cosh\beta, \cos\phi, \sin\phi, \sinh\beta). \tag{3.4}$$

The choice of Γ for the C system is $\xi_0^2 - \xi_3^2 = 1$, and the consequent invariant measure is $d^2\xi = d\phi d\beta$.

In the realization on the cone the generators of the Lorentz group corresponding to a "photon" of discrete helicity λ are[9]

$$M_1 = -i(\xi, \nabla)_1 + \lambda[\xi_1/(\xi_0 + \xi_3)],$$
$$M_2 = -i(\xi, \nabla)_2 + \lambda[\xi_2/(\xi_0 + \xi_3)],$$
$$M_3 = -i(\xi, \nabla)_3 + \lambda,$$
$$N_1 = -i\xi_0\frac{\partial}{\partial\xi_1} - \lambda\frac{\xi_2}{\xi_0 + \xi_3},$$
$$N_2 = -i\xi_0\frac{\partial}{\partial\xi_2} + \lambda\frac{\xi_1}{\xi_0 + \xi_3},$$
$$N_3 = -i\xi_0\frac{\partial}{\partial\xi_3}. \tag{3.5}$$

For the parametrization (3.4) of ξ, the Casimir invariants have the form

$$\mathbf{M}^2 - \mathbf{N}^2 = \frac{d^2}{dc^2} + 2\frac{d}{dc} + \lambda^2, \quad \mathbf{M}\cdot\mathbf{N} = i\lambda\left(1 + \frac{d}{dc}\right). \tag{3.6}$$

From (3.4) and (3.6) it is not hard to show that the simultaneous eigenfunctions of $\mathbf{M}^2 - \mathbf{N}^2$, $\mathbf{M}\cdot\mathbf{N}$, M_3 and N_3 have the form

$$\mathcal{C}_{p\lambda}(\tau,\rho) = e^{-(1-i\rho)c} e^{ip\phi} e^{i\tau\beta} e^{-i\lambda\phi}, \tag{3.7}$$

in particular, on the C system contour

$$\mathcal{C}_{p\lambda}(\tau,\rho) = e^{ip\phi} e^{i\tau\beta} e^{-i\lambda\phi}. \tag{3.8}$$

The function $\Phi_{j_0\rho}(\xi)$ is now expanded in terms of the $\mathcal{C}_{p\lambda}(\tau,\rho)$ functions according to

$$\Phi_{j_0\rho}(\xi) = \sum_{p,\tau} a_p^{j_0}(\tau,\rho)\mathcal{C}_{pj_0}(\tau,\rho). \tag{3.9}$$

For evaluation of the integral over $d^2\xi$ in (3.1), it is most convenient to assume u in the form

$$u = u_0 = (\cosh a, \sinh a, 0, 0); \tag{3.10}$$

the required expansion for the more general form of u can be obtained by using the simple group properties of the $O(1,1) \otimes O(2)$ matrix elements. So combining (3.9) and (3.1) requires the calculation of the following integral:

$$I = \int_0^{2\pi} d\phi \int_{-\infty}^{\infty} d\beta\, (\cosh a \cosh\beta - \sinh a \cos\phi)^{-1-i\rho} \times D^s_{\lambda j_0}(R)\mathcal{C}_{pj_0}(\tau,\rho). \tag{3.11}$$

We now turn our attention to the explicit form of $D^s_{\lambda j_0}(R)$. For this it is convenient to write

$$\mathbf{n} = ((\cos\phi/\cosh\beta), (\sin\phi/\cosh\beta), \tanh\beta), \tag{3.12}$$

the direction vector of the photon 3-momentum. Now if $\mathbf{n}$ is rotated by $-\phi$ about the z axis, $\mathbf{n}$ becomes

$$\mathbf{n} \to \mathbf{n}_0 = ((1/\cosh\beta), 0, \tanh\beta). \tag{3.13}$$

According to the prescription of Ref. 7, the remaining rotation is a rotation in the xz plane by an amount η given by

$$\cos\eta = \frac{u_0\cos\theta - |\mathbf{u}|}{u_0 - \mathbf{u}\cdot\mathbf{n}}, \tag{3.14}$$

where θ is the angle between $\mathbf{n}_0$ and $\mathbf{u} = (\sinh a, 0, 0)$. In our case

$$\cos\theta = 1/\cosh\beta$$

and

$$\cos\eta = \frac{\cosh a - \cosh\beta \sinh a}{\cosh a \cosh\beta - \sinh a}, \tag{3.15}$$

so that we finally have

$$R = M_3(\tfrac{1}{2}\pi - \phi)M_1(\eta)M_3(-\tfrac{1}{2}\pi). \tag{3.16}$$

The integral I can now be evaluated. It is found to be given by

$$I = \frac{e^{i\pi(\lambda - j_0)}}{\Gamma(1+i\rho)} \sum_{r_i} \frac{\Gamma(1+2r_1+i\rho)}{\Gamma(r_1+1-\frac{1}{2}\bar{p})\Gamma(r_1+1+\frac{1}{2}\bar{p})} \times \Lambda_{sr_4,\lambda j_0} \frac{(2r_1+1+i\rho)_{r_2}(-i\tau)_{r_3}(-i\tau+\frac{1}{2})_{r_3}}{(\frac{1}{2})_{r_3} r_2!\, r_3!} \times \frac{\Gamma(s)\Gamma(c-b)}{\Gamma(c)} e^{a(\lambda - j_0 + 2r_4)} \times (\tfrac{1}{2}\tanh a)^{2r_1} (\cosh a)^{-1-i\rho}\, {}_2F_1(s,b;c;-e^{2a}), \tag{3.17}$$

where

$$b = \tfrac{1}{2}(\lambda - j_0) + r_2 + r_3 + r_4 + \tfrac{1}{2},$$

$$c = 2r_1 + i\rho - i\tau + \tfrac{1}{2}(\lambda - j_0) + r_2 + r_3 + r_4 + \tfrac{3}{2},$$

$$\Lambda_{sr_4,\lambda j_0} = [\Gamma(s+\lambda+1)\Gamma(s-\lambda+1)\Gamma(s+j_0+1) \times \Gamma(s-j_0+1)]^{1/2}[\Gamma(s-\lambda-j_0+1) \times \Gamma(s+j_0-r_4+1)\Gamma(r_4+\lambda-j_0+1) \times \Gamma(r^4+1)]^{-1}$$

$$(d)_n = \Gamma(d+n)/\Gamma(d), \qquad \bar{p} = p - \lambda.$$

We now identify I with the mixed basis matrix element in the following way:

$$\langle \rho j_0; s\lambda | N_1(a) | \rho j_0; \tau p\rangle = C^{\rho j_0}_{s\lambda,\tau p}(a) = I. \tag{3.18}$$

The expansion of a single particle helicity state in terms of C system matrix elements is then

$$|u; s, \lambda\rangle = \frac{1}{2(2\pi)^3} \sum_{j_0=-s}^{s} \int_{-\infty}^{\infty} d\rho(\rho^2 + j_0^2) \times \sum_{p=-\infty}^{\infty} \int_{-\infty}^{\infty} d\tau a_p^{j_0}(\tau,\rho) C^{\rho j_0}_{s\lambda;\tau M}(a)\, e^{i\tau b}\, e^{ip\psi}. \tag{3.19}$$

4. RECURRENCE RELATIONS FOR THE MIXED BASIS MATRIX ELEMENTS

In this section we use the infinitesmal operator method[10,11] to establish recurrence relations and differential equations for the mixed basis matrix elements. For this method we use a fixed column of the mixed basis matrix element $\langle \rho j_0; JM | L | \rho j_0; \tau p\rangle$ (i.e., τ and p fixed) as a set of $SU(2)$ basis vectors spanning the UIR$\{j_0, \rho\}$ of $SL(2,C)$. L is a general lorentz transformation. The generators M_i, N_i are then differential operators acting on the six parameters needed to specify L. Now using Eqs. (1.10) and (1.9) and making a particular choice for L we can derive the relations we need. For the C system we parametrize L as follows

$$L = M_3(\phi)M_1(\theta)M_3(\alpha)N_1(a)N_3(b)M_3(\psi), \tag{4.1}$$

so that the mixed basis matrix element is

$$\langle \rho j_0; JM | L | \rho j_0; \tau p\rangle = \bar{C}^{\rho j_0}_{JM,\tau p} = \sum_{\lambda} D^J_{M\lambda}(\phi,\theta,\alpha) C^{\rho j_0}_{J\lambda,\tau p}(a) e^{i\tau b} e^{ip\psi}. \tag{4.2}$$

The generators M_i, N_i corresponding to the parametrization (4.1) are

$$M_1 = -\cot\theta \sin\phi \frac{\partial}{\partial\phi} + \cos\phi \frac{\partial}{\partial\phi} + \frac{\sin\phi}{\sin\theta}\frac{\partial}{\partial\alpha},$$

$$M_3 = \frac{\partial}{\partial\phi},$$

$$N_1 = -\sin\phi\cos\alpha\tanh a \frac{\partial}{\partial\phi} + \tanh a \sin\phi \sin\theta \sin\alpha \frac{\partial}{\partial\theta} - \frac{1}{\sinh a \cosh a}(\sin\phi\cos\theta\cos\alpha + \cos\phi\sin\alpha\cosh^2 a) \times \frac{\partial}{\partial\alpha} + (\cos\phi\cos\alpha - \sin\phi\sin\alpha\cos\theta)\frac{\partial}{\partial a} + \frac{\sin\phi\sin\theta}{\cosh a} \times \frac{\partial}{\partial b} + \frac{1}{\sinh a}(\cos\phi\sin\alpha + \sin\phi\cos\alpha\cos\theta)\frac{\partial}{\partial\psi},$$

$$N_3 = -\cot\theta\cos\alpha\tanh a \frac{\partial}{\partial\phi} + \tanh a\cos\theta\sin\alpha\frac{\partial}{\partial\theta} + \cos\alpha(\tanh a\cot\theta\cos\theta + \sin\theta\coth a)\frac{\partial}{\partial a} + \sin\theta\sin\alpha\frac{\partial}{\partial\alpha}. \tag{4.3}$$

M_2 and N_2 can be obtained from M_1 and N_1, respectively, via the transformation $\phi \to -\frac{1}{2}\pi + \phi$.

In the $SU(2)$ basis we have chosen, the Casimir invariant equations have the form

$$(\mathbf{M}^2 - \mathbf{N}^2)\,\bar{C}^{\rho j_0}_{JM,\tau p} = (1 + \rho^2 - j_0^2)\;\bar{C}^{\rho j_0}_{JM,\tau p},$$
$$\mathbf{M}\cdot\mathbf{N}\,\bar{C}^{\rho j_0}_{JM,\tau p} = \rho j_0 \bar{C}^{\rho j_0}_{JM,\tau p}. \quad (4.4)$$

The explicit expression of the Casimir invariants in terms of differential operators is found from (4.3) to be

$$\mathbf{N}^2 - \mathbf{M}^2 = \frac{\partial^2}{\partial a^2} + (\tanh a + \coth a)\frac{\partial}{\partial a} + \frac{1}{\cosh^2 a}\frac{\partial^2}{\partial b^2} + \frac{1}{\sinh^2 a}\frac{\partial^2}{\partial \psi^2} - 2\frac{\coth a}{\sinh a}\frac{\partial^2}{\partial\alpha\partial\psi} \tanh^2 a\,\bar{M}_2^2 - \mathbf{M}^2 + 2\frac{\tanh a}{\cosh a}\bar{M}_2\frac{\partial}{\partial b} + \coth^2 a\frac{\partial^2}{\partial\alpha^2} \quad (4.5)$$

$$\mathbf{M}\cdot\mathbf{N} = \bar{M}_1\left(\frac{\partial}{\partial a} + \tanh a\right) + \bar{M}_2\left((\tanh a - \coth a)\frac{\partial}{\partial\alpha} + \frac{1}{\sinh a}\frac{\partial}{\partial\psi}\right) + \frac{1}{\coth a}\frac{\partial^2}{\partial b\,\partial\alpha}, \quad (4.6)$$

where

$$\bar{M}_2 = \cot\theta\cos\alpha\frac{\partial}{\partial\alpha} + \sin\alpha\frac{\partial}{\partial\theta} - \frac{\cos\alpha}{\sin\theta}\frac{\partial}{\partial\phi},$$

$$\bar{M}_1 = -\cot\theta\sin\alpha\frac{\partial}{\partial\alpha} + \cos\alpha\frac{\partial}{\partial\theta} + \frac{\sin\alpha}{\sin\theta}\frac{\partial}{\partial\phi}.$$

So applying the Casimir invariants (4.5) to the $\bar{C}^{\rho j_0}_{JM,\tau p}$ functions and separating out all but the a dependence, using known recurrence relations of the $SU(2)$ matrix elements[12] and the orthogonality properties of the $O(1,1)\otimes O(2)$ matrix elements, we get the relations

$$\alpha^J_\lambda\left(\frac{d}{da} + \lambda\tanh a + (1-\lambda)\coth a + \frac{p}{\sinh a}\right)C^{\rho j_0}_{J,\lambda+1;\tau p} + \alpha^J_{\lambda+1}\left(\frac{d}{da} - \lambda\tanh a + (1+\lambda)\coth a - \frac{p}{\sinh a}\right) \times C^{\rho j_0}_{J,\lambda-1;\tau p} + i\left(\frac{2\lambda\tau}{\cosh a} + \rho j_0\right)C^{\rho j_0}_{J\lambda;\tau p} = 0, \quad (4.7)$$

$$\left(\frac{d^2}{da^2} + (\tanh a + \coth a)\frac{d}{da} - \frac{\tau^2}{\cosh^2 a} - \frac{p^2}{\sinh^2 a} + 2\lambda p\frac{\coth a}{\sinh a} + J(J+1) + \tfrac{1}{2}\tanh^2 a[J(J+1) - \lambda^2\coth^2 a + (1 - j_0^2 + \rho^2)\right)C^{\rho j_0}_{J\lambda;\tau p} + \tfrac{1}{4}\tanh^2 a\,[\alpha^J_{\lambda+1}\alpha^J_{\lambda+2}C^{\rho j_0}_{J,\lambda+2;\tau p} + \alpha^J_\lambda\alpha^J_{\lambda-1}C^{\rho j_0}_{J,\lambda-2;\tau p}] + i\tau\,(\tanh a/\cosh a)\,(\alpha^J_\lambda C^{\rho j_0}_{J,\lambda-1;\tau p} - \alpha^J_{\lambda+1}C^{\rho j_0}_{J,\lambda+1;\tau p}) = 0. \quad (4.8)$$

The remaining recurrence relations are determined from the known action of the generators $N_\pm$ in an $SU(2)$ basis [Eqs. (1, 10)]. They are

$$[(J+\lambda)(J+\lambda+1)]^{1/2}\left(\frac{d}{da} + \frac{p}{\sinh a} + (1-\lambda)\coth a + (J-\lambda+1)\tanh a\right)C^{\rho j_0}_{J,\lambda-1;\tau p} - [(J-\lambda)(J-\lambda+1)]^{1/2}\left(\frac{d}{da} - \frac{p}{\sinh a} + (1+\lambda)\coth a + (J+\lambda+1)\tanh a\right)C^{\rho j_0}_{J,\lambda+1;\tau p} - (2i\tau/\cosh a)\,[(J+1)^2 - \lambda^2]^{1/2}\,C^{\rho j_0}_{J\lambda,\tau p} \quad (4.9)$$

$$= 2\{[(J+1)^2 - j_0^2]\,[(J+1)^2 + \rho^2]\,[(2J+1)/(2J+3)]\}^{1/2}\,C^{\rho j_0}_{J+1,\lambda;\tau p} - [(J-\lambda)(J-\lambda+1)]^{1/2}\left(\frac{d}{da} + \frac{p}{\sinh a} + (1-\lambda)\coth a + (J+\lambda)\tanh a\right)C^{\rho j_0}_{J,\lambda-1;\tau p} + [(J+\lambda)(J+\lambda+1)]^{1/2} \times\left(\frac{d}{da} - \frac{p}{\sinh a} + (1+\lambda)\coth a + (J-\lambda)\tanh a\right)C^{\rho j_0}_{J,\lambda+1;\tau p} + (J^2-\lambda^2)^{1/2}\,(2i\tau/\cosh a)\,C^{\rho j_0}_{J\lambda;\tau p}$$
$$= 2\{(J^2 - j_0^2)(J^2+\rho^2)[(2J+1)/(2J-1)]\}^{1/2}\;C^{\rho j_0}_{J-1;\lambda,\tau p}. \quad (4.10)$$

These relations we have developed here are the ones we will use in the next section in our analysis of the Proca and Dirac fields.

5. SOLUTION OF THE DIRAC AND PROCA FREE FIELD EQUATIONS IN THE C SYSTEM

As an application of the previous three sections we derive invariant expansions of solutions of the Dirac and Proca equations in terms of the functions

$$D^{\rho j_0}_{J\lambda;\tau p}(a,b,\phi) = C^{\rho j_0}_{J\lambda;\tau p}(a)\,e^{i\tau b}e^{ip\phi}. \quad (5.1)$$

This has already been done in the S system for these equations[13] and more general ones.[14,15]

An outline of the general method is as follows. In order to achieve an invariant expansion of an arbitrary field $F^{\rho j_0}_{JM}(x)$, it is convenient to go over into a coordinate system in which each component transforms independently. The components of $F^{\rho j_0}_{JM}(x)$ in this new coordinate system are

$$\bar{F}^{\rho j_0}_{JM}(g) = U(g)F^{\rho j_0}_{JM}(x) = D^{\rho j_0}_{J'M',JM}(g)F^{\rho j_0}_{J'M'}(g^{-1}x). \quad (5.2)$$

From this definition it follows that each component does indeed transform independently:

$$U(g_0)\bar{F}^{\rho j_0}_{JM}(g) = \bar{F}^{\rho j_0}_{JM}(g_0 g) \quad (5.3)$$

so that each component of $\bar{F}^{\rho j_0}_{JM}(g)$ constitutes a representation space for the Lorentz group and can, therefore, be expanded in terms of matrix elements of that group.

We now turn our attention to the Proca field $A_K(x)$ of mass μ, i.e.,

$$(\Box - \mu^2)A_K(x) = \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} - \frac{\partial^2}{\partial x_0^2} - \mu^2\right) \times A_K(x) = 0,\quad \frac{\partial A_K}{\partial x_K} = 0. \quad (5.4)$$

We seek a solution for this equation inside the light

cone, so in the C system we choose x to be parametrized by

$$x = (s\cosh a\cosh b, s\sinh a\cos\phi, s\sinh a\sin\phi, s\cosh a\sinh b). \quad (5.5)$$

The operators $\partial/\partial x_i$ have the form

$$\begin{aligned}
\frac{\partial}{\partial x_0} &= \cosh a\cosh b\,\frac{\partial}{\partial s} - \frac{\sinh a\cosh b}{s}\,\frac{\partial}{\partial a} - \frac{\sinh b}{s\cosh a}\,\frac{\partial}{\partial b},\\
\frac{\partial}{\partial x_1} &= -\sinh a\cos\phi\,\frac{\partial}{\partial s} + \frac{\cosh a\cos\phi}{s}\,\frac{\partial}{\partial a} - \frac{\sin\phi}{s\sinh a}\,\frac{\partial}{\partial\phi},\\
\frac{\partial}{\partial x_2} &= -\sinh a\sin\phi\,\frac{\partial}{\partial s} + \frac{\cosh a\sin\phi}{s}\,\frac{\partial}{\partial a} + \frac{\cos\phi}{s\cosh a}\,\frac{\partial}{\partial b}\\
\frac{\partial}{\partial x_3} &= -\sinh b\cosh a\,\frac{\partial}{\partial s} + \frac{\sinh a\sinh b}{s}\,\frac{\partial}{\partial a}\\
&\quad + \frac{\cosh b}{s\cosh a}\,\frac{\partial}{\partial b}. \quad (5.6)
\end{aligned}$$

The transformation to the independent variables changes the 4-vector x as if at the point (a,b,ϕ) the space has been subjected to the Lorentz transformation

$$\Omega = N_1(-a)N_3(-b)M_3(-\phi). \quad (5.7)$$

Under this transformation $\partial/\partial x_K$ and $A_K(x)$ are transformed according to

$$\frac{\partial}{\partial\bar{x}_K} = \Omega_{kn}\,\frac{\partial}{\partial x_n}, \qquad \bar{A}_K(\bar{x}) = \Omega_{kl}A_l(x), \quad (5.8)$$

where

$$\frac{\partial}{\partial\bar{x}_0} = \frac{\partial}{\partial s}, \quad \frac{\partial}{\partial\bar{x}_1} = \frac{1}{s}\,\frac{\partial}{\partial a},$$
$$\frac{\partial}{\partial\bar{x}_2} = \frac{1}{s\sinh a}\,\frac{\partial}{\partial\phi}, \quad \frac{\partial}{\partial\bar{x}_3} = \frac{1}{s\cosh a}\,\frac{\partial}{\partial b}, \quad (5.9)$$

The transformed Proca equation now becomes

$$\begin{aligned}
&(\Box - \mu^2)\bar{A}_k(\bar{x}) - D_i\,\frac{\partial\bar{A}_k(\bar{x})}{\partial\bar{x}_i} - \Omega_{kl}\\
&\quad\times\left(D_i\,\frac{\partial\Omega^{-1}_{lv}}{\partial\bar{x}_i} + 2\,\frac{\partial\Omega^{-1}_{lv}}{\partial\bar{x}_i}\,\frac{\partial}{\partial\bar{x}_i} + \frac{\partial^2\Omega^{-1}_{lv}}{\partial\bar{x}_i^2}\right)\bar{A}_v(\bar{x}) = 0,\\
&\frac{\partial\bar{A}_i(\bar{x})}{\partial\bar{x}_i} + D_i\bar{A}_i(\bar{x}) = 0,
\end{aligned} \quad (5.10)$$

where

$$D_i = \frac{\partial\Omega_{iv}}{\partial\bar{x}_k}\,\Omega^{-1}_{vk}$$

passing to the canonical basis

$$f_0 = A_0, \qquad \sqrt{2}f_\pm = iA_2 \mp A_1, \qquad f_1 = A_3 \quad (5.11)$$

and expanding f_0, f_1 and $f_\pm$ according to

$$\begin{aligned}
f_0 &= \sum \chi_0^{(\rho,j_0)}(s)C^{\rho j_0}_{00;\tau p}\,e^{i\tau b}\,e^{ip\phi},\\
f_\pm &= \sum \chi_\pm^{(\rho,j_0)}(s)C^{\rho j_0}_{1,\pm1;\tau p}\,e^{i\tau b}e^{ip\phi},\\
f_1 &= \sum \chi_1^{(\rho,j_0)}(s)C_{1,0;\tau p}\,e^{i\tau b}e^{ip\phi},
\end{aligned} \quad (5.12)$$

where the summation is over j_0, p, τ, ρ, the system of equations (5.10) becomes

$$\begin{aligned}
&\left(\frac{\partial\chi_0}{\partial s} + \frac{3}{s}\,\chi_0\right)C_0 + \frac{1}{\sqrt{2}s}\left[\left(\frac{\partial C_-}{\partial a} + (\tanh a + \coth a)\,C_- + \frac{p}{\sinh a}\,C_-\right)\chi_-\right.\\
&\quad\left. - \left(\frac{\partial C_+}{\partial a} + (\tanh a + \coth a)\,C_+ - \frac{p}{\sinh a}\,C_+\right)\chi_+ + \frac{\sqrt{2}i\tau}{\cosh a}\,\chi_1C_1\right] = 0,\\
&\left(\frac{\partial^2\chi_0}{\partial s} + \frac{3}{s}\,\frac{\partial\chi_0}{\partial s} - \frac{3\chi_0}{s^2} + u^2\chi_0\right)C_0 - \frac{1}{s^2}\left[\left(\frac{\partial^2C_0}{\partial a^2} + (\tanh a + \coth a)\,\frac{\partial C_0}{\partial a}\right.\right.\\
&\quad\left. - \frac{\tau^2}{\cosh^2 a}\,C_0 - \frac{p^2}{\sinh^2 a}\,C_0\right)\chi_0 + \sqrt{2}\left(\frac{\partial C_-}{\partial a} + (\tanh a + \coth a)\,C_- + \frac{p}{\sinh a}\,C_-\right)\chi_-\\
&\quad\left. - \sqrt{2}\left(\frac{\partial C_+}{\partial a} + (\tanh a + \coth a)\,C_+ - \frac{p}{\sinh a}\,C_+\right)\chi_+ + \frac{2i\tau}{\cosh a}\,\chi_1C_1\right] = 0,\\
&\left(\frac{\partial^2\chi_1}{\partial s^2} + \frac{3}{s}\,\frac{\partial\chi_1}{\partial s} + \mu^2\chi_1\right)C_1 - \frac{1}{s^2}\left[\left(\frac{\partial^2C_1}{\partial a^2} + (\tanh a + \coth a)\,\frac{\partial C_1}{\partial a} - \frac{\tau^2}{\coth^2 a}\,C_1\right.\right.\\
&\quad\left. - \frac{p^2}{\sinh^2 a}C_1 + \frac{1}{\cosh^2 a}C_1\right)\chi_1 + i\tau\sqrt{2}\,\frac{\tanh a}{\cosh a}\,(\chi_-C_- - \chi_+C_+) + \frac{2i\tau}{\cosh a}\,\chi_0C_0\Big] = 0,\\
&\left(\frac{\partial^2\chi_\pm}{\partial s^2} + \frac{3}{s}\,\frac{\partial\chi_\pm}{\partial s} + \mu^2\chi_\pm\right)C_\pm - \frac{1}{s^2}\left[\left(\frac{\partial^2C_\pm}{\partial a^2} + (\tanh a + \coth a)\,\frac{\partial C_\pm}{\partial a} - \frac{\tau^2}{\cosh^2 a}C_\pm - \frac{p}{\sinh^2 a}C_\pm - \frac{1}{\sinh^2 a}C_\pm\right.\right.\\
&\quad\left. - \frac{1}{2}\tanh^2 a\,C_\pm \pm 2p\,\frac{\coth a}{\sinh a}\,C_\pm\right)\chi_\pm \pm i\tau\sqrt{2}\,\frac{\tanh a}{\cosh a}\,\chi_1C_1 + \frac{1}{2}\tanh^2 a\,\chi_\mp C_\mp + 2\left(\mp\frac{\partial C_0}{\partial a} - \frac{p}{\sinh a}\,C_0\right)\chi_0\Big] = 0,
\end{aligned} \quad (5.13)$$

where we have used the shorthand

$$C_0 = C^{\rho j_0}_{00;\tau\rho}, \qquad C_\pm = C^{\rho j_0}_{1,\pm1;\tau p}, \qquad C_1 = C^{\rho j_0}_{1,0;\tau p}.$$

From the recurrence relations (4.7)–(4.10) we see that the variables separate if

$$\chi_+ = \chi_- = -\chi_1. \quad (5.14)$$

We then arrive at the same system of equations as in Ref. 13 viz.

$$\left(\frac{d}{ds} + \frac{3}{s}\right)\chi_0^{(\rho,0)}(s) + \frac{[3(1+\rho^2)]^{1/2}}{s}\,\chi_1^{(\rho,0)}(s) = 0,$$

$$\left(\frac{d}{ds^2}+\frac{5}{s}\frac{d}{ds}+\frac{4+\rho^2}{s^2}+\mu^2\right)\chi_0^{(\rho,0)}=0, \qquad (5.15)$$
$$\left(\frac{d^2}{ds^2}+\frac{3}{s}\frac{d}{ds}+\frac{1+\rho^2}{s^2}+\mu^2\right)\chi_1^{(\rho\pm1)}(s)=0$$

(remember the summation on j_0 consists of $j_0=0$ only, for f_0). These equations have the solution

$$\chi_1^{(\rho,\pm1)}(s)=(1/\mu s)[c_1H_{i\rho}^{(2)}(\mu s)+c_2H_{-i\rho}^{(2)}(\mu s)],$$
$$\chi_0^{(\rho,0)}(s)=[1/(\mu s)^2][c_3H_{i\rho}^{(2)}(\mu s)+c_4H_{-i\rho}^{(2)}(\mu s)]. \qquad (5.16)$$

So the solutions to the Proca equation have the form

$$f_0=\sum_{p=-\infty}^{\infty}\int_{-\infty}^{\infty}d\tau\int_0^{\infty}d\rho\,\chi_0^{(\rho,0)}(s)C_{00,\tau p}^{\rho 0}e^{i\tau b}e^{ip\phi},$$
$$h_\lambda=\sum_{j_0=-1}^{+1}\sum_{p=-\infty}^{+\infty}\int_0^{\infty}d\tau\int_0^{\infty}d\rho\,\chi_1^{(\rho,j_0)}(s)C_{1,\lambda;\tau p}^{\rho j_0}e^{i\tau b}e^{ip\phi}, \qquad (5.17)$$

where $h_{\pm1}=-f_\pm$, $h_0=f_1$.

This then completes the derivation of an invariant expansion of solutions of the Proca equation inside the light cone.

We now turn our attention to the Dirac equation. In order to obtain an invariant decomposition of a solution of the Dirac equation, we write the equation in a canonical basis

$$\left(i\gamma^n\frac{\partial}{\partial x^n}-\mu\right)\psi(x)=0, \qquad (5.18)$$

where

$$\gamma^0=\begin{pmatrix}0&I\\I&0\end{pmatrix},\qquad \gamma^\alpha=\begin{pmatrix}0&-\sigma_\alpha\\\sigma_\alpha&0\end{pmatrix}, \qquad (5.19)$$

($\alpha=1,2,3$), where σ_α are the Pauli spin matrices and I the 2×2 identity matrix. Under the transformation Ω of (5.7), Eq. (5.18) changes to

$$i\gamma^n\frac{\partial\overline{\psi}(\bar{x})}{\partial\bar{x}^n}+i\gamma^n\Lambda\frac{\partial\Lambda^{-1}}{\partial\bar{x}^n}\overline{\psi}(\bar{x})-\mu\overline{\psi}(\bar{x})=0, \qquad (5.20)$$

where $\overline{\psi}(\bar{x})=\Lambda\psi(x)$, i.e., Λ is the 4×4 matrix according to which the spinor ψ transforms under the Lorentz transformation Ω.

In the C system we have that

$$\gamma^n\Lambda\frac{\partial\Lambda^{-1}}{\partial\bar{x}^n}=\frac{i}{2s}[(\tanh a+\coth a)\gamma^1-3\gamma^0]. \qquad (5.21)$$

If we now look for solutions of the form

$$\psi_i=\sum f_i(s)C_i(a)e^{i\tau b}e^{ip\phi},\qquad i=1,3,$$
$$\psi_j=\sum f_j(s)C_j(a)e^{i\tau b}e^{ip\phi},\qquad j=2,4,$$

the system of equations (5.20) becomes

$$i\left(\frac{\partial f_1}{\partial s}+\frac{3}{2s}f_1\right)C_1+\frac{\tau}{s\cosh a}f_1C_1-\frac{i}{s}\left(\frac{\partial C_2}{\partial a}+\frac{p}{\sinh a}C_2+\frac{1}{2}(\tanh a+\coth a)C_2\right)f_2-\mu f_3C_3=0,$$
$$i\left(\frac{\partial f_2}{\partial s}+\frac{3}{2s}f_2\right)C_2-\frac{\tau}{s\cosh a}f_2C_2-\frac{i}{s}\left(\frac{\partial C_1}{\partial a}-\frac{p}{\sinh a}C_1+\frac{1}{2}(\tanh a+\coth a)C_1\right)f_1-\mu f_4C_4=0,$$
$$i\left(\frac{\partial f_3}{\partial s}+\frac{3}{2s}f_3\right)C_3-\frac{\tau}{s\cosh a}f_3C_3+\frac{i}{s}\left(\frac{\partial C_4}{\partial a}+\frac{p}{\sinh a}C_4+\frac{1}{2}(\tanh a+\coth a)C_4\right)f_4-\mu f_1C_1=0,$$
$$i\left(\frac{\partial f_4}{\partial s}+\frac{3}{2s}f_4\right)C_4+\frac{\tau}{s\cosh a}f_4C_4+\left(\frac{\partial C_3}{\partial a}-\frac{p}{\sinh a}C_3+\frac{1}{2}(\tanh a+\coth a)C_3\right)f_3-\mu f_2C_2=0, \qquad (5.22)$$

from which we see that the variables separate if we take

$$f_1(s)=f_2(s),\qquad f_3(s)=f_4(s),$$
$$C_i(a)=C_{1/2,1/2;\tau p}^{\rho j_0}(a),\qquad i=1,3, \qquad (5.23)$$
$$C_j(a)=C_{1/2,-1/2;\tau p}^{\rho j_0}(a),\qquad j=2,4.$$

The form of $f_1(s)$ and $f_3(s)$ is now determined by the pair of coupled equations

$$\left(\frac{d}{ds}+\frac{3}{2s}-2ij_0\rho\right)f_3+i\mu f_1=0, \qquad (5.24)$$
$$\left(\frac{d}{ds}+\frac{3}{2s}+2ij_0\rho\right)f_1+i\mu f_3=0,$$

which have solutions of the form[16]

$$f_1(s)=(1/\mu s)[c_1J_\nu(\mu s)+c_2J_\nu(\mu s)],$$
$$f_3(s)=(1/\mu s)[c_2J_\nu(\mu s)-c_1J_{-\nu}(\mu s)] \qquad (5.25)$$

with $\nu=\frac{1}{2}+2ij_0\rho$.

So the solutions of the Dirac equation are

$$\psi_i=\sum_{j_0=-1/2}^{+1/2}\sum_{p=-\infty}^{\infty}\int_{-\infty}^{\infty}d\tau\int_0^{\infty}d\rho f_i(s)C_{1/2,1/2;\tau p}^{\rho j_0}(a)\times e^{i\tau b}e^{ip\phi},\qquad i=1,3,$$
$$\psi_j=\sum_{j_0=-1/2}^{+1/2}\sum_{p=-\infty}^{\infty}\int_{-\infty}^{\infty}d\tau\int_0^{\infty}d\rho f_j(s)C_{1/2,1/2;\tau p}^{\rho j_0}(a)\times e^{i\tau b}e^{ip\phi},\qquad j=2,4.$$

This then completes this section on the solution of the Proca and Dirac equation in the C system.

6. DIFFERENTIAL EQUATIONS SATISFIED BY THE EXPANSION MATRIX ELEMENTS AND THE NONRELATIVISTIC LIMIT

From the recurrence relations derived in Sec. 4 we deduce that the matrix elements used in the expansions of Sec. 5 satisfy the following differential equations:

(i) Using the shorthand

$$C_{J\lambda;\tau\rho}^{\rho j_0}(a)=C_{J\lambda}^{\rho j_0},$$

we have for $j_0=J=\lambda=0$ the differential equation

$$\left(\frac{d^2}{da^2}+(\tanh a+\coth a)\frac{d}{da}-\frac{\tau^2}{\cosh^2 a}-\frac{p^2}{\sinh^2 a}(1+\rho^2)\right)C_{00}^0=0. \qquad (6.1)$$

$C^0_{1\lambda}$ may be calculated from C^0_{00} by using

$$-i\tau/\cosh a\ C^0_{00} = [\tfrac{4}{3}(1+\rho^2)]^{-1/2} C^0_{10},$$
$$\pm\sqrt{2}\left(\frac{d}{da} \pm \frac{p}{\sinh a}\right) C^0_{00} = [\tfrac{4}{3}(1+\rho^2)]^{1/2} C^0_{1,\pm 1}. \tag{6.2}$$

(ii) $j_0 = 1$; C^1_{10} satisfies the equation

$$\left[\frac{d^2}{da^2} + \left((\tanh a + \coth a) + \frac{4\tau}{\tau^2 - \rho^2\cosh^2 a}\right)\frac{d}{da} - \frac{\tau^2}{\cosh^2 a} - \frac{p^2}{\sinh^2 a} + 2 + p^2 + \tanh^2 a + \frac{4\tau\tanh a}{\tau^2 - \rho^2\cosh^2 a}(2\tau\tanh a - \rho p\coth a)\right] C^1_{10} = 0; \tag{6.3}$$

the other $j_0 = 1$ matrix elements may be deduced from the relations

$$i\left(\mp\frac{2\tau}{\cosh a} - \rho\right) C^1_{1,\pm 1} = \sqrt{2}\left(\frac{d}{da} + \tanh a \pm \frac{p}{\sinh a}\right) C^1_{10}. \tag{6.4}$$

(iii) $j_0 = \frac{1}{2}$; $C^{1/2}_{1/2,1/2}$ satisfies the equation

$$\left[\frac{d^2}{da^2} + \left((\tanh a + \coth a) + \frac{\tau\tanh a}{\tau \pm \rho\cosh a}\right)\frac{d}{da} - \frac{\tau^2}{\cosh^2 a} - \frac{p^2}{\sinh^2 a} \pm p\frac{\coth a}{\sinh a} + \tfrac{1}{4}(\tanh^2 a - \coth^2 a) + \frac{\tau}{\tau \pm \rho\cosh a}\left(\tfrac{1}{2}(\tanh a + \coth a) \mp \frac{p}{\sinh a}\right) + \rho^2\right] \times C^{1/2}_{1/2,\mp 1/2} = 0. \tag{6.5}$$

Similar equations to those of (ii) and (iii) hold for the cases $j_0 = -1, j_0 = -\frac{1}{2}$, respectively.

These equations are useful in the passage to the nonrelativistic limit.[13-17] In this limit we have

$$a \to 0, \quad s \to \infty \quad \text{s.t.} \quad sa = r, \tag{6.6}$$

where r is the polar radius in the xy plane in nonrelativistic 3-space

$$b \to 0, \quad s \to \infty \quad \text{s.t.} \quad sb = z. \tag{6.7}$$

In addition we must require that

$$\tau \to \infty \quad \text{in such a way that}$$
$$\tau/s \to \tau', \quad -\infty < \tau' < \infty; \tag{6.8}$$

finally

$$\rho \to |\mathbf{p}|s.$$

In this limit Eq. (6.1) becomes

$$\left(\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} + |\mathbf{p}|^2 - \tau'^2 - \frac{p^2}{r^2}\right) C^0_{00}(r) = 0; \tag{6.9}$$

so taking the regular solution at $r = 0$, we have

$$C^0_{00} \to c_1 J_p(\alpha r), \qquad \alpha^2 = |\mathbf{p}|^2 - \tau'^2.$$

From relations (6.2) we see that

$$C^0_{10} \to c_2 J_p(\alpha r), \qquad C^0_{1,\pm 1} \to C_{\pm} J_{p\pm 1}(\alpha r). \tag{6.10}$$

Similar results hold in the $j_0 = 1$ case as C^1_{10} then satisfies Eq. (6.9)

This then gives the correct set of functions in 3-space corresponding to the expansion of Maxwell's equations in cylindrical coordinates,[18] viz.,

$$\bar{C}_\lambda(r, z, \phi) = J_{p+\lambda}(\alpha r) e^{i\tau' z} e^{ip\phi},$$
$$\lambda = \pm 1, 0, \quad p = 0, \pm 1, \pm 2, \ldots, \quad -\infty < \tau' < \infty. \tag{6.11}$$

We note that the solution in cylindrical coordinates is an expansion invariant with respect to the group $O(2) \otimes T_3$, the direct product of rotations about Oz, and translations along Oz. So the reduction $O(1,1) \otimes O(2) \subset O(3,1)$ becomes in the nonrelativistic limit the reduction $O(2) \otimes T_3 \subset E(3)$.

For the nonrelativistic limit of the functions used in the Dirac equation solution we have the following differential equations

$$\left(\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} + |\mathbf{p}|^2 - \tau'^2 - \frac{(p \pm \frac{1}{2})^2}{r^2}\right) C^{1/2}_{1/2,\pm 1/2} = 0, \tag{6.12}$$

so that this corresponds to a nonrelativistic solution of the Dirac equation in terms of the complete set of functions

$$P_{\pm 1/2}(r, a, \phi) = J_{p\pm 1/2}(\alpha r) e^{i\tau' z} e^{ip\phi}. \tag{6.13}$$

This coincides with the solution in cylindrical coordinates in 3-space.

7. CONCLUSION

In this paper we have carried out the reduction of the principal series of $O(3,1)$ in an $O(1,1) \otimes O(2)$ basis and examined the properties of the $O(3) \leftrightarrow O(1,1) \otimes O(2)$ mixed basis matrix elements. It was shown that the expansion of solutions of the Proca and Dirac free fields (inside the light cone) corresponds to the relativistic generalization of cylindrical coordinates in 3-space. In future developments we propose to study the solution of other wave equations (both inside and outside the light cone) using these mixed basis matrix elements. Other related problems of interest include the reduction of the supplementary series of $O(3,1)$ with respect to $O(1,1) \otimes O(2)$[19] and a study of the matrix elements in an $O(1,1) \otimes O(2)$ basis.

ACKNOWLEDGMENTS

The author would like to thank Professor I. A. Sakmar for his constant encouragement throughout the duration of this work. This work has been partially supported by the National Research Council research grant of Professor I. A. Sakmar for which we are grateful.

---

* Work supported by a National Research Council of Canada graduate scholarship.

[1] A partial list of references is the following: N. Mukunda, J. Math. Phys. **9**, 50, 417 (1968); K. Szegö and K. Tóth, *ibid.* **12**, 846, 853 (1971); N. W. Macfadyen, *ibid.* 492 (1971); A. Sciarrino and M. Toller, *ibid.* **8**, 1252 (1967); P. Winternitz, I. Lukač, and Ya. A. Smoro-

dinski, Yad. Fiz. **7**, 192 (1968) [Sov. J. Nucl. Phys. **7**, 139, (1968)]; S. Chang and L. O'Raifeartaigh, J. Math. Phys. **10**, 21 (1969); Ya. A. Smorodinski and M. Huszar, Theoret. Math. Phys. **4**, 867 (1970).
[2] N. Ya. Vilenkin and Ya. A. Smorodinski, Zh. Eksp. Teor. Fiz. **46**, 1793 (1964); [Soviet Phys. JETP **19**, 1209 (1964)].
[3] For a discussion of Poincaré and Lorentz invariance in the scalar case, see P. Winternitz, Ya. A. Smorodinski, and M. B. Sheftel, Yad. Fiz. 7, 1325 (1968); [Sov. J. Nucl. Phys. **7**, 785 (1969)].
[4] In connection with little group Partial wave analysis (this is only a partial list), see H. Joos, Fortschr. Physik, **10**, 65 (1962); J. F. Boyce, R. Delbourgo, A. Salam, and J. Strathdee, 1C/67/9 Trieste (1967); R. Delbourgo, A. Salam, and J. Strathdee, Phys. Letters **25B**, 230 (1967); M. Toller, Nuovo Cimento **53A**, 671 (1968).
[5] For a detailed treatment of the Lorentz group, see M. A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon, New York, 1964).
[6] See I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic, New York, 1964), Vol. 1, p. 31.
[7] G. I. Kuznetsov, M. A. Liberman, A. A. Makarov, and Ya. A. Smorodinski, Yad. Fiz. **10**, 644 (1969) [Sov. J. Nucl. Phys. **10**, 370 (1970)].
[8] I. M. Gel'fand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5.
[9] J. S. Lomont and Harry E. Moses, J. Math. Phys. **3**, 405 (1963).
[10] For details of this method in an $O(3)$ basis see I. A. Verdiev and L. A. Dadashev, Yad. Fiz. **6**, 1094 (1967) [Sov. J. Nucl. Phys. **6**, 795 (1968)].
[11] S. Strom, Arkiv Fysik **29**, 467 (1965).
[12] I. M. Gel'fand, R. M. Minlos, and Z. Ya. Shapiro, *Representations of the Rotation and Lorentz Groups and Their Applications* (Pergamon, New York, 1963), p. 98.
[13] A complete analysis of these equations in terms of $SU(2) \leftrightarrow SU(1, 1)$ mixed basis elements has been given by the author, E. G. Kalnins, in Ref. 18.
[14] A. K. Agamaliev, N. M. Atakishiev, and I. A. Verdiev, Yad. Fiz. **9**, 201 (1969) [Sov. J. Nucl. Phys. **9**, 120 (1969)] **10**, 187 (1969) [**10**, 106 (1970)].
[15] N. M. Atakishiev and I. A. Verdiev, Theoret. Math. Phys. **4**, 833 (1970).
[16] This agrees with Ref. 14.
[17] V. G. Kadyshevsky, R. M. Mir-Kasimov, and N. B. Shakov, 1C/67/72, Trieste (1967).
[18] E. G. Kalnins, Ph.D. thesis (University of Western Ontario).
[19] For a treatment of the supplementary series reduced with respect to $O(2, 1)$ see V. I. Kolomytsev, Theoret. Math. Phys. **2**, 153 (1970).

Dynamics of Harmonically Bound Semi-Infinite and Infinite Chains with Friction and Applied Forces

K. H. Lee
Department of Physics, University of Missouri, Columbia, Missouri 65201
(Received 31 March 1972)

The dynamics of semi-infinite and infinite linear chains of identical masses and ideal springs is studied. In addition to the harmonic coupling between nearest neighbors, each particle is harmonically bound to ts equilibrium position and is subject to friction and time-dependent applied forces. The Laplace transform method is used to express the motion of all the particles. The exact solutions are found and discussed for four different cases: (a) an infinite chain, (b) a semi-infinite chain, (c) a semi-infinite chain with the position of the end particle specified as a function of time, and (d) an infinite chain with the position of one particle specified as a function of time. By specializing some results of the present work, those of previous calculations on simpler systems by other authors are recovered.

There are two main approaches to the mathematical description of physical phenomena. One sometimes tries to study as exactly as possible a simplified model with only the main features of a real system, while some are more interested in an approximate solution of a realistic model. The one-dimensional systems have been favorite models for the first approach.[1] One such system extensively studied is the infinite chain of point masses and ideal massless springs[1,2] because it is one of the very few many-body systems in which exact calculations are possible. However, there has not been much study of an exact treatment of a semi-infinite chain. Although there have been many calculations treating semi-infinite lattices in conjunction with studies on surface phenomena,[3] most of them can be classified under the second approach above.

The present work studies the exact dynamics of semi-infinite and infinite linear chains of identical masses and ideal massless springs with identical force constants. In addition to the harmonic coupling between nearest neighbors, each mass is harmonically bound to its equilibrium position and is subject to friction and time-dependent applied forces. The motion of each of the particles is expressed exactly in terms of the given quantities and initial conditions. Four different systems are studied: (a) an infinite chain, (b) a semi-infinite chain, (c) a semi-infinite chain with the position of the end particle specified as a function of time, and (d) an infinite chain with the position of one particle specified as a function of time. By specializing some of the results, those of previous calculations on simpler systems by other authors are recovered.

Let $x_n(t)$ represent the displacement of the nth particle measured from its equilibrium position. The integer n is restricted to $n \geq 0$ for the semi-infinite systems (b) and (c). The coupled equations for the system are

$$m\ddot{x}_n = -k(x_n - x_{n+1}) - k(x_n - x_{n-1}) \begin{bmatrix} 1 \\ (1 - \delta_{n0}) \end{bmatrix} - Kx_n - \beta\dot{x}_n + \phi_n, \quad \text{(1a, 1d)(1b, 1c)}$$

where m is the particle mass, k and K are the spring constants, β is the friction coefficient, δ is the Kronecker delta, $\phi_n(t)$ represents the external force applied to the nth particle and is assumed to be a known function of time. This system of equations is to be solved for $x_n(t)$ subject to the initial conditions

$$x_n(0) = d_n, \quad \dot{x}_n(0) = v_n. \tag{2}$$

For cases (c) and (d), in which $x_0(t)$ is specified, Eq. (1) for $n = 0$ determines the applied force $\phi_0(t)$ required to achieve such a specified motion for the particle $n = 0$.

If one assumes that x_n and ϕ_n have the Laplace transforms

$$X_n(s) = L\{x_n(t)\} = \int_0^\infty dt x_n(t) \exp(-st), \tag{3}$$

$$\Phi_n(s) = L\{\phi_n(t)/k\}, \tag{4}$$

then Eqs. (1) and (2) lead to an inhomogeneous linear difference equation of second order

$$X_{n+1} - 2\left(2\sigma^2 + 4\mu\sigma + 2\alpha^2 - (1/2)\begin{bmatrix} 0 \\ \delta_{n0} \end{bmatrix}\right)X_n + X_{n-1}\begin{bmatrix} 1 \\ (1 - \delta_{n0}) \end{bmatrix} = -H_n, \quad \text{(5a, 5d)(5b, 5c)}$$

dinski, Yad. Fiz. **7**, 192 (1968) [Sov. J. Nucl. Phys. **7**, 139, (1968)]; S. Chang and L. O'Raifeartaigh, J. Math. Phys. **10**, 21 (1969); Ya. A. Smorodinski and M. Huszar, Theoret. Math. Phys. **4**, 867 (1970).
[2] N. Ya. Vilenkin and Ya. A. Smorodinski, Zh. Eksp. Teor. Fiz. **46**, 1793 (1964); [Soviet Phys. JETP **19**, 1209 (1964)].
[3] For a discussion of Poincaré and Lorentz invariance in the scalar case, see P. Winternitz, Ya. A. Smorodinski, and M. B. Sheftel, Yad. Fiz. 7, 1325 (1968); [Sov. J. Nucl. Phys. **7**, 785 (1969)].
[4] In connection with little group Partial wave analysis (this is only a partial list), see H. Joos, Fortschr. Physik, **10**, 65 (1962); J. F. Boyce, R. Delbourgo, A. Salam, and J. Strathdee, 1C/67/9 Trieste (1967); R. Delbourgo, A. Salam, and J. Strathdee, Phys. Letters **25B**, 230 (1967); M. Toller, Nuovo Cimento **53A**, 671 (1968).
[5] For a detailed treatment of the Lorentz group, see M. A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon, New York, 1964).
[6] See I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic, New York, 1964), Vol. 1, p. 31.
[7] G. I. Kuznetsov, M. A. Liberman, A. A. Makarov, and Ya. A. Smorodinski, Yad. Fiz. **10**, 644 (1969) [Sov. J. Nucl. Phys. **10**, 370 (1970)].
[8] I. M. Gel'fand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5.
[9] J. S. Lomont and Harry E. Moses, J. Math. Phys. **3**, 405 (1963).
[10] For details of this method in an $O(3)$ basis see I. A. Verdiev and L. A. Dadashev, Yad. Fiz. **6**, 1094 (1967) [Sov. J. Nucl. Phys. **6**, 795 (1968)].
[11] S. Strom, Arkiv Fysik **29**, 467 (1965).
[12] I. M. Gel'fand, R. M. Minlos, and Z. Ya. Shapiro, *Representations of the Rotation and Lorentz Groups and Their Applications* (Pergamon, New York, 1963), p. 98.
[13] A complete analysis of these equations in terms of $SU(2) \leftrightarrow SU(1, 1)$ mixed basis elements has been given by the author, E. G. Kalnins, in Ref. 18.
[14] A. K. Agamaliev, N. M. Atakishiev, and I. A. Verdiev, Yad. Fiz. **9**, 201 (1969) [Sov. J. Nucl. Phys. **9**, 120 (1969)] **10**, 187 (1969) [**10**, 106 (1970)].
[15] N. M. Atakishiev and I. A. Verdiev, Theoret. Math. Phys. **4**, 833 (1970).
[16] This agrees with Ref. 14.
[17] V. G. Kadyshevsky, R. M. Mir-Kasimov, and N. B. Shakov, 1C/67/72, Trieste (1967).
[18] E. G. Kalnins, Ph.D. thesis (University of Western Ontario).
[19] For a treatment of the supplementary series reduced with respect to $O(2, 1)$ see V. I. Kolomytsev, Theoret. Math. Phys. **2**, 153 (1970).

Dynamics of Harmonically Bound Semi-Infinite and Infinite Chains with Friction and Applied Forces

K. H. Lee
Department of Physics, University of Missouri, Columbia, Missouri 65201
(Received 31 March 1972)

The dynamics of semi-infinite and infinite linear chains of identical masses and ideal springs is studied. In addition to the harmonic coupling between nearest neighbors, each particle is harmonically bound to ts equilibrium position and is subject to friction and time-dependent applied forces. The Laplace transform method is used to express the motion of all the particles. The exact solutions are found and discussed for four different cases: (a) an infinite chain, (b) a semi-infinite chain, (c) a semi-infinite chain with the position of the end particle specified as a function of time, and (d) an infinite chain with the position of one particle specified as a function of time. By specializing some results of the present work, those of previous calculations on simpler systems by other authors are recovered.

There are two main approaches to the mathematical description of physical phenomena. One sometimes tries to study as exactly as possible a simplified model with only the main features of a real system, while some are more interested in an approximate solution of a realistic model. The one-dimensional systems have been favorite models for the first approach.[1] One such system extensively studied is the infinite chain of point masses and ideal massless springs[1,2] because it is one of the very few many-body systems in which exact calculations are possible. However, there has not been much study of an exact treatment of a semi-infinite chain. Although there have been many calculations treating semi-infinite lattices in conjunction with studies on surface phenomena,[3] most of them can be classified under the second approach above.

The present work studies the exact dynamics of semi-infinite and infinite linear chains of identical masses and ideal massless springs with identical force constants. In addition to the harmonic coupling between nearest neighbors, each mass is harmonically bound to its equilibrium position and is subject to friction and time-dependent applied forces. The motion of each of the particles is expressed exactly in terms of the given quantities and initial conditions. Four different systems are studied: (a) an infinite chain, (b) a semi-infinite chain, (c) a semi-infinite chain with the position of the end particle specified as a function of time, and (d) an infinite chain with the position of one particle specified as a function of time. By specializing some of the results, those of previous calculations on simpler systems by other authors are recovered.

Let $x_n(t)$ represent the displacement of the nth particle measured from its equilibrium position. The integer n is restricted to $n \geq 0$ for the semi-infinite systems (b) and (c). The coupled equations for the system are

$$m\ddot{x}_n = -k(x_n - x_{n+1}) - k(x_n - x_{n-1}) \begin{bmatrix} 1 \\ (1 - \delta_{n0}) \end{bmatrix} - Kx_n - \beta\dot{x}_n + \phi_n, \quad \text{(1a, 1d)(1b, 1c)}$$

where m is the particle mass, k and K are the spring constants, β is the friction coefficient, δ is the Kronecker delta, $\phi_n(t)$ represents the external force applied to the nth particle and is assumed to be a known function of time. This system of equations is to be solved for $x_n(t)$ subject to the initial conditions

$$x_n(0) = d_n, \quad \dot{x}_n(0) = v_n. \tag{2}$$

For cases (c) and (d), in which $x_0(t)$ is specified, Eq. (1) for $n = 0$ determines the applied force $\phi_0(t)$ required to achieve such a specified motion for the particle $n = 0$.

If one assumes that x_n and ϕ_n have the Laplace transforms

$$X_n(s) = L\{x_n(t)\} = \int_0^\infty dt x_n(t) \exp(-st), \tag{3}$$

$$\Phi_n(s) = L\{\phi_n(t)/k\}, \tag{4}$$

then Eqs. (1) and (2) lead to an inhomogeneous linear difference equation of second order

$$X_{n+1} - 2\left(2\sigma^2 + 4\mu\sigma + 2\alpha^2 - (1/2)\begin{bmatrix} 0 \\ \delta_{n0} \end{bmatrix}\right)X_n + X_{n-1}\begin{bmatrix} 1 \\ (1 - \delta_{n0}) \end{bmatrix} = -H_n, \quad \text{(5a, 5d)(5b, 5c)}$$

where

$$\sigma = (s/2\omega), \quad \omega = (k/m)^{1/2}, \tag{6}$$

$$\mu = \beta(16km)^{-1/2}, \tag{7}$$

$$\alpha = [2 + (K/k)]^{1/2}/2, \tag{8}$$

$$H_n = (\sigma + 2\mu)D_n + V_n + \Phi_n, \quad D_n = (2d_n/\omega), \quad V_n = (v_n/\omega^2). \tag{9}$$

One can show (Appendix A) that Eq. (5) has the following solutions:

$$X_n = \frac{1}{D}\sum_{-\infty}^{\infty} H_r \zeta^{2|n-r|}, \tag{10a}$$

$$X_n = \frac{1}{D}\sum_{0}^{\infty} H_r[\zeta^{2|n-r|} + \zeta^{2(n+r+1)}], \tag{10b}$$

$$X_n = X_0\zeta^{2n} + \frac{1}{D}\sum_{1}^{\infty} H_r[\zeta^{2|n-r|} - \zeta^{2(n+r)}], \tag{10c}$$

$$X_n = X_0\zeta^{|2n|} + \frac{1}{D}\left(\sum_{\substack{1\\(n\geq 0)}}^{\infty} \text{ or } \sum_{\substack{-1\\(n\leq 0)}}^{-\infty}\right) H_r[\zeta^{2|n-r|} - \zeta^{2|n+r|}], \tag{10d}$$

where

$$D = 4[(\sigma^2 + 2\mu\sigma + \alpha^2)^2 - (\tfrac{1}{2})^2]^{1/2}, \tag{11}$$

$$\zeta = (\sigma^2 + 2\mu\sigma + \alpha^2 + \tfrac{1}{2})^{1/2} - (\sigma^2 + 2\mu\sigma + \alpha^2 - \tfrac{1}{2})^{1/2}. \tag{12}$$

It can be shown (Appendix B) that the inverse transforms of X_n are

$$x_n(t) = \sum_{-\infty}^{\infty} \{K_r\} G_{n-r}(t), \tag{13a}$$

$$x_n(t) = \sum_{0}^{\infty} \{K_r\}[G_{n-r}(t) + G_{n+r+1}(t)], \tag{13b}$$

$$x_n(t) = \delta_{n0}x_0(t) + \omega^2 x_0(t)^*[G_{n-1}(t) - G_{n+1}(t)] + \sum_{1}^{\infty} \{K_r\}[G_{n-r}(t) - G_{n+r}(t)], \tag{13c}$$

$$x_n(t) = \delta_{n0}x_0(t) + \omega^2 x_0(t)^*[G_{|n|-1}(t) - G_{|n|+1}(t)] + \left(\sum_{\substack{1\\(n\geq 0)}}^{\infty} \text{ or } \sum_{\substack{-1\\(n\leq 0)}}^{-\infty}\right)\{K_r\}[G_{n-r}(t) - G_{n+r}(t)], \tag{13d}$$

where

$$\{K_r\} = d_r(d/dt) + (4\mu\omega d_r + v_r) + (1/m)\phi_r(t)^*, \tag{14}$$

$$G_n(t) = G_{-n}(t) = \exp(-2\mu\omega t)\int_0^t dt' J_0[2b\omega(t^2 - t'^2)^{1/2}] \times J_{2n}(2\omega t'), \tag{15}$$

$$b = (\alpha^2 - \mu^2 - \tfrac{1}{2})^{1/2}. \tag{16}$$

The J's are the ordinary Bessel functions with the property $J_{-2n} = J_{2n}$ and * stands for the convolution. It is to be noted that the inverse Laplace transforms for the case

$$\alpha^2 - \mu^2 - \tfrac{1}{2} < 0, \quad b' = (-\alpha^2 + \mu^2 + \tfrac{1}{2})^{1/2} \tag{17}$$

lead to

$$G_n(t) = \exp(-2\mu\omega t)\int_0^t dt' I_0[2b'\omega(t^2 - t'^2)^{1/2}] \times J_{2n}(2\omega t') \tag{18}$$

with the modified Bessel function I_0, a result contained in Eqs. (15) and (16) because $J_0(iz) = I_0(z)$.

Study of the results given in Eq. (13) can be made most naturally by examining the properties of G_n. To this end, one defines

$$F_n(t) \equiv \dot{G}_n(t) = F_{-n}(t), \tag{19}$$

$$g_n(t) \equiv \exp(2\mu\omega t)G_n(t) = g_{-n}(t) = -g_n(-t), \tag{20}$$

$$f_n(t) \equiv \dot{g}_n(t) = f_{-n}(t) = f_n(-t), \tag{21}$$

reducing the problem to the study of g_n. It can be shown (Appendix C) that

$$g_n(t) = \int_0^t dt' J_0[2b\omega(t^2 - t'^2)^{1/2}] J_{2n}(2\omega t') \tag{22}$$

$$= \frac{1}{\pi}\int_0^{\pi} d\phi \cos(n\phi)\left(\frac{\sin\{2\omega t[b^2 + \sin^2(\phi/2)]^{1/2}\}}{2\omega[b^2 + \sin^2(\phi/2)]^{1/2}}\right), \tag{23}$$

$$f_n(t) = J_{2n}(2\omega t) - t(2b\omega)^2[g_n(t) + g_n^1(t)]/2 \tag{24}$$

$$= \frac{1}{\pi}\int_0^{\pi} d\phi \cos(n\phi)\cos\{2\omega t[b^2 + \sin^2(\phi/2)]^{1/2}\}, \tag{25}$$

$$\dot{g}_n = (t\omega^2/2n)(g_{n-1} - g_{n+1}), \quad n \neq 0, \tag{26}$$

$$\dot{f}_n = f_n/t + (t\omega^2/2n)(f_{n-1} - f_{n+1}), \quad n \neq 0, \tag{27}$$

$$\ddot{h}_n = -4\omega^2(b^2 + 1/2)h_n + \omega^2(h_{n-1} + h_{n+1}), \quad h = g \text{ or } f, \tag{28}$$

$$g_n(0) = \ddot{g}_n(0) = \dot{f}_n(0) = 0, \tag{29}$$

$$\dot{g}_n(0) = f_n(0) = \delta_{n0}, \tag{30}$$

$$g_n(\infty) = \delta_{b0}/(2\omega), \tag{31}$$

$$\dot{g}_n(\infty) = \ddot{g}_n(\infty) = f_n(\infty) = \dot{f}_n(\infty) = 0, \tag{32}$$

where g_n^1 in Eq. (24) is an integral of the form (22) with J_2 instead of J_0. It is straightforward to write expressions similar to Eqs. (22)–(32) for G and F. For example, one obtains

$$G_n(0) = \dot{G}_n(\infty) = \ddot{G}_n(\infty) = F_n(\infty) = \dot{F}_n(\infty) = 0, \tag{33}$$

$$\dot{G}_n(0) = F_n(0) = \delta_{n0}, \tag{34}$$

$$\ddot{G}_n(0) = \dot{F}_n(0) = -4\mu\omega\delta_{n0}, \tag{35}$$

$$G_n(\infty) = \delta_{b0}\delta_{\mu 0}/(2\omega). \tag{36}$$

By use of the properties (33)–(35), one easily verifies that the solutions (13a)–(13d) satisfy the initial and boundary conditions.

Eqs. (13) and (14) provide the physical interpretation of G and F as "propagators." For example, one observes that $[F_{n-r}(t) + 4\mu\omega G_{n-r}(t)]d_r$ and $G_{n-r}(t)v_r$ represent the displacement components in $x_n(t)$ due to $d_r = x_r(0)$ and $v_r = \dot{x}_r(0)$. It is to be noted that the second propagators in Eqs. (13b)–(13d) represent the following reflections: (13b) a stiff-to-soft reflection

without phase reversal, (13c) and (13d) a soft-to-stiff reflection with phase reversal. The second terms of Eqs. (13c) and (13d) combined with the last term of Eq. (14) imply that the effect of specifying $x_0(t)$ is equivalent to an effective force $m\omega^2 x_0(t) = kx_0(t)$ applied to the particle $|n| = 1$, an expected result.

Special cases of (a) and (c) have been studied, [4,5] and those results can be recovered very easily by specializing the present results. For systems without friction and applied forces, one sets

$$\beta(16km)^{-1/2} = \mu = 0 = \phi_n(t) \tag{37}$$

for all n, whence

$$b = (\alpha^2 - 1/2)^{1/2}. \tag{38}$$

Then Eq. (25) reduces to

$$f_n(t) = \frac{1}{\pi}\int_0^{\pi} d\phi \cos(n\phi)\cos[\Omega t(1 - 2\gamma\cos\phi)^{1/2}], \tag{39}$$

where

$$\Omega = 2\alpha\omega = [(K + 2k)/m]^{1/2}, \quad \gamma = (\omega/\Omega)^2. \tag{40}$$

Huetra et al.[4] obtained Eq. (39) for this special case of (a), and their $g_n(t)/\Omega$ is identical to $g_n(t)$ of the present work with (38). For a special case of (c), namely the simple semi-infinite chain, one sets

$$K = \beta = \phi_n(t) = 0 \tag{41}$$

to get

$$\mu = 0, \quad \alpha = 1/\sqrt{2}, \quad b = 0. \tag{42}$$

Morse and Ingard[5] studied the propagation of a disturbance along such a chain. Their initial and boundary conditions were

$$d_n = x_n(0) = 0, \quad v_n = \dot{x}_n(0) = \delta_{n0}v_0, \tag{43}$$

$$x_0(t) = v_0 t. \tag{44}$$

For these conditions, Eq. (13c) becomes

$$\begin{aligned} x_n(t) &= \delta_{n0}v_0 t + \omega^2(v_0 t)^* \left(\int_0^t dt'[J_{2n-2}(2\omega t') - J_{2n+2}(2\omega t')]\right) \\ &= \delta_{n0}v_0 t + \frac{v_0}{\omega}\sum_{r=0}^{\infty}(2r+1)J_{2n+2r+1}(2\omega t), \end{aligned} \tag{45}$$

which is the result of Morse and Ingard.

APPENDIX A

The highlights of the procedure for (c) are sketched below. The other cases can be treated similarly. Write the solution of Eq. (5c) as

$$X_n = A_n\xi_+^n + B_n\xi_-^n \tag{A1}$$

$$= A_n\zeta_+^{2n} + B_n\zeta_-^{2n}, \tag{A2}$$

where $\xi_\pm^n$ are the solutions of the homogeneous equation and

$$\zeta_\pm = (\sigma^2 + 2\mu\sigma + \alpha^2 + 1/2)^{1/2} \pm (\sigma^2 + 2\mu\sigma + \alpha^2 - 1/2)^{1/2}. \tag{A3}$$

Using the variation of parameter method for the difference equation,[6] one finds

$$X_n = \left(A_0 - \frac{1}{D}\sum_1^n H_r\xi_-^r\right)\xi_+^n + \left(B_0 + \frac{1}{D}\sum_1^n H_r\xi_+^r\right)\xi_-^n, \tag{A4}$$

where A_0 and B_0 are independent of n but dependent on ξ and are to be determined by the boundary conditions. Since $\xi_\pm^m$ and $\xi_\pm^m/D$ for $m > 0$ are unacceptable solutions[7] and $\xi_+\xi_- = 1$, the choice of

$$A_0 = \frac{1}{D}\sum_1^{\infty} H_r\xi_-^r \tag{A5}$$

is made. Imposing the boundary condition, namely the specified X_0, one gets

$$B_0 = X_0 - \frac{1}{D}\sum_1^{\infty} H_r\xi_-^r \tag{A6}$$

and the final result, Eq. (10c) with $\zeta = \zeta_-$.

APPENDIX B

If one writes

$$\begin{aligned} p &= (\sigma^2 + 2\mu\sigma + \alpha^2 - 1/2)^{1/2} \\ &= [(\sigma + \mu)^2 + b^2]^{1/2}, \end{aligned} \tag{B1}$$

with b given in Eq. (16), it follows from Eqs. (11) and (12) that

$$\zeta = (p^2 + 1)^{1/2} - p, \tag{B2}$$

$$D = 4p(p^2 + 1)^{1/2}. \tag{B3}$$

With these, all the terms in Eq. (10) become easily recognizable from tables.[8]

APPENDIX C

In Eq. (22) use the integral representation

$$J_{2n}(2\omega t') = \frac{1}{\pi}\int_0^{\pi} d\phi \cos(n\phi)\cos[2\omega t'\sin(\phi/2)] \tag{C1}$$

for J_{2n} and make the power series expansions for J_0 and for the second factor of the above integrand. The resulting series after the t' integration reduces to the second factor of the integrand in Eq. (23). For (31) and (32), use the relationship

$$\lim_{t\to\infty} y(t) = \lim_{s\to 0}[sY(s)], \quad \text{where} \quad Y(s) = L\{y(t)\}. \tag{C2}$$

---

[1] See, for example, E. H. Lieb and D. C. Mattis, *Mathematical Physics in One Dimension* (Academic, New York, 1966).

[2] See, for example, the reviews and references in J. Hori, *Spectral Properties of Disordered Chains and Lattices* (Pergamon, Oxford, 1968) and A. A. Maradudin, E. W. Montroll, G. H. Weiss, and I. P. Ipatova, *Theory of Lattice Dynamics in the Harmonic Approximation* (Academic, New York, 1971), 2nd ed., Chap. 10. M. A. Huetra and H. S. Robertson, J. Stat. Phys. 1, 393 (1969) contains a short historical note.

[3] See, for example, A. A. Maradudin *et al.*, see Ref. 2, Chap. 9.

[4] M. A. Huetra, H. S. Robertson, and J. C. Nearing, J. Math. Phys. 12, 2305 (1971) and the papers quoted therein.

[5] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968), pp. 80–91.

[6] F. B. Hildebrand, *Finite-Difference Equations and Simulations* (Prentice Hall, Englewood Cliffs, N.J., 1968), p. 33.

[7] R. V. Churchill, *Operational Mathematics* (McGraw-Hill, New York, 1958), p. 178.
[8] A. Erdélyi *et al.*, *Tables of Integral Transforms* (McGraw-Hill, New York, 1954), Vol. I [p. 227–(6), p. 237–(43), p. 240–(23), p. 228–(13) instead of p. 227–(6) when $b^2 < 0$, etc.].

Lattice Wind-Tree Models. II. Analytic Property

D. J. Gates*
Mathematics Department, Rockefeller University, New York, New York 10021
(Received 10 March 1972)

Three lattice versions of the wind-tree model of Ehrenfest are studied. It is shown that various moments, including the recurrence time and the Cesaro limits $\lim_{T\to\infty}(1/T)\sum_{t=1}^{T}\Delta(t)$ of the mean-square displacement $\Delta(t)$ and of the one-particle distribution $\rho(t,\mathbf{x})$ at time t, are analytic functions of the reciprocal of the fugacity of the trees, or equivalently of the deviation $1-\rho$ of the density ρ of the trees from their close packing density 1, in certain disks in the complex plane. Two of the models were considered in Paper I, but the third is new.

1. INTRODUCTION

In Paper I[1] we introduced four lattice versions of the wind-tree model of Ehrenfest, and showed that various moments, including the mean-square displacement and the recurrence time, are bounded above if the density of the trees is sufficiently high. We begin by introducing another model (V) and prove the analogous results. We shall then prove a new analytic property of models I, II, and V.

We consider a system of square particles, called trees, with diagonal of length 1, centered on the points ($\mathbf{Z}^2$) of a simple square lattice with unit spacing. A typical configuration is illustrated in Fig. 1. We suppose that a particle, called a wind particle, starts at some point with unit velocity to the right, and is deflected through a right angle whenever it strikes a tree while the tree remains fixed, as shown in Fig. 1. It is clear that in this model a wind particle is deflected always in the same direction by a tree, the actual direction depending on the starting point. It is therefore essentially equivalent to a model[2] (Fig. 2) in which the wind particles move only on the lattice lines and are deflected always to the left by point trees. We henceforth confine our attention to this latter model, and suppose for definiteness that the wind particle starts at the origin.

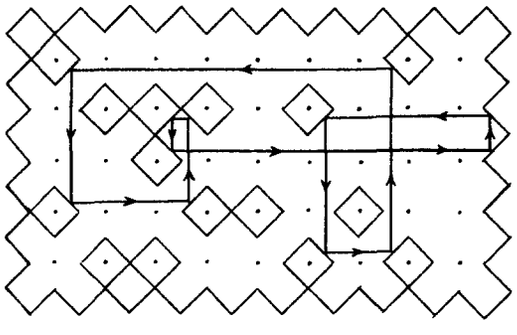


FIG. 1. A typical trajectory in model V. Note that trajectories bend always in the same direction; in the case shown, to the left.

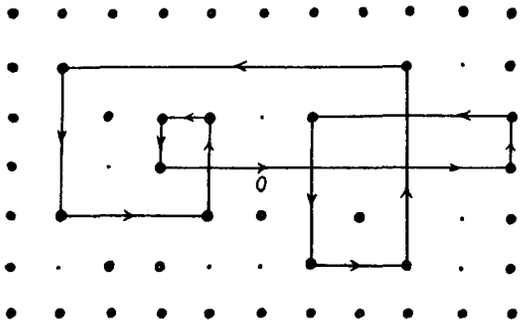


FIG. 2. The model, which is essentially model V, in which trajectories are confined to the lattice bonds.

Only a certain set $R(t)$ of trees can be reached in time t. The position $\mathbf{q}_t(C)$ and velocity $\mathbf{p}_t(C)$ at time t depend only on t and the set $C \subset R(t)$ of sites occupied by trees. The probability of a set C in $R(t)$ is taken to be ($|\ |$ indicates number of points)

$$z^{|C|}(1+z)^{-|R(t)|}. \tag{1.1}$$

This means that the probability of finding a tree on any site (i.e., the density of the trees) is just $z/(1+z)$. We define the moment $m_\alpha(t)$ as the average with respect to tree configurations of $|\mathbf{q}_t(C)|^\alpha$, viz.

$$m_\alpha(t) = (1+z)^{-|R(t)|}\sum_{C\subset R(t)} z^{|C|}|\mathbf{q}_t(C)|^\alpha. \tag{1.2}$$

If the system is enclosed in a box, formed by a close-packed array of trees, there will be a set Λ of sites available to trees. We define a trajectory J as the complete continuation of the path of the wind particle for all t. It is clear that all J in Λ are closed. The probability $P_\Lambda(J)$ of a trajectory is just the total probability of all sets C which result in J (see Paper I for details). The *recurrence time* $\tau(\Lambda)$ is defined as the average time taken for the particle to return to the origin with its initial velocity, namely

$$\tau(\Lambda) = \sum_{J\in[0]} P_\Lambda(J)|J|, \tag{1.3}$$

where $|J|$ is the length of J and $[0]$ is the set of J's which pass through 0 in the horizontal direction. We also define the moments

$$M_\alpha(\Lambda) = \sum_{J\in[0]} P_\Lambda(J)|J|^\alpha. \tag{1.4}$$

The $P_\Lambda(J)$ are independent of Λ if J does not touch the boundary trees: They therefore have well-defined limits $P(J)$ as $\Lambda\to\infty$, given by

$$P(J) = z^{F(J)}(1+z)^{-F(J)-E(J)}, \tag{1.5}$$

where $F(J)$ is the number of occupied (i.e., filled) lattice sites which J meets and $E(J)$ is the number of empty sites through J passes.

2. MODEL V: ABSENCE OF DIFFUSION

Our first result is

Theorem 1: If $z > 2^{16} - 1$ in model V, then the $m_\alpha(t)$ are bounded uniformly in t, and the limits

[7] R. V. Churchill, *Operational Mathematics* (McGraw-Hill, New York, 1958), p. 178.
[8] A. Erdélyi *et al.*, *Tables of Integral Transforms* (McGraw-Hill, New York, 1954), Vol. I [p. 227–(6), p. 237–(43), p. 240–(23), p. 228–(13) instead of p. 227–(6) when $b^2 < 0$, etc.].

Lattice Wind-Tree Models. II. Analytic Property

D. J. Gates*
Mathematics Department, Rockefeller University, New York, New York 10021
(Received 10 March 1972)

Three lattice versions of the wind-tree model of Ehrenfest are studied. It is shown that various moments, including the recurrence time and the Cesaro limits $\lim_{T\to\infty}(1/T)\sum_{t=1}^{T}\Delta(t)$ of the mean-square displacement $\Delta(t)$ and of the one-particle distribution $\rho(t,\mathbf{x})$ at time t, are analytic functions of the reciprocal of the fugacity of the trees, or equivalently of the deviation $1-\rho$ of the density ρ of the trees from their close packing density 1, in certain disks in the complex plane. Two of the models were considered in Paper I, but the third is new.

1. INTRODUCTION

In Paper I[1] we introduced four lattice versions of the wind-tree model of Ehrenfest, and showed that various moments, including the mean-square displacement and the recurrence time, are bounded above if the density of the trees is sufficiently high. We begin by introducing another model (V) and prove the analogous results. We shall then prove a new analytic property of models I, II, and V.

We consider a system of square particles, called trees, with diagonal of length 1, centered on the points ($\mathbf{Z}^2$) of a simple square lattice with unit spacing. A typical configuration is illustrated in Fig. 1. We suppose that a particle, called a wind particle, starts at some point with unit velocity to the right, and is deflected through a right angle whenever it strikes a tree while the tree remains fixed, as shown in Fig. 1. It is clear that in this model a wind particle is deflected always in the same direction by a tree, the actual direction depending on the starting point. It is therefore essentially equivalent to a model[2] (Fig. 2) in which the wind particles move only on the lattice lines and are deflected always to the left by point trees. We henceforth confine our attention to this latter model, and suppose for definiteness that the wind particle starts at the origin.

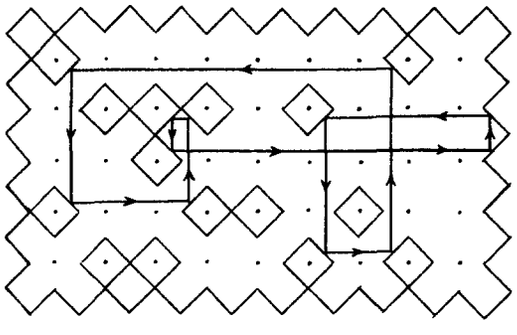


FIG. 1. A typical trajectory in model V. Note that trajectories bend always in the same direction; in the case shown, to the left.

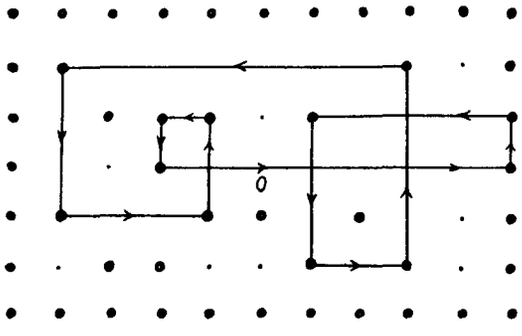


FIG. 2. The model, which is essentially model V, in which trajectories are confined to the lattice bonds.

Only a certain set $R(t)$ of trees can be reached in time t. The position $\mathbf{q}_t(C)$ and velocity $\mathbf{p}_t(C)$ at time t depend only on t and the set $C \subset R(t)$ of sites occupied by trees. The probability of a set C in $R(t)$ is taken to be ($|\ |$ indicates number of points)

$$z^{|C|}(1+z)^{-|R(t)|}. \tag{1.1}$$

This means that the probability of finding a tree on any site (i.e., the density of the trees) is just $z/(1+z)$. We define the moment $m_\alpha(t)$ as the average with respect to tree configurations of $|\mathbf{q}_t(C)|^\alpha$, viz.

$$m_\alpha(t) = (1+z)^{-|R(t)|}\sum_{C\subset R(t)} z^{|C|}|\mathbf{q}_t(C)|^\alpha. \tag{1.2}$$

If the system is enclosed in a box, formed by a close-packed array of trees, there will be a set Λ of sites available to trees. We define a trajectory J as the complete continuation of the path of the wind particle for all t. It is clear that all J in Λ are closed. The probability $P_\Lambda(J)$ of a trajectory is just the total probability of all sets C which result in J (see Paper I for details). The *recurrence time* $\tau(\Lambda)$ is defined as the average time taken for the particle to return to the origin with its initial velocity, namely

$$\tau(\Lambda) = \sum_{J\in[0]} P_\Lambda(J)|J|, \tag{1.3}$$

where $|J|$ is the length of J and $[0]$ is the set of J's which pass through 0 in the horizontal direction. We also define the moments

$$M_\alpha(\Lambda) = \sum_{J\in[0]} P_\Lambda(J)|J|^\alpha. \tag{1.4}$$

The $P_\Lambda(J)$ are independent of Λ if J does not touch the boundary trees: They therefore have well-defined limits $P(J)$ as $\Lambda\to\infty$, given by

$$P(J) = z^{F(J)}(1+z)^{-F(J)-E(J)}, \tag{1.5}$$

where $F(J)$ is the number of occupied (i.e., filled) lattice sites which J meets and $E(J)$ is the number of empty sites through J passes.

2. MODEL V: ABSENCE OF DIFFUSION

Our first result is

Theorem 1: If $z > 2^{16} - 1$ in model V, then the $m_\alpha(t)$ are bounded uniformly in t, and the limits

$$M_\alpha \equiv \lim_{\Lambda\to\infty} M_\alpha(\Lambda) \tag{2.1}$$

and

$$\overline{m}_\alpha \equiv \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} m_\alpha(t) \tag{2.2}$$

exist and are given by

$$M_\alpha = \sum_{J\in[0]} P(J)|J|^\alpha \tag{2.3}$$

and

$$\overline{m}_\alpha = \sum_{J\in[0]} P(J)\frac{1}{|J|}\sum_{t=1}^{|J|} |\mathbf{q}_t'(J)|^\alpha, \tag{2.4}$$

where $\mathbf{q}_t'(J)$ is the position at time of a particle moving on J.

One can also replace the sums over t in (2.2) and (2.4) by integrals. The theorem implies in particular a finite recurrence time, absence of diffusion, and randomization of the initial velocity as described in Paper I. Note that only closed J's appear in (2.3) and (2.4) even though the box is infinitely large.

To prove the theorem, we first need an upper bound on $P_\Lambda(J)$. We note that a trajectory of length $j > 4$ cannot meet trees on more than three successive sites, since it cannot have more than three successive bends. If we call the absence of a bend a *continuation*, we deduce that there must be at least $j/4$ continuations in a trajectory of length $j > 4$. However, at most four different continuations in the same trajectory may occur at the same lattice site, as illustrated in Fig. 3. There are therefore at least $j/16$ different empty lattice sites traversed by the trajectory. This implies (see Paper I) that

$$P_\Lambda(J) \leq (1+z)^{-j/16}. \tag{2.5}$$

Next we note that the number of trajectories of length j is less than 2^{j-1} because a trajectory has only two possibilities at each lattice point: a bend to the left or a continuation. The probability of the trajectory of length 4 is just $z^4/(1+z)^4$. It follows that

$$M_\alpha(\Lambda) < \sum_{j>4} 2^{j-1}(1+z)^{-j/16} j^\alpha + z^4(1+z)^{-4}4^\alpha. \tag{2.6}$$

If we let the sum extend to infinity, we have a bound for all Λ, and this bound is finite for $2(1+z)^{-1/16} < 1$. The boundedness of $m_\alpha(t)$ follows from the inequality

$$m_\alpha(t) < 2^{-\alpha} M_\alpha,$$

which follows in turn from Sec. 2 of Paper I. The existence of M_α and $\overline{m}_\alpha$ and the formulas (2.3) and (2.4) follow from the arguments of Sec. 6 in Paper I. These arguments also prove the existence of the Cesaro limit

$$\overline{\rho}(x) \equiv \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \rho(t,x) \tag{2.7}$$

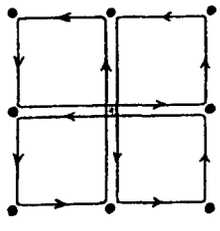


FIG. 3. A trajectory in model V for which empty and occupied sites are traversed more than once.

of the one-particle distribution function $\rho(t,x)$ defined in I.

3. MODEL V: ANALYTICITY

The density ρ of the trees is related to their fugacity z by $\rho = z/(1+z)$ so that $z = \infty$ corresponds to the close-packing density $\rho = 1$. The natural variables for a series expansion of the moments M_α and $\overline{m}_\alpha$ at high density are therefore

$$\zeta \equiv 1 - \rho = 1/(1+z) \quad \text{and} \quad \xi \equiv 1/z. \tag{3.1}$$

The purpose of this section is to show that the moments are analytic functions of these variables in a certain disk in the complex plane, so that such series expansions are possible and, hopefully, may be useful for extrapolation to lower densities. We shall prove

Theorem 2: The moments M_α and $\overline{m}_\alpha$ for model V exist as functions of the complex variables ζ or ξ, and are analytic in the disks

$$|\zeta| < 2^{-16}(1+2^{-16})^{-15} \tag{3.2}$$

or

$$|\xi| < 2^{-16}(1+2^{-16})^{-15}/[1-2^{-16}(1+2^{-16})^{-15}]. \tag{3.3}$$

To prove the theorem, we note from (1.5) and (3.1) that

$$P(J) = \zeta^E(1-\zeta)^F, \tag{3.4}$$

which is an entire function of ζ in the complex plane. If J has length $j > 4$, then $E > j/16$ as before, and also

$$F \leq j - E < \tfrac{15}{16} j. \tag{3.5}$$

These yield, provided $|\zeta| < 1$,

$$\begin{aligned} |P(J)| &\leq |\zeta|^E |1-\zeta|^F \\ &\leq |\zeta|^E(1+|\zeta|)^F \\ &< |\zeta|^{j/16}(1+|\zeta|)^{15j/16}. \end{aligned} \tag{3.6}$$

Now (2.3) can be written

$$M_\alpha = \sum_j p(j) j^\alpha,$$

where

$$p(j) \equiv \sum_{J\in[0]:|J|=j} P(J) \tag{3.7}$$

is the probability of a trajectory in [0] of length j. We deduce that for $j > 4$

$$|p(j)| < 2^{j-1}|\zeta|^{j/16}(1+|\zeta|)^{15j/16}. \tag{3.8}$$

It follows from the Weierstrass M-test that (2.3) converges uniformly in ζ provided

$$2|\zeta|^{1/16}(1+|\zeta|)^{15/16} < 1. \tag{3.9}$$

It is clearly necessary that $|\zeta| < 2^{-16}$. This is not sufficient, but substituting it in (3.9) yields the sufficient condition

$$2|\zeta|^{1/16}(1+2^{-16})^{15/16} < 1,$$

which is just (3.2). Every term in (2.3) is analytic, so that the uniform convergence in the disk (3.2) implies that M_α is analytic in this disk. The same argu-

ment applied to (2.4) proves the analyticity of the $\bar{m}_\alpha$. Finally we use (3.1) to obtain

$$|\xi| = |\zeta|/|1-\zeta| \leq |\zeta|/(1-|\zeta|),$$

which in combination with (3.2) yields (3.3).

The same argument can be used to prove the analyticity of $\bar{\rho}(\mathbf{x})$ defined by (2.7).

4. ANALYTICITY IN MODEL I

For a description of model I we refer the reader to Paper I. It too has only closed trajectories. The analog of Theorem 1 was proved in Secs. 2 and 6 of I. Formula (1.5) for $P(J)$ also holds if we reinterpret $F(J)$ and $E(J)$ as the numbers of occupied and empty squares respectively, of one sublattice of $\mathbf{Z}^2$, which are contained in a "sausage" formed by the squares which border j (see Fig. 4). Now $E \geq j/4$ as shown in Sec. 2 of I. Also we can have at most one tree per unit length of J, so that $F \leq j$. This yields by the argument of the previous section

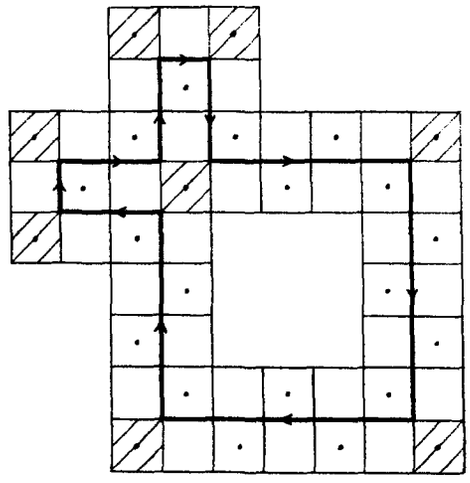


FIG. 4. The "sausage" formed by the squares which border a trajectory in model I. Only the sublattice of squares containing dots are possible sites for trees. Trees are shown by shaded squares. The number of occupied sites in the sausage is $F(J) = 8$, and the number of empty sites is $E(J) = 18$.

$$|p(j)| < 2^{j-1}|\zeta|^{j/4}(1+|\zeta|)^j \tag{4.1}$$

for $|\zeta| < 1$, which in turn leads to

Theorem 3: The moments M_α and $\bar{m}_\alpha$ for model I exist as functions of the complex variables ζ and ξ, and are analytic in the disks

$$|\zeta| < \tfrac{1}{16}(1+\tfrac{1}{16})^{-4} \tag{4.2}$$

or

$$|\xi| < \tfrac{1}{16}(1+\tfrac{1}{16})^{-4}/[1-\tfrac{1}{16}(1+\tfrac{1}{16})^{-4}]. \tag{4.3}$$

5. ANALYTICITY IN MODEL II

Model II is described fully in Paper I. It differs from model I in that all squares of the lattice may contain trees. Again formula (1.5) for $P(J)$ holds with the same definition of $F(J)$ and $E(J)$ as for model I. Now we find instead that $E \geq j/2$, while again $F \leq j$. Thus the argument of Sec. 3 yields

$$|p(j)| < 3^{j-1}|\zeta|^{j/2}(1+|\zeta|)^j, \tag{5.1}$$

which in turn leads to

Theorem 4: The moments M_α and $\bar{m}_\alpha$ for model II exist as complex functions of the variables ζ or ξ and are analytic in the disks

$$|\zeta| < \tfrac{1}{9}(1+\tfrac{1}{9})^{-2} \tag{5.2}$$

or

$$|\xi| < \tfrac{1}{9}(1+\tfrac{1}{9})^{-2}/[1-\tfrac{1}{9}(1+\tfrac{1}{9})^{-2}]. \tag{5.3}$$

Again the function $\bar{\rho}(\mathbf{x})$ defined by (2.7) is analytic for models I and II in the appropriate disks. An open problem is to extend the results of this paper to the models III and IV defined in Paper I.

ACKNOWLEDGMENTS

I am grateful to Professor E. G. D. Cohen, Professor M. Kac, Professor J. L. Lebowitz, and particularly Professor J. T. Cannon for helpful discussions.

* Supported by the U.S. Air Force under grant AFOSR 72-2187

1 D. J. Gates, J. Math. Phys. 13, 1005 (1972).

2 This model was suggested to the author by Professor M. Kac.

Exact Nearest Neighbor Statistics for One-Dimensional Lattice Spaces

R. B. McQuistan

Department of Physics and Laboratory for Surface Studies, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201

(Received 22 February 1972)

It is shown that $A(n_{11}, q, N)$, the number of ways of arranging q indistinguishable particles on a one-dimensional lattice space of N compartments in such a way as to create n_{11} nearest neighbor pairs is $A(n_{11}, q, N) = \binom{N-q+1}{q-n_{11}} \times \binom{q-1}{n_{11}}$. A similar expression is also derived for n_{00}, the number of pairs of vacant nearest neighbors. The normalization, first moment, and most probable value of these statistics are also discussed.

I. INTRODUCTION

A complete statistical mechanical treatment of cooperative phenomena based on the nearest neighbor approximation requires knowledge of the degeneracy associated with pairs of nearest neighbors which have

(1) both sites vacant (0-0),
(2) one site vacant and one occupied (0-1),
(3) both sites occupied (1-1).

If n_{00}, n_{01}, and n_{11} are the numbers of these pairs, respectively, in an arrangement, then they are related

ment applied to (2.4) proves the analyticity of the $\bar{m}_\alpha$. Finally we use (3.1) to obtain

$$|\xi| = |\zeta|/|1-\zeta| \leq |\zeta|/(1-|\zeta|),$$

which in combination with (3.2) yields (3.3).

The same argument can be used to prove the analyticity of $\bar{\rho}(\mathbf{x})$ defined by (2.7).

4. ANALYTICITY IN MODEL I

For a description of model I we refer the reader to Paper I. It too has only closed trajectories. The analog of Theorem 1 was proved in Secs. 2 and 6 of I. Formula (1.5) for $P(J)$ also holds if we reinterpret $F(J)$ and $E(J)$ as the numbers of occupied and empty squares respectively, of one sublattice of $\mathbf{Z}^2$, which are contained in a "sausage" formed by the squares which border j (see Fig. 4). Now $E \geq j/4$ as shown in Sec. 2 of I. Also we can have at most one tree per unit length of J, so that $F \leq j$. This yields by the argument of the previous section

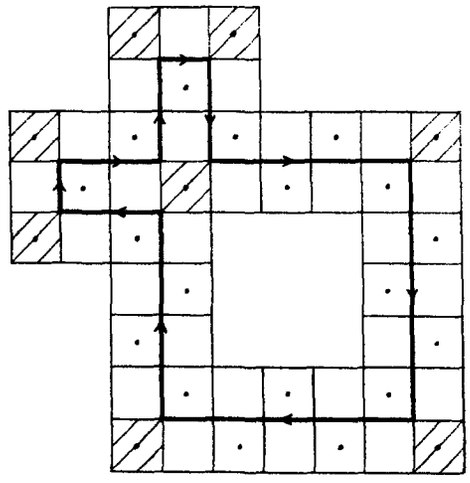


FIG. 4. The "sausage" formed by the squares which border a trajectory in model I. Only the sublattice of squares containing dots are possible sites for trees. Trees are shown by shaded squares. The number of occupied sites in the sausage is $F(J) = 8$, and the number of empty sites is $E(J) = 18$.

$$|p(j)| < 2^{j-1}|\zeta|^{j/4}(1+|\zeta|)^j \tag{4.1}$$

for $|\zeta| < 1$, which in turn leads to

Theorem 3: The moments M_α and $\bar{m}_\alpha$ for model I exist as functions of the complex variables ζ and ξ, and are analytic in the disks

$$|\zeta| < \tfrac{1}{16}(1+\tfrac{1}{16})^{-4} \tag{4.2}$$

or

$$|\xi| < \tfrac{1}{16}(1+\tfrac{1}{16})^{-4}/[1-\tfrac{1}{16}(1+\tfrac{1}{16})^{-4}]. \tag{4.3}$$

5. ANALYTICITY IN MODEL II

Model II is described fully in Paper I. It differs from model I in that all squares of the lattice may contain trees. Again formula (1.5) for $P(J)$ holds with the same definition of $F(J)$ and $E(J)$ as for model I. Now we find instead that $E \geq j/2$, while again $F \leq j$. Thus the argument of Sec. 3 yields

$$|p(j)| < 3^{j-1}|\zeta|^{j/2}(1+|\zeta|)^j, \tag{5.1}$$

which in turn leads to

Theorem 4: The moments M_α and $\bar{m}_\alpha$ for model II exist as complex functions of the variables ζ or ξ and are analytic in the disks

$$|\zeta| < \tfrac{1}{9}(1+\tfrac{1}{9})^{-2} \tag{5.2}$$

or

$$|\xi| < \tfrac{1}{9}(1+\tfrac{1}{9})^{-2}/[1-\tfrac{1}{9}(1+\tfrac{1}{9})^{-2}]. \tag{5.3}$$

Again the function $\bar{\rho}(\mathbf{x})$ defined by (2.7) is analytic for models I and II in the appropriate disks. An open problem is to extend the results of this paper to the models III and IV defined in Paper I.

ACKNOWLEDGMENTS

I am grateful to Professor E. G. D. Cohen, Professor M. Kac, Professor J. L. Lebowitz, and particularly Professor J. T. Cannon for helpful discussions.

* Supported by the U.S. Air Force under grant AFOSR 72-2187

1 D. J. Gates, J. Math. Phys. 13, 1005 (1972).

2 This model was suggested to the author by Professor M. Kac.

Exact Nearest Neighbor Statistics for One-Dimensional Lattice Spaces

R. B. McQuistan

Department of Physics and Laboratory for Surface Studies, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201

(Received 22 February 1972)

It is shown that $A(n_{11}, q, N)$, the number of ways of arranging q indistinguishable particles on a one-dimensional lattice space of N compartments in such a way as to create n_{11} nearest neighbor pairs is $A(n_{11}, q, N) = \binom{N-q+1}{q-n_{11}} \times \binom{q-1}{n_{11}}$. A similar expression is also derived for n_{00}, the number of pairs of vacant nearest neighbors. The normalization, first moment, and most probable value of these statistics are also discussed.

I. INTRODUCTION

A complete statistical mechanical treatment of cooperative phenomena based on the nearest neighbor approximation requires knowledge of the degeneracy associated with pairs of nearest neighbors which have

(1) both sites vacant (0-0),
(2) one site vacant and one occupied (0-1),
(3) both sites occupied (1-1).

If n_{00}, n_{01}, and n_{11} are the numbers of these pairs, respectively, in an arrangement, then they are related

by[1]

$$2q = 2n_{11} + n_{01}, \tag{1}$$

$$2(N-q) = 2n_{00} + n_{01}, \tag{2}$$

where N is the number of sites and q the number of particles.

The degeneracy associated with 0–1 pairs has been considered in connection with the statistical mechanical treatment of the one-dimensional Ising model of magnetism.[2] The present paper is concerned with a determination of the exact degeneracy of 0–0 and 1–1 nearest neighbor pairs. Specifically, we first wish to calculate $A(n_{11}, q, N)$, the number of arrangements containing n_{11} occupied nearest neighbor pairs, created when q indistinguishable particles are arranged on a one-dimensional lattice space of N sites.

In previous papers the exact statistics of one-dimensional random arrays of dumbbells,[3] λ-bells,[4] and the exact ensemble average[5] of the number of nearest, next nearest and third nearest occupied neighbor pairs for simple particles on a two-dimensional lattice space have been treated. To a considerable degree these papers serve as a point of departure for the present calculation.

II. DETERMINATION OF $A(n_{11}, q, N)$

If we consider the number of arrangements possible when q indistinguishable particles are arranged in all possible ways on a one-dimensional lattice space having N sites and select those which contain n_{11} occupied nearest neighbor pairs, we find that the selected arrangements always contain $q - n_{11}$ "units" (see Fig. 1). These "units" consist of one or more contiguous occupied sites together with a vacancy if one is needed to isolate a "unit" from other particles or other vacancies. Thus, the number of separating vacancy is one less than the number of "units."

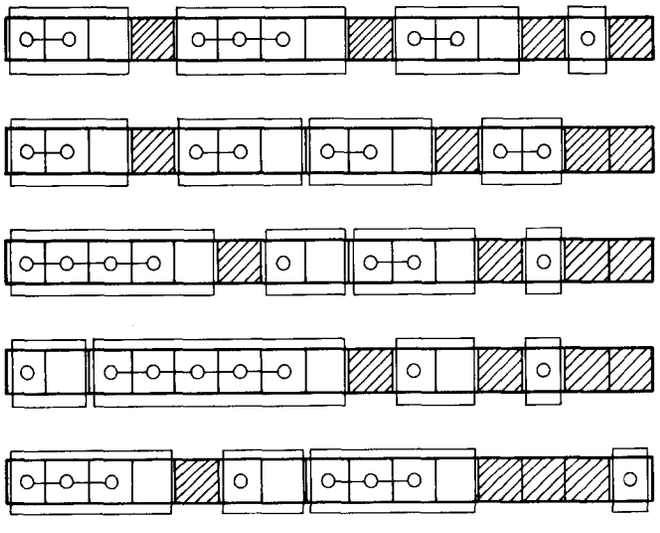


FIG. 1. Eight indistinguishable particles are arranged on a one-dimensional lattice space of $N = 15$ sites to yield four nearest neighbor pairs. Regardless of the configuration, there are always $q - n_{11} = 4$ "units". (The "units" are shown as the unshaded sites). Each of these "units" is separated from other "units" by $q - n_{11} - 1$ vacancies which cannot be permuted, i.e., the occupied nearest neighbor groups and their terminating vacancy (if needed) form an indistinguishable unit. There are $N - q - (q - n_{11} - 1) = N - 2q + n_{11} + 1 = 4$ permutable vacancies (shaded). Thus there are $\binom{8}{4}$ ways of arranging the four "units" (including their separating vacancies) and the four permutable vacancies. This figure shows five possible arrangements in which the indistinguishable "units" are composed of all the possible various groups of nearest neighbor pairs.

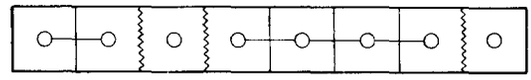


FIG. 2. The eight particles illustrated in Fig. 1 have $q - 1 = 7$ partition between them. Of the seven partitions $n_{11} = 4$ separate occupied nearest neighbor pairs (short horizontal lines) and $q - n_{11} - 1 = 3$ (the jagged lines) do not. Thus there are $\binom{7}{4}$ ways of arranging the separating partitions to form the $q - n_{11} = 4$ indistinguishable "units" discussed in Fig. 1.

There are $N - q$ vacancies, but not all of these are permutable, i.e., not all of the $N - q$ vacancies can be interchanged to form additional arrangements. There are $q - n_{11} - 1$ vacancies which must be utilized to separate the $q - n_{11}$ "units." Thus the number of indistinguishable, permutable vacancies is $N - q - (q - n_{11} - 1) = N - 2q + n_{11} + 1$.

It follows then that the total number of individual things to be permuted is the sum of the "units" and permutable vacancies, $(q - n_{11}) + (N - 2q + n_{11} + 1) = N - q + 1$. The number of ways of arranging $N - q + 1$ things of which $q - n_{11}$ are one kind and $N - 2q + n_{11}$ are another is

$$\binom{N-q+1}{q-n_{11}} = \binom{N-q+1}{N-2q+n_{11}+1}. \tag{3}$$

Equation (3) describes the number of ways the "units" and permutable vacancies may be arranged. To determine $A(n_{11}, q, N)$, we must multiply Eq. (3) by the number of ways the particles can be arranged within the indistinguishable "units." There are $q - 1$ partitions separating the q particles in the "units" (see Fig. 2); n_{11} of these partitions separate occupied nearest neighbor pairs and $q - n_{11} - 1$ do not. There are

$$\binom{q-1}{n_{11}} = \binom{q-1}{q-n_{11}-1} \tag{4}$$

ways of arranging the $q - 1$ partitions, where n_{11} are partitions separating occupied nearest neighbor pairs and $q - n_{11} - 1$ are the number of partitions which do not.

Thus $A(n_{11}, q, N)$ is the product of Eqs. (3) and (4), i.e.,

$$A(n_{11}, q, N) = \binom{N-q+1}{q-n_{11}} \binom{q-1}{n_{11}}. \tag{5}$$

III. DETERMINATION OF $A(n_{00}, q, N)$

To determine $A(n_{00}, q, N)$, the number of arrangements containing n_{00} vacant nearest neighbor pairs which are created when q particles are arranged on a one-dimensional array of N sites, we consider the "units" to consist of one or more contiguous vacant sites together with a filled site if one is needed to isolate the "unit" from other particles or vacancies (see Fig. 3). There are always $N - q - n_{00}$ "units," and they are initially considered to be indistinguishable.

There are q occupied sites; but not all of these can be permuted to create additional arrangements, i.e., some of the particles are utilized to isolate the

"units." More specifically, $N-q-n_{00}-1$ occupied sites must be used to isolate the $N-q-n_{00}$ "units." Thus the number of indistinguishable, permutable occupied sites is

$$q-(N-q-n_{00}-1)=2q+n_{00}-N+1.$$

There are $q+1$ things to be permuted, $N-q-n_{00}$ "units," and $2q+n_{00}-N+1$ permutable vacancies.

These can be arranged in

$$\binom{q+1}{N-q-n_{00}}=\binom{q+1}{2q+n_{00}-N+1} \tag{6}$$

ways.

The positions of the vacancies of which the "units" are composed can be changed to other units to yield additional arrangements. There are $N-q-1$ partitions separating these vacancies (see Fig. 4); of these partitions n_{00} separate vacant nearest neighbors and $N-q-1-n_{00}$ do not. These partitions may be permuted in

$$\binom{N-q-1}{n_{00}}=\binom{N-q-1}{N-q-n_{00}-1} \tag{7}$$

independent ways.

$A(n_{00},q,N)$, the total number of arrangements each containing n_{00} vacant nearest neighbor pairs created when q particles are arranged on a one-dimensional lattice space of N sites, is the product of Eq. (6) with Eq. (7), i.e.,

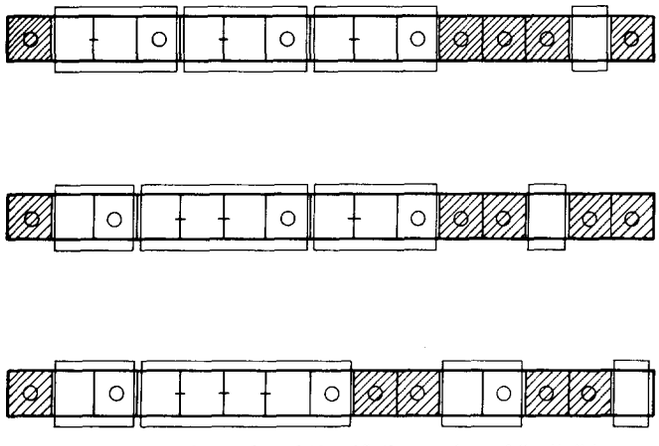


FIG. 3. For $n_{00}=3, q=8$, and $N=15$, the number of "units" is $N-q-n_{00}=4$ (the unshaded sites) and the number of permutable vacancies (shaded) is $q-(N-q-n_{00}-1)=2q+n_{00}-N+1=5$. There are $\binom{9}{5}=126$ ways of arranging the "units" and the permutable vacancies.

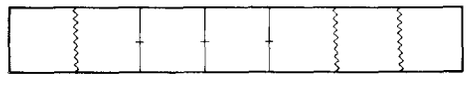


FIG. 4. The seven vacancies shown in Fig. 3 have $N-q-1=6$ partitions between them. $n_{00}=3$ of these partitions separate vacant nearest neighbor pairs (short horizontal lines) and three do not (jagged lines). There are $\binom{6}{3}=20$ ways of arranging the separating partitions to form the $N-q-n_{00}=4$ "units" illustrated in Fig. 3.

$$A(n_{00},q,N)=\binom{q+1}{N-q-n_{00}}\binom{N-q-1}{n_{00}}. \tag{8}$$

IV. NORMALIZATION

The zeroth moment of the statistics characterized by $A(n_{11},q,N)$ is obtained by summing $A(n_{11},q,N)$ over all possible values of n_{11}, i.e., for $n_{11}=0$ up to $n_{11}=q-1$. The Vandermonde theorem[6] shows this sum to be

$$\sum_{n=0}^{q-1}\binom{N-q+1}{q-n_{11}}\binom{q-1}{n_{11}}=\binom{N}{q}. \tag{9}$$

The result explicitly stated in Eq. (9) is to be expected because $\binom{N}{q}$, the totality of all arrangements is composed of those arrangements having no nearest neighbor pairs (nnp's), one nnp, two nnp's, etc.

A similar result for $A(n_{00},q,N)$ is obtained, i.e.,

$$\sum_{n_{00}=0}^{N-q-1}\binom{q+1}{N-q-n_{00}}\binom{N-q-1}{n_{00}}=\binom{N}{q}. \tag{10}$$

V. FIRST MOMENT AND MOST PROBABLE VALUE

The mean value of these statistics, i.e., the ensemble average number of occupied nearest neighbor pairs per arrangement is

$$\begin{aligned}\langle n_{11}\rangle &= \sum_{n_{11}=0}^{q-1} n_{11}\binom{N-q+1}{q-n_{11}}\binom{q-1}{n_{11}}\binom{N}{q}\\ &= (q-1)\sum_{n_{11}=0}^{q-1}\binom{N-q+1}{q-n_{11}}\binom{q-2}{n_{11}-1}\binom{N}{q}\\ &= (N-1)\binom{N-2}{q-2}\binom{N}{q}\\ &= \frac{q(q-1)}{N},\end{aligned} \tag{11}$$

a relationship consistent with previously published results[5] which indicate that the average number of nearest neighbors that a particle has varies as the coverage, $\theta\equiv q/N$, for a one-dimensional lattice space. A similar result,

$$n_{00}=(N-q)(N-q-1)/N, \tag{12}$$

is obtained for the average number of vacant nearest neighbor pairs per arrangement.

The most probable value of $A(n_{11},q,N)$, i.e., the value of n_{11} at which $A(n_{11},q,N)$ is a maximum, can be determined by using the Stirling approximation. It is found to be $q(q-1)/N$, so that the most probable value and the average value are the same.

ACKNOWLEDGMENTS

The author wishes to thank S. J. Lichtman for many helpful discussions and D. Bork for doing the drawings.

1 K. Huang, *Statistical Mechanics* (Wiley, New York, 1963).
2 E. Ising, Z. Physik **31**, 253 (1925).
3 D. Lichtman and R. B. McQuistan, J. Math Phys. 8, 2441 (1967).
4 R. B. McQuistan, Nuovo Cimento **58**, 86 (1968).
5 R. B. McQuistan, J. Math. Phys. **12**, 2113 (1971).
6 J. Riordan, *Combinational Identities* (Wiley, New York, 1968).

Anharmonic Oscillator with Polynomial Self-Interaction

Thomas I. Banks*
Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

and

Carl M. Bender†
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 7 February 1972; Revised Manuscript Received 18 April 1972)

A quantum anharmonic oscillator with a polynomial self-interaction is defined in coordinate space by a Hamiltonian of the form $H = -d^2/dx^2 + \frac{1}{4}x^2 + g[(\frac{1}{2}x^2)^N + a(\frac{1}{2}x^2)^{N-1} + b(\frac{1}{2}x^2)^{N-2} + \cdots]$. Using WKB techniques we derive a secular equation which determines the eigenvalues of H for small $|g|$. We find that the qualitative analytic structure of these eigenvalues as functions of complex g remains unchanged for all fixed values of a, $b, \ldots$, including $a = b = \cdots = 0$. The secular equation also implies an elegant theorem which predicts how the $a, b, \cdots$ terms in H affect the large-order growth of perturbation theory. We use this theorem to compare the perturbative behavior of non-Wick-ordered and Wick-ordered field theories in one-dimensional space-time. In particular, we show that the perturbation series $\sum A_n g^n$ and $\sum B_n g^n$ for the energy levels of the $(g\psi^{2N})_1$ and $(:g\psi^{2N}:)_1$ field theories differ in large order by an over-all multiplicative constant $\lim_{n\to\infty} A_n/B_n = \exp[N(2N-1)/(2N-2)]$.

I. INTRODUCTION

This study was originally motivated by some remarkable computer output. We compared the first 100 Rayleigh–Schrödinger coefficients[1] in the perturbation series for the ground-state energy of two very different theories, the non-Wick-ordered and the Wick-ordered x^6 anharmonic oscillators. We observed that as n, the order of perturbation theory, gets large, the ratio of the nth coefficients for each oscillator approaches a constant. Numerical analysis of successive ratios[2] determined that the value of this constant is $42.521\,082\cdots$. We then recognized that this number was probably equal to $e^{15/4}$. A subsequent search for a theoretical explanation of this result led us to formulate a completely general procedure for analyzing anharmonic oscillators with polynomial self-interactions. (The Wick-ordered x^6 oscillator is one such example).

A quantum anharmonic oscillator with a polynomial self-interaction is defined by the differential equation

$$\left\{-\frac{d^2}{dx^2} + \frac{x^2}{4} + g\left[\left(\frac{x^2}{2}\right)^N + a\left(\frac{x^2}{2}\right)^{N-1} + b\left(\frac{x^2}{2}\right)^{N-2} + \cdots\right] - E(g, a, b, \cdots)\right\}\psi(x) = 0 \quad (1)$$

and the boundary condition

$$\lim_{|x|\to\infty} \psi(x) = 0. \quad (2)$$

The special case of an anharmonic oscillator with a monomial self-interaction ($a = b = \cdots = 0$) has already been investigated.[3] It was shown that WKB techniques lead to an approximate relation (secular equation) between the eigenvalues E and the coupling constant g for small $|g|$.

In Sec. II we extend the WKB techniques of Ref. 3 and derive Eq. (32), the corresponding secular equation for the general problem ($a, b, \cdots \neq 0$). Then, following the approach of Ref. 3, we use the general secular equation to ascertain the analytic structure of E as a function of complex g. We find that for fixed $a, b, \ldots$, the array of singularities that $E(g)$ exhibits (an infinite sequence of square-root-type branch points with a limit point at $g = 0$) is qualitatively independent of the choice (zero or nonzero) of $a, b, \cdots$. As $a, b, \cdots$ vary, the locations of the singularities shift but the number and nature of the singularities remain constant. This result strongly supports the conjecture made in Ref. 3 that the singularity structure and the associated phenomenon of level crossing[4] are *very general and model independent characteristics of singular perturbation theory.*

In Sec. III we turn our attention to the large-order behavior of perturbation theory. Applying some recently published dispersion techniques[5] to the general secular equation derived in Sec. II gives the following theorem: Let $E^{K,N,a,b,\cdots}(g)$, the Kth energy eigenvalue of the differential equation (1), have a perturbation expansion of the form

$$E^{K,N,a,b,\cdots} = K + \tfrac{1}{2} + \sum_{n=1}^{\infty} A_n^{K,N,a,b,\cdots} g^n. \quad (3)$$

Then, for large n,

$$\frac{A_n^{K,N,a,b,\cdots}}{A_n^{K,N,0,0,\cdots}} = e^{a/(N-1)}\left(1 + C_1(K,N,a,b)\frac{1}{n} + C_2(K,N,a,b,c)\frac{1}{n^2} + C_3(K,N,a,b,c,d)\frac{1}{n^3} + \cdots\right). \quad (4)$$

(The leading large-n behavior of $A_n^{K,N,0,0,\cdots}$ is given in Ref. 5).

From Eq. (4) and Ref. 5 we observe that the dominant growth of $A_n^{K,N,a,b,\cdots}$ is controlled by the x^{2N} term in the Hamiltonian H in Eq. (1). The x^{2N-2} term in H only affects the over-all constant. The x^{2N-4} term can at most influence the $O(1/n)$ correction to the growth; the x^{2N-6} term can only affect the $O(1/n^2)$ correction, and so on. Thus, when we consider the large-order behavior of perturbation theory, the $a, b, \cdots$ terms in the Hamiltonian may be considered "small." In this sense, Eq. (4) tells how to "perturb about infinite order" in perturbation theory.[6]

In Sec. IV we show that the quantum anharmonic oscillator defined in Eq. (1) is equivalent to a self-interacting Bose quantum field theory in one-dimensional space–time. We then define Wick ordering and discuss its effect upon perturbation theory. In particular, we show that a Wick-ordered monomial $(g\psi^{2N})_1$ field theory is equivalent to a non-Wick-ordered field theory with a Hermite polynomial self-interaction. We then apply the theorem of Sec. III and show that as $n \to \infty$ the ratio of the nth Rayleigh–Schrödinger coefficients for the perturbation expansion of the energy levels of the $(g\psi^{2N})_1$ and $(:g\psi^{2N}:)_1$ field theories is $\exp[N(2N-1)/(2N-2)]$. When $N = 2$, this expression attains the value e^3, a result already derived in Ref. 5. When $N = 3$, it re-

duces to $e^{15/4}$, which is precisely the result of our computer calculation.

II. DERIVATION OF THE SECULAR EQUATION

In this section we follow Ref. 3 closely. We summarize the techniques needed to obtain the secular equation and emphasize the important new difficulties encountered in treating polynomial interactions.

A. Analytic Continuation of $E(g)$

As in Ref. 3 we define $E(g)$ for *complex* g by Eq. (1) and the general boundary condition

$$\lim_{|x|\to\infty} \psi(x) = 0 \text{ for } |\arg(\pm x) + (2N+2)^{-1}\arg g| < \pi(2N+2)^{-1}. \quad (5)$$

[The sector in Eq. (5) is unchanged by the x^{2N-2}, $x^{2N-4}, \cdots$ terms in Eq. (1) because for $|x|$ sufficiently large, only the x^{2N} term determines the size and location of the sector.] Equation (5) tells us that as g rotates into the complex g plane, the differential equation [Eq. (1)] and its associated boundary condition [Eq. (5)] must be rotated into the complex x plane.

B. The Turning Points

The turning points for the zeroth-order WKB solution to Eq. (1) are solutions of the equation

$$\tfrac{1}{4}x^2 + g[(\tfrac{1}{2}x^2)^N + a(\tfrac{1}{2}x^2)^{N-1} + b(\tfrac{1}{2}x^2)^{N-2} + \cdots] - E = 0. \quad (6)$$

Two of the turning points lie near the origin at a distance of order 1:

$$x \sim \pm(2E)^{1/2}. \quad (7)$$

The others lie approximately equally spaced on a circle of radius $g^{-1/(2N-2)}$:

$$x \sim g^{-1(2N-2)} e^{\pi i/(2N-2)} 2^{(N-2)/(2N-2)} e^{\pi i m/(N-1)}, \quad m = 0, 1, \ldots, 2N-3. \quad (8)$$

As g rotates into the complex plane, the circle of turning points rotates in the same direction as, but *faster* than, the sector in Eq. (5). Thus, turning points periodically enter on one side and leave from the other side of the rotating sector. The sector is so narrow that there is either zero or one, but never more than one, turning point lying inside it.[7]

C. The WKB Approach

There are two methods for solving Eq. (1) approximately using matched asymptotic expansions, the choice of method depending upon whether or not a distant turning point lies in the rotating sector. When there is no turning point in the sector (for example, when g is real and positive), (a) Eq. (1) is solved near $|x| \sim 0$, where it is approximated by a parabolic cylinder (Weber) equation; (b) it is solved again near $|x| \sim \infty$ using WKB theory; (c) the two solutions are matched asymptotically in the intermediate region; (d) the matching places a condition on E, which in zeroth-order WKB gives the expected result

$$E = K + \tfrac{1}{2} + O(g). \quad (9)$$

When a turning point lies in the sector, it interferes with the above connection by distorting the approximations to the function $\psi(x)$. To treat this configuration (a) Eq. (1) is approximated for $|x| \sim 0$ by a parabolic cylinder equation as above; (b) the approximation in part (a) is matched asymptotically to a WKB solution valid between the origin and the turning point; (c) the WKB solution in part (b) is matched asymptotically to the Airy equation solution valid near the turning point; (d) the Airy solution from part (c) is in turn matched asymptotically to a WKB solution valid as $|x| \to \infty$; (e) the condition for matching is the desired secular equation obeyed by g and E.

D. Derivation of the Secular Equation

To simplify the derivation of the secular equation we introduce new notation

$$g = \lambda^{N-1}, \quad (10)$$

$$\rho = \lambda \exp[-i\pi(N+1)/(2N-2)], \quad (11)$$

$$r = xe^{i\pi/4} 2^{-(N-2)/(2N-2)}, \quad (12)$$

$$\epsilon = iE2^{N/(N-1)}, \quad (13)$$

$$T = 2^{-1/(N-1)}. \quad (14)$$

In terms of the new notation, Eq. (1) becomes

$$\left[\frac{d^2}{dr^2} + T^2\left(r^2 - r^{2N}\rho^{N-1} - \frac{ia}{T}r^{2N-2}\rho^{N-1} + \frac{b}{T^2}\rho^{N-1}r^{2N-4} + \cdots - \epsilon\right)\right]\psi(r) = 0. \quad (15)$$

Keeping in mind that $|\rho|$ is small, we proceed with the zeroth-order WKB solution outlined above. We will retain terms to lowest order in powers of ρ. ϵ is of order 1.

The locations of the turning points are

$$r_0 \sim \epsilon^{1/2}[1 + O(\rho^{N-1})] \quad (16)$$

and

$$r_1 \sim \rho^{-1/2}\{1 + \rho/(2N-2)[-\epsilon - (ia/T)] + O(\rho^2)\}. \quad (17)$$

Note that the first-order correction to r_1 depends only on a, the second-order involves a and b, and so on.

Careful analysis shows that when only lowest-order powers in ρ are kept, the parabolic cylinder function and Airy solutions do not depend on $a, b, \cdots$. Moreover, although the WKB solution does depend on a, $b, \ldots$, the expression for the asymptotic matching between the parabolic cylinder and WKB functions does not contain $a, b, \cdots$. The *only* dependence on $a, b, \cdots$ comes from matching the WKB to the Airy function, and this is all contained in the WKB function

$$\psi_{\mathrm{WKB}}(r) \sim [f(r)]^{-1/4} K \exp\left(\pm iT\int_{r_0}^{r}[f(r)]^{1/2}\,dr\right), \quad (18)$$

where

$$f(r) = r^2 - r^{2N}\rho^{N-1} - (ia/T)r^{2N-2}\rho^{N-1} + (b/T^2)\rho^{N-1}r^{2N-4} + \cdots - \epsilon. \quad (19)$$

To evaluate the asymptotic behavior of $\psi_{\mathrm{WKB}}(r)$ in Eq. (18) for $r \sim r_1$, we break the integral in the expo-

nent into two pieces:

$$\int_{r_0}^{r}[f(r)]^{1/2}\,dr = A + B(r), \tag{20}$$

where

$$A = \int_{r_0}^{r_1} dr[f(r)]^{1/2} \tag{21}$$

and

$$B(r) = -\int_{r}^{r_1} dr[f(r)]^{1/2}. \tag{22}$$

For r very near r_1, we let $R = r_1 - r$. Then it is easy to show that

$$B(R) \sim -\rho^{-1/4} R^{3/2}\, \tfrac{2}{3}(2N-2)^{1/2}, \tag{23}$$

which again does not depend on $a, b, \cdots$. Hence, the entire $a, b, \cdots$ dependence resides in A.

To approximate A, we introduce an intermediate value $\bar{r}$, say $\bar{r} = (r_0 r_1)^{1/2}$,[8] and decompose A by

$$A = A_1 + A_2. \tag{24}$$

Using Eq. (16), we can appoximate

$$A_1 = \int_{r_0}^{\bar{r}} dr[f(r)]^{1/2} \tag{25}$$

$$\sim \int_{r_0}^{\bar{r}} dr(r^2 - \epsilon)^{1/2}$$

$$= \tfrac{1}{2}r^{-2} - \tfrac{1}{2}\epsilon \log(2\bar{r}/\sqrt{\epsilon}) - \tfrac{1}{4}\epsilon. \tag{26}$$

To evaluate A_2, we let $x = r/r_1$:

$$A_2 = \int_{\bar{r}}^{r_1} dr[f(r)]^{1/2}$$

$$= r_1^2 \int_{\bar{r}/r_1}^{1} dx\left(x^2 - \rho^{N-1} r_1^{2N-2} x^{2N} - \frac{\epsilon}{r_1^2} - \frac{ia}{T}\rho^{N-1} x^{2N-2} r_1^{2N-4} + \cdots\right)^{1/2}. \tag{27}$$

The first two terms in the brackets in Eq. (27) are large compared with the others, so we expand the integrand using the binomial theorem and then Eq. (17):

$$A_2 \sim r_1^2 \int_{\bar{r}/r_1}^{1} x\,dx(1 - x^{2N-2})^{1/2} + \frac{r_1^2\rho}{2}\int_{\bar{r}/r_1}^{1} dx\,\frac{\epsilon x^{2N} + (ia/T)x^{2N} - (ia/T)x^{2N-2} - \epsilon}{x(1 - x^{2N-2})^{1/2}}. \tag{28}$$

Equation (28) no longer contains $b, c, \cdots$ because those terms contain more and more powers of ρ which we neglect.

To do the first integral in Eq. (28) we use

$$r_1^2 \int_0^{\bar{r}/r_1} x\,dx(1 - x^{2N-2})^{1/2} \sim \tfrac{1}{2}\bar{r}^2.$$

Subtracting and adding this quantity gives a new integral which, after one substitutes

$$t = x^{2N-2}, \tag{29}$$

reduces to an Eulerian integral of the first kind. The result for the first integral in Eq. (28) is thus

$$-\frac{1}{2}\bar{r}^2 + \frac{r_1^2}{N+1}\frac{\Gamma[(N-1)^{-1}]\Gamma(\frac{3}{2})}{\Gamma[\frac{1}{2} + (N-1)^{-1}]}. \tag{30}$$

There are three different integrals to perform in the second term of Eq. (28). The first,

$$\int_{\bar{r}/r_1}^{1} dx\, x^{2N-1}(1 - x^{2N-2})^{-2},$$

is done by replacing the lower limit with 0 and using Eq. (29). The second is an exact differential and may be done directly. The third is also an exact differential. Equation (29) reduces it to an integral of the form $\int dt\, t^{-1}(1-t)^{-1/2}$ which may be done easily.

Performing all of the indicated integrals and using Eq. (17) repeatedly gives a rather complicated expression which then simplifies markedly to

$$A \sim -\frac{\epsilon}{4} + \frac{ia}{T(2N-2)} + \frac{\epsilon}{4}\log\left(T^2\frac{\rho\epsilon}{4}\right) + \frac{1}{\rho(N+1)}\frac{\Gamma[(N-1)^{-1}]\Gamma(\frac{3}{2})}{\Gamma[\frac{1}{2} + (N-1)^{-1}]}. \tag{31}$$

Equation (31) contains no reference to $\bar{r}$, which verifies the correctness of the above lengthy sequence of approximations. This completes the evaluation of A.

The expression for A in Eq. (31) is now used to complete the matching of the WKB and Airy functions as was done in Ref. 3. The resulting general secular equation is

$$\frac{\Gamma(\frac{1}{4} + \frac{1}{2}E)}{\Gamma(\frac{1}{4} - \frac{1}{2}E)} = \exp\left(\frac{5\pi i}{4} - E\log(T\rho) + \frac{iT\Gamma[(N-1)^{-1}]\Gamma(\frac{1}{2})}{(N+1)\rho\Gamma[\frac{1}{2} + (N-1)^{-1}]} + \frac{a}{N-1}\right) \tag{32a}$$

for even-parity energy levels and

$$\frac{\Gamma(\frac{3}{4} + \frac{1}{2}E)}{\Gamma(\frac{3}{4} - \frac{1}{2}E)} = \exp\left(-\frac{5\pi i}{4} - E\log(T\rho) + \frac{iT\Gamma[(N-1)^{-1}]\Gamma(\frac{1}{2})}{(N+1)\rho\Gamma[\frac{1}{2} + (N-1)^{-1}]} + \frac{a}{N-1}\right) \tag{32b}$$

for odd-parity energy levels.

The parameter a only enters Eq. (32) in an over-all multiplicative constant. When we set $a = 0$, we recover the results of Ref. 3. Had we carried out the above analysis using higher-order WKB (following the procedure of Appendix F of Ref. 2, for example), the parameter b would have appeared as an additional term in the secular equation multiplied by ρ, c would have appeared multiplied by ρ^2, and so on.

The implications of the secular equation [Eq. (32)] are the same regardless of the choice of a. It predicts the existence of an infinite sequence of square-root-type branch points in the complex g plane with a limit point at $g = 0$. Level crossing occurs at the branch points. Quantitative descriptions of these phenomena may be found in Refs. 2 and 3 and need not be repeated here. However, we strongly emphasize the apparent model independence and universality of these phenomena. It has now been demonstrated that these same features are exhibited by an extraordinarily wide class of singular perturbation theories.

III. LARGE-ORDER BEHAVIOR OF PERTURBATION THEORY

The perturbation series for the Kth energy eigenvalue for Eq. (1) is given in Eq. (3). To determine how $A_n^{K,N,a,b,\cdots}$ behaves for large n, we use the dispersion techniques introduced in Ref. 5. There it was shown that

$$A_n^{K,N,a,b,\cdots} = \frac{1}{2\pi i}\int_{-\infty}^{0} dx\, x^{-n} D^{K,N,a,b,\cdots}(x), \quad (33)$$

where

$$D^{K,N,a,b,\cdots}(x) \equiv \lim_{\epsilon\to 0}\left[F^{K,N,a,b,\cdots}(x+i\epsilon) - F^{K,N,a,b,\cdots}(x-i\epsilon)\right] \quad (34)$$

and

$$F^{K,N,a,b,\cdots}(x) \equiv x^{-1}\left[E^{K,N,a,b,\cdots}(x) - K - \tfrac{1}{2}\right]. \quad (35)$$

The discontinuity in Eq. (34) may be computed directly from the secular equation [Eq. (32)]. We do not present the details of this calculation here as it will appear elsewhere.[9] However, it is clear that the only dependence of $D^{K,N,a,b,\cdots}(x)$ on the parameter a is contained in an over-all multiplicative constant. Specifically, we find that in zeroth-order WKB

$$D^{K,N,a,b,\cdots}(x) = e^{a/(N-1)}\,\frac{2i2^K\pi^{1/2}}{K!}\left(-\frac{x}{2}\right)^{-(K+1/2)/(N-1)} \times \exp\left(\frac{-\Gamma^2(N/N-1)}{\Gamma(2N/N-1)(-\frac{1}{2}x)^{1/(N-1)}}\right). \quad (36)$$

As $n\to\infty$, we may insert the expression for $D^{K,N,a,b,\cdots}(x)$ in Eq. (36) into the integral in Eq. (33) because for large n the integral is only sensitive to the small $|x|$ behavior of D. (Recall that the WKB techniques in Sec. II become accurate in the limit as $|g|\to 0$). Therefore, using Eq. (33) we establish that in lowest-order WKB

$$\lim_{n\to\infty}\frac{A_n^{K,N,a,b,\cdots}}{A_n^{K,N,0,0,\cdots}} = e^{a/(N-1)}. \quad (37)$$

As was seen in Ref. 5, the corrections to Eq. (37) that arise from higher-order WKB calculations of $D^{K,N,a,b,\cdots}$ are of order n^{-1}, n^{-2}, and so on. These large-n corrections take the form

$$\frac{A_n^{K,N,a,b,\cdots}}{A_n^{K,N,0,0,\cdots}} = e^{a/(N-1)}\left[1 + C_1(K,N,a,b)n^{-1} + C_2(K,N,a,b,c)n^{-2} + C_3(K,N,a,b,c,d)n^{-3} + \cdots\right]. \quad (38)$$

The specific functions $C_1, C_2, \cdots$ have not been determined because we have not carried out the higher-order WKB calculations. However, to verify the theorem in Eq. (38), it *is* important to establish that C_1 just depends on a and b, C_2 on a, b, and c, and so on. We do this by noticing that b enters the secular equation multiplied by ρ, c is multiplied by ρ^2, d by ρ^3, and so on. Combining this observation with the rather complicated arguments given to establish Eq. (4) of Ref. 5 proves the above assertions on the structures of $C_1, C_2, \cdots$.

This theorem [Eq. (38)] on the large-order behavior of the Rayleigh–Schrödinger coefficients is most unusual. It tells how to "perturb about infinite order" in perturbation theory, which is indeed a strange concept. In the next section we present a straightforward application of this theorem. We ascertain the effect of Wick-ordering on one-dimensional field-theoretic perturbation theory.

IV. WICK-ORDERING IN ONE-DIMENSIONAL FIELD THEORY

The $(\psi^{2N})_1$ quantum field theory in one-dimensional space-time is defined by the Hamiltonian and commutation relation

$$H = \tfrac{1}{2}\dot{\psi}^2 + \tfrac{1}{2}m^2\psi^2 + g\psi^{2N}, \quad (39)$$

$$[\psi, \dot{\psi}] = i. \quad (40)$$

To Wick-order the Hamiltonian, one rewrites it in terms of creation and annihilation operators where the annihilation operators stand to the right of the creation operators. In terms of fields

$$\begin{aligned} :\psi^2: &= \psi^2 - (1/2m),\\ :\psi^4: &= \psi^4 - (3/m)\psi^2 + (3/4m^2),\\ :\psi^6: &= \psi^6 - (15/2m)\psi^4 + (45/4m^2)\psi^2 - (15/8m^3). \end{aligned} \quad (41)$$

The polynomials in Eq. (41) are the Hermite polynomials[10]:

$$\begin{aligned} H_2(\tfrac{1}{2}x) &= x^2 - 2,\\ H_4(\tfrac{1}{2}x) &= x^4 - 12x^2 + 12,\\ H_6(\tfrac{1}{2}x) &= x^6 - 30x^4 + 180x^2 - 120. \end{aligned} \quad (42)$$

When $m = \frac{1}{4}$, the coefficients in Eqs. (41) and (42) become identical.

The Feynman diagrammatic expansion of a one-dimensional field theory is topologically identical to that of a higher-dimensional theory. (The Feynman rules may be found in Ref. 3.) Wick-ordering the perturbation expansion removes all Feynman diagrams having self-loops (lines with both ends connected to the same vertex). The energy levels of a field theory are the eigenvalues of the Hamiltonian H:

$$H|E\rangle = E|E\rangle. \quad (43)$$

These eigenvalues may be computed perturbatively in terms of diagrams. The Kth energy level is the K-particle pole of the $2K$-point Green's function. The ground-state energy is the sum of all diagrams with no external legs.

How does Wick-ordering the Hamiltonian in Eq. (39) affect the large-order behavior of the perturbation expansions of the energy levels? To answer this question we transform from a creation–annihilation operator to a coordinate space representation of H:

$$\psi \to 2^{-1/2}x, \qquad \dot{\psi} \to -i2^{1/2}\frac{d}{dx}, \quad (44)$$

and we set $m = 1$. Then Eq. (43) becomes

$$\left[-\frac{d^2}{dx^2} + \frac{x^2}{4} + g\left(\frac{x^2}{2}\right)^N - E\right]\psi(x) = 0. \quad (45)$$

If H is Wick-ordered before using Eq. (44), we have

$$\left\{-\frac{d^2}{dx^2}+\frac{x^2}{4}-\frac{1}{2}+g\left[\left(\frac{x^2}{2}\right)^N-\frac{N(2N-1)}{2}\left(\frac{x^2}{2}\right)^{N-1}+\cdots\right]-E\right\}\psi(x)=0, \quad (46)$$

using the formula for the coefficients of a Hermite polynomial.

If we define the perturbation expansions of the eigenvalues of Eqs. (45) and (46) to be $\sum A_n g^n$ and $\sum B_n g^n$, then for large n the ratio A_n/B_n is given by the theorem in Eq. (38):

$$\lim_{n\to\infty}\frac{A_n}{B_n}=\exp\frac{N(2N-1)}{2N-2}. \quad (47)$$

When $N=3$, the right-hand side of Eq. (47) reduces to $e^{15/4}$, in spectacular agreement with the computer result mentioned at the beginning of this paper.

* This work is supported in part through funds provided by the Atomic Energy Commission under contract At(11-1) 3069. NDEA Predoctoral Fellow.
† Supported in part by National Science Foundation Grant GP29463. A Sloan Fellow.

1 The nth Rayleigh-Schrödinger coefficient is the coefficient of g^n in the perturbation series.

2 A discussion of the numerical methods used may be found in C. M. Bender and T. T. Wu, Phys. Rev. **184**, 1231 (1969), Appendix E.

3 C. M. Bender, J. Math. Phys. **11**, 796 (1970).

4 For a complete description of level crossing in the anharmonic oscillator see Ref. 2, Sec. VI. For a numerical study of level crossing in other models see Th. W. Ruijgrok, Ref. TH. 1393—CERN (August, 1971).

5 C. M. Bender and T. T. Wu, Phys. Rev. Letters **27**, 461 (1971). These dispersion techniques were discovered independently by B. Simon, Ann. Phys. (N.Y.) **58**, 79 (1970).

6 Although the x^{2N} term in Eq. (1) is a *singular* perturbation of the harmonic oscillator (the reason for the divergence of perturbation theory), the $x^{2N-2}, x^{2N-4}, \ldots$, terms are *regular* perturbations of the x^{2N} term and, therefore, cause only small changes in large-order growth of the Rayleigh-Schrödinger coefficients.

7 For a schematic representation of this phenomenon, see Ref. 3.

8 The particular choice of $\bar{r}$ is not significant. What is important is that $\bar{r}$ must be very far from both r_0 and r_1. The geometric mean between r_0 and r_1 is a satisfactory choice for $\bar{r}$.

9 C. M. Bender and T. T. Wu, "Anharmonic Oscillator. II" and "Generalized Anharmonic Oscillator. II" (submitted to Phys. Rev.).

10 M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Nat. Bur. Stds., Washington, D.C., 1964), p. 801.

Quantum Electromagnetic Zero-Point Energy of a Conducting Spherical Shell

B. Davies
Department of Applied Mathematics, The Australian National University, Canberra, Australia
(Received 29 March 1971)

The quantum zero-point energy of a conducting spherical shell was first calculated by Boyer [Phys. Rev. **174**, 1764 (1968)]. Because of the importance of this calculation and also of Boyer's uncertainty about the analytical dependence of the energy on the cutoff function, we have checked the calculation independently. We determine an analytic continuation of the energy function using the Mellin transform, and thereby show how an exact value of the self-energy can be obtained from the divergent series. We also compute an approximate value of the self-energy by extrapolating a direct numerical evaluation of the cutoff integrals. These calculations confirm Boyer's result.

1. INTRODUCTION

Recently there has been a considerable amount of interest in calculating the quantum zero-point energy of various objects.[1-3] In particular, the calculation of quantum electromagnetic zero-point energies has turned out to be a useful way of evaluating long range electromagnetic forces in some cases of interest.[3] Casimir evaluated the attractive force between two conducting parallel plates separated by vacuum more than twenty years ago,[4] and went on to suggest an intriguing model for a charged particle on the basis of this result.[5] The idea is that the electron is a charged, perfectly conducting sphere, and that its Coulomb self-energy is balanced by the quantum electromagnetic zero-point energy, which Casimir considered would be negative. Since the electrostatic energy is proportional to e^2, and the quantum zero-point energy to $\hbar c$, this raises the exciting possibility of being able to calculate a rough value for the fine structure constant and of gaining some insight into the structure of the electron. The calculation of the zero-point energy was performed by Boyer,[1] who found that it is in fact positive, so that Casimir's model cannot account for the stability of a charged particle. Since Boyer's evaluation of this self-energy involves delicate numerical calculations including a limiting procedure which he could not show to be valid (although it appears to be so), we have thought it worthwhile to perform the calculation by alternate means. It is disappointing to report that the result of our calculations is to confirm Boyer's result; however, we think that the methods which we have applied, and the confirmation of a tricky but important calculation, are of some interest in their own right. Furthermore, we intend to extend these methods to investigate the quantum zero-point energy of other systems which have spherical symmetry.

In talking about the zero-point energy of a conducting spherical shell of radius a, we mean the difference between the zero-point energy of the "universe" when the sphere is present and when it is absent. In order to make the calculation feasible, we take the "universe" to be a sphere of radius $R \gg a$ and the two spheres to be concentric. We have three regions to consider: the interior of the small sphere (I), the concentric shell between the two spheres (II), and the interior of the large sphere when the small one is absent (III).[6] Each of these regions has a set of normal electromagnetic modes of frequency ω_k, and the zero-point energy of each region is defined by

$$E=\sum_k \tfrac{1}{2}\hbar\omega_k. \quad (1)$$

The zero-point energy of the conducting shell of radius a is then

$$\left\{-\frac{d^2}{dx^2}+\frac{x^2}{4}-\frac{1}{2}+g\left[\left(\frac{x^2}{2}\right)^N-\frac{N(2N-1)}{2}\left(\frac{x^2}{2}\right)^{N-1}+\cdots\right]-E\right\}\psi(x)=0, \quad (46)$$

using the formula for the coefficients of a Hermite polynomial.

If we define the perturbation expansions of the eigenvalues of Eqs. (45) and (46) to be $\sum A_n g^n$ and $\sum B_n g^n$, then for large n the ratio A_n/B_n is given by the theorem in Eq. (38):

$$\lim_{n\to\infty}\frac{A_n}{B_n}=\exp\frac{N(2N-1)}{2N-2}. \quad (47)$$

When $N=3$, the right-hand side of Eq. (47) reduces to $e^{15/4}$, in spectacular agreement with the computer result mentioned at the beginning of this paper.

* This work is supported in part through funds provided by the Atomic Energy Commission under contract At(11-1) 3069. NDEA Predoctoral Fellow.
† Supported in part by National Science Foundation Grant GP29463. A Sloan Fellow.

1 The nth Rayleigh-Schrödinger coefficient is the coefficient of g^n in the perturbation series.

2 A discussion of the numerical methods used may be found in C. M. Bender and T. T. Wu, Phys. Rev. **184**, 1231 (1969), Appendix E.

3 C. M. Bender, J. Math. Phys. **11**, 796 (1970).

4 For a complete description of level crossing in the anharmonic oscillator see Ref. 2, Sec. VI. For a numerical study of level crossing in other models see Th. W. Ruijgrok, Ref. TH. 1393—CERN (August, 1971).

5 C. M. Bender and T. T. Wu, Phys. Rev. Letters **27**, 461 (1971). These dispersion techniques were discovered independently by B. Simon, Ann. Phys. (N.Y.) **58**, 79 (1970).

6 Although the x^{2N} term in Eq. (1) is a *singular* perturbation of the harmonic oscillator (the reason for the divergence of perturbation theory), the $x^{2N-2}, x^{2N-4}, \ldots$, terms are *regular* perturbations of the x^{2N} term and, therefore, cause only small changes in large-order growth of the Rayleigh-Schrödinger coefficients.

7 For a schematic representation of this phenomenon, see Ref. 3.

8 The particular choice of $\bar{r}$ is not significant. What is important is that $\bar{r}$ must be very far from both r_0 and r_1. The geometric mean between r_0 and r_1 is a satisfactory choice for $\bar{r}$.

9 C. M. Bender and T. T. Wu, "Anharmonic Oscillator. II" and "Generalized Anharmonic Oscillator. II" (submitted to Phys. Rev.).

10 M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Nat. Bur. Stds., Washington, D.C., 1964), p. 801.

Quantum Electromagnetic Zero-Point Energy of a Conducting Spherical Shell

B. Davies
Department of Applied Mathematics, The Australian National University, Canberra, Australia
(Received 29 March 1971)

The quantum zero-point energy of a conducting spherical shell was first calculated by Boyer [Phys. Rev. **174**, 1764 (1968)]. Because of the importance of this calculation and also of Boyer's uncertainty about the analytical dependence of the energy on the cutoff function, we have checked the calculation independently. We determine an analytic continuation of the energy function using the Mellin transform, and thereby show how an exact value of the self-energy can be obtained from the divergent series. We also compute an approximate value of the self-energy by extrapolating a direct numerical evaluation of the cutoff integrals. These calculations confirm Boyer's result.

1. INTRODUCTION

Recently there has been a considerable amount of interest in calculating the quantum zero-point energy of various objects.[1-3] In particular, the calculation of quantum electromagnetic zero-point energies has turned out to be a useful way of evaluating long range electromagnetic forces in some cases of interest.[3] Casimir evaluated the attractive force between two conducting parallel plates separated by vacuum more than twenty years ago,[4] and went on to suggest an intriguing model for a charged particle on the basis of this result.[5] The idea is that the electron is a charged, perfectly conducting sphere, and that its Coulomb self-energy is balanced by the quantum electromagnetic zero-point energy, which Casimir considered would be negative. Since the electrostatic energy is proportional to e^2, and the quantum zero-point energy to $\hbar c$, this raises the exciting possibility of being able to calculate a rough value for the fine structure constant and of gaining some insight into the structure of the electron. The calculation of the zero-point energy was performed by Boyer,[1] who found that it is in fact positive, so that Casimir's model cannot account for the stability of a charged particle. Since Boyer's evaluation of this self-energy involves delicate numerical calculations including a limiting procedure which he could not show to be valid (although it appears to be so), we have thought it worthwhile to perform the calculation by alternate means. It is disappointing to report that the result of our calculations is to confirm Boyer's result; however, we think that the methods which we have applied, and the confirmation of a tricky but important calculation, are of some interest in their own right. Furthermore, we intend to extend these methods to investigate the quantum zero-point energy of other systems which have spherical symmetry.

In talking about the zero-point energy of a conducting spherical shell of radius a, we mean the difference between the zero-point energy of the "universe" when the sphere is present and when it is absent. In order to make the calculation feasible, we take the "universe" to be a sphere of radius $R \gg a$ and the two spheres to be concentric. We have three regions to consider: the interior of the small sphere (I), the concentric shell between the two spheres (II), and the interior of the large sphere when the small one is absent (III).[6] Each of these regions has a set of normal electromagnetic modes of frequency ω_k, and the zero-point energy of each region is defined by

$$E=\sum_k \tfrac{1}{2}\hbar\omega_k. \quad (1)$$

The zero-point energy of the conducting shell of radius a is then

$$\Delta E(a) = \lim_{R\to\infty}(E_{\mathrm{I}} + E_{\mathrm{II}} - E_{\mathrm{III}}). \tag{2}$$

Of course, each of the sums involved in (2) is infinite and this represents an aspect of the problem which is not yet well understood. In Boyer's calculation, this problem is overcome by introducing a cutoff function into each of the sums, performing the subtraction, and then removing the cut-off functions again. By suitable choice of these cutoff functions, we could recover any finite or infinite answer that we please; but Boyer argues on physical grounds that the cutoff functions for each sum must be the same function, and that it must depend only on the ratio of the frequency to a cutoff frequency. This then makes the answer unique for a wide range of cutoff functions. We must emphasize that although the energy is unique in this sense, the decision to base the cutoff on frequency, and to apply the same function to each region, is a part of the physical model.

2. APPLICATION OF THE MELLIN TRANSFORM

We show in the Appendix how the self-energy (2) may be evaluated by contour integrals. The cutoff function which we shall use in this section is $\exp(-\lambda k)$ and we denote the corresponding energy by $\Delta E(a,\lambda)$. It is given by

$$\Delta E(a,\lambda) = \sum_{l=1}^{\infty} \Delta E_l(a,\lambda), \tag{3}$$

where

$$\Delta E_l(a,\lambda) = \mathfrak{D}_l(e^{-\lambda k}) + \overline{\mathfrak{D}}_l(e^{-\lambda k}) \tag{4}$$

and the linear functionals $\mathfrak{D}_l$ and $\overline{\mathfrak{D}}_l$ are defined in the Appendix. Numerical evaluations, which are presented in Sec. 3, indicate that $\Delta E(a,\lambda)$ is finite for $\lambda > 0$, but that the sum diverges for $\lambda = 0$. This behavior shows that the series is not uniformly convergent around $\lambda = 0$, and also raises the possibility that $\Delta E(a,0)$ is infinite.

This ambiguity in the value of $\Delta E(a,0)$ has led us to apply the Mellin transform to investigate more precisely the analytic behavior of $\Delta E(a,\lambda)$ around $\lambda = 0$. The Mellin transform of $\Delta E(a,\lambda)$ is

$$\begin{aligned}\Delta\tilde{E}(a,p) &= \int_0^{\infty}\lambda^{p-1}\Delta E(a,\lambda)d\lambda \\ &= \sum_{l=1}^{\infty}\int_0^{\infty}\lambda^{p-1}[\mathfrak{D}_l(e^{-\lambda k}) + \overline{\mathfrak{D}}_l(e^{-\lambda k})]d\lambda \qquad (5)\\ &= \Gamma(p)\sum_{l=1}^{\infty}[\mathfrak{D}_l(k^{-p}) + \overline{\mathfrak{D}}_l(k^{-p})], \qquad \mathrm{Re}(p) > c \geq 0,\end{aligned}$$

where the constant c should be equal to zero if $\Delta E(a,0)$ is finite or $+1$ if it diverges as λ^{-1}. Now the residues at the poles of $\Delta\tilde{E}(a,p)$ determine the coefficients of various powers of λ (not necessarily integer powers) in an asymptotic expansion of $\Delta E(a,\lambda)$. In particular, $\Gamma(p)$ has a pole of unit residue at $p = 0$, so that we have

$$\Delta E(a,0) = \sum_{l=1}^{\infty}[\mathfrak{D}_l(k^{-0}) + \overline{\mathfrak{D}}_l(k^{-0})]. \tag{6}$$

However, each of the terms in the sum (6) is equal to $\Delta E_l(a,0)$, so that it diverges. Now this is precisely what we should expect if $\Delta\tilde{E}(a,p)$ had a pole at $p = 1$, for then we would restrict (5) to $\mathrm{Re}(p) > 1$. What we must do, therefore, is to investigate $\Delta\tilde{E}(a,p)$ for $\mathrm{Re}(p) > 1$, and find an analytic continuation to $p = 0$.

Application of (A10) and (A12) gives

$$\begin{aligned}\Delta\tilde{E}_l(a,p) = \frac{\Gamma(p)\hbar c(l+\frac{1}{2})}{a^{1-p}}\Bigg(&-\frac{\cos(\pi p/2)}{\pi}\\ &\times\int_0^{\infty}\xi^{-p}[\phi_l(\xi) + \overline{\phi}_l(\xi)]d\xi\\ &+\sum{}'[\gamma_{li}^{1-p} + (\gamma_{li}^{*})^{1-p}] + \sum{}'[\bar{\gamma}_{li}^{1-p} + (\bar{\gamma}_{li}^{*})^{1-p}]\Bigg),\end{aligned} \tag{7}$$

where the notation is defined in the Appendix. Now the integrals in (7) diverge if $\mathrm{Re}(p) \geq 1$, so we need to proceed carefully. An alternative expression for $\Delta\tilde{E}_l(a,p)$ is given by using (A7) and we see that if we use this expression, the factor ξ^{-p}, which causes (7) to diverge at the origin, gives no trouble for $\mathrm{Re}(p) < 4$. Hence to use (7) for $\mathrm{Re}(p) \geq 1$, we analytically continue by treating the divergent part of the integral directly. To do this, we need the following expansions of the functions ϕ_l and $\overline{\phi}_l$ defined in (A13):

$$\begin{aligned}\phi_l(\xi) &= (2l+1)\{1 + \alpha_2\xi^2 + \alpha_4\xi^4 + \cdots\},\\ \overline{\phi}_l(\xi) &= (2l+1)\{1 + \overline{\alpha}_2\xi^2 + \overline{\alpha}_4\xi^4 + \cdots\}.\end{aligned} \tag{8}$$

Now we write

$$\begin{aligned}\int_0^{\infty}\xi^{-p}[\phi_l(\xi) + \overline{\phi}_l(\xi)]d\xi &= \frac{4l+2}{1-p}\\ &+\int_0^1\xi^{-p}[\phi_l(\xi) + \overline{\phi}_l(\xi) - 4l - 2]d\xi\\ &+\int_1^{\infty}\xi^{-1}[\phi_l(\xi) + \overline{\phi}_l(\xi)]d\xi\end{aligned} \tag{9}$$

and this performs the necessary extension for $\mathrm{Re}(p) < 3$. Note that there is no pole at $p = 1$, as we would expect from our comments above on the possible use of (A7). At $p = 1$, the factor $\cos(\pi p/2)$ cancels the finite integrals in (9), so that (7) becomes

$$\Delta\tilde{E}_l(a,1) = \hbar c(l+\tfrac{1}{2})[-(2l+1) + \sum{}'2 + \sum{}'2]. \tag{10}$$

The two sums are equal to the total number of complex zeroes of $h_l^{(2)}$ and $(\gamma h_l^{(2)})'$; that is[7] to $(2l+1)$, so that $\Delta\tilde{E}_l(a,1) = 0$.

We have evaluated (7) numerically using (9) for $0 \leq p \leq 2$, and we find that for large l we have the asymptotic form

$$\Delta\tilde{E}_l(a,p)/\Gamma(p) = f(p)[(2l+1)^{-p} + O(2l+1)^{-p-2}]. \tag{11}$$

We show in Fig. 1 a graph of $f(p)$, which is seen to have a zero at $p = 1$, as we have just proved. It is

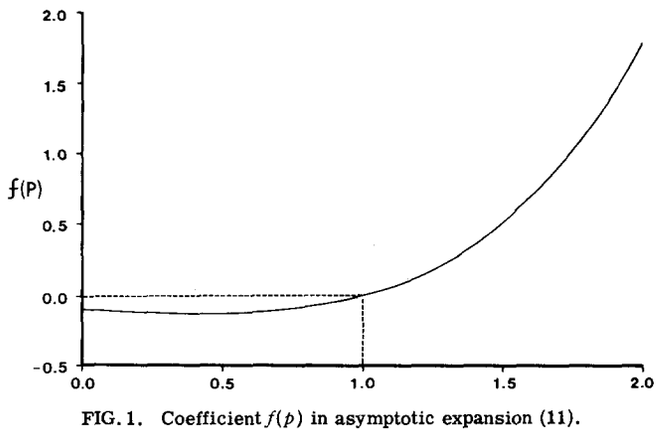


FIG. 1. Coefficient $f(p)$ in asymptotic expansion (11).

TABLE I. Computed values of $\Delta E_l(a,\lambda)$ defined by (4).

λ	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
l										
1	−0.094 542	−0.093 440	−0.090 233	−0.085 173	−0.078 603	−0.070 904	−0.062 456	−0.053 614	−0.044 690	−0.035 946
	−0.093 970	−0.091 048	−0.082 923	−0.071 046	−0.057 043	−0.042 373	−0.028 177	−0.015 242	−0.004 038	0.005 231
	−0.093 852	−0.088 286	−0.073 715	−0.054 333	−0.034 006	−0.015 434	−0.000 092	0.011 505	0.019 480	0.024 309
	−0.093 809	−0.084 863	−0.063 032	−0.036 975	−0.013 110	0.005 373	0.017 738	0.024 648	0.027 342	0.027 130
5	−0.093 788	−0.080 821	−0.051 570	−0.020 623	0.003 628	0.018 861	0.026 144	0.027 751	0.025 936	0.022 445
	−0.093 777	−0.076 247	−0.039 954	−0.006 327	0.015 620	0.025 863	0.027 844	0.025 233	0.020 746	0.016 007
	−0.093 770	−0.071 239	−0.028 686	0.005 396	0.023 172	0.028 006	0.025 620	0.020 443	0.015 005	0.010 409
	−0.093 766	−0.065 894	−0.018 141	0.014 419	0.027 039	0.026 968	0.021 606	0.015 371	0.010 139	0.006 348
	−0.093 762	−0.060 307	−0.008 575	0.020 883	0.028 117	0.024 143	0.017 171	0.010 960	0.006 519	0.003 692
10	−0.093 760	−0.054 563	−0.000 139	0.025 084	0.027 256	0.020 545	0.013 064	0.007 507	0.004 035	0.002 069
	−0.093 758	−0.048 744	0.007 096	0.027 389	0.025 177	0.016 835	0.009 610	0.004 982	0.002 422	0.001 126
	−0.093 757	−0.042 921	0.013 125	0.028 182	0.022 441	0.013 393	0.006 882	0.003 223	0.001 419	0.000 598
	−0.093 756	−0.037 158	0.017 994	0.027 828	0.019 455	0.010 404	0.004 820	0.002 042	0.000 814	0.000 311
	−0.093 755	−0.031 509	0.021 779	0.026 647	0.016 494	0.007 925	0.003 314	0.001 270	0.000 459	0.000 159
15	−0.093 755	−0.026 023	0.024 580	0.024 912	0.013 728	0.005 937	0.002 243	0.000 778	0.000 255	0.000 080
	−0.093 754	−0.020 738	0.026 506	0.022 844	0.011 250	0.004 385	0.001 498	0.000 471	0.000 140	0.000 040
	−0.093 754	−0.015 685	0.027 674	0.020 615	0.009 097	0.003 199	0.000 988	0.000 281	0.000 076	0.000 020
	−0.093 753	−0.010 891	0.028 196	0.018 352	0.007 272	0.002 308	0.000 645	0.000 166	0.000 041	0.000 010
	−0.093 753	−0.006 374	0.028 181	0.016 149	0.005 754	0.001 650	0.000 417	0.000 098	0.000 022	0.000 005
20	−0.093 753	−0.002 148	0.027 730	0.014 068	0.004 513	0.001 169	0.000 268	0.000 057	0.000 011	0.000 002
	−0.093 752	0.001 779	0.026 934	0.012 145	0.003 511	0.000 822	0.000 171	0.000 033	0.000 006	0.000 001
	−0.093 752	0.005 403	0.025 874	0.010 403	0.002 712	0.000 574	0.000 108	0.000 019	0.000 003	0.000 001
	−0.093 752	0.008 723	0.024 621	0.008 848	0.002 081	0.000 399	0.000 068	0.000 011	0.000 002	0.000 000
	−0.093 752	0.011 743	0.023 234	0.007 477	0.001 588	0.000 275	0.000 042	0.000 006	0.000 001	0.000 000
25	−0.093 752	0.014 467	0.021 766	0.006 283	0.001 205	0.000 189	0.000 026	0.000 003	0.000 000	0.000 000
	−0.093 752	0.016 904	0.020 258	0.005 251	0.000 910	0.000 129	0.000 016	0.000 002	0.000 000	0.000 000
	−0.093 752	0.019 064	0.018 744	0.004 368	0.000 684	0.000 088	0.000 010	0.000 001	0.000 000	0.000 000
	−0.093 752	0.020 958	0.017 252	0.003 617	0.000 512	0.000 059	0.000 006	0.000 001	0.000 000	0.000 000
	−0.093 752	0.022 598	0.015 802	0.002 983	0.000 382	0.000 040	0.000 004	0.000 001	0.000 000	0.000 000
30	−0.093 752	0.023 998	0.014 411	0.002 451	0.000 284	0.000 026	0.000 003	0.000 001	0.000 000	0.000 000

also possible to evaluate analytically the derivative of (11) with respect to p at $p = 1$, to get

$$\left[\frac{\partial}{\partial p}\frac{\Delta\tilde{E}_l(a,p)}{\Gamma(p)}\right]_{p=1} = \tfrac{1}{2}\hbar c(l+\tfrac{1}{2})\ln\left(\frac{(l+\frac{1}{2})^2}{l(l+1)}\right) = \tfrac{1}{4}\hbar c\,[(2l+1)^{-1} + O(2l+1)^{-3}], \quad (12)$$

which provides added strength to the numerical evidence in support of (11). With the information which we now have, we can perform an analytic continuation of $\Delta\tilde{E}(a,p)$ to $\mathrm{Re}(p) \leq 1$. For $\mathrm{Re}(p) > 1$, (11) shows that the series converges, so that we can rearrange the terms to write

$$\begin{aligned}\Delta\tilde{E}(a,p) &= \Gamma(p)f(p)\sum_{l=1}^{\infty}(2l+1)^{-p} + \sum_{l=1}^{\infty}[\Delta\tilde{E}_l(a,p) - \Gamma(p)f(p)(2l+1)^{-p}] \\ &= \Gamma(p)f(p)[(1-2^{-p})\zeta(p) - 1] + \sum_{l=1}^{\infty}[\Delta\tilde{E}_l(a,p) - \Gamma(p)f(p)(2l+1)^{-p}]. \quad (13)\end{aligned}$$

The series which now appears in (13) is convergent for $\mathrm{Re}(p) > -1$, so that we have an analytic continuation of $\Delta\tilde{E}(a,p)$ into this region. Note that the conjunction of a pole in $\zeta(p)$ and a zero in $f(p)$ at $p = 1$ causes $\Delta\tilde{E}(a,p)$ to be finite there, even though each term in the series expansion is zero at $p = 1$. This shows why it is necessary to use an analytic continuation to reach $p = 0$; the situation is entirely analogous to the computation of the series

$$\sum_{n=1}^{\infty}(p-1)n^{-p} \quad (14)$$

for $p = 0$. This series defines the function $(p-1)\zeta(p)$, which is analytic in the entire plane, and which has the value $+1$ at $p = 1$. Nevertheless, each term is zero at $p = 1$ and the series diverges for $\mathrm{Re}(p) < 1$.

The upshot of this is that $\Delta E(a,\lambda)$ approaches a finite limit as $\lambda \to 0$, because (13) defines a function which has no poles for $0 < \mathrm{Re}(p) \leq 1$. At $p = 0$, there is a pole due to the function $\Gamma(p)$, and the residue of the inverse Mellin transform at this pole gives the value of $\Delta E(a,0)$: It is

$$\Delta E(a,0) = -f(0) + \sum_{l=1}^{\infty}\{\Delta E_l(a,0) - f(0)\}. \quad (15)$$

It is interesting to note that this formula was suggested by Boyer to be a reasonable approximation to the self energy, although not used by him for the actual evaluation. However, the steps leading to the derivation in Boyer's paper are incorrect, as we shall show in the next section. Numerical values of the quantities needed in (15) are given in Table I for l up to 30; inserting them into this formula gives the result

$$\Delta E(a,0) = +0.092\,43. \quad (16)$$

3. NUMERICAL EVALUATION OF $\Delta E(a,\lambda)$

We have evaluated the quantities $\Delta E_l(a,\lambda)$ for various values of λ and l, and the results are given in Table I. Each of these numbers is the difference between the value of the integrals and the residues, both of which are of order $l^3\exp(-\lambda l)$. It is therefore apparent that computer round-off errors limit the maximum value of l for which the error is acceptable. Our experience leads us to believe that the values given in Table I are correct to within one in the sixth decimal place; but that we could not extend the results to higher l values or higher accuracy for $l = 30$, with the present computing facilities. Since the series

TABLE II. Values of $\Delta E(a,\lambda)$ obtained by application of summation formula (18).

λ	0.3	0.4	0.5	0.6	0.7	0.8	0.9
l							
11	−0.5067	0.2721	0.1143	0.0948	0.0894	0.0865	0.0842
	−1.1540	0.1737	0.1034	0.0923	0.0887	0.0863	0.0841
	2.0573	0.1359	0.0979	0.0910	0.0883	0.0862	0.0841
	0.5026	0.1175	0.0949	0.0904	0.0882	0.0862	0.0841
15	0.2840	0.1075	0.0933	0.0900	0.0881	0.0862	0.0841
	0.2014	0.1017	0.0923	0.0898	0.0880	0.0862	0.0840
	0.1603	0.0982	0.0918	0.0897	0.0880	0.0861	0.0840
	0.1370	0.0960	0.0914	0.0897	0.0880	0.0861	0.0840
	0.1226	0.0946	0.0912	0.0896	0.0880	0.0861	0.0840
20	0.1133	0.0937	0.0911	0.0896	0.0880	0.0861	0.0840
	0.1072	0.0931	0.0910	0.0896	0.0880	0.0861	0.0840
	0.1030	0.0927	0.0910	0.0896	0.0880	0.0861	0.0840
	0.1000	0.0924	0.0910	0.0896	0.0880	0.0861	0.0840
	0.0980	0.0922	0.0909	0.0896	0.0880	0.0861	0.0840
25	0.0965	0.0921	0.0909	0.0896	0.0880	0.0861	0.0840
	0.0954	0.0921	0.0909	0.0896	0.0880	0.0861	0.0840
	0.0947	0.0920	0.0909	0.0896	0.0880	0.0861	0.0840
	0.0941	0.0920	0.0909	0.0896	0.0880	0.0861	0.0840
	0.0937	0.0920	0.0909	0.0896	0.0880	0.0861	0.0840
30	0.0934	0.0919	0.0909	0.0896	0.0880	0.0861	0.0840

does not converge uniformly, we have to compute $\Delta E(a,\lambda)$ for various values of λ and extrapolate to $\lambda = 0$. We are limited to a maximum value of $l = 30$ in forming the infinite sums and we see from Table I that as a result of this we will not be able to evaluate $\Delta E(a,\lambda)$ for values of λ smaller than about 0.3. Even for this value of λ, there is an appreciable contribution from terms with $l > 30$, so that we have to use a summation formula to estimate the tail of the series. Any such formula depends on an assumed asymptotic form for the terms in the series. Boyer has conjectured that this asymptotic form is

$$\Delta E_l(a,\lambda) \approx q e^{-[l+(1/2)]\lambda}; \tag{17}$$

but it is obvious from our results that this cannot hold for small values of $(l + \frac{1}{2})\lambda$. In fact, if (17) did hold, then $\Delta E(a,\lambda)$ would diverge as λ^{-1} for small λ, which is not the case. The trouble with (17) is that it appears to be an asymptotic form for large l and fixed λ, and in fact our numerical results show that it is not too bad if $l\lambda \gtrsim 10$. However, Boyer uses this asymptotic form to sum the series and then takes $\lambda \to 0$, which violates the condition for the validity of the approximation. Now (17) expresses the fact that the terms become close to those of a geometric progression for large l, and using this fact, we have made the approximation

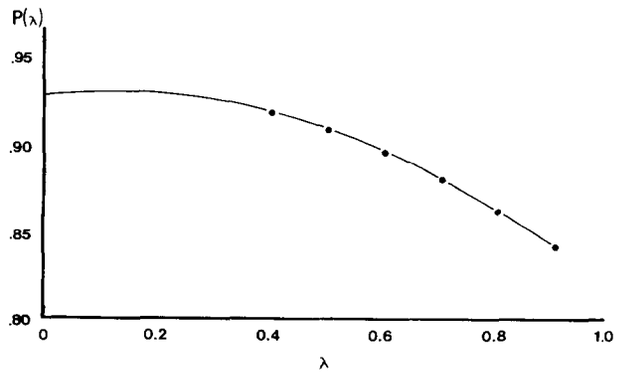


FIG. 2. Extrapolation of $\Delta E(a,\lambda)$ to $\lambda = 0$ given by quadratic $p(\lambda) = 0.0929 + 0.0035\lambda - 0.0150\lambda^2$. Data points from Table II.

$$\Delta E(a,\lambda) \cong \sum_{l=1}^{L-2} \Delta E_l(a,\lambda) - \frac{\Delta E_{L-1}^2(a,\lambda)}{\Delta E_{L-1}(a,\lambda) - \Delta E_L(a,\lambda)}, \tag{18}$$

which would be exact if the series were exactly a geometric progression for $l \geq L - 1$. The values of this approximation to $\Delta E(a,\lambda)$ are shown in Table II. For $\lambda \geq 0.3$, it would seem that we have values of $\Delta E(a,\lambda)$ to within one digit in the fourth decimal place, except for $\lambda = 0.3$, where the calculation has obviously not yet converged, and the value is too high. We have fitted these results to the quadratic expression $p(\lambda) = 0.0929 + 0.0035\lambda - 0.0150\lambda^2$, which fits the values from $\lambda = 0.4$ to 0.8 exactly. These results are displayed in Fig. 2, where it is seen that the point for $\lambda = 0.3$ is above the curve. The extrapolation given by $p(\lambda)$ leads to

$$\Delta E(a,0) \approx 0.0929, \tag{19}$$

which is in reasonable agreement with the exact result (16). In theory, the extrapolation could be improved by including higher powers of λ, but this would make it necessary to obtain values of $\Delta E_l(a,\lambda)$ to higher accuracy and for higher values of l. In view of the ease with which the value of $\Delta E(a,0)$ can be computed using the Mellin transform, there is no point in improving the accuracy of this extrapolation. The most important aspect of these numerical calculations is that they demonstrate the nature of the conditional convergence of the series (3) around $\lambda = 0$. From Table I we see that the initial negative terms in the series, of which there are an increasing number as λ decreases, almost cancel the positive terms in the tail. An example of a simple series with these properties is

$$a_n(\lambda) = e^{-n\lambda} - 2e^{-2n\lambda}, \tag{20}$$

which has the properties

$$a_n(0) = -1, \qquad \lim_{\lambda\to 0} \sum_{n=1}^{\infty} a_n(\lambda) = +\tfrac{1}{2}, \tag{21}$$

which are very similar to those which we have already encountered.

4. CONCLUSIONS

We have determined the necessary analytic properties of $\Delta E(a,\lambda)$ to allow an accurate determination of $\Delta E(a,0)$. The physical reason why this energy is positive is not clear; and the situation is made more puzzling by the following considerations

(a) If the sphere is flattened, it must eventually approximate a pair of parallel plates, for which the energy is negative. In fact, it was the fact that the force is attractive for parallel plates which led Casimir to suggest his model of the electron.

(b) If we consider a sphere of dielectric material, we would expect a negative energy, because the model is also appropriate for a macroscopic sphere of material held together by Van der Waals' forces and these are attractive.

We are at present attempting to elucidate these problems, and in particular to develop an effective way of evaluating the quantum electromagnetic energy of a dielectric sphere. It is too early to say what the results of this investigation will be, or what light (if any) will be shed on the present problem.

ACKNOWLEDGMENT

We wish to thank Professor B. W. Ninham for stimulating discussions and for his encouragement.

APPENDIX

In this appendix we derive an expression for $\Delta E_l(a)$ as a contour integral. In order to avoid unnecessary duplication of technical details contained in Boyer's paper,[1] we use his notation for the frequencies of the various normal modes. The reader should therefore refer to this paper for definitions and details of the various symbols which we use in the following. For the TE modes, we need to evaluate the quantities

$$\mathfrak{D}_l(\mathfrak{F}) = \hbar c(l+\tfrac{1}{2})[\sum_s k_{ls}(a)\mathfrak{F}(k_{ls}(a)) + \sum_S K_{lS}(a,R)\mathfrak{F}(K_{lS}(a,R)) - \sum_s k_{ls}(R)\mathfrak{F}(k_{ls}(R))]. \tag{A1}$$

It is important to our analysis in Sec. 2 that this sum is a linear functional in the cut-off function, and so we have denoted it by $\mathfrak{D}_l(\mathfrak{F})$. For large R, we have[8]

$$\sum_S K_{lS}(a,R)\mathfrak{F}(K_{lS}(a,R)) \approx \int_1^\infty K_{lS}\mathfrak{F}(K_{lS})\,dS \approx \int_0^\infty K\mathfrak{F}(K)\left(\frac{\partial S_l(a,R,K)}{\partial K}\right)dK \tag{A2}$$

$$\sum_s k_{ls}(R)\mathfrak{F}(k_{ls}(R)) \approx \int_1^\infty k_{ls}\mathfrak{F}(k_{ls})\,ds \approx \int_0^\infty k\mathfrak{F}(k)\left(\frac{\partial s_l(R,k)}{\partial k}\right)dk$$

and, using the relation $S_l(a,R,K) = s_l(R,K) - s_l(a,K)$, we get

$$\mathfrak{D}_l(\mathfrak{F}) = \frac{\hbar c(l+\tfrac{1}{2})}{a}\left[\sum_s \gamma_{ls}\mathfrak{F}\left(\frac{\gamma_{ls}}{a}\right) - \int_0^\infty \gamma\mathfrak{F}\left(\frac{\gamma}{a}\right)\left(\frac{\partial s_l(\gamma)}{\partial\gamma}\right)d\gamma\right], \tag{A3}$$

where γ_{ls} and $s_l(\gamma)$ are related by

$$g_l(s,\gamma) = \cos(\pi s)j_l(x) + \sin(\pi s)y_l(x) = 0. \tag{A4}$$

Note that we use $y_l(x)$ for the spherical Neumann functions in place of Boyer's $n_l(x)$.

For the TM modes the analysis is very similar, giving the result

$$\bar{\mathfrak{D}}_l(\mathfrak{F}) = \frac{\hbar c(l+\tfrac{1}{2})}{a}\left[\sum_{\bar{s}} \bar{\gamma}_{l\bar{s}}\mathfrak{F}\left(\frac{\bar{\gamma}_{l\bar{s}}}{a}\right) - \int_0^\infty \bar{\gamma}\mathfrak{F}\left(\frac{\bar{\gamma}}{a}\right)\left(\frac{\partial\bar{s}_l(\bar{\gamma})}{\partial\bar{\gamma}}\right)d\bar{\gamma}\right], \tag{A5}$$

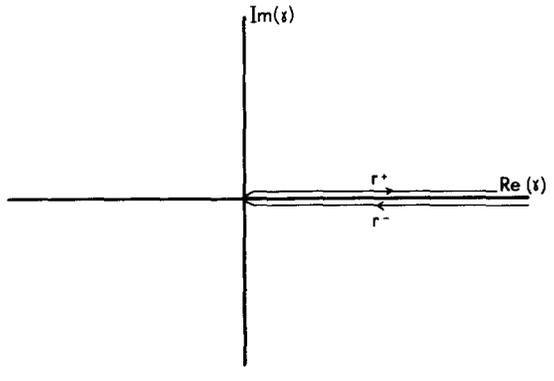


FIG. 3. Contour of integration for (A9).

where

$$\bar{g}_l(\bar{s},\bar{\gamma}) = \cos(\pi\bar{s})(\bar{\gamma}j_l(\bar{\gamma}))' + \sin(\pi\bar{s})(\bar{\gamma}y_l(\bar{\gamma}))' = 0. \tag{A6}$$

By using the Euler–Mclaurin summation formula and integrating by parts, we get the final expression given in Boyer:

$$\Delta E(a) = \lim_{\mathfrak{F}} \sum_{l=1}^\infty [\mathfrak{D}_l(\mathfrak{F}) + \bar{\mathfrak{D}}_l(\mathfrak{F})] = \lim_{\mathfrak{F}} \sum_{l=1}^\infty \frac{\hbar c(l+\tfrac{1}{2})}{a}\int_0^\infty \{s_l(x) - [s_l(x)] + \bar{s}_l(x) - [\bar{s}_l(x)] - 1\}\frac{d}{dx}\left[x\mathfrak{F}\left(\frac{x}{a}\right)\right]. \tag{A7}$$

We may now use (A4) to obtain an expression for the quantity $\partial s_l/\partial\gamma$ which appears in (A3), as follows:

$$\frac{\partial s}{\partial\gamma} = \frac{1}{2\pi i}\left(\frac{h'^{(1)}(\gamma)}{h^{(1)}(\gamma)} - \frac{h'^{(2)}(\gamma)}{h^{(2)}(\gamma)}\right),$$

where

$$h_l^{(1,2)} = j_l \pm iy_l. \tag{A8}$$

Equation (A3) may now be rewritten as a contour integral involving various spherical Bessel functions. The contour is chosen so as to encircle all the real zeroes of $j_l(x)$ and to avoid all the complex zeroes of the spherical Hankel functions. This is shown in Fig. 3. Using (A8), we may therefore write (A3) as

$$\mathfrak{D}_l(\mathfrak{F}) = -\frac{\hbar c(l+\tfrac{1}{2})}{2\pi i a}\times\left[\int_{\Gamma^-}\gamma\mathfrak{F}(\gamma/a)\left(\frac{j_l'(\gamma)}{j_l(\gamma)} - \frac{h_l'^{(1)}(\gamma)}{h_l^{(1)}(\gamma)}\right)d\gamma + \int_{\Gamma^+}\gamma\mathfrak{F}(\gamma/a)\left(\frac{j_l'(\gamma)}{j_l(\gamma)} - \frac{h_l'^{(2)}(\gamma)}{h_l^{(2)}(\gamma)}\right)d\gamma\right]. \tag{A9}$$

Finally we deform this contour so that the integrations are taken along the imaginary axis. In the process, we must pick up the residues at the complex poles of the Hankel functions, $j_l(x)$ having no complex poles.[9] Furthermore, the Hankel functions have a zero on the imaginary axis for odd values of l; we therefore have to pick up only half of the residue at this pole and evaluate the integral as a principal value integral. The integrand itself can be simplified by using the Wronskian relations on the square braces, changing variables to $\xi = i\gamma$, and using the parity relations for spherical Bessel functions to express all quantities in the lower half plane in terms of quantities in the upper half plane. The result of all these manipulations is

$$\mathfrak{D}_l(\mathfrak{F}) = \frac{-\hbar c(l+\tfrac{1}{2})}{2\pi a}\int_0^\infty\left[\mathfrak{F}\left(\frac{i\xi}{a}\right) + \mathfrak{F}\left(\frac{-i\xi}{a}\right)\right]\phi_l(\xi)d\xi + \frac{\hbar c(l+\tfrac{1}{2})}{a}\sum{}'\left[\gamma_{li}\mathfrak{F}\left(\frac{\gamma_{li}}{a}\right) + \gamma_{li}^*\mathfrak{F}\left(\frac{\gamma_{li}^*}{a}\right)\right]. \tag{A10}$$

Here the stroke through the integral sign indicates that a principal value is to be taken if there is a pole in the integrand. The quantities γ_{li} and γ_{li}^* are those complex zeroes of the functions $h_l^{(2)}$ and $h_l^{(1)}$, respectively, which fall in the right-hand half-plane. The prime on the summation sign indicates that if one of the roots γ_{li} is pure imaginary, only half of the contribution is to be taken. $\phi_l(\xi)$ will be defined below.

Now a similar series of manipulations yield a similar result for the TM modes. Corresponding to (A9), we get

$$\overline{\mathfrak{D}}_l(\mathfrak{F}) = -\frac{\hbar c(l+\frac{1}{2})}{2\pi i a} \times\left[\int_{\Gamma^-}\gamma\mathfrak{F}\left(\frac{\gamma}{a}\right)\left(\frac{(\gamma j_l(\gamma))''}{(\gamma j_l(\gamma))'}-\frac{(\gamma h_l^{(1)}(\gamma))''}{(\gamma h_l^{(1)}(\gamma))'}\right)d\gamma + \int_{\Gamma^+}\gamma\mathfrak{F}\left(\frac{\gamma}{a}\right)\left(\frac{(\gamma j_l(\gamma))''}{(\gamma j_l(\gamma))'}-\frac{(\gamma h_l^{(2)}(\gamma))''}{(\gamma h_l^{(2)}(\gamma))'}\right)d\gamma\right], \quad \text{(A11)}$$

and corresponding to (A10) we get

$$\overline{\mathfrak{D}}_l(\mathfrak{F}) = \frac{-\hbar c(l+\frac{1}{2})}{2\pi a}\int_0^\infty\left[\mathfrak{F}\left(\frac{i\xi}{a}\right)+\mathfrak{F}\left(\frac{-i\xi}{a}\right)\right]\overline{\phi}_l(\xi)d\xi + \frac{\hbar c(l+\frac{1}{2})}{a}\sum{}'\left[\bar{\gamma}_{li}\mathfrak{F}\left(\frac{\bar{\gamma}_{li}}{a}\right)+\bar{\gamma}_{li}^*\mathfrak{F}\left(\frac{\bar{\gamma}_{li}^*}{a}\right)\right]. \quad \text{(A12)}$$

Here the quantities $\bar{\gamma}_{li}$ and $\bar{\gamma}_{li}^*$ are the complex roots of the functions $(\gamma h_l^{(2)})'$ and $(\gamma h_l^{(1)})'$, respectively, which fall in the right-hand half plane. The functions ϕ_l and $\overline{\phi}_l$ are defined by

$$\phi_l(\xi) = [\xi j_l(i\xi)h_l^{(2)}(i\xi)]^{-1},$$
$$\overline{\phi}_l(\xi) = \frac{-[l(l+1)+\xi^2]}{[\xi(\xi j_l(i\xi))'(\xi h_l^{(2)}(i\xi))']}, \quad \text{(A13)}$$

where the prime denotes $d/d\xi$.

Equations (A10) and (A12) represent a convenient form for computation, compared with an expression of the type (A7). The integrals converge rapidly because of the exponential rate of decrease of the functions ϕ_l and $\overline{\phi}_l$ for large ξ, and for each value of l a knowledge of the $(2l+1)$ complex zeroes involved in (A10) and (A12) is sufficient to allow determination of these quantities to arbitrary accuracy. With (A7), however, the rate of convergence of the integral is determined by the cutoff function, and increasing the accuracy of a particular evaluation involves the evaluation of more of the infinite number of real zeroes of the Bessel functions.

1 T. H. Boyer, Phys. Rev. **174**, 1764 (1968).

2 E. M. Lifshitz, Zh. Eksp. Teor. Fiz. **29**, 94 (1955)[Sov. Phys. JETP **2**, 73 (1956)]; T. H. Boyer, Phys. Rev. **182**, 1374 (1969); **186**, 1304 (1969); D **1**, 1526 (1970); D 1, 2257 (1970). T. H. Boyer, Ann. Phys. (N.Y.) **56**, 474 (1970); D. Langbein, Phys. Rev. B 2, 3371 (1970); M. Fierz, Helv. Phys. Acta. **33**, 855 (1960); N. G. Van Kampen, B. R. A. Nijboer, and K. Schram, Phys. Letters **26A**, 307 (1968); C. M. Hargreaves, Proc. Koninkl. Ned. Akad. Wetenschap. **B68**, 231 (1965).

3 T. H. Boyer, Phys. Rev. **174**, 1631 (1968); **180**, 19 (1969). B. W. Ninham and V. A. Parsegian, J. Chem. Phys. **52**, 4578 (1970); Biophys. J. **10**, 646 (1970); J. Chem. Phys. **53**, 3398 (1970); B. W. Ninham, V. A. Parsegian, and G. H. Weiss, J. Stat. Mech. **2**, 323 (1970); V. A. Parsegian and B. W. Ninham, Nature **224**, 1197 (1969); Biophys. J. **10**, 664 (1970); J. Colloid Interface Sci. (to be published). I. E. Dzyaloshinskii, E. M. Lifshitz, and L. P. Pitaevskii, Advan. Phys. 10, 165 (1961).

4 H. B. G. Casimir, Proc. Koninkl. Ned. Akad. Wetenschap **51**, 793 (1948).

5 H. B. G. Casimir, Physica **19**, 846 (1953).

6 Note that Boyer performs his calculation in a slightly different way, taking the difference in energy between two pairs of concentric spheres; the first pair of radii a and R, the second pair of radii R/η and R. His expression for the energy therefore contains pairs of terms identical except for the replacement of a by R/η. Our method eliminates the unwanted terms containing R/η.

7 B. Davies (to be published).

8 A detailed discussion of the validity of this transformation is the subject of much of Boyer's paper. See Ref. 1.

9 M. Abrimowitz and I. A. Stegun, *Handbook of Mathematical Functions* (U.S. Natl. Bur. Stds., Washington, D.C., 1964).

On the Symmetric Tensor Operators of the Unitary Groups

S. J. Ališauskas, A.-A. A. Jucys, and A. P. Jucys

Institute of Physics and Mathematics of the Academy of Sciences of Lithuanian SSR, USSR, Vilnius, 232600

(Received 5 April 1971)

The algebraic expressions for the matrix elements of symmetric tensor operators (the powers of infinitesimal operators) of the unitary groups in the Gel'fand basis have been studied. The expressions for the isoscalar factors of the related Clebsch–Gordan coefficients, one of the two representations to be coupled being symmetric, as well as the elements of a special recoupling matrix have been found. The supplementary symmetry properties of the isoscalar factors corresponding to the Regge symmetries of the Wigner and $6j$ coefficients of SU_2 have been examined.

1. INTRODUCTION

The mathematical apparatus of the irreducible tensor operators of the unitary groups[1] is a very important generalization of the theory of angular momentum of contemporary theoretical physics. The main problem of this apparatus is to obtain the algebraic expressions for Clebsch–Gordan (CG) coefficients, recoupling matrices and matrix elements of irreducible tensor operators. It is useful at first to solve the simpler special problems, for example, to consider the matrix elements of the extremal tensor operators of the unitary groups.[2]

An interesting and more difficult problem is to obtain the expressions for the matrix elements of the symmetric tensor operators, which enables one to find the expressions for the general tensor operators. The main aim of this paper is to consider these symmetric operators. We take the product of powers of commuting infinitesimal operators (generators) of U_{n+1} of the type E_{in+1} $(i = 1, 2, \ldots, n)$, as a realization of the symmetric tensor operator of U_n. The matrix elements of such an operator can be expressed as a product of the reduced matrix elements and isoscalar factors (i.f.) of CG coefficients with one of the two representations symmetric. The combinatorial-graphical techniques for calculating such a special i.f. has been found by one[3] of the authors by the use of the Young operators of the symmetric groups as projection operators. The expressions of Ref. 3 are not optimal ones with respect to the number of terms in the sum, because the summation is taking place over the permutations of labeled squares of the Young tableau. Here we are going to obtain the

Now a similar series of manipulations yield a similar result for the TM modes. Corresponding to (A9), we get

$$\overline{\mathfrak{D}}_l(\mathfrak{F}) = -\frac{\hbar c(l+\frac{1}{2})}{2\pi i a} \times\left[\int_{\Gamma^-}\gamma\mathfrak{F}\left(\frac{\gamma}{a}\right)\left(\frac{(\gamma j_l(\gamma))''}{(\gamma j_l(\gamma))'}-\frac{(\gamma h_l^{(1)}(\gamma))''}{(\gamma h_l^{(1)}(\gamma))'}\right)d\gamma + \int_{\Gamma^+}\gamma\mathfrak{F}\left(\frac{\gamma}{a}\right)\left(\frac{(\gamma j_l(\gamma))''}{(\gamma j_l(\gamma))'}-\frac{(\gamma h_l^{(2)}(\gamma))''}{(\gamma h_l^{(2)}(\gamma))'}\right)d\gamma\right], \quad \text{(A11)}$$

and corresponding to (A10) we get

$$\overline{\mathfrak{D}}_l(\mathfrak{F}) = \frac{-\hbar c(l+\frac{1}{2})}{2\pi a}\int_0^\infty\left[\mathfrak{F}\left(\frac{i\xi}{a}\right)+\mathfrak{F}\left(\frac{-i\xi}{a}\right)\right]\overline{\phi}_l(\xi)d\xi + \frac{\hbar c(l+\frac{1}{2})}{a}\sum{}'\left[\bar{\gamma}_{li}\mathfrak{F}\left(\frac{\bar{\gamma}_{li}}{a}\right)+\bar{\gamma}_{li}^*\mathfrak{F}\left(\frac{\bar{\gamma}_{li}^*}{a}\right)\right]. \quad \text{(A12)}$$

Here the quantities $\bar{\gamma}_{li}$ and $\bar{\gamma}_{li}^*$ are the complex roots of the functions $(\gamma h_l^{(2)})'$ and $(\gamma h_l^{(1)})'$, respectively, which fall in the right-hand half plane. The functions ϕ_l and $\overline{\phi}_l$ are defined by

$$\phi_l(\xi) = [\xi j_l(i\xi)h_l^{(2)}(i\xi)]^{-1},$$
$$\overline{\phi}_l(\xi) = \frac{-[l(l+1)+\xi^2]}{[\xi(\xi j_l(i\xi))'(\xi h_l^{(2)}(i\xi))']}, \quad \text{(A13)}$$

where the prime denotes $d/d\xi$.

Equations (A10) and (A12) represent a convenient form for computation, compared with an expression of the type (A7). The integrals converge rapidly because of the exponential rate of decrease of the functions ϕ_l and $\overline{\phi}_l$ for large ξ, and for each value of l a knowledge of the $(2l+1)$ complex zeroes involved in (A10) and (A12) is sufficient to allow determination of these quantities to arbitrary accuracy. With (A7), however, the rate of convergence of the integral is determined by the cutoff function, and increasing the accuracy of a particular evaluation involves the evaluation of more of the infinite number of real zeroes of the Bessel functions.

1 T. H. Boyer, Phys. Rev. **174**, 1764 (1968).

2 E. M. Lifshitz, Zh. Eksp. Teor. Fiz. **29**, 94 (1955)[Sov. Phys. JETP **2**, 73 (1956)]; T. H. Boyer, Phys. Rev. **182**, 1374 (1969); **186**, 1304 (1969); D **1**, 1526 (1970); D 1, 2257 (1970). T. H. Boyer, Ann. Phys. (N.Y.) **56**, 474 (1970); D. Langbein, Phys. Rev. B 2, 3371 (1970); M. Fierz, Helv. Phys. Acta. **33**, 855 (1960); N. G. Van Kampen, B. R. A. Nijboer, and K. Schram, Phys. Letters **26A**, 307 (1968); C. M. Hargreaves, Proc. Koninkl. Ned. Akad. Wetenschap. **B68**, 231 (1965).

3 T. H. Boyer, Phys. Rev. **174**, 1631 (1968); **180**, 19 (1969). B. W. Ninham and V. A. Parsegian, J. Chem. Phys. **52**, 4578 (1970); Biophys. J. **10**, 646 (1970); J. Chem. Phys. **53**, 3398 (1970); B. W. Ninham, V. A. Parsegian, and G. H. Weiss, J. Stat. Mech. **2**, 323 (1970); V. A. Parsegian and B. W. Ninham, Nature **224**, 1197 (1969); Biophys. J. **10**, 664 (1970); J. Colloid Interface Sci. (to be published). I. E. Dzyaloshinskii, E. M. Lifshitz, and L. P. Pitaevskii, Advan. Phys. 10, 165 (1961).

4 H. B. G. Casimir, Proc. Koninkl. Ned. Akad. Wetenschap **51**, 793 (1948).

5 H. B. G. Casimir, Physica **19**, 846 (1953).

6 Note that Boyer performs his calculation in a slightly different way, taking the difference in energy between two pairs of concentric spheres; the first pair of radii a and R, the second pair of radii R/η and R. His expression for the energy therefore contains pairs of terms identical except for the replacement of a by R/η. Our method eliminates the unwanted terms containing R/η.

7 B. Davies (to be published).

8 A detailed discussion of the validity of this transformation is the subject of much of Boyer's paper. See Ref. 1.

9 M. Abrimowitz and I. A. Stegun, *Handbook of Mathematical Functions* (U.S. Natl. Bur. Stds., Washington, D.C., 1964).

On the Symmetric Tensor Operators of the Unitary Groups

S. J. Ališauskas, A.-A. A. Jucys, and A. P. Jucys

Institute of Physics and Mathematics of the Academy of Sciences of Lithuanian SSR, USSR, Vilnius, 232600

(Received 5 April 1971)

The algebraic expressions for the matrix elements of symmetric tensor operators (the powers of infinitesimal operators) of the unitary groups in the Gel'fand basis have been studied. The expressions for the isoscalar factors of the related Clebsch–Gordan coefficients, one of the two representations to be coupled being symmetric, as well as the elements of a special recoupling matrix have been found. The supplementary symmetry properties of the isoscalar factors corresponding to the Regge symmetries of the Wigner and $6j$ coefficients of SU_2 have been examined.

1. INTRODUCTION

The mathematical apparatus of the irreducible tensor operators of the unitary groups[1] is a very important generalization of the theory of angular momentum of contemporary theoretical physics. The main problem of this apparatus is to obtain the algebraic expressions for Clebsch–Gordan (CG) coefficients, recoupling matrices and matrix elements of irreducible tensor operators. It is useful at first to solve the simpler special problems, for example, to consider the matrix elements of the extremal tensor operators of the unitary groups.[2]

An interesting and more difficult problem is to obtain the expressions for the matrix elements of the symmetric tensor operators, which enables one to find the expressions for the general tensor operators. The main aim of this paper is to consider these symmetric operators. We take the product of powers of commuting infinitesimal operators (generators) of U_{n+1} of the type E_{in+1} $(i = 1, 2, \ldots, n)$, as a realization of the symmetric tensor operator of U_n. The matrix elements of such an operator can be expressed as a product of the reduced matrix elements and isoscalar factors (i.f.) of CG coefficients with one of the two representations symmetric. The combinatorial-graphical techniques for calculating such a special i.f. has been found by one[3] of the authors by the use of the Young operators of the symmetric groups as projection operators. The expressions of Ref. 3 are not optimal ones with respect to the number of terms in the sum, because the summation is taking place over the permutations of labeled squares of the Young tableau. Here we are going to obtain the

corresponding expressions with a greatly reduced number of terms in the sums, the summation taking place over integers only. Furthermore, a rather simple recurrence method is described for obtaining the expressions under consideration. The technique resembles that used by one of the authors[4] in the case of SU_3. The two classes of expressions are given and their symmetry properties, partly given in Ref. 3, discussed.

In Appendix A it is shown how the corresponding formulas can be obtained from the results of Ref. 3.

In Appendix B there is given a useful relation between the isoscalar factors of the Gel'fand basis and the elements of the recoupling matrix described in Ref. 5.

The results of this paper can be used for obtaining the general expressions for the CG coefficients of U_n. For this purpose one can use the projection operators in the form of polynomials of infinitesimal operators[6] as has been done in the case of SU_3.[7] Alternatively, one can use the recurrence relations obtained by forming the general tensor operators from the symmetric ones. The normalization procedure in the second case can be carried out using the relation of Ref. 5.

2. THE MATRIX ELEMENTS OF POWERS OF INFINITESIMAL OPERATORS

At first we obtain the expression for the i.f. of a special kind

$$\begin{bmatrix} [m]_n & p & [m']_n \\ [m]_{n-1} & 0 & [m]_{n-1} \end{bmatrix} = [p! \prod_{1\leqslant j\leqslant k\leqslant n} (m'_{jn} - m'_{kn} - j + k)]^{1/2} \times \frac{D([m]_n, [m']_n)\Gamma([m']_n, [m]_{n-1})}{\Gamma([m]_n, [m]_{n-1})}. \quad (1)$$

Here $[m]_k \equiv [m_{1k}, m_{2k}, \ldots, m_{kk}]$ means the corresponding row of the Gel'fand pattern of the representation of U_n (c.f. Ref. 1), p is the single parameter of the symmetric representation. In (1), and in what follows, we use the notations

$$\Gamma([m]_n, [m]_{n-1}) = \left(\frac{\prod_{1\leqslant i\leqslant j\leqslant n-1} (m_{in} - m_{jn-1} - i + j)!}{\prod_{1\leqslant j<i\leqslant n} (m_{jn-1} - m_{in} + i - j - 1)!} \right)^{1/2}; \quad (2)$$

$$D([m]_n, [m']_n) = \left(\frac{\prod_{1\leqslant i<j\leqslant n} (m_{in} - m'_{jn} - i + j - 1)!}{\prod_{1\leqslant j\leqslant i\leqslant n} (m'_{jn} - m_{in} + i - j)!} \right)^{1/2}. \quad (3)$$

The dependence of the i.f. of Eq. (1) on the parameters of the representations of the subgroup U_{n-1} is confined to Γ, this dependence being deduced by factorizing the simpler i.f.[8] The remaining part of the expression on the right-hand side of Eq. (1) is a normalization factor. This factor can be deduced by equating the particular case of Eq. (1) ($[m]_{n-1} = [m]_n$, $m_{nn} = 0$) to the one calculated with the help of projection operators (Refs 6 and 7). It must be noted that we use the general weight lowering operators of the form

$$F_- \begin{bmatrix} [h]_n \\ [h']_{n-1} \end{bmatrix} = \prod_{i=1}^{n-1} \left(\frac{1}{(h_i - h'_i)} \prod_{s=i+1}^{n-1} \frac{(h_i - h'_s - i + s)!}{(h'_i - h'_s - i + s)!} \times \prod_{s=i+1}^{n} \frac{(h'_i - h_s - i + s - 1)!}{(h_i - h_s - i + s - 1)!} \right)^{1/2} \times P^{n-1}_{max} [h']_{n-1} \prod_{i=1}^{n-1} E_{ni}^{\,h_i - h'_i}, \quad (4)$$

rather than that of Ref. 9. Here $P^{n-1}_{max}[h']_{n-1}$ is a projection operator of maximal weight as defined in Ref. 6.

The easiest way to obtain (1) is to use the results of Ref. 3 [these last ones being contained in Eq. (A2) of Appendix A]. The sum $F_{p_l}(x_i - y_k)$ in this case reduces to one term equal to 1.

Let us now consider the calculation of the matrix elements of powers of generators. The simplest cases of them are obtained by factorizing the matrix elements of individual generators of the group. To these cases belong, in the first place, the matrix elements stretched with respect to the parameters of the representations of subgroups.

The simplest of powers of generators seems to be $(E_{n-1,n})^p$. With respect to the subgroup U_{n-1}, this operator corresponds to the scalar component U_{n-2} of the symmetric tensor T^p. Hence this matrix element is proportional to the i.f. of U_{n-1} being calculated with the help of (1). The corresponding reduced matrix elements are to be obtained from the relation

$$\left\langle \begin{bmatrix} [m]_n \\ [m'']_{n-1} \end{bmatrix} \middle\| E_{n-1n}{}^p \middle\| \begin{bmatrix} [m]_n \\ [m']_{n-1} \end{bmatrix} \right\rangle = \begin{bmatrix} [m']_{n-1} & p & [m'']_{n-1} \\ [m']_{n-2} & 0 & [m']_{n-2} \end{bmatrix}^{-1} \times \left\langle \begin{bmatrix} [m]_n \\ [m'']_{n-1} \end{bmatrix}_{\max} \middle| F_-^{\dagger} \begin{bmatrix} [m'']_{n-1} \\ [m']_{n-2} \end{bmatrix} (E_{n-1n})^p \middle| \begin{bmatrix} [m]_n \\ [m']_{n-1} \end{bmatrix}_{\max} \right\rangle. \quad (5)$$

Here $m'_{in-2} = m'_{in-1}$ $(i \leqslant n-2)$ is the maximal weight of U_{n-2}. The operator $P^{n-2}_{\max}[m']_{n-2}$ in F_- gives unity in this case. With the help of Eq. (2.11) of Ref. 7 we transpose E_{n-1n} with $F_-^{\dagger}$. All the powers of E_{in-1} with nonvanishing exponents acting on the maximal state of U_{n-1} give zero. For this reason, summations arising in the process of the transposition disappear and we are left with the matrix element of the operator

$$\frac{p!}{(m''_{n-1n-1} - m'_{n-1n-1})!} \prod_{i=1}^{n-1} E_{in}{}^{m''_{in-1} - m'_{in-1}},$$

which is stretched in this case.

The operations described above give the following expression for the reduced matrix element under consideration:

$$\left\langle \begin{bmatrix} [m]_n \\ [m']_{n-1} \end{bmatrix} \middle\| E_{n-1n}{}^p \middle\| \begin{bmatrix} [m]_n \\ [m]_{n-1} \end{bmatrix} \right\rangle = \delta\left(\sum_{i=1}^{n-1} m'_{in-1}, p + \sum_{i=1}^{n-1} m_{in-1} \right) [p! \times \prod_{1\leqslant i<j\leqslant n-1} (m_{in-1} - m_{jn-1} - i + j)]^{1/2} \times \frac{D([m]_{n-1}, [m']_{n-1})\Gamma([m]_n, [m]_{n-1})}{\Gamma([m]_n, [m']_{n-1})}. \quad (6)$$

It is to be noted that this reduced matrix element coincides with those of the operators

$$\left(\frac{p!}{\prod_{i=1}^{n-1}\alpha_i!}\right)^{1/2}\prod_{i=1}^{n-1}E_{in}{}^{\alpha_i},\left(\sum_{i=1}^{n-1}\alpha_i=p\right), \tag{7}$$

because they form the basis of the representation p of U_{n-1}. The corresponding matrix elements themselves are obtained by multiplying the reduced matrix elements obtained above by the products of i.f.'s to be dealt with in what follows.

3. ISOSCALAR FACTOR WITH ONE OF THE REPRESENTATIONS SYMMETRIC

In generalizing and simplifying the method of Ref. 4, we take the lowest weight component T_0^p of the symmetric unit operator of U_n. Its matrix element with respect to the corresponding basis is equal to the i.f. given by Eq. (1). More general components of the same operator are

$$\begin{aligned}T_q^p &= \left(\frac{(p-q)!}{p!q!}\right)^{1/2}\\ &\times [E_{n-1n}[E_{n-1n}[\cdots[E_{n-1n}T_0^p]\cdots]]]\\ &\xleftarrow{\qquad q\text{ times}\qquad}\\ &= \left(\frac{(p-q)!q!}{p!}\right)^{1/2}\sum_x\frac{(-1)^x}{x!(q-x)!}E_{n-1n}{}^{q-x}T_0^pE_{n-1n}{}^x.\end{aligned} \tag{8}$$

The reduced matrix element of this operator with respect to the subgroup U_{n-1} is the i.f. under consideration; it is

$$\begin{aligned}&\begin{bmatrix}[m]_n & p & [m']_n\\ [m]_{n-1} & q & [m']_{n-1}\end{bmatrix}\\ &= [(p-q)!\prod_{1\leqslant i<j\leqslant n-1}(m_{in-1}-m_{jn-1}-i+j)\\ &\times\prod_{1\leqslant i<j\leqslant n}(m'_{in}-m'_{jn}-i+j)]^{1/2}\\ &\times\frac{D([m]_n,[m']_n)\Gamma([m]_n,[m]_{n-1})}{D([m]_{n-1},[m']_{n-1})\Gamma([m']_n,[m']_{n-1})}\\ &\times\sum_{[r]_{n-1}}(-1)^{\sum_{i=1}^{n-1}(r_{in-1}-m_{in-1})}\\ &\times\prod_{1\leqslant i<j\leqslant n-1}(r_{in-1}-r_{jn-1}-i+j)\\ &\times D^2([m]_{n-1},[r]_{n-1})D^2([r]_{n-1},[m']_{n-1})\\ &\times\frac{\Gamma^2([m']_n,[r]_{n-1})}{\Gamma^2([m]_n,[r]_{n-1})},\end{aligned}$$

$$p=\sum_{i=1}^{n}(m'_{in}-m_{in}),\quad q=\sum_{i=1}^{n-1}(m'_{in-1}-m_{in-1}), \tag{9}$$

$[r]_{n-1}$ being the Young scheme and the summation taking place over $n-1$ parameter r_{in-1}. When both of the two representations to be coupled are symmetric, expression (9) reduces to the CG coefficient of SU_2 [the second of Eq. (13.1) of Ref. 10]. On the other hand, when $n=3$ it turns into Eq. (3.14) of Ref. 4. Furthermore, we can limit ourselves to the case $m_{nn}=0$ which does not influence the value of the i.f., as pointed out in Ref. 1.

It is easy to see that (9), after omitting the square root, possesses the high symmetry indicated in Ref. 3. For example, it is possible to transpose the parameters m_{in-1} and m'_{i+1n} with the appearance of the phase factor $(-1)^{m_{in-1}-m'_{i+1n}}$. The other kind of Regge symmetry gives the transposition of m_{in} with m'_{in-1}, without any phase factor. For the tabulation of the symmetric part of (9), it is useful to apply the following scheme of $4n-2$ parameters:

$$\begin{aligned}&m'_{1n},\ \max(m'_{1n-1},m_{1n}),\min(m'_{1n-1},m_{1n}),\ \max(m'_{2n},m_{1n-1}),\\ &\quad\min(m'_{2n},m_{1n-1}),\ \max(m'_{2n-1},m_{2n}),\ \ldots,\ m_{nn}=0\end{aligned} \tag{10}$$

arranged in a lexical order and using specified phase relations for the transpositions of the first Regge symmetry type.

Another symmetry property follows from the contragredience relations.[8] This procedure gives $(-1)^q$ as a phase factor, and the set of parameters (10) turns into the set obtained from this one by changing the signs and writing in inverted order, all the parameters becoming positive after adding m'_{1n}.

In this way one obtains 2^{2n-1} symmetry properties for the quantity (9). It stands to reason that not all the Regge symmetry properties[11] of quantities of SU_2 can be generalized to SU_n with $n>2$.

We observe that the symmetry property of Ref. 3 allowing one to interchange the rows in the skew scheme belongs to the substitution group symmetry[12] rather than to one of the Regge type. Equation (9) is invariant with respect to this group which is equivalent to partial hook permutations (c.f. Ref. 1).

It is to be noted that the relation between i.f.'s which couple the bases of two symmetric representations (of equal or different contragrediency) and SU_2 CG coefficients[13] follows immediately from the Regge and substitution symmetry properties.

Expression (9) does not allow one to carry out the summations even for particular cases. Thus, it is worthwhile to use other methods to obtain different expressions for the same i.f. We can obtain one such expression by the use of the operator

$$\begin{aligned}&\left(\frac{p!}{(p-q)!q!}\right)^{1/2}E_{nn+1}{}^{p-q}E_{n-1n+1}{}^q\\ &=\left(\frac{p!}{(p-q)!q!}\right)^{1/2}\sum_y\frac{(-1)^{y-\alpha}}{(y-\alpha)!(q-y+\alpha)!}\\ &\times E_{nn+1}{}^yE_{n-1n}{}^qE_{nn+1}{}^{p-y}\end{aligned} \tag{11}$$

instead of (8). After dividing its matrix element by the reduced matrix element of the operator $E_{nn+1}{}^p$ and the i.f. of U_{n-1}, one obtains

$$\begin{aligned}&\begin{bmatrix}[m]_n & p & [m']_n\\ [m]_{n-1} & q & [m']_{n-1}\end{bmatrix}\\ &= [(p-q)!]^{-1/2}[\prod_{1\leqslant i<j\leqslant n}(m'_{in}-m'_{jn}-i+j)\\ &\times\prod_{1\leqslant i<j\leqslant n-1}(m_{in-1}-m_{jn-1}-i+j)]^{1/2}\\ &\times\frac{D([m]_{n-1},[m']_{n-1})\Gamma([m']_n,[m']_{n-1})}{D([m]_n,[m']_n)\Gamma([m]_n,[m]_{n-1})}\\ &\times R\begin{pmatrix}[m]_n & [m']_n\\ [m]_{n-1} & [m']_{n-1}\end{pmatrix},\end{aligned} \tag{12}$$

where

$$R\begin{pmatrix}[m]_n & [m']_n\\ [m]_{n-1} & [m']_{n-1}\end{pmatrix} = \sum_{[r]_n} \frac{(-1)^{y-\alpha} y!\,(p-y)!\,\Gamma^2([r]_n,[m]_{n-1})}{(y-\alpha)!\,(q-y+\alpha)!\,\Gamma^2([r],[m']_{n-1})} \times \prod_{1\leqslant i<j\leqslant n}(r_{in}-r_{jn}-i+j)D^2([r]_n,[m']_n) \times D^2([m]_n,[r]_n), \quad y=\sum_{i=1}^{n}(m'_{in}-r_{in}). \tag{13}$$

The number of the summation parameters in (13) is n. The terms of this sum depend on $\alpha(0\leqslant\alpha\leqslant p-q)$. However, the final result must be independent of this parameter. It turns out that in expression (13) the summation with respect to one of parameters r_{in} can be carried out by the use of the summation formula

$$\sum_x \frac{(-1)^x \prod_{i=1}^{a}(x+A_i)\prod_{i=1}^{b}(B_i-x)}{x!\,(a+b+c-x)!} = (-1)^a\,\delta(c,0), \tag{14}$$

a,b,c being nonnegative integers. Equation (14) can be proved by induction starting from Eq. (14.3) of Ref. 10.

In order to use this summation formula for the purpose indicated in Eq. (13), we transform the factorials depending on r_{in} (i fixed) into the quasipowers (c.f. Ref. 10), all these being brought into the numerator. The factors left in the denominator are

$$(m'_{1n}-r_{in}+i-1)!\,(r_{in}-m_{nn}-i+n)!.$$

It is evident that the sum in (13) in this new form has a much wider summation region, because it involves $n-1$ new regions. However, this procedure does not change the value of the sum (13), because nonvanishing terms in these new regions are compensated by a set of terms equal in absolute value and opposite in sign to the first ones. These terms can be found by renumerating the summation parameters $r_{jn}-j \longleftrightarrow r_{in}-i$, $j\neq i$ labeling the newly appearing regions.

The above mentioned summation with respect to r_{in} leads us to the expression

$$R_i\begin{pmatrix}[m]_n & [m']_n\\ [m]_{n-1} & [m']_{n-1}\end{pmatrix} = \sum_{r_{jn},j\neq i}(-1)^{\varphi_i} \times \prod_{\substack{1\leqslant k<l\leqslant n\\ k\neq i,\,l\neq i}}(r_{kn}-r_{ln}-k+l) \times D^2_{i,0}([r]_n,[m']_n)D^2_{0,i}([m]_n,[r]_n)\frac{\Gamma^2_{i,0}([r]_n,[m]_{n-1})}{\Gamma^2_{i,0}([r]_n,[m']_{n-1})}$$

$$\varphi_i=\sum_{j=1}^{i-1}(m'_{jn-1}-m_{jn-1}+m_{jn})+\sum_{j=i+n}^{n}m'_{jn}-\sum_{\substack{j=1\\ j\neq i}}^{n}r_{jn}. \tag{15}$$

The quantities $D_{i,0}$, $D_{0,i}$, and $\Gamma_{i,0}$ are obtained from those of Eqs. (2) and (3) by removing those factors involving parameters with subscripts i, out of $[r]_n$.

Since all R_i ($i=1,2,\ldots,n$) in (15) are equivalent, they are connected by the elements of the substitution group of Ref. 12. R_1 and R_n are more convenient for some problems. For example, in the semistretched case ($m'_{nn}=m_{nn}$), it is useful to take R_1. On the other hand, in the case of the maximal state ($m'_{in-1}=m'_{in}$, $i=1,2,\ldots n-1$), it is more convenient to use R_n.

In the case of SU_2, our expression turns into the third of Eqs. (13.1) of Ref. 10. On the other hand, in the case of the semistretched coupling of representations of SU_3 (Ref. 4) it leads us to the doubly stretched $9j$ coefficient of SU_2 given by Eq. (25.17) of Ref. 10.

It is worth noting that our expressions (13) and (15) have n regions for n summation parameters in the case of Eq. (13) and $n-1$ summation parameters in the case of Eq. (15). In the second case, one of the summation regions is free. This occurs because Eq. (15) does not possess the full Regge symmetry exhibited by Eq. (9), which has $n-1$ summation parameters as well as regions.

APPENDIX A: AN ALTERNATIVE APPROACH TO THE PROBLEM

Let $[\lambda^1],[\lambda^2],[\lambda^3],[\lambda^4]$ be the Young schemes such that $\lambda_i^1\leqslant\lambda_i^2\leqslant\lambda_i^3\leqslant\lambda_i^4$ ($i=1,2,\ldots,n$ labeling the rows, λ_i^k being the lengths of the corresponding rows). We define the quantity

$$U\begin{bmatrix}[\lambda^3] & [\lambda^4]\\ [\lambda^1] & [\lambda^2]\end{bmatrix} = \prod_{i,j=1}^{n}\prod_{k_i=\lambda_i^3+1}^{\lambda_i^4}\prod_{l_j=\lambda_j^1+1}^{\lambda_j^2}\left(1+\frac{1}{k_i-i-l_j+j}\right) = \prod_{1=i\leq j}^{n}\frac{(\lambda_i^4-\lambda_j^1-i+j)!\,(\lambda_i^3-\lambda_j^2-i+j)!}{(\lambda_i^4-\lambda_j^2-i+j)!\,(\lambda_i^3-\lambda_j^1-i+j)!} \times \prod_{1=i<j}^{n}\frac{(\lambda_i^2-\lambda_j^4-i+j-1)!\,(\lambda_i^1-\lambda_j^3-i+j-1)!}{(\lambda_i^1-\lambda_j^4-i+j-1)!\,(\lambda_i^2-\lambda_j^3-i+j-1)!}. \tag{A1}$$

Let, further, $[\lambda^2]$ and $[\lambda^3]$ be the Young schemes with $\lambda_i^2=\min(m_{in},m'_{in-1})$, $\lambda_i^3=\max(m_{in},m'_{in-1})$. In the notations of this paper, the result of Ref. 3 allows us to write

$$\begin{bmatrix}[m]_n & p & [m']_n\\ [m]_{n-1} & q & [m']_{n-1}\end{bmatrix} = \left(\frac{A[m]_n\,A[m']_{n-1}(p-q)!}{A[m']_nA[m]_{n-1}}\right)^{1/2}\left(U\begin{bmatrix}[\lambda^3] & [m']_n\\ [m]_{n-1} & [\lambda^2]\end{bmatrix}\right)^{-1} \times \left(U\begin{bmatrix}[\lambda^2] & [m']_n\\ [m]_{n-1} & [\lambda^2]\end{bmatrix}U\begin{bmatrix}[\lambda^3] & [m']_n\\ [m]_{n-1} & [\lambda^3]\end{bmatrix}\right)^{1/2}F_{p_1}, \tag{A2}$$

where

$$A[\lambda]=(\textstyle\sum\lambda_i)!/f_{[\lambda]},$$

$f_{[\lambda]}$ being the dimension of the representation $[\lambda]$ of the symmetric group on $\sum_i\lambda_i$ symbols. F_{p_1} is the sum of the coefficients of those permutations in the expression

$$\prod_{k=p_1+1}^{p_1+p_2}\uparrow\prod_{l=1}^{p_1}\left(\mathcal{E}+\frac{(kl)}{x_k-y_l}\right), \tag{A3}$$

which the symbols $k\leqslant p_1$ substitutes by the symbols $l>p_1$. $\mathcal{E}$ in (A3) is the unit element of the symmetric group and (lk), the transposition of symbols l and k. $\uparrow$ indicates that the order of multipliers with respect to label l is the same for each k. F_{p_1} is the symmetric function on two sets of variables x_k and y_l. For calculating the i.f. according to (A2), we must substitute the values of the function F_{p_1} with the x equal to

$k_i - i\,(i = 1, 2, \ldots, n; k_i = \lambda_i^3 + 1, \lambda_i^3 + 2, \ldots, m'_{in}$; $p_2 = \sum_{i=1}^{n}(m'_{in} - \lambda_i^3)$ and the y to the $l_j - j\,(j = 1, 2, \ldots, n-1; l_j = m_{jn-1} + 1, m_{jn-1} + 2, \ldots, \lambda_j^2; p_1 = \sum_{j=1}^{n-1}(\lambda_j^2 - m_{jn-1}))$.

Extending the definition of F_{p_1}, we can define the set of bisymmetric functions $F_s\,(s = 0, 1, \ldots, p_1), F_s$ being the sum of the coefficients of those permutations in (A3) in which s arbitrary symbols from the set $1, 2, \ldots, p_1$ are substituted by the symbols from another set $p_1 + 1, p_1 + 2, \ldots, p_1 + p_2$. It can be shown, that the following set of equations hold for the F_s:

$$\sum_s \frac{(p_1 - s)!\,(p_2 - s)!}{(p_1 - v_1)!\,(v_1 - s)!\,(p_2 - v_2)!\,(v_2 - s)!}\, F_s = V_{v_1 v_2}^{p_1 p_2}, \tag{A4}$$

where

$$V_{v_1 v_2}^{p_1 p_2} = \sum_{(1)} \sum_{(2)} \prod_{i=p_1+v_2+1}^{p_1+p_2} \prod_{k=p_1+1}^{p_1+v_2} \prod_{l=1}^{v_1} \prod_{j=v_1+1}^{p_1} \times \left(1 + \frac{1}{x_i - x_k}\right)\left(1 + \frac{1}{x_k - y_l}\right)\left(1 + \frac{1}{y_l - y_j}\right). \tag{A5}$$

v_1 and v_2 can take on arbitrary values from the intervals $0 \leq v_1 \leq p_1$ and $0 \leq v_2 \leq p_2$, respectively. The first summation in (A5) is taken with respect to permutations, one from each left coset of the group of permutations of indices $1, 2, \ldots, p_1$ with respect to the subgroup of permutations of indices $1, \ldots, v_1$ and $v_1 + 1, \ldots, p_1$, within the two sets. The second summation is analogous to the first one, the group being the permutation group of the symbols $p_1 + 1, \ldots, p_1 + p_2$, and the subgroup having as its elements the permutations within the two sets $p_1 + 1, \ldots, p_1 + v_2$ and $p_1 + v_2 + 1, \ldots, p_1 + p_2$.

Taking the different sets of $(p_1 + 1)$ equations from the extended set (A4), we obtain different expressions for F_{p_1}. Thus, if we take the equations with $v_2 = p_2$ and v_1 varying from 0 to p_1, we have

$$F_{p_1} = \sum_{v_1=0}^{p_1} (-1)^{p_1 - v_1} V_{v_1 p_2}^{p_1 p_2}. \tag{A6}$$

The value of the bisymmetric function $V_{v_1 p_2}^{p_1 p_2}$, the arguments taking the mentioned values, is equal to

$$\sum_{[r]} U\begin{bmatrix}[\lambda^3] & [m']_n \\ [r] & [\lambda^2]\end{bmatrix} U\begin{bmatrix}[r] & [\lambda^2] \\ [m]_{n-1} & [r]\end{bmatrix}, \quad \sum_i (\lambda_i^2 - r_i) = v_1. \tag{A7}$$

Using (A6)–(A7) for the F_{p_1}, Young's expression for the dimensions $f_{[\lambda]}$, and performing the simplifications needed, we obtain formula (9) for the i.f. under consideration.

On the other hand, solving equations (A4) with $v_1 = p_1$ and $v_2 = p_2 - p_1 - \alpha, p_2 - p_1 - \alpha + 1, \ldots, p_2 - \alpha$ using the values of x_i and y_k indicated above, one obtains

$$F_{p_2} = \sum_{[r]_n} \frac{(-1)^{y-\alpha}\, y!\,(p_2 - y)!}{(p - q)!\,(y - \alpha)!\,(p_1 + \alpha - y)!} \times U\begin{bmatrix}[\lambda^3] & [r]_n \\ [m]_{n-1} & [\lambda^2]\end{bmatrix} U\begin{bmatrix}[r]_n & [m']_n \\ [\lambda^3] & [r]_n\end{bmatrix}. \tag{A8}$$

Formulas (A2) and (A8) may be brought into the form equivalent to the result given by the Eq. (12) and (13).

APPENDIX B: RELATION BETWEEN RECOUPLING MATRICES AND ISOSCALAR FACTORS

According to the results of Ref. 5, the following relation holds between the elements of the recoupling matrix of four representations of U_n with three of them symmetric and the i.f.:

$$\begin{aligned}&\langle [m]_{n-1}\, q([m']_{n-1}), rp - q(r'); [m']_n \mid \\ &\quad \times \mid [m]_{n-1}\, r([m]_n), qp - q(p); [m']_n\rangle \\ &= \left(\frac{r!\,q!\,(p-q)!\,A[m]_{n-1}A[m']_n}{r'!\,p!\,A[m]_n A[m]_{n-1}}\right)^{1/2} \\ &\quad \times \begin{bmatrix}[m]_n & p & [m']_n \\ [m]_{n-1} & q & [m']_{n-1}\end{bmatrix}, \quad r = \sum_{i=1}^{n} m_{in} - \sum_{i=1}^{n-1} m_{in-1}, \\ &\quad r' = p - q + r = \sum_{i=1}^{n} m'_{in} - \sum_{i=1}^{n-1} m'_{in-1}\,,\end{aligned} \tag{B1}$$

$A[\lambda]$ is given in (A2).

The particular cases of this relation (when $p = q$ for U_n and for the semistretched i.f. of SU_3) have been obtained in Refs. 3 and 4. It can be seen that in the semistretched case of the i.f. ($m'_{nn} = m_{nn}$), the recoupling matrix goes over into the one of U_{n-1}. A particular case of this matrix (with $p = q$ and $m'_{nn} = m_{nn}$) gives the matrix changing the canonical chains of subgroups in U_n.[3, 14] Equations (12) and (15) for the i.f. on the right-hand side of (B1) are more convenient to use than Eq. (9), because in the first case there remain only $n - 2$ summation parameters, instead of $n - 1$ as is in the second case.

1 J. D. Louck, Am. J. Phys. **38**, 3 (1970).
2 L. C. Biedenharn and J. D. Louck, Commun. Math. Phys. **8**, 89 (1968).
3 A.-A. A. Jucys, Lietuvos. Mat. Rinkinys **8**, 597 (1968); Lietuvos Fiz. Rinkinys **9**, 629 (1969); **10**, 5 (1970).
4 S. J. Ališauskas, Lietuvos Fiz. Rinkinys **9**, 641 (1969).
5 S. J. Ališauskas, V. V. Vanagas, and A. P. Jucys, Dokl. Akad. Nauk SSSR **197**, 804 (1971).
6 R. M. Asherova and Yu. F. Smirnov, Usp. Mat. Nauk **24**, No. 3, 227 (1969).
7 R. M. Asherova and Yu. F. Smirnov, Nucl. Phys. **B4**, 399 (1968).
8 G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).
9 J. G. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965).
10 A. P. Jucys and A. A. Bandzaitis, *The Theory of Angular Momentum in Quantum Mechanics* (in Russian) (Mintis, Vilnius, 1965).
11 T. Regge, Nuovo Cimento **10**, 544 (1958); **11**, 116 (1959); A. M. Bincer, J. Math. Phys. **11**, 1835 (1970).
12 S. J. Ališauskas and A. P. Jucys, J. Math. Phys. **8**, 2250 (1967).
13 A. DiLeva and G. Ponzano, Nuovo Cimento **A51**, 1107 (1967); S. J. Ališauskas and P. A. Jucys, Dokl. Akad. Nauk SSSR **177**, 61 (1967).
14 M. Moshinsky and E. Chacon, *Spectroscopic and Group Theoretical Methods in Physics* (Racah Memorial Volume) (North-Holland, Amsterdam, 1968), p. 99; J. Flores and J. Niederle, Czech. J. Phys. **B20**, 1241 (1970).

Deformable Magnetically Saturated Media. II. Constitutive Theory*

Gérard A. Maugin and A. Cemal Eringen
Princeton University, Princeton, New Jersey
(Received 23 July 1971; Revised Manuscript Received 30 August 1971)

This article is devoted to the development of constitutive equations of deformable magnetically saturated media in three dimensions. In Sec. 1 we recapitulate the local balance laws and jump conditions derived previously. A thorough study of the consequences of the objectivity requirement is given in Sec. 2. In the following sections, the material symmetry restrictions are examined and exact and approximate constitutive equations are obtained for a variety of material classes.

1. RECAPITULATION

In this section we set forth in one place the complete system of differential equations [valid in the body (B)], boundary conditions [on the surface (∂B)], jump conditions [across a *steady* surface of discontinuity (Γ)] and constitutive relations (the nonlinear theory) of the theory of magnetomechanical materials developed in Part I.[1] We limit the presentation to the *nondissipative* case and note that the "spin rotation" equation derived in Part I[1] is to be used *only* when the material is magnetically saturated. With the recognition of certain quantities and use of the objectivity requirement other forms obtained before will result in the form given below.

(a) The set of field equations for *saturated* media consists of:

(i) Continuity equation:

$$\frac{\partial \rho}{\partial t} + (\rho v^k)_{;k} = 0 \text{ in } (B-\Gamma), \quad [\rho v^k]\, n_k = 0 \text{ on } (\Gamma). \tag{1.1}$$

(ii) Balance of linear momentum:

$$\rho \dot{v}^k = \rho f^k + \rho B^{l;k}\mu_l + t^{kl}{}_{;l} \text{ in } (B-\Gamma),$$
$$[t^{kl} + t^{kl}_{(\mathrm{em})} - \rho v^k v^l] n_l = 0 \text{ on } (\Gamma). \tag{1.2}$$

(iii) Balance of moment of momentum:

$$t^{[kl]} = \rho_L B^{[k}\mu^{l]} \quad \text{in } (B-\Gamma). \tag{1.3}$$

(iv) Balance of spin angular momentum:

$$(\rho/\Gamma)\dot{\mu}^k = \epsilon^{klm} B_{(\mathrm{eff})l} M_m \quad \text{in } (B-\Gamma),$$
$$\mu^{[m} t^{(\mu)k]l} n_l = 0 \quad \text{on } (\Gamma). \tag{1.4}$$

(v) Maxwell's equations: In the case of quasi-magnetostatics, the set of Maxwell's equations with $|\boldsymbol{\mu}| = \text{const.}$ in (B) and on (∂B), reduces to:

$$\nabla \times \mathbf{B} = \nabla \times \mathbf{M}, \quad \nabla \cdot \mathbf{B} = 0 \quad \text{in } (B-\Gamma),$$
$$\nabla \times \mathbf{B} = 0, \quad \nabla \cdot \mathbf{B} = 0 \quad \text{outside (B)}, \tag{1.5}$$
$$\mathbf{n} \times [\mathbf{B}] = \mathbf{n} \times [\mathbf{M}], \quad \mathbf{n} \cdot [\mathbf{B}] = 0 \quad \text{on } (\Gamma),$$

where we have defined

$$\mathbf{M} = \rho\boldsymbol{\mu}, \quad B_{(\mathrm{eff.})k} = B_k + {}_L B_k + \rho^{-1} t^{(\mu)l}_{kl;}{}^{l},$$
$$t^{kl}_{(\mathrm{em})} = H^k B^l - (\tfrac{1}{2}\mathbf{B}^2 - \mathbf{B}\cdot\mathbf{M}) g^{kl}. \tag{1.6}$$

Alternatively, (1.4) may be written in the "couple stress" form:

$$(\rho/2\Gamma)\, \dot{S}^{kl} = \rho L^{kl} + t^{[lk]} + M^{klm}{}_{;m} \quad \text{in } (B-\Gamma),$$
$$[M^{klm}] n_m = 0 \quad \text{on } (\Gamma), \tag{1.4'}$$

with

$$S^{kl} \equiv \epsilon^{klm}\mu_m, \quad L^{kl} \equiv B^{[k}\mu^{l]}, \quad M^{klm} \equiv -\mu^{[k} t^{(\mu)l]m}, \tag{1.7}$$

and the definitions

$\rho \equiv$ mass density,

$B^k \equiv$ magnetic field intensity,

$M^k \equiv$ magnetization per unit volume,

$v^k \equiv$ velocity field,

$H^k \equiv$ magnetic field,

$\mu^k \equiv$ magnetization per unit mass

$f^k \equiv$ nonmagnetic body force per unit mass (e.g., gravity),

$t^{kl} \equiv$ stress tensor,

${}_L B^k \equiv$ local magnetic field intensity,

$B^k_{(\mathrm{eff.})} \equiv$ effective magnetic field intensity,

$L^{kl} \equiv$ ponderomotive magnetic couple,

$\Gamma \equiv$ gyromagnetic ratio $= -e/m_0 c$.

$t^{kl}_{(\mathrm{em})} \equiv$ Maxwell's stress tensor,

$t^{(\mu)kl} \equiv$ "spin interaction" stress tensor

$S^{kl} \equiv$ spin bivector

$M^{klm} \equiv$ "magnetic" couple stress tensor,

With a strain energy function of the form

$$\mathfrak{F} = \mathfrak{F}(\mathbf{F}_x, \boldsymbol{\mu}, \mathbf{F}_{(\mu)}, \mathbf{X}), \quad (\mathbf{F}_x)^k_K \equiv x^k{}_{,K},$$
$$(\mathbf{F}_{(\mu)})^k{}_K \equiv \mu^k{}_{,K} \tag{1.8}$$

where x^k and X^K are the spatial and material coordinates, the constitutive equations are

$$t_k{}^l = \rho \frac{\partial \mathfrak{F}}{\partial x^k{}_{,K}} x^l{}_{,K}, \quad {}_L B_k = -\frac{\partial \mathfrak{F}}{\partial \mu^k},$$
$$t^{(\mu)}{}_k{}^l = \rho \frac{\partial \mathfrak{F}}{\partial \mu^k{}_{,K}} x^l{}_{,K} \quad \text{or} \quad M^{pqr} = \rho \frac{\partial \mathfrak{F}}{\partial \mu_{[p,K}} \mu^{q]} x^r{}_{,K}. \tag{1.9}$$

The field equations (1.1)–(1.5) are supplemented with the mixed boundary conditions:

$$[\rho v^k] n_k = 0 \quad \text{across } (\partial B - \Gamma),$$
$$t^k = t^{kl} n_l \text{ on } (\partial B_t - \Gamma),$$
$$u^k = u^k{}_{(\partial B)} \quad \text{on } (\partial B - \partial B_t - \Gamma),$$
$$\epsilon_{ikm} t^{(\mu)kl} \mu^m n_l = 0 \quad \text{on } (\partial B - \Gamma), \tag{1.10}$$
$$\mathbf{n} \times [\mathbf{B}] = \mathbf{M}_{(in)} \times \mathbf{n}, \quad \mathbf{n}\cdot[\mathbf{B}] = 0 \quad \text{on } (\partial B - \Gamma),$$

with, [cf., Eq. (6.1) of Part I]

$$t^k = \tau^k - (\mathbf{M}_{(\mathrm{in})} \cdot [\mathbf{B}])\, \tfrac{1}{2} n^k,$$

where t^k is the stress vector, τ^k is the mechanical surface traction prescribed on $(\partial B_t - \Gamma)$ with $\mathbf{M}_{(\mathrm{in})}$ the inside value of $\mathbf{M}$ on $(\partial B - \Gamma)$; u^k is the displacement.

The number of unknowns $(\rho, v^k, t^{kl}, \mu^k, B^k, {}_L B^k, t^{(\mu)kl})$ amounts to $1 + 3 + 9 + 3 + 3 + 3 + 9 = 31$, and the number of components of Eqs. (1.1) to (1.5) and (1.9) is precisely: $1 + 3 + 3 + 3 + 3 + 6 + 3 + 9 = 31$.

With the following Cauchy's data at $t = t_0$

$$\rho(t = t_0) = \rho_R, \quad x^k(t = t_0) = \delta^k{}_K X^K,$$
$$v\ (\mathbf{x}, t = t_0) = v_o^k(\mathbf{x}), \quad \mu^k(t = t_0) = \delta^k{}_K \mu^K, \qquad (1.11)$$

problems relevant to the *nonlinear theory of magnetically saturated elastic media* can be solved, the magnetic part being statically treated. Ultimately, theorems of existence and uniqueness for the system of partial differential equations given above have to be proved.

(b) For *nonsaturated media*, an equation such as (1.4) cannot be written since this equation describes the rotation of a vector constant in magnitude. Equations (1.1), (1.2), and (1.5) are still valid. With the form of $\mathfrak{F}$ given by (1.8), the balance of moment of momentum cannot be reduced to (1.3), but reads

$$t^{[kl]} = {}_L B^{[k} M^{l]} - t^{(\mu)[k}{}_m \mu^{l];m}. \qquad (1.12)$$

This follows from

$$\rho \frac{\partial \mathfrak{F}}{\partial x_{[l,K}} x^{k]}{}_{,K} + \rho \frac{\partial \mathfrak{F}}{\partial \mu_{[l}} \mu^{k]} + \rho \frac{\partial \mathfrak{F}}{\partial \mu_{[l,K}} \mu^{k]}{}_{,K} = 0 \qquad (1.13)$$

which is the Euclidean invariance requirement equation of Sec. 7 of Part I. Alternatively, (1.13) is the partial differential equation to be verified by $\mathfrak{F}$ if $\mathfrak{F}$ is to be objective (invariant under time-dependent rotations in E^3). We can thus write the stress tensor t^{kl} in the form

$$t^{kl} = {}_E t^{kl} + t^{[kl]}, \qquad (1.14)$$

where ${}_E t^{kl}$ is the elastic stress tensor, the constitutive equation of which is given by

$${}_E t^{kl} = \rho \frac{\partial \mathfrak{F}}{\partial x_{(k,K}} x^{l)}{}_{,K}. \qquad (1.15)$$

In summary, for nonsaturated media, no equation describes the motion of μ which is solely determined by the solution of Maxwell's equations. The field equations are (1.1), (1.2), (1.5), and (1.12). The constitutive equations are given by (1.15) and the second and third of Eqs. (1.9). A constitutive equation must be given in the form $\mathbf{M} = \mathbf{M}(\mathbf{B})$ if one wants to solve (1.5). The boundary conditions and the Cauchy's data at $t = t_0$ are

$$\begin{aligned}
&[\rho v^k] n_k = 0 \quad \text{across } (\partial B - \Gamma)\\
&t^k = t^{kl} n_l \quad \text{on } (\partial B_t - \Gamma),\\
&u^k = u^k_{(\partial B)} \quad \text{on } (\partial B - \partial B_t - \Gamma),\\
&\mathbf{n} \times [\mathbf{B}] = \mathbf{M}_{(\text{in})} \times \mathbf{n},\\
&\mathbf{n} \cdot [\mathbf{B}] = 0 \quad \text{on } (\partial B - \Gamma),
\end{aligned} \qquad (1.16)$$

$$\rho(t = t_0) = \rho_R, \quad x^k(t = t_0) = \delta^k{}_K X^K,$$
$$v^k(\mathbf{x}, t = t_0) = v_0^k(\mathbf{x}), \qquad (1.17)$$

with

$$t^k = \tau^k - (\mathbf{M}_{(\text{in})} \cdot [\mathbf{B}]) \tfrac{1}{2} n^k.$$

This set of equations is adequate for describing magnetoelastic effects (such as magnetostriction) in nonsaturated bodies. We must however remark that in this case, theories such as those of Jordan and Eringen[2], Dixon and Eringen[3], and Grot and Eringen[4] are more suited for a constitutive theory since, then, the needed equation $\mathbf{M} = \mathbf{M}(\mathbf{B})$ would be included in the set of constitutive equations as well.

2. OBJECTIVITY REQUIREMENT[5]

In determining the restrictions arising from the axiom of objectivity, we distinguish between two cases: (i) no saturation of the magnetization, (ii) saturation of the magnetization, i.e., $|\boldsymbol{\mu}| = \text{const}$. For each case it must be understood that *only constitutive equations labelled for the saturation case* (or *nonsaturation*) must be used as companions of the set of corresponding field equations given in the preceding paragraph.

According to Part I, the Lagrangian density $\mathfrak{L}$ is expressed in terms of the strain energy function $\mathfrak{F}$ by

$$\mathfrak{L} = \mathfrak{K} - \rho_R \mathfrak{F}, \quad \mathfrak{K} = \tfrac{1}{2} \rho_R \dot{\mathbf{x}}^2, \quad \mathfrak{F} = \mathfrak{F}(\mathbf{F}_x, \boldsymbol{\mu}, \mathbf{F}_{(\mu)}, \mathbf{X}).$$

Making no hypothesis as to the magnetization magnitude per unit mass, we shall require $\mathfrak{F}$ to be invariant under the orthogonal group, i.e.,

$$\mathfrak{F}(\boldsymbol{\mu}, \mathbf{F}_x, \mathbf{F}_{(\mu)}, \mathbf{X}) = \mathfrak{F}(\mathfrak{Q}\boldsymbol{\mu}, \mathfrak{Q}\mathbf{F}_x, \mathfrak{Q}\mathbf{F}_{(\mu)}, \mathbf{X}) \qquad (2.1)$$

for all orthogonal constant tensors $\mathfrak{Q}$. In particular, we select

$$\mathfrak{Q} = \mathbf{R}^T, \qquad (2.2)$$

where $\mathbf{R}^T$ is the transposed of the rotation tensor. Thus, using the polar decomposition of $\mathbf{F}_x$,[6]

$$\mathbf{F}_x = \mathbf{R} \cdot \mathbf{U}, \qquad (2.3)$$

$$\mathfrak{F} = \mathfrak{F}(\mathbf{R}^T \boldsymbol{\mu}, \mathbf{R}^T \mathbf{F}_x, \mathbf{R}^T \mathbf{F}_{(\mu)}, \mathbf{X}). \qquad (2.4)$$

Since

$$\mathbf{R}^T \mathbf{F}_x = \mathbf{U} = \mathbf{C}^{1/2}, \qquad (2.5)$$

we can write

$$\mathfrak{F} = \mathfrak{F}(\boldsymbol{\Pi}, \mathbf{C}, \mathbf{D}, \mathbf{X}), \qquad (2.6)$$

where

$$\begin{aligned}
&\mathbf{C} = \mathbf{F}_x^T \mathbf{F}_x, \quad C_{AB} = g_{ik}\, x^i{}_{,A} x^k{}_{,B},\\
&\boldsymbol{\Pi} = \mathbf{F}_x^T \boldsymbol{\mu}, \quad \Pi_A = x^i{}_{,A} \mu_i,\\
&\mathbf{D} = \mathbf{F}_x^T \mathbf{F}_{(\mu)}, \quad D_{AB} = g_{ik}\, x^i{}_{,A} \mu^k{}_{,B}
\end{aligned} \qquad (2.7)$$

form a set of 18 independent quantities forming a *single-valued minimal integrity basis* for the arguments involved in $\mathfrak{F}$. The same result may be arrived at in applying Cauchy's theorem (see Weyl[7]), according to which $\mathfrak{F}$ is objective if it is a function of the following 57 quantities:

$$\begin{aligned}
&\mu^i \mu_i = \mu^2, && J = (1/3!)\, \epsilon_{ijk}\, \epsilon^{KLM} x^i{}_{,K} x^j{}_{,L} x^k{}_{,M},\\
&x^i{}_{,A} x_{i,B} = C_{AB}, && L = (1/3!)\, \epsilon_{ijk}\, \epsilon^{KLM} \mu^i{}_{,K} \mu^j{}_{,L} \mu^k{}_{,M},\\
&\mu^i{}_{,A} \mu_{i,B} = G_{AB}, && P_{RN} = \tfrac{1}{2} \epsilon_{ijk}\, \epsilon^{KL}{}_R\, x^i{}_{,K} x^j{}_{,L} \mu^k{}_{,N},\\
&x^i{}_{,A} \mu_i = \Pi_A, && Q_{RN} = \tfrac{1}{2} \epsilon_{ijk}\, \epsilon^{LM}{}_R\, x^i{}_{,L} \mu^j{}_{,M} \mu^k{}_{,N},\\
&x^i{}_{,A} \mu_{i,B} = D_{AB}, && W_{LM} = \epsilon_{ijk}\, x^i{}_{,L} x^j{}_{,M} \mu^k = -W_{ML},\\
&\mu^i \mu_{i,A} = J_A, && K_N = \tfrac{1}{2} \epsilon_{ijk}\, \epsilon^{LM}{}_N\, x^i{}_{,L} x^j{}_{,M} \mu^k,\\
& && R_S = \tfrac{1}{2} \epsilon_{ij}\ \epsilon^{LM}{}_S \mu^i{}_{,L} \mu^j{}_{,M} \mu^k.
\end{aligned} \qquad (2.8)$$

But one can show that only 18 (C_{AB}, D_{AB} and Π_A) of the 57 quantities are independent. Thus

$$\mathcal{L} = \rho_R[\tfrac{1}{2}\dot{\mathbf{x}}^2 - \mathcal{F}(C_{AB}, \Pi_A, D_{AB}, \mathbf{X})]. \tag{2.9}$$

It follows that the constitutive equations (1.9) for an anisotropic inhomogeneous elastic material with electronic spin and unsaturated magnetization are:

$$t^{l} = \rho\left(2\frac{\partial \mathcal{F}}{\partial C_{KB}} x_{k,B} + \frac{\partial \mathcal{F}}{\partial \Pi_K}\mu_k + \frac{\partial \mathcal{F}}{\partial D_{KB}}\mu_{,B}\right)x^l{}_{,K}, \tag{2.10}$$

$${}_L B_k = -\frac{\partial \mathcal{F}}{\partial \Pi_A} x_{k,A}, \tag{2.11}$$

$$t^{(\mu)kl} = \rho\frac{\partial \mathcal{F}}{\partial D_{AK}} x^k{}_{,A} x^l{}_{,K}, \qquad M^{pqr} = -\rho\frac{\partial \mathcal{F}}{\partial D_{AK}}\mu^{[p}x^{q]}{}_{,A}\, x^r{}_{,K}. \tag{2.12}$$

Magnetically saturated media: When the material is magnetically saturated, we must impose the constraints

$$\mu^i\mu_i = \mu_S^2 = \text{const}, \qquad \mu^k\mu_{k,K} = 0. \tag{2.13}$$

The minimal integrity basis is now reduced to fourteen members since C_{AB}, Π_A, and D_{AB} are no longer independent and must be consistent with (2.13). In fact, using (2.13), we have

$$\mu_S^2 = \Pi_A \overset{-1}{C}{}^{AB}\Pi_B, \qquad J_M = \Pi_A \overset{-1}{C}{}^{AB}D_{BM} = 0, \tag{2.14}$$

where $\overset{-1}{C}{}^{AB}$ is the reciprocal of C_{AB} introduced by

$$C_{AB}\overset{-1}{C}{}^{BM} = \delta_A^M.$$

We must take account of (2.4) when we differentiate $\mathcal{F}$ with respect to the different arguments. In fact, depending on the choice of fourteen independent variables among the list (2.8), there exist many objective forms for the strain energy $\mathcal{F}$. Of course all these forms are equivalent to each other. For the approximate theories however, a form may be preferred over the others depending on the class of problems under consideration. Below, we give four interesting forms:

Theorem: For magnetically saturated media, the strain energy is objective if it has any one of the following forms:

$$\mathcal{F} = \mathcal{F}(C_{AB}, \Pi_A, D_{AB}, \mathbf{X}) \quad \text{subject to} \quad \Pi_A\overset{-1}{C}{}^{AB}\Pi_B = \text{const and } \Pi_A\overset{-1}{C}{}^{AB}D_{BM} = 0, \tag{2.15}$$

or

$$\mathcal{F} = \hat{\mathcal{F}}(E_{AB}, \Pi_A, G_{AB}, \mathbf{X}) \quad \text{subject to} \quad \Pi_A\overset{-1}{C}{}^{AB}\Pi_B = \text{const}, \tag{2.16}$$

or

$$\mathcal{F} = \bar{\mathcal{F}}(C_{AB}, \hat{\Pi}^A, \hat{D}^{AB}, \mathbf{X}) \quad \text{subject to} \quad \hat{\Pi}_A\hat{\Pi}^A = \mu_S^2 = \text{const and } \hat{\Pi}_A\hat{D}^{AB} = 0, \tag{2.17}$$

or

$$\mathcal{F} = \check{\mathcal{F}}(C_{AB}, \hat{D}^{AB}, \mathbf{X}) \quad \text{subject to} \quad \det(\hat{D}^A{}_B) = 0, \tag{2.18}$$

where

$$E_{AB} \equiv \tfrac{1}{2}(C_{AB} - \delta_{AB}), \qquad G_{AB} \equiv \mu^i{}_{,A}\mu_{i,B} = D_{LA}\overset{-1}{C}{}^{LM}D_{MB}, \tag{2.19}$$

$$\hat{\Pi}^A \equiv \overset{-\frac{1}{2}}{C}{}^{AB}\Pi_B, \qquad \hat{D}^{AB} \equiv \overset{-\frac{1}{2}}{C}{}^{AM}D_M{}^B. \tag{2.20}$$

Here E_{AB} is the classical Lagrangian strain tensor.

The equivalence of the first three forms is clear, the last one, eq. (2.18), can be seen as follows:

Since $\hat{\mathbf{\Pi}}$ is a left null vector for $\hat{\mathbf{D}}$ (alternatively $\hat{\mathbf{D}}$ a right zero operator for $\hat{\mathbf{\Pi}}$) the determinant of $\hat{\mathbf{D}}$ must vanish. Thus $\hat{\mathbf{\Pi}}$ will no longer appear explicitly in $\mathcal{F}$ since it is determined through $\hat{\mathbf{D}}$. One can verify that the number of independent components for the arguments of $\mathcal{F}$ amounts to $6 + 9 - 1 = 14$ [-1 due to restriction (2.18)]. We take notice of the inconvenience that the quantity $\partial\hat{\mathbf{D}}/\partial\hat{\Pi}$ appearing in the constitutive equations must be computed from $\det\hat{\mathbf{D}} = 0$.

3. CONSTITUTIVE EQUATIONS FOR ANISOTROPIC MAGNETICALLY SATURATED MEDIA

Different functional forms of the strain energy function $\mathcal{F}$ lead to constitutive equations involving different variables. Some of these expressions are simpler than others. Below, we give the constitutive equations for the four cases enumerated in Sec. 2. To this end, we consider the material tensors

$$T^{LK} = X^L{}_{,k}T^{kK} = X^L{}_{,k}\frac{\partial\Sigma}{\partial x_{k,K}}, \qquad {}_L B^K = \rho_R J X^K{}_{,k}B^k = -JX^K{}_{,k}\frac{\partial\Sigma}{\partial\mu_k}, \qquad T^{(\mu)LK} = X^L{}_{,k}T^{(\mu)kK} = X^L{}_{,k}\frac{\partial\Sigma}{\partial\mu_{k,K}}, \tag{3.1}$$

and we introduce the strain energy function per unit of undeformed volume by

$$\Sigma = \rho_R\mathcal{F}. \tag{3.2}$$

The mixed tensors T^{kK} and $T^{(\mu)kK}$ were encountered in Part I.

Case (i), Strain energy given by (2.15): We introduce the strain energy function

$$\Sigma^*_{(1)} = \Sigma_{(1)}(C_{AB}, \Pi_A, D^{AB}, \mathbf{X}) - \wp_{(1)}[\Pi_A\overset{-1}{C}{}^{AB}\Pi_B - \mu_S^2] - \wp_{(1)}{}^M\Pi_A\overset{-1}{C}{}^{AB}D_{BM}, \tag{3.3}$$

where $\wp_{(1)}$ and $\wp^M_{(1)}$ are Lagrange's multipliers.

The following relations are useful in differentiating $\Sigma^*_{(1)}$

$$\frac{\partial\overset{-1}{C}{}^{AJ}}{\partial C_{MN}} = -\overset{-1}{C}{}^{AM}\overset{-1}{C}{}^{NJ}, \qquad \frac{\partial C_{MN}}{\partial x_{k,K}} = 2\delta^K_M x^k{}_{,N}, \tag{3.4}$$

$$\frac{\partial\overset{-1}{C}{}^{AJ}}{\partial x^k{}_{,K}} = -2\overset{-1}{C}{}^{AK}X^J{}_{,k}, \tag{3.5}$$

$$\frac{\partial E_{AB}}{\partial x_{k,K}} = \delta^K_A x^k{}_{,B}, \qquad \frac{\partial\Pi_A}{\partial x_{k,K}} = \delta^K_A\mu^k, \qquad \frac{\partial\Pi_A}{\partial\mu_k} = x^k{}_{,A}, \tag{3.6}$$

$$\frac{\partial D_{AB}}{\partial x_{k,K}} = \delta_A^K \mu^k{}_{,B}, \quad \frac{\partial D_{AB}}{\partial \mu_{k,K}} = \delta_B^K x^k{}_{,A}, \quad \frac{\partial G_{AB}}{\partial \mu_{k,K}} = 2\delta_A^K \mu^k{}_{,B}, \tag{3.7}$$

which follow by differentiating appropriate expressions in (2.8). We also introduce the *material* form of the magnetization vector by

$$\mu^M \overset{\text{DEF}}{=} X^M{}_{,i}\mu^i, \text{ conversely, } \mu^i = x^i{}_{,M}\mu^M. \tag{3.8}$$

By using (2.8) it can be shown that

$$\mu^B = \overset{-1}{C}{}^{BA}\Pi_A, \quad \Pi_A = C_{AM}\mu^M, \tag{3.9}$$

$$\mu^k{}_{,B}X^L{}_{,k} = \overset{-1}{C}{}^{LA}D_{AB}. \tag{3.10}$$

Upon carrying (3.3) into (3.1) and using (3.4)–(3.10), we obtain the constitutive equations:

$$\begin{aligned} T^{LK} &= 2\frac{\partial \Sigma_{(1)}}{\partial C_{KL}} + \frac{\partial \Sigma_{(1)}}{\partial \Pi_K}\overset{-1}{C}{}^{LA}\Pi_A + \frac{\partial \Sigma_{(1)}}{\partial D_{KB}}\overset{-1}{C}{}^{LA}D_{AB} \\ &\quad - 2\mathcal{P}_{(1)}^M \overset{-1}{C}{}^{L[A}\overset{-1}{C}{}^{B]K}\Pi_A D_{BM}, \\ {}_L B^K &= -J\left(\frac{\partial \Sigma_{(1)}}{\partial \Pi_K} - 2\mathcal{P}_{(1)}\overset{-1}{C}{}^{KB}\Pi_B - \mathcal{P}_{(1)}^M\overset{-1}{C}{}^{KB}D_{BM}\right), \\ T^{(\mu)LK} &= \frac{\partial \Sigma_{(1)}}{\partial D_{KL}} - \mathcal{P}_{(1)}^K\overset{-1}{C}{}^{AL}\Pi_A. \end{aligned} \tag{3.11}$$

Note that $\mathcal{P}_{(1)}$ does not appear in the first of Eqs. (3.11). The corresponding spatial tensors or vectors, when no saturation occurs are given by (2.10)–(2.12).

Case (ii), Strain energy given by (2.16): In this case only one scalar Lagrange's multiplier $\mathcal{P}_{(2)}$ is necessary. Thus introducing

$$\Sigma_{(2)}^* = \Sigma_{(2)}(E_{AB}, \Pi_A, G_{AB}, \mathbf{X}) - \mathcal{P}_{(2)}\left[\Pi_A\overset{-1}{C}{}^{AB}\Pi_B - \mu_S^2\right], \tag{3.12}$$

we find the simple equations

$$\begin{aligned} T^{LK} &= \frac{\partial \Sigma_{(2)}}{\partial E_{KL}} + \frac{\partial \Sigma_{(2)}}{\partial \Pi_K}\overset{-1}{C}{}^{LA}\Pi_A, \\ {}_L B^K &= -J\left(\frac{\partial \Sigma_{(2)}}{\partial \Pi_K} - 2\mathcal{P}_{(2)}\overset{-1}{C}{}^{KB}\Pi_B\right), \\ T^{(\mu)LK} &= 2\frac{\partial \Sigma_{(2)}}{\partial G_{KB}}\overset{-1}{C}{}^{LJ}D_{JB}. \end{aligned} \tag{3.13}$$

Corresponding to these are the mixed material-spatial expressions

$$\begin{aligned} T^{kK} &= \frac{\partial \Sigma_{(2)}}{\partial E_{KL}}x^k{}_{,L} + \frac{\partial \Sigma_{(2)}}{\partial \Pi_K}\mu^k, \\ {}_L B^k &= -\frac{\partial \Sigma_{(2)}}{\partial \Pi_K}x^k{}_{,K} + 2\mathcal{P}_{(2)}\mu^k, \\ T^{(\mu)kK} &= 2\frac{\partial \Sigma_{(2)}}{\partial G_{KL}}\mu^k{}_{,L}, \end{aligned} \tag{3.14}$$

which shows that $\mathcal{P}_{(2)}$ is irrelevant for the spin equation. Nevertheless, the form (3.14) can only be employed for the case of saturation because of the use of the G_{KL}'s. This remark holds equally for the second of Eqs. (3.11), (3.13) and (3.15) given below. The spatial forms corresponding to (2.10)–(2.12) are

$$\begin{aligned} t^{km} &= \frac{\rho}{\rho_R}\left(\frac{\partial \Sigma_{(2)}}{\partial E_{KL}}x^k{}_{,L} + \frac{\partial \Sigma_{(2)}}{\partial \Pi_K}\mu^k\right)x^m{}_{,K}, \\ {}_L B^k &= -\frac{1}{\rho_R}\left(\frac{\partial \Sigma_{(2)}}{\partial \Pi_K}x^k{}_{,K} - 2\mathcal{P}_{(2)}\mu^k\right), \\ t^{(\mu)mp} &= 2\frac{\rho}{\rho_R}\frac{\partial \Sigma_{(2)}}{\partial G_{KB}}\mu^m{}_{,B}x^p{}_{,K}, \\ M^{pqr} &= -2\frac{\rho}{\rho_R}\frac{\partial \Sigma_{(2)}}{\partial G_{KB}}\mu^{[p}\mu^{q]}{}_{,B}x^r{}_{,K}. \end{aligned} \tag{3.15}$$

At this point it is of interest to remark the following result

$$\epsilon_{imk}t^{(\mu)mp}\mu^k{}_{;p} \equiv 0, \text{ or equivalently, } \frac{\partial \mathfrak{F}}{\partial \mu^{[m}{}_{,K}}\mu_{k],K} = 0, \tag{3.16}$$

which follows immediately from the third of Eqs. (3.15) and the symmetry of G_{KB}. Hence the last term of Eq. (7.36) of Part I vanishes and Eqs. (7.36) and (8.1) of Part I become identical for nondissipative media.

Case (iii), Strain energy of the form (2.17): In this case we introduce

$$\Sigma_{(3)}^* = \Sigma_{(3)}(C_{KL}, \hat{\Pi}_A, \hat{D}^{AB}, \mathbf{X}) - \mathcal{P}_{(3)}(\hat{\Pi}_A\hat{\Pi}^A - \mu_S^2) - \mathcal{P}_{(3)M}\hat{\Pi}_A\hat{D}^{AM} \tag{3.17}$$

and obtain the constitutive equations

$$\begin{aligned} T^{LK} &= 2\frac{\partial \Sigma_{(3)}}{\partial C_{KL}} + \left(\frac{\partial \Sigma_{(3)}}{\partial \hat{\Pi}^A}\overset{-1}{C}{}^{LN}\Pi_N + \frac{\partial \Sigma_{(3)}}{\partial \hat{D}^A{}_B}\overset{-1}{C}{}^{LM}D_{MB}\right. \\ &\quad - 2\mathcal{P}_{(3)}\hat{\Pi}_A\Pi_B\overset{-1}{C}{}^{LB} - \mathcal{P}_{(3)}{}^M\hat{D}_{AM}\overset{-1}{C}{}^{LB}\Pi_B \\ &\quad \left. - \mathcal{P}_{(3)}{}^M\hat{\Pi}_A\hat{C}^{LN}D_{NM}\right)\left(\overset{-1/2}{C}{}^{AK} - \overset{-1}{C}{}^{AK}\right), \\ {}_L B^K &= -J\left(\frac{\partial \Sigma_{(3)}}{\partial \hat{\Pi}^A} - 2\mathcal{P}_{(3)}\hat{\Pi}_A - \mathcal{P}_{(3)}{}^M\hat{D}_{AM}\right)\overset{-1/2}{C}{}^{AK}, \\ T^{(\mu)LK} &= \left(\frac{\partial \Sigma_{(3)}}{\partial \hat{D}^A{}_K} - \mathcal{P}_{(3)}{}^K\hat{\Pi}_A\right)\overset{-1/2}{C}{}^{AL}. \end{aligned} \tag{3.18}$$

For the computation of T^{LK}, the following expressions are used:

$$\frac{\partial \overset{-1/2}{C}{}^{AB}}{\partial x^k{}_{,K}} = \frac{\partial \overset{-1/2}{C}{}^{AB}}{\partial \overset{-1}{C}{}^{MN}}\frac{\partial \overset{-1}{C}{}^{MN}}{\partial x^k{}_{,K}} = -\overset{-1}{C}{}^{AK}X^B{}_{,k}, \tag{3.19}$$

$$\frac{\partial \hat{\Pi}^A}{\partial \mu^k} = \overset{-1/2}{C}{}^{AB}x_k{}_{,B}, \quad \frac{\partial \hat{\Pi}_A}{\partial x^k{}_{,K}} = -\overset{-1}{C}{}^{AK}X^B{}_{,k}\Pi_B + \overset{-1/2}{C}{}^{AK}\mu_k, \tag{3.20}$$

$$\frac{\partial \hat{D}^{AB}}{\partial \mu^k} = 0, \quad \frac{\partial \hat{D}^{AB}}{\partial \mu^k{}_{,K}} = \overset{-1/2}{C}{}^{AM}x_{k,M}g^{KB}, \tag{3.21}$$

$$\frac{\partial \hat{D}^A{}_B}{\partial x^k{}_{,K}} = -\overset{-1}{C}{}^{AK}X^M{}_{,k}D_{MB} + \overset{-1/2}{C}{}^{AK}\mu_k{}_{,B}. \tag{3.22}$$

Case (iv), Strain energy of the form (2.18): Here we write

$$\Sigma^{*}_{(4)} = \Sigma_{(4)}(C_{KL}, \hat{D}^{AB}, \mathbf{X}) - \mathcal{P}_{(4)} \det(\hat{\mathbf{D}}). \tag{3.23}$$

It follows that

$$\begin{aligned} T^{LK} &= 2\frac{\partial \Sigma_{(4)}}{\partial C_{KL}} - \mathcal{K}_M{}^N \overset{-1}{C}{}^{LJ} D_{JN}\left(\overset{-1/2}{C}{}^{MK} - \overset{-1}{C}{}^{MK}\right), \\ {}_{L}B^{K} &= -J\mathcal{K}_M{}^N \overset{-1}{C}{}^{AK}\frac{\partial \hat{D}^M{}_N}{\partial \hat{\Pi}^A}, \\ T^{(\mu)LK} &= \overset{-1}{C}{}^{ML}\mathcal{K}_M{}^K, \end{aligned} \tag{3.24}$$

where we have set

$$\begin{aligned} \mathcal{K}_{MN} &= \frac{\partial \Sigma_{(4)}}{\partial \hat{D}^{MN}} - \mathcal{P}_{(4)}\frac{\partial(\det\hat{\mathbf{D}})}{\partial \hat{D}^{MN}}, \\ \mathcal{K}_M{}^N &= \frac{\partial \Sigma_{(4)}}{\partial \hat{D}^M{}_N} - \tfrac{1}{2}\mathcal{P}_{(4)}\epsilon_{MBC}\epsilon^{NPQ}\hat{D}^B{}_P\hat{D}^C{}_Q. \end{aligned} \tag{3.25}$$

For this case we have $\hat{\Pi} = \hat{\Pi}(\hat{D}^{AB})$ and, therefore, one must compute $\partial \hat{D}^{MN}/\partial \hat{\Pi}^A$.

An examination of various forms obtained above indicates that the simplest form is provided by (3.13) which corresponds to the strain energy of the form (2.16). This is the form we shall employ in the rest of this article.

4. MATERIAL SYMMETRY

Materials in their natural states may possess certain symmetry regulations in their properties. The geometrical symmetry conditions in the physical properties of materials can be expressed by the form-invariance of the constitutive equations (hence the strain energy function) under a group of orthogonal transformations $\{\mathbf{S}\}$ and translations $\{\mathbf{B}\}$ at the material frame of reference, i.e.,

$$\bar{\mathbf{X}} = \mathbf{S}\mathbf{X} = \mathbf{B}, \tag{4.1}$$

where

$$\mathbf{S}\mathbf{S}^T = \mathbf{S}^T\mathbf{S} = \mathbf{I}, \quad \det \mathbf{S} = \pm 1. \tag{4.2}$$

The invariance under all members of $\{\mathbf{B}\}$ provides restrictions on the inhomogeneities and under all members of the group $\{\mathbf{S}\}$ places restrictions on the type of anisotropy present in the material at its natural state. Presently all known thirty-two classes of crystallographic elastic solids are obtained by use of twelve members of the group $\{\mathbf{S}\}$. However, here we are concerned mostly with *hemitropic* materials for which the symmetry group is the proper orthogonal group ($\det\mathbf{S} = +1$) and *isotropic* materials for which the symmetry group is the *full* group of orthogonal transformations.

The material symmetries in all classes of magnetic materials cannot be determined by use of the crystallographic $\{\mathfrak{G}\}$ included in $\{\mathbf{S}\}$. In addition to the geometrical symmetries present in the lattice structure of the crystals, the atoms of the lattice in magnetic materials are endowed with atomic magnetic moments (spins). It may turn out that the usual spatial symmetry operations, rotation and rotation-reflection, while preserving the geometrical properties of the lattice may reverse the orientation of spins. Thus there is need for the enlargement of the three-dimensional crystal group $\{\mathfrak{G}\}$. A separate argument in this regard may be made from the basic premises in the solid state theory[8]: The atoms of a crystal lattice are in constant oscillating motions and the time average of the positions of the component parts of the lattice constitutes the basis for the crystallographic symmetry. If there exists a microscopic nonzero time average for each component part, then this information is not contained in any specification of the average positions. If the nonzero time average of the position recurs in an identical manner in each unit cell of the crystal, then this information is not contained in the geometrical symmetry of the crystal. It is known that ferromagnetic, ferrimagnetic and antiferromagnetic crystals are characterized by orderly distributions of spin magnetic moments.

The foregoing argument suggests that in the treatment of the symmetry properties we must consider the time symmetry along with the spatial symmetry regulations. Thus a four-dimensional (space–time) group is needed for a satisfactory discussion of the physical properties of magnetic materials. For example, the effect of time reversal must be taken into account. Fortunately the time reversal adjoined to spatial symmetry operations such as rotation and rotation-reflection is the only other relevant operation which affects the symmetry properties of magnetic materials. Properties of diamagnetic and paramagnetic crystals are invariant under the time reversal so that these materials are governed solely by the three-dimensional group $\{\mathfrak{G}\}$. However, ferromagnetic, ferrimagnetic, and certain antiferromagnetic crystals are not time symmetric.

The space–time symmetry operators correspond to proper and improper spatial rotations combined with time inversion. Zheludev[9] introduced the notion of *complementary operation* $\mathfrak{R}$ *referred to as reversal of the spins* (or time inversion) whose product with R_i, an element of the conventional crystallographic point group, gives an element of the magnetic group $\{\mathfrak{M}\}$. The composition with $\mathfrak{R}$ is noted as

$$\mathfrak{R}\cdot R_i = \underline{R}_i. \tag{4.3}$$

For instance, we have

$$\begin{aligned} &R_1\cdot R_2 = R_3 \ (\text{since } R_1, R_2 \in \{\mathfrak{G}\} \Rightarrow R_3 \in \{\mathfrak{G}\}), \\ &\underline{R}_1\cdot\underline{R}_2 = R_3, \quad \underline{R}_1\cdot R_2 = R_1\cdot\underline{R}_2 = \underline{R}_3, \quad \underline{R}_3 \in \{\mathfrak{M}\}. \end{aligned} \tag{4.4}$$

The recognition of these new complementary symmetry operations, as explained above in the context of magnetic structures, results in a number of *ninety* possible crystallographic groups $\{\mathfrak{M}'\}$ (32 classical groups $\{\mathfrak{G}\}$ + 58 additional groups $\{\mathfrak{M}\}$; $\{\mathfrak{M}'\} = \{\mathfrak{G}\} \oplus \{\mathfrak{M}\}$) for short referred to as *magnetic point groups*. For the classical 32 groups it means that it is possible to orient magnetic moments in a crystal such that there is no deterioration of spatial symmetry even if invariance of magnetic moment orientation under a symmetry operation is demanded. The distinct variants $\{\mathfrak{M}'\}$ are obtained from the 32 ordinary groups [cf., (4.3)]. For example, for a cubic system $m3m$ (O_h in Schönflies' classification) we get $m3m$, $\underline{m}3\underline{m}$, $\underline{m}3m$, $m3\underline{m}$.

The distinctive feature of magnetic (material) *tensors* (i.e., the material coefficients appearing, for instance, in the expansion of the strain energy function)

consists therefore in their transformation properties under $\mathcal{R}$.

Finally we recognize the *Neumann's principle* of far-reaching insight[10]:

Every physical property of a crystal must possess at least the symmetry of the point group of the crystal.

In conclusion, the macroscopic symmetry properties of ferromagnetic crystals must be classified under the ninety magnetic classes. It is interesting to note that certain magnetic properties have been shown to be possible after recognizing the magnetic symmetry classes, and their existence has received experimental confirmation, e.g., piezomagnetism in CoF_2 and MnF_2 (antiferromagnetic crystals $\in \underline{4}/mm\underline{m}$ (Borovik-Romanov[11]). If the operation $\mathcal{R}$ has no effect on material properties, then the material falls in one of the classical 32 classes. For example, diamagnetic and paramagnetic materials, in fact, all centrosymmetric materials, fall in this category. Therefore there are no tensorial coefficients of odd rank in the expansion of the free energy and, hence, the magnetic properties are not influenced by $\mathcal{R}$. Similarly, the second order magnetostriction effect does not require the consideration of magnetic symmetries of crystals (the first order, yes). For ferromagnetism this, of course, implies the presence of a magnetic moment even in the absence of an external magnetic field.

5. HEMITROPIC AND ISOTROPIC MATERIALS

Suppose that η is the number of independent components of the tensorial arguments of the strain energy function $\mathfrak{F}$. If $\mathfrak{F}$ is form invariant under the orthogonal transformations $\mathbf{S}$ in an n-dimensional space then, according to a theorem of Cauchy, $\mathfrak{F}$ is a function of $\eta - p$ scalars I which form a minimal function basis of $\mathfrak{F}$, where $p \equiv n(n-1)/2$ (cf. Smith[12], Kafadar and Eringen[13]). If $I_1, \cdots, I_q$ $(0 \leq q \leq \eta - p)$ are absolute scalars and $I_{q+1}, \cdots, I_{\eta-p}$ are axial invariants, then $\mathfrak{F}$ is invariant under the *full* group if

$$\mathfrak{F}(I_1, \cdots, I_q, I_{q+1}, \cdots, I_{\eta-p}) = \mathfrak{F}(I_1, \cdots, I_q, -I_{q+1}, \cdots, -I_{\eta-p}). \quad (5.1)$$

This means that we only need to select $\eta - p$ independent members of the *minimal function basis* among the q independent members of the *integrity basis* (in the sense of Spencer[14]).

There remains the question of single-valuedness of the minimal function basis which has been the basis of certain recent controversies.[15]

Suppose that a minimal function basis is built up of $\eta - p$ members chosen from the q members of the minimal integrity basis. Suppose that m members of the remaining $q - (\eta - p)$ members cannot be expressed as single-valued functions (*not* always polynomials) of the $\eta - p$ function basis members, i.e., the equations

$$f(I_\alpha, J_\beta) = 0, \quad \alpha = 1, \ldots, \eta - p, \quad \beta = \eta - p + 1, \cdots, \eta - p + m \quad (5.2)$$

cannot be solved for J_β to yield a single-valued solution

$$J_\beta = J_\beta(I_\alpha).$$

In that case we cannot construct a minimal function basis only on $\eta - p$ members (cf. Smith[15]) but we can construct one based on $\eta - p + m$ members. But as has been pointed out by C. B. Kafadar (private communication), the above mentioned result becomes irrelevant for the formulation of nondissipative constitutive equations derivable from a strain energy function. Indeed, consider a second order tensor-valued function $\mathbf{t}$ derived from a potential $\mathfrak{F}$:

$$\mathbf{t} = \sum_{(K)=1}^{\eta} \frac{\partial \mathfrak{F}}{\partial U^{(K)}} \mathbf{B}^{(K)}, \quad \mathfrak{F} = \mathfrak{F}(U^{(K)}), \quad K = 1, 2, \ldots, \eta, \quad (5.3)$$

where $U^{(K)}$ is the indexed series of independent components of the arguments involved in $\mathfrak{F}$, $\mathbf{B}^{(K)}$ are functions of the $U^{(K)}$'s. Let $\{I_\alpha, \alpha = 1, 2, \cdots, \eta - p\}$ be the minimal function basis and $\{J_\beta, \beta = \eta - p + 1, \cdots, \eta - p + m\}$ the supplementary invariants that cannot be solved uniquely from the m relations

$$f_\gamma(I_\alpha, J_\beta) = 0, \quad \gamma = 1, 2, \cdots, m. \quad (5.4)$$

Following the preceding remark, we might consider $\mathfrak{F} = \hat{\mathfrak{F}}(I_\alpha, J_\beta)$ and (5.3) would be read as

$$\mathbf{t} = \sum_{(K)=1}^{\eta} \left(\frac{\partial \hat{\mathfrak{F}}}{\partial I_\alpha} \frac{\partial I_\alpha}{\partial U^{(K)}} + \frac{\partial \hat{\mathfrak{F}}}{\partial J_\beta} \frac{\partial J_\beta}{\partial U^{(K)}} \right) \mathbf{B}^{(K)}. \quad (5.5)$$

Differentiating (5.4) with respect to I_α we obtain

$$\frac{\partial f_\gamma}{\partial I_\alpha} + \frac{\partial f_\gamma}{\partial J_\beta} \frac{\partial J_\beta}{\partial I_\alpha} = 0. \quad (5.6)$$

This relation can be solved for $\partial J_\beta / \partial I_\alpha$ since $\partial f_\gamma / \partial J_\beta$ is well defined, hence (5.5) may be written as

$$\mathbf{t} = \sum_{(K)=1}^{\eta} \sum_{\alpha=1}^{\eta-p} \lambda^\alpha \frac{\partial I_\alpha}{\partial U^{(K)}} \mathbf{B}^{(K)}, \quad (5.7)$$

with

$$\lambda^\alpha = \lambda^\alpha(I_\sigma, J_\beta) = \frac{\partial \hat{\mathfrak{F}}}{\partial I_\alpha} + \frac{\partial \hat{\mathfrak{F}}}{\partial J_\beta} \frac{\partial J_\beta}{\partial I_\alpha}, \quad \alpha = 1, 2, \cdots, \eta - p, \quad \sigma = 1, 2, \cdots, \eta - p. \quad (5.8)$$

It follows that, in this formulation, though we take account of the nonsingle valuedness of the solutions of (5.4), there are still $\eta - p$ coefficients λ^α in the constitutive equation. Furthermore, an astute choice of the I_α may lead each constitutive equation not to contain all these coefficients, the value of which follows from experiment results.

For tranversely isotropic materials, there exists a single preferred direction $\mathbf{N}$ which is the same at all points of the body at its natural state. In the special case when the direction of the spin coincides with $\mathbf{N}$, then we need no special case for the magnetic group symmetries. If we select $\mathbf{N}$ in the direction of the X_3 axis then the group of transformations under which the strain energy function $\mathfrak{F}$ is invariant may be represented by the matrices

$$\mathbf{M}_\theta = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{R}_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad 0 \leq \theta \leq 2\pi. \quad (5.9)$$

In this case $\mathbf{R}_1 \cdot \mathbf{M}_\theta$ preserves the orientation of the spins.

There exists a class of materials which appear to satisfy such transverse isotropy conditions. These are the rubber bonded barium ferrite composite materials (see Kafadar and Eringen[13]). In the initial state of these materials, all magnetic moments are alined along a direction which is also a privileged direction as far as the mechanical properties are studied. Under application of an external magnetic field, the embedded ferrite particles tend in general to rotate, a behavior that can be described by the "*spin rotation*" equation of the above developed theory.[16] Generally these materials are viscoelastic, but one may assume for the sake of simplicity that these media are nonlinear elastic, able to support large deformations, and obviously transversely isotropic.

6. CONSTITUTIVE EQUATIONS OF NONLINEAR HEMITROPIC MAGNETO-ELASTIC SOLIDS

A. Material Expressions

Here we obtain the constitutive equations of nonlinear magneto-elastic solids based on the strain energy function (2.16), i.e.,

$$\Sigma_{(2)} = \Sigma_{(2)}(C_{KL}, \Pi_A, G_{AB}) \tag{6.1}$$

subject to the saturation condition

$$\Pi_A \overset{-1}{C}{}^{AB} \Pi_B = \mu_S^2 = \text{const.} \tag{6.2}$$

For the present case, $n = 3$ (space is three-dimensional), $p = n(n-1)/2 = 3, \eta = 18$ for nonsaturated media and, for saturated media, $\eta = 14$. Thus

$$\eta - p = 15 \text{ for nonsaturated media,}$$
$$\eta - p = 11 \text{ for saturated media.}$$

For hemitropic materials Σ must be invariant under the *proper* group. The axial vector Π_A may be replaced by its dual Σ^{AB}. We have the identities

$$\text{tr}\boldsymbol{\Sigma} = \text{tr}\mathbf{C}\boldsymbol{\Sigma} = \text{tr}\mathbf{C}^2\boldsymbol{\Sigma} = \text{tr}\mathbf{G}\boldsymbol{\Sigma} = 0, \tag{6.3}$$

and the fact that $\text{tr}\boldsymbol{\Sigma}^3$ is not independent of $\text{tr}\boldsymbol{\Sigma}^2$. Among the minimal integrity basis,[14] for the case of nonsaturated media, we select the following 15 invariants $I_\beta, \beta = 1, 2, \cdots, 15$,

$$\begin{aligned}
&I_1 = \text{tr}\mathbf{C}, \quad I_2 = \tfrac{1}{2}\text{tr}\mathbf{C}^2, \quad I_3 = \tfrac{1}{3}\text{tr}\mathbf{C}^3,\\
&I_4 = \tfrac{1}{2}\text{tr}\boldsymbol{\Sigma}^2, \quad I_5 = \text{tr}\boldsymbol{\Sigma}^2\mathbf{C}, \quad I_6 = \tfrac{1}{2}\text{tr}\boldsymbol{\Sigma}^2\mathbf{C}^2,\\
&I_7 = \text{tr}\mathbf{G}, \quad I_8 = \tfrac{1}{2}\text{tr}\mathbf{G}^2, \quad I_9 = \tfrac{1}{3}\text{tr}\mathbf{G}^3,\\
&I_{10} = \text{tr}\mathbf{C}\mathbf{G}, \quad I_{11} = \text{tr}\mathbf{C}\mathbf{G}^2, \quad I_{12} = \text{tr}\mathbf{G}\mathbf{C}^2,\\
&I_{13} = \tfrac{1}{2}\text{tr}\mathbf{C}^2\mathbf{G}^2, \quad I_{14} = \tfrac{1}{2}\text{tr}\boldsymbol{\Sigma}^2\mathbf{G}^2, \quad I_{15} = \text{tr}\boldsymbol{\Sigma}^2\mathbf{G}.
\end{aligned} \tag{6.4}$$

For the case of saturation we may choose the first eleven invariants I_β $\beta = 1, 2, \cdots, 11$. The constitutive equations (3.1) are

$$\begin{aligned}
T^{LK} &= 2(\alpha^1 - \alpha^5 \Pi_P \Pi^P) g^{KL} + \alpha^2 C^{KL} + \alpha^3 C^K{}_M C^{ML}\\
&\quad - \alpha^5 \Pi^K \Pi^L - \alpha^4 \Pi^K \overset{-1}{C}{}^{LA} \Pi_A\\
&\quad + \alpha^6 (\Sigma^M{}_N \Sigma^{NK} C^L{}_M - \Sigma^L{}_N \Sigma^N{}_P C^{PK})\\
&\quad + 2\alpha^{10} G^{LK} + 2\alpha^{11} G^L{}_P G^{PK}\\
&\quad + 2\alpha^5 \epsilon_{PQ}{}^K \Sigma^Q{}_J C^{PJ} \overset{-1}{C}{}^{LA} \Pi_A\\
&\quad + \alpha^6 \epsilon_{PQ}{}^K \Sigma^Q{}_J C^J{}_I C^{PI} \overset{-1}{C}{}^{LA} \Pi_A,\\
{}_L B^K &= 2J\alpha^4 \Pi^K - 2J\alpha^5 \epsilon_{PQ}{}^K \Sigma^Q{}_J C^{PJ}\\
&\quad - J\alpha^6 \epsilon_{PQ}{}^K \Sigma^Q{}_J C^J{}_I C^{PI} + 2J\mathcal{P}_{(2)} \overset{-1}{C}{}^{KB} \Pi_B,\\
T^{(\mu)LK} &= 2[\alpha^7 g^{KB} + \alpha^8 G^{KB} + \alpha^9 G^K{}_M G^{MB} + \alpha^{10} C^{KB}\\
&\quad + \alpha^{11}(C^{KM} G_M{}^B + G^{KN} C_N{}^B)] D_{JB} \overset{-1}{C}{}^{LJ},
\end{aligned} \tag{6.5}$$

where g^{KB} is the metric tensor in material coordinates and $\alpha^\beta \equiv \partial \Sigma_{(2)} / \partial I_\beta$.

It is interesting to see what happens when one starts with the expressions (2.17). Calling $\hat{\boldsymbol{\Sigma}}$ the dual of $\hat{\boldsymbol{\Pi}}$ and noting that $\hat{\mathbf{D}}$ is a general second order tensor, we can select 15 invariants for the nonsaturated case:

$$\begin{aligned}
&I_1 = \text{tr}\mathbf{C}, \quad I_2 = \tfrac{1}{2}\text{tr}\mathbf{C}^2, \quad I_3 = \tfrac{1}{3}\text{tr}\mathbf{C}^3,\\
&I_4 = \text{tr}\mathbf{C}\hat{\mathbf{D}}, \quad I_5 = \text{tr}\mathbf{C}^2\hat{\mathbf{D}}, \quad I_6 = \text{tr}\hat{\mathbf{D}},\\
&I_7 = \tfrac{1}{2}\text{tr}\hat{\mathbf{D}}^2, \quad I_8 = \tfrac{1}{3}\text{tr}\hat{\mathbf{D}}^3, \quad I_9 = \tfrac{1}{2}\text{tr}\hat{\mathbf{D}}\hat{\mathbf{D}}^T,\\
&I_{10} = \text{tr}\hat{\mathbf{D}}^2\hat{\mathbf{D}}^T, \quad I_{11} = \tfrac{1}{2}\text{tr}\hat{\mathbf{D}}^2(\hat{\mathbf{D}}^T)^2, \quad I_{12} = \tfrac{1}{2}\text{tr}\hat{\boldsymbol{\Sigma}}^2\\
&I_{13} = \text{tr}\hat{\boldsymbol{\Sigma}}\hat{\mathbf{D}}, \quad I_{14} = \text{tr}\hat{\boldsymbol{\Sigma}}\hat{\mathbf{D}}^2, \quad I_{15} = \text{tr}\hat{\boldsymbol{\Sigma}}\hat{\mathbf{D}}\mathbf{C},
\end{aligned} \tag{6.6}$$

where the superscript T stands for transposition.

For the case of saturation we may take account of the constraints $\hat{\Pi}_A \hat{D}^{AB} = 0$ by choosing to discard those invariants which vanish as a consequence of these constraints. For instance, we notice that

$$I_{13} = I_{14} = I_{15} = 0. \tag{6.7}$$

Hence the multipliers $\mathcal{P}_{(3)}{}^M$ are unimportant and we keep the eleven invariants $I_\beta, \beta = 1, 2, \cdots, 11$, of the list (6.6). Alternatively, since $I_{12} = \text{const}$, we may take the twelve invariants $I_\beta, \beta = 1, 2, \cdots, 12$, and the unknown $\mathcal{P}_{(3)}$ may be preserved. It follows that the constitutive equations (3.1) are

$$\begin{aligned}
T^{LK} &= 2\sum_{\beta=1}^{5} \alpha^\beta \frac{\partial I_\beta}{\partial C_{KL}} + \left(\alpha^{12} \frac{\partial I_{12}}{\partial \hat{\Pi}^A} \overset{-1}{C}{}^{LN} \Pi_N \right.\\
&\quad \left. + \sum_{\beta=4}^{11} \alpha^\beta \frac{\partial I_\beta}{\partial \hat{D}^A{}_B} \overset{-1}{C}{}^{LM} D_{MB} - 2\mathcal{P}_{(3)} \hat{\Pi}_A \Pi_B \overset{-1}{C}{}^{LB}\right)\\
&\quad \times (\overset{-1/2}{C}{}^{AK} - \overset{-1}{C}{}^{AK}),\\
{}_L B^K &= -J\left(\alpha^{12} \frac{\partial I_{12}}{\partial \hat{\Pi}^A} - 2\mathcal{P}_{(3)} \hat{\Pi}_A\right) \overset{-1/2}{C}{}^{AK},\\
T^{(\mu)LK} &= \left(\sum_{\beta=4}^{11} \alpha^\beta \frac{\partial I_\beta}{\partial \hat{D}^A{}_K}\right) \overset{-1/2}{C}{}^{AL}.
\end{aligned} \tag{6.8}$$

Twelve, one, and eight material coefficients are, respectively, present in Eqs. (6.8). The full expressions corresponding to (6.8), especially to the first of (6.8) are quite involved and we choose not to give them here. We note that minor simplifications are still possible in formulas (6.5) an (6.8) by use of the Cayley–Hamilton theorem.

B. Spatial Expressions

In many applications the spatial forms of the constitutive equations are needed. To this end we set

$$t_{kl} = {}_E t_{kl} + \rho_L B_k \mu_l - t^{(\mu)}{}_k{}^m \mu_{l;m}, \tag{6.9}$$

where

$$ {}_E t_{kl} = \frac{\rho}{\rho_R}\frac{\partial \Sigma}{\partial x^{(k}{}_{,K}}\, x_{l),K}, \qquad {}_L B_k = -\frac{1}{\rho_R}\frac{\partial \Sigma}{\partial \mu^k}, \tag{6.10} $$

$$ t^{(\mu)}{}_k{}^m = \frac{\rho}{\rho_R}\frac{\partial \Sigma}{\partial \mu^k{}_{,K}}\, x^m{}_{,K}. \tag{6.11} $$

Here ${}_E\mathbf{t}$ is the elastic stress tensor and $\mathbf{t}^{(\mu)}$ is the "*spin interaction*" stress tensor. In addition, the following *spatial* strain measures are needed:

$$ \overset{-1}{c}{}^{kl} \equiv g^{KL} x^k{}_{,K} x^l{}_{,L}, \qquad \mu^k, \qquad \overset{-1}{d}{}^{kl} \equiv g^{KL} x^k{}_{,K} \mu^l{}_{,L}. \tag{6.12} $$

Thus, we may consider

$$ \Sigma = \Sigma(\overset{-1}{\mathbf{c}}, \boldsymbol{\mu}, \overset{-1}{\mathbf{d}}). \tag{6.13} $$

The constitutive equations (2.10)–(2.12) in *spatial* form are

$$ \mathbf{t} = \frac{\rho}{\rho_R}\left[2\frac{\partial \Sigma}{\partial \overset{-1}{\mathbf{c}}}\cdot\overset{-1}{\mathbf{c}} + \frac{\partial \Sigma}{\partial \overset{-1}{\mathbf{d}}}\left(\overset{-1}{\mathbf{d}}\right)^T + \boldsymbol{\mu}\otimes\frac{\partial \Sigma}{\partial \boldsymbol{\mu}}\right], \tag{6.14} $$

$$ {}_L\mathbf{B} = -\frac{1}{\rho_R}\frac{\partial \Sigma}{\partial \boldsymbol{\mu}}, \qquad \mathbf{t}^{(\mu)} = \frac{\rho}{\rho_R}\frac{\partial \Sigma}{\partial \left(\overset{-1}{\mathbf{d}}\right)^T}\cdot\overset{-1}{\mathbf{c}}. \tag{6.15} $$

Moreover, one may verify that the form (6.13) constitutes an integral of Eq. (7.35) of Part I, thus satisfying the Euclidean invariance requirement. In obtaining (6.14) and (6.15) we have made use of the results

$$ \frac{\partial \overset{-1}{c}{}^{mn}}{\partial x^k{}_{,K}} = 2g^{KN}\delta^m{}_k x^n{}_{,N}, \qquad \frac{\partial \overset{-1}{d}{}^{mn}}{\partial x^k{}_{,K}} = g^{KN}\delta^m{}_k \mu^n{}_{,N}, \tag{6.16} $$

$$ \frac{\partial \overset{-1}{d}{}^{mn}}{\partial \mu^k{}_{,K}} = g^{MK}\delta^n_k x^m{}_{,M}. \tag{6.17} $$

For the *saturation case*, we may take

$$ \Sigma = \Sigma\left(\overset{-1}{\mathbf{c}}, \boldsymbol{\mu}, \overset{-1}{\boldsymbol{\gamma}}\right) \tag{6.18} $$

subject to

$$ \boldsymbol{\mu}\cdot\boldsymbol{\mu} = \mu_s^2 = \text{const.} $$

In (6.18), we have defined

$$ \overset{-1}{\gamma}{}^{kl} \equiv g^{KL}\mu^k{}_{,K}\mu^l{}_{,L}, \qquad \overset{-1}{\gamma}{}^{kl} = \overset{-1}{\gamma}{}^{lk}, \tag{6.19} $$

that is related to $\overset{-1}{\mathbf{d}}$ and $\overset{-1}{\mathbf{c}}$ through the relation

$$ \overset{-1}{\boldsymbol{\gamma}} = \left(\overset{-1}{\mathbf{d}}\right)^T\cdot\overset{-1}{\mathbf{c}}\cdot\overset{-1}{\mathbf{d}}, \tag{6.20} $$

where

$$ \mathbf{c} = \left(\overset{-1}{\mathbf{c}}\right)^{-1}, \text{ i.e., } c_{kl} \equiv g_{KL} X^K_{,k} X^L_{,l}. $$

Setting:

$$ \Sigma^* = \Sigma - \mathcal{P}(\mu^k\mu_k - \mu_s^2) \tag{6.21} $$

to take account of the constraint, we get the constitutive equations

$$ \mathbf{t} = 2\frac{\rho}{\rho_R}\frac{\partial \Sigma}{\partial \overset{-1}{\mathbf{c}}}\cdot\overset{-1}{\mathbf{c}} + \frac{\rho}{\rho_R}\,\boldsymbol{\mu}\otimes\frac{\partial \Sigma}{\partial \boldsymbol{\mu}}, \tag{6.22} $$

$$ {}_L\mathbf{B} = -\frac{1}{\rho_R}\left(\frac{\partial \Sigma}{\partial \boldsymbol{\mu}} - 2\mathcal{P}\boldsymbol{\mu}\right), \qquad \mathbf{t}^{(\mu)} = 2\frac{\rho}{\rho_R}\frac{\partial \Sigma}{\partial \overset{-1}{\boldsymbol{\gamma}}}\left(\overset{-1}{\mathbf{d}}\right)^T. \tag{6.23} $$

In the last relation we have used the result

$$ \frac{\partial \overset{-1}{\gamma}{}^{mn}}{\partial \mu^k{}_{,K}} = 2g^{KN}\delta^m_k \mu^n{}_{,N}. $$

Note that the unknown $\mathcal{P}$ appearing in (6.23) is irrelevant to the spin rotation equation.

If we note the relation

$$ \overset{-1}{\mathbf{d}} = \overset{-1}{\mathbf{c}}\cdot\nabla\boldsymbol{\mu}, \tag{6.24} $$

then the second of Eqs. (6.23) can be written in the alternative form

$$ \mathbf{t}^{(\mu)} = 2\frac{\rho}{\rho_R}\frac{\partial \Sigma}{\partial \overset{-1}{\boldsymbol{\gamma}}}\left(\overset{-1}{\mathbf{c}}\cdot\nabla\boldsymbol{\mu}\right)^T. \tag{6.25} $$

Analogous to (6.4), spatial forms of the minimal function basis are

$$ \begin{aligned} &I_1 = \operatorname{tr}\overset{-1}{\mathbf{c}}, \quad I_2 = \tfrac{1}{2}\operatorname{tr}\overset{-2}{\mathbf{c}}, \quad I_3 = \tfrac{1}{3}\operatorname{tr}\overset{-3}{\mathbf{c}}, \\ &I_4 = \tfrac{1}{2}\operatorname{tr}\boldsymbol{\sigma}^2, \quad I_5 = \operatorname{tr}\boldsymbol{\sigma}^2\overset{-1}{\mathbf{c}}, \quad I_6 = \tfrac{1}{2}\operatorname{tr}\boldsymbol{\sigma}^2\overset{-2}{\mathbf{c}}, \\ &I_7 = \operatorname{tr}\overset{-1}{\boldsymbol{\gamma}}, \quad I_8 = \tfrac{1}{2}\operatorname{tr}\overset{-2}{\boldsymbol{\gamma}}, \quad I_9 = \tfrac{1}{3}\operatorname{tr}\overset{-3}{\boldsymbol{\gamma}}, \\ &I_{10} = \operatorname{tr}\overset{-1-1}{\mathbf{c}\,\boldsymbol{\gamma}}, \quad I_{11} = \operatorname{tr}\overset{-1-2}{\mathbf{c}\,\boldsymbol{\gamma}}, \quad I_{12} = \operatorname{tr}\overset{-1-2}{\boldsymbol{\gamma}\,\mathbf{c}}, \\ &I_{13} = \tfrac{1}{2}\operatorname{tr}\overset{-2-2}{\mathbf{c}\,\boldsymbol{\gamma}}, \quad I_{14} = \tfrac{1}{2}\operatorname{tr}\boldsymbol{\sigma}^2\overset{-2}{\boldsymbol{\gamma}}, \quad I_{15} = \operatorname{tr}\boldsymbol{\sigma}^2\overset{-1}{\boldsymbol{\gamma}}, \end{aligned} \tag{6.26} $$

where $\boldsymbol{\sigma} \equiv$ dual $\boldsymbol{\mu}$. Only the first eleven invariants need be considered for saturated magnetization. In that case, the spatial equivalent of (6.5) is

$$ \begin{aligned} t_{kl} = {}&(\rho/\rho_R)[2(\alpha^1 - \alpha^5\mu_s^2)\overset{-1}{c}{}_{kl} + 2\alpha^2\overset{-1}{c}{}_{km}\overset{-1}{c}{}^m{}_l \\ &+ 2\alpha^3\overset{-1}{c}{}_{km}\overset{-1}{c}{}^{mn}\overset{-1}{c}{}_{nl} + 2\alpha^5\mu_k\mu_m\overset{-1}{c}{}^m{}_l \\ &+ 4\alpha^6\sigma^{pq}\sigma_{n[k}\overset{-1}{c}{}_{m]p}\overset{-1}{c}{}^m{}_l + 2\alpha^{10}\overset{-1}{\boldsymbol{\gamma}}{}_{km}\overset{-1}{c}{}^m{}_l \\ &+ 2\alpha^{11}\overset{-1}{\gamma}{}_{kn}\overset{-1}{\gamma}{}^{nm}\overset{-1}{c}{}_{ml} - 2\alpha^4\mu_l\mu_k \\ &+ 2\alpha^5\epsilon_{pql}\mu_k\sigma^{qn}\overset{-1}{c}{}^p{}_n + 2\alpha^6\epsilon_{pql}\sigma^{qn}\overset{-1}{c}{}^{pr}\overset{-1}{c}{}_{nr}\mu_k, \\ {}_L B_k = {}&-(2/\rho_R)\Big[-(\alpha^4 + \mathcal{P})\mu_k + \alpha^5\epsilon_{pqk}\sigma^{qn}\overset{-1}{c}{}^p{}_n \\ &+ \alpha^6\epsilon_{pqk}\sigma^{qn}\overset{-1}{c}{}^{pr}\overset{-1}{c}{}_{nr}\Big], \\ t^{(\mu)}{}_{kl} = {}&2(\rho/\rho_R)\Big[\alpha^7\overset{-1}{c}{}^q_k + \alpha^8\overset{-1}{\gamma}{}_{km}\overset{-1}{c}{}^{mq} \\ &+ \alpha^9\overset{-1}{\gamma}{}_{mn}\overset{-1}{\gamma}{}^n_k\overset{-1}{c}{}^{mq} + \alpha^{10}\overset{-1}{c}{}_{km}\overset{-1}{c}{}^{mq} \\ &+ \alpha^{11}\left(\overset{-1}{c}{}^p{}_k\overset{-1}{\gamma}{}_{mp} + \overset{-1}{c}{}_{mn}\overset{-1}{\gamma}{}^n_k\right)\overset{-1}{c}{}^{mq}\Big]\mu_{l;q}. \end{aligned} \tag{6.27} $$

7. APPROXIMATIONS, INFINITESIMAL DEFORMATIONS

For infinitesimal deformations, we can take[17]

$$ \begin{aligned} &\overset{-1}{c}{}^{kl} \approx g^{kl} + 2e^{kl}, \qquad \rho/\rho_R \approx 1 - e, \\ &\overset{-1}{d}{}^{kl} \approx \mu^{l;k} = g^{KL}g^k{}_K\mu^l{}_{,L}, \qquad \overset{-1}{\gamma}{}^{kl} \approx \mu^{(l;k)}, \end{aligned} \tag{7.1} $$

where e_{kl} is the Eulerian strain tensor, e is its trace and $g^k{}_K$ are the shifters. Therefore, the set of constitutive equations (6.14)–(6.15) for the *nonsaturated case*, reduces to

$$t_{kl} = \frac{\partial \Sigma}{\partial e^{kl}} + \frac{\partial \Sigma}{\partial \mu^l}\mu_k + \frac{\partial \Sigma}{\partial \mu^{n;k}}\mu^n{}_{;l},$$
$$_LB_k = -\frac{1}{\rho}\frac{\partial \Sigma}{\partial \mu^k}, \qquad t^{(\mu)}{}_{kl} = \frac{\partial \Sigma}{\partial \mu^{k;l}}. \tag{7.2}$$

For the *saturated case*, we get

$$t_{kl} = \frac{\partial \Sigma}{\partial e^{kl}} + \frac{\partial \Sigma}{\partial \mu^l}\mu_k$$
$$_LB_k = -\frac{1}{\rho}\left(\frac{\partial \Sigma}{\partial \mu^k} - 2\mathcal{P}\mu_k\right), \qquad t^{(\mu)kl} = 2\frac{\partial \Sigma}{\partial \mu_{(l;k)}}. \tag{7.3}$$

The equations (7.2) constitute the magnetic counterpart of the equations (4.1) in Suhubi.[18]

For the linearized theory, one needs a strain energy function expanded as a function of the independent variables, i.e.,

$$\Sigma = a^k\mu_k + a^{kl}\mu_{l;k} + \tfrac{1}{2}a^{klmn}\mu_{l;k}\mu_{n;m} + \tfrac{1}{2}b^{mn}\mu_m\mu_n + j^{klm}\mu_k\mu_{m;l} + d^{kl}e_{kl} + \tfrac{1}{2}d^{klmn}e_{kl}e_{mn} + f^{klm}\mu_k e_{lm} + c^{klmn}e_{kl}\mu_{n;m} + \tfrac{1}{2}\lambda^{klmn}e_{kl}\mu_m\mu_n, \tag{7.4}$$

where the coefficients satisfy the symmetry relations

$$b^{mn} = b^{nm}, \quad a^{klmn} = a^{mnkl}, \quad c^{klmn} = c^{lkmn},$$
$$d^{kl} = d^{lk}, \quad d^{klmn} = d^{lkmn} = d^{klnm} = d^{mnkl}, \tag{7.5}$$
$$f^{klm} = f^{kml}, \quad \lambda^{klmn} = \lambda^{lkmn} = \lambda^{klnm}.$$

The strain energy function is subject to some restrictions:

(a) If in the natural state the initial stress, the initial local magnetic field intensity and the initial "spin interaction" vanish, then

$$d^{kl} = 0, \quad a^k = 0, \quad a^{kl} = 0; \tag{7.6}$$

(b) For *centrosymmetric* materials, there exists no odd rank material tensors, i.e.,

$$a^k = j^{klm} = f^{klm} = 0; \tag{7.7}$$

(c) The strain energy is subject to the symmetry operation $\mathcal{R}$ (cf. Sec. 4.). Since μ is an axial tensor, this implies that

$$a^k = a^{kl} = f^{klm} = c^{klmn} = 0. \tag{7.8}$$

In the case for which (7.7) and (7.8) hold true, the coupling between magnetization and deformation fields subsist only in the form of the magnetostrictive effect due to the material coefficient λ^{klmn}. This material coefficient, which was first introduced by Akulov,[19] admits only two independent components for a centrosymmetric cubic crystal. Thus the strain energy function assumes the simple form

$$\Sigma = \tfrac{1}{2}a^{klmn}\mu_{l;k}\mu_{n;m} + \tfrac{1}{2}b^{mn}\mu_m\mu_n + \tfrac{1}{2}d^{klmn}e_{kl}e_{mn} + \tfrac{1}{2}\lambda^{klmn}e_{kl}\mu_m\mu_n \tag{7.9}$$

of which the different terms are often referred to as the *exchange* energy, the *magnetic anisotropy* energy, the *elastic* energy and the *magnetostrictive* energy.

Without restrictions (7.6)–(7.8), from (7.4) and (7.2) we obtain the constitutive relations for *nonsaturated* magnetization:

$$\begin{aligned} t_{kl} &= d_{kl} + d_{klmn}e^{mn} + f_{mkl}\mu^m + c_{klmn}\mu^{n;m} \\ &\quad + \tfrac{1}{2}\lambda_{klmn}\mu^m\mu^n + a_l\mu_k + b_{ml}\mu^m\mu_k \\ &\quad + (j_{lmn}\mu^{n;m} + f_{lmn}e^{mn} + \lambda_{mnlp}e^{mn}\mu^p)\mu_k \\ &\quad + a_{kn}\mu^n{}_{;l} + a_{nkpq}\mu^{q;p}\mu^n{}_{;l} + j_{mkn}\mu^n{}_{;l}\mu^m \\ &\quad + c_{pqkn}e^{pq}\mu^n{}_{;l}, \\ {}_LB_k &= -\rho^{-1}(a_k + b_{mk}\mu^m + j_{klm}\mu^{m;l} + f_{klm}e^{lm} \\ &\quad + \lambda_{mnkl}e^{mn}\mu^l), \\ t^{(\mu)}{}_{kl} &= a_{lk} + a_{lkmn}\mu^{n;m} + j_{mkl}\mu^m + c_{mnlk}e^{mn} \end{aligned} \tag{7.10}$$

and the following ones for the case of saturation:

$$\begin{aligned} t_{kl} &= d_{kl} + d_{klmn}e^{mn} + f_{mkl}\mu^m + c_{klmn}\mu^{n;m} \\ &\quad + \tfrac{1}{2}\lambda_{klmn}\mu^m\mu^n + a_l\mu_k + b_{ml}\mu^m\mu_k \\ &\quad + (j_{lmn}\mu^{n;m} + f_{lmn}e^{mn} + \lambda_{mnlp}e^{mn}\mu^p)\mu_k, \\ {}_LB_k &= -\rho^{-1}(a_k + b_{mk}\mu^m + j_{klm}\mu^{m;l} + f_{klm}e^{lm} \\ &\quad + \lambda_{mnkl}e^{mn}\mu^l - 2\mathcal{P}\mu_k), \\ t^{(\mu)}{}_{kl} &= 2(a_{(kl)} + a_{(lk)mn}\mu^{n;m} + j_{m(kl)}\mu^m + c_{mn(kl)}e^{mn}). \end{aligned} \tag{7.11}$$

We note that the "spin interaction" stress tensor is symmetric in the latter case while the stress tensor t_{kl} is not.

With the hypotheses (7.6)–(7.8), the *nonsaturated* case reduces to

$$\begin{aligned} t_{kl} &= d_{klmn}e^{mn} + c_{klmn}\mu^{n;m} + \tfrac{1}{2}\lambda_{klmn}\mu^m\mu^n \\ &\quad + b_{ml}\mu^m\mu_k + \lambda_{mnlp}e^{mn}\mu^p\mu_k, \\ {}_LB_k &= -\rho^{-1}(b_{mk}\mu^m + \lambda_{mnkl}e^{mn}\mu^l) \\ t^{(\mu)}{}_{kl} &= a_{lkmn}\mu^{n;m}, \end{aligned} \tag{7.12}$$

and the *saturated* case to

$$t_{kl} = d_{klmn}e^{mn} + \tfrac{1}{2}\lambda_{klmn}\mu^m\mu^n + b_{ml}\mu^m\mu_k + \lambda_{mnlp}e^{mn}\mu^p\mu_k \tag{7.13a}$$
$$_LB_k = -\rho^{-1}(b_{mk}\mu^m + \lambda_{mnkl}e^{mn}\mu^l - 2\mathcal{P}\mu_k), \tag{7.13b}$$
$$t^{(\mu)}{}_{kl} = 2a_{(lk)mn}\mu^{n;m}. \tag{7.13c}$$

If the material is noncentrosymmetric (for instance if it is transversely isotropic according to the definition given in Sec. 5), we must keep $j^{klm} \neq 0$. Thus, for example, in the case of saturation, we must add $j_{lmn}\mu^{n;m}\mu_k$, $-\rho^{-1}j_{klm}\mu^{m;l}$ and $2j_{m(kl)}\mu^m$ to (7.13a), (7.13b), and (7.13c) respectively.

Finally, if we disregard terms of order higher than that of e_{kl}, the constitutive relations (7.13) provide the set

$$t_{kl} = d_{klmn}e^{mn} + \tfrac{1}{2}\lambda_{klmn}\mu^m\mu^n + b_{ml}\mu^m\mu_k, \tag{7.14a}$$
$$_LB_k = -\rho^{-1}b_{kn}\mu^n + (2/\rho)\mathcal{P}\mu_k, \tag{7.14b}$$
$$t^{(\mu)}{}_{kl} = 2a_{(lk)mn}\mu^{n;m}. \tag{7.14c}$$

The first term in (7.14a) is none other than the Hooke's law for infinitesimal deformations in Elas-

ticity theory; the second term that represents the magnetostrictive effect does not perturb the symmetry of t_{kl}. The last term is due to the magnetic anisotropy.

For ferromagnetic materials, the following remark is in order. We recall that $\mathbf{B} = \mathbf{H} + \rho\boldsymbol{\mu}$. For very strong fields, we may assume that $\mathbf{M} = \rho\boldsymbol{\mu} \gg \mathbf{H}$, hence there are no magnetic domains in the magnetically saturated crystal and the magnetic anisotropy energy may be considered to be negligible. It follows that $\mathbf{M}$ and $\mathbf{B}$ are parallel and the only remaining coupling is through the magnetostrictive effect. If the ferromagnetic material is uniaxial with $\mathbf{H} \approx \rho\boldsymbol{\mu}$, one must take account of the magnetic anisotropy effect the influence of which is comparable to that of the magnetostrictive term.

Finally we give a useful form of the strain energy function which has applications in certain further studies, e.g., magnetization surface energy if there is any, study of the nonnegativeness of the strain energy, uniqueness theorem for the linearized theory.

Starting from (7.11) which is valid only for saturated media, considering a stress free initial state and neglecting terms of order higher than that of e_{kl}, we can write

$$\begin{aligned}\Sigma = {} & \tfrac{1}{2} t_{kl} e^{kl} - (\rho/2)\, {}_L B^k \mu^k + \tfrac{1}{2} t^{(\mu)}{}_{kl} \mu^{k;l} + \mathcal{P}\mu_s^2 \\ & + (\tfrac{1}{2} a_k \mu^k + \tfrac{1}{2} j_{m[kl]} \mu^m \mu^{k;l} + \tfrac{1}{2} a_{[lk]mn} \mu^{n;m} \mu^{k;l} \\ & - \tfrac{1}{2} a_l \mu_k e^{kl} - \tfrac{1}{2} b_{ml} \mu^m \mu_k e^{kl} - \tfrac{1}{4} \lambda_{klmn} \mu^m \mu^n e^{kl} \\ & - \tfrac{1}{2} c_{mn[lk]} e^{mn} \mu^{k;l} + a_{[lk]} \mu^{k;l} + \tfrac{1}{2} a_{(lk)} \mu^{k;l}).\end{aligned} \tag{7.15}$$

In the simpler case of *saturated, centrosymmetric, initially stress free* materials corresponding to the form (7.9), this reduces to

$$\begin{aligned}\Sigma = {} & \tfrac{1}{2} t_{kl} e^{kl} - (\rho/2)\, {}_L B_k \mu^k + \tfrac{1}{2} t^{(\mu)}{}_{kl} \mu^{k;l} + \mathcal{P}\mu_s^2 \\ & + (\tfrac{1}{2} a_{[lk]mn} \mu^{n;m} \mu^{k;l} - \tfrac{1}{2} b_{ml} \mu^m \mu_k e^{kl} \\ & - \tfrac{1}{4} \lambda_{klmn} \mu^m \mu^n e^{kl}).\end{aligned} \tag{7.16}$$

The quantity $\mathcal{P}\mu_s^2$ being a constant can be dropped without loss of generality.

In the following paragraphs, we present three commonly encountered material structures likely of applications.

8. LINEAR ISOTROPIC MATERIALS

For *fully isotropic materials*, we have (7.7) and (7.8) valid.

Hence,

$$\begin{aligned}\Sigma = {} & d^{kl} e_{kl} + \tfrac{1}{2} d^{klmn} e_{kl} e_{mn} + \tfrac{1}{2} a^{klmn} \mu_{l;k} \mu_{n;m} \\ & + \tfrac{1}{2} b^{mn} \mu_m \mu_n + \tfrac{1}{2} \lambda^{klmn} e_{kl} \mu_m \mu_n.\end{aligned} \tag{8.1}$$

The tensorial coefficients have their isotropic values:

$$\begin{aligned} & d_{kl} = d\delta_{kl}, \qquad b_{mn} = b\delta_{mn}, \\ & d_{klmn} = \lambda \delta_{kl}\delta_{mn} + (\mu + \mathcal{K})\delta_{km}\delta_{ln} + \mu\delta_{kn}\delta_{lm}, \\ & a_{lkmn} = \sigma\delta_{lk}\delta_{mn} + \nu\delta_{lm}\delta_{kn} + \bar{\omega}\delta_{ln}\delta_{km}, \\ & \lambda_{klmn} = \alpha\delta_{kl}\delta_{mn} + \beta\delta_{km}\delta_{ln} + \gamma\delta_{kn}\delta_{lm}.\end{aligned} \tag{8.2}$$

For nonsaturated media, neglecting terms of order higher than first for e_{kl}, we obtain the constitutive equations

$$\begin{aligned} t_{kl} = {} & d\delta_{kl} + \lambda e^m{}_m \delta_{kl} + (2\mu + \mathcal{K}) e_{kl} + (\alpha/2)\mu^2 \delta_{kl} \\ & + [b + \tfrac{1}{2}(\gamma + \beta)]\mu_k \mu_l, \\ {}_L B_k = {} & -(b/\rho)\mu_k, \\ t^{(\mu)}{}_{kl} = {} & \sigma \mu^m{}_{;m} \delta_{kl} + \nu \mu_{k;l} + \bar{\omega} \mu_{l;k}. \end{aligned} \tag{8.3}$$

Using (8.1) and (8.2) and assuming the existence of a stress free state ($d \equiv 0$), the strain energy can be written in a form similar to that of Eq. (21.2) of Eringen[20] for micropolar elasticity, i.e.,

$$\begin{aligned}\Sigma = {} & \tfrac{1}{2}[\lambda(\mathrm{tr}\,\mathbf{e})^2 + (2\mu + \mathcal{K})\mathrm{tr}\,\mathbf{e}^2] + \tfrac{1}{2} b\mu^2 \\ & + \tfrac{1}{2}\{\tfrac{1}{3}(3\sigma + \nu + \bar{\omega})(\mathrm{tr}\,\mathfrak{M})^2 + (\nu - \bar{\omega})\,\mathrm{tr}(\{\mathfrak{M}\} : \{\mathfrak{M}\}^T) \\ & + (\nu + \bar{\omega})\mathrm{tr}[{}_d(\mathfrak{M}) : {}_d(\mathfrak{M})^T]\} \\ & + \tfrac{1}{2}\{\tfrac{1}{3}(3\alpha + \beta + \gamma)(\mathrm{tr}\,\mathbf{e})\mu^2 \\ & + (\beta + \gamma)\mathrm{tr}[{}_d\mathbf{e} : {}_d(\boldsymbol{\mu} \otimes \boldsymbol{\mu})]\},\end{aligned} \tag{8.4}$$

where we have introduced the definitions

$$\mathfrak{M}_{kl} \equiv \mu_{k;l} \approx \mathbf{F}_{(\mu)kK} \delta^K{}_l, \qquad \mathfrak{M} = \{\mathfrak{M}\} + (\mathfrak{M}), \tag{8.5}$$

with

$$\mathrm{tr}(\mathfrak{M} : \mathfrak{M}^T) = \mu_{l;m}\mu^{l;m}, \qquad \mathrm{tr}\,\mathfrak{M}^2 = \mu_{l;m}\mu^{m;l}, \tag{8.6}$$

and the deviators according to the relations

$${}_d(\mathfrak{M}) = (\mathfrak{M}) - \tfrac{1}{3}(\mathrm{tr}\,\mathfrak{M})\mathbf{I}, \qquad {}_d(\boldsymbol{\mu} \otimes \boldsymbol{\mu}) = \boldsymbol{\mu} \otimes \boldsymbol{\mu} - \tfrac{1}{3}\mu^2\mathbf{I}. \tag{8.7}$$

From classical elasticity, it is known that

$$3\lambda + 2\mu + \mathcal{K} \geq 0, \qquad 2\mu + \mathcal{K} \geq 0, \qquad \mathcal{K} \geq 0 \tag{8.8}$$

are *sufficient* for the first term in (8.4) to be nonnegative.

Similarly the second and third terms of (8.4) are nonnegative if

$$b \geq 0, \qquad 3\sigma + \nu + \bar{\omega} \geq 0, \qquad \bar{\omega} + \nu \geq 0 \geq \bar{\omega} - \nu. \tag{8.9}$$

It is easily verified that the third term of (8.4) can never be made nonnegative for all independent variations of $\mathbf{e}$ and $\boldsymbol{\mu}$. Thus,

$$\alpha = \beta = \gamma = 0. \tag{8.10}$$

There is no magnetostrictive effect in isotropic magnetized materials. Using a method similar to that used by Eringen[20], one may prove that (8.8), (8.9), and (8.10) are also *necessary* conditions. The only effect of the magnetization field upon the stress is due to the "magnetic anisotropy" coefficient b. Therefore, Eq. (8.3) can be written

$$\begin{aligned} t_{kl} = {} & d\delta_{kl} + \lambda e^m{}_m \delta_{kl} + (2\mu + \mathcal{K}) e_{kl} + b\mu_k\mu_l, \\ {}_L B_k = {} & -(b/\rho)\mu_k, \\ t^{(\mu)}{}_{kl} = {} & \sigma\mu^m{}_{;m}\delta_{kl} + \nu\mu_{k;l} + \bar{\omega}\mu_{l;k}. \end{aligned} \tag{8.11}$$

In (8.11) we can set $d = -p$ when the material is incompressible, where p is the *mechanical* pressure, an unknown to be determined upon solving a peculiar problem with ad hoc boundary conditions.

For *saturation*, the strain energy function may be dependent on the direction of μ but not on its magnitude. Hence the term $\frac{1}{2}b\mu^2$ of (8.4) may be ignored and we obtain from (7.3) the constitutive equations for *incompressible isotropic linear elastic magnetically saturated solids:*

$$\begin{aligned} t_{kl} &= (-p + \lambda e^m{}_m)\delta_{kl} + (2\mu + \mathcal{K})e_{kl}, \\ {}_L B_k &= 2(\mathcal{P}/\rho)\mu_k, \\ t^{(\mu)}{}_{kl} &= 2\sigma\mu^m{}_{;m}\delta_{kl} + 2(\nu + \bar{\omega})\mu_{(l;k)}, \end{aligned} \tag{8.12}$$

in which $\mathcal{P}$ and p are two unknowns ($\mathcal{P}$ can be set equal to zero without loss of generality, given the form of the "spin rotation" equation). Only four material coefficients are necessary to describe the behavior of this material.

We see that numerous effects disappear for an isotropic body. There are neither magnetostrictive effects nor magnetic anisotropic effects. The rotation of the spin is solely due to the Maxwell's magnetic field and the interaction of neighboring spins through the gradients of μ. It is then certainly more instructive to consider materials with less degrees of symmetry, for instance transversely isotropic materials or centrosymmetric cubic materials of which the latter is commonly encountered in micromagnetism theory.

9. LINEAR TRANSVERSELY ISOTROPIC MATERIALS

Here we are not concerned with the establishment of constitutive equations for transversely isotropic materials in the frame of the nonlinear theory. The result would certainly be rich in effects but the treatment would require finding a minimal function basis which is unfortunately not yet at our disposal. The purely elastic case has been studied by Ericksen and Rivlin.[21] For magnetized materials, the question of single valuedness would be quite intractable. We thus restrict the present work to the approximate theory of infinitesimal deformations.

These materials are not centrosymmetric, the hypotheses (7.8) are therefore valid but not (7.7). Hence we take $j_{klm} \neq 0$ and write the strain energy function in the form

$$\begin{aligned} \Sigma = {} & d^{kl}e_{kl} + \tfrac{1}{2}d^{klmn}e_{kl}e_{mn} + \tfrac{1}{2}\lambda^{klmn}e_{kl}\mu_m\mu_n \\ & + \tfrac{1}{2}b^{mn}\mu_m\mu_n + \tfrac{1}{2}a^{klmn}\mu_{l;k}\mu_{n;m} + j^{klm}\mu_k\mu_{m;l}. \end{aligned} \tag{9.1}$$

We call $\mathbf{h}$ the preferred direction with $\mathbf{h}$ parallel to μ in the reference configuration. $\mathbf{h}$ is normalized, thus

$$h^k h_k = 1, \tag{9.2}$$

$$\mu_k = |\mu| h_k \text{ in the reference configuration } (t = t_0). \tag{9.3}$$

The tensorial coefficients appearing in (9.1) must be transversely isotropic tensors. These tensors may be expressed as linear combinations of outer products of h_i and δ_{ij} (see Smith and Rivlin[22]), i.e.,

$$\begin{aligned} & d_{kl} = d_0\delta_{kl} + d_1 h_k h_l, \qquad b_{kn} = b_0\delta_{kn} + b_1 h_k h_n, \\ & j_{klm} = j_1\delta_{kl}h_m + j_2\delta_{km}h_l + j_3\delta_{lm}h_k + j_4 h_k h_l h_m, \\ & d_{klmn} = \alpha_0\delta_{kl}\delta_{mn} + \alpha_1\delta_{km}\delta_{ln} \\ & \quad + \alpha_2\delta_{kn}\delta_{lm} + \alpha_3\delta_{mn}h_k h_l \\ & \quad + \alpha_4\delta_{kl}h_m h_n + \alpha_5\delta_{km}h_l h_n + \alpha_6\delta_{ln}h_k h_m \\ & \quad + \alpha_7\delta_{lm}h_k h_n + \alpha_8\delta_{kn}h_l h_m + \alpha_9 h_k h_l h_m h_n. \end{aligned} \tag{9.4}$$

Similar expressions can be written for a_{klmn} and λ_{klmn} with the scalar coefficients β_i and γ_i, $i = 0, 1, \cdots, 9$, replacing α_i. In general, the coefficients $d_i, b_i, j_i, \alpha_i, \beta_i$ and γ_i may depend on the temperature θ and the magnitude of μ. For the sake of simplicity, however, the material coefficient will be considered as pure constant.

We introduce the following notations:

$$\begin{aligned} & e \equiv e^m{}_m, \quad \mathcal{E} \equiv h_m h_n e^{mn}, \quad \mathcal{K} \equiv e^{mn}\mu_m h_n, \\ & \nabla\cdot\mu = \mu^k{}_{;k}, \quad \mathcal{R} \equiv h_m h_n \mu^{m;n}, \quad \mathcal{U} \equiv h_m\mu^m, \\ & \mathcal{V} \equiv h^m h^n \mu_m \mu_n = \mathcal{U}^2. \end{aligned} \tag{9.5}$$

Using the notations (8.5)–(8.6) and assuming the existence of a stress free state ($d_0 = d_1 = 0$), we can write (9.1) in the form

$$\begin{aligned} \Sigma = {} & \tfrac{1}{2}b_0\mu^2 + \tfrac{1}{2}b_1\mathcal{U}^2 + \tfrac{1}{2}[\bar{\alpha}_0 e^2 + \bar{\alpha}_1 \operatorname{tr}\mathbf{e}^2 + \bar{\alpha}_2 e\mathcal{E} \\ & + \bar{\alpha}_3 \operatorname{tr}(\mathbf{h}\otimes\mathbf{h}:\mathbf{e}^2) + \bar{\alpha}_4\mathcal{E}^2] + j_1 \operatorname{tr}(\mu\otimes\mathbf{h}:\mathfrak{M}) \\ & + j_2 \operatorname{tr}(\mathfrak{M}:\mathbf{h}\otimes\mu) + j_3\mathcal{U}\nabla\cdot\mu + j_4\mathcal{U}\operatorname{tr}(\mathbf{h}\otimes\mathbf{h}:\mathfrak{M}) \\ & + \tfrac{1}{2}[\beta_0\nabla\cdot\mu + \beta_1 \operatorname{tr}(\mathfrak{M}\mathfrak{M}^T) + \beta_2 \operatorname{tr}\mathfrak{M}^2 + \bar{\beta}_3\mathcal{R}\nabla\cdot\mu \\ & + \beta_5 \operatorname{tr}(\mathfrak{M}\mathfrak{M}^T:\mathbf{h}\otimes\mathbf{h}) + \beta_6 \operatorname{tr}(\mathfrak{M}^T\mathfrak{M}:\mathbf{h}\otimes\mathbf{h}) \\ & + \beta_7 \operatorname{tr}(\mathfrak{M}:\mathbf{h}\otimes\mathbf{h}:\mathfrak{M}^T) + \beta_8 \operatorname{tr}(\mathfrak{M}^2:\mathbf{h}\otimes\mathbf{h}) + \beta_9\mathcal{R}^2] \\ & + \tfrac{1}{2}[\bar{\gamma}_0 e\mu^2 + \bar{\gamma}_1 \operatorname{tr}(\mathbf{e}:\mu\otimes\mu) + \bar{\gamma}_3\mathcal{E}\mu^2 \\ & + \bar{\gamma}_4\mathcal{U}^2 e + \bar{\gamma}_2 \operatorname{tr}(\mathbf{e}:\mu\otimes\mathbf{h}) + \bar{\gamma}_5\mathcal{E}\mathcal{U}^2], \end{aligned} \tag{9.6}$$

with the new material coefficients:

$$\begin{aligned} & \bar{\alpha}_0 \equiv \alpha_0, \quad \bar{\alpha}_1 \equiv \alpha_1 + \alpha_2, \quad \bar{\alpha}_2 \equiv \alpha_3 + \alpha_4, \\ & \bar{\alpha}_3 \equiv \alpha_5 + \alpha_6 + \alpha_7 + \alpha_8, \quad \bar{\alpha}_4 \equiv \alpha_9, \\ & \bar{\beta}_3 \equiv \beta_3 + \beta_4, \\ & \bar{\gamma}_0 \equiv \gamma_0, \bar{\gamma}_1 \equiv \gamma_1 + \gamma_2, \bar{\gamma}_2 \equiv \gamma_5 + \gamma_6 + \gamma_7 + \gamma_8, \\ & \bar{\gamma}_3 \equiv \gamma_3, \bar{\gamma}_4 \equiv \gamma_4, \bar{\gamma}_5 \equiv \gamma_9. \end{aligned}$$

For the *nonsaturated* case the constitutive equations result in long expressions that we shall not give here. For the saturated case, neglecting terms of order higher than that of $|e_{kl}|\cdot|\mu_m|$, we get

$$\begin{aligned} t_{kl} = {} & (\bar{\alpha}_0 e + \tfrac{1}{2}\bar{\alpha}_2\mathcal{E})\delta_{kl} + (\tfrac{1}{2}\bar{\alpha}_2 e + \bar{\alpha}_4\mathcal{E})h_k h_l \\ & + \bar{\alpha}_1 e_{kl} + \bar{\alpha}_3 e^m{}_{(l}h_{k)}h_m + \tfrac{1}{2}\bar{\gamma}_2\mu_k h_l \\ & + j_1\mu_k\mu^m{}_{;l}h_m + j_2\mu_k\mu_{l;m}h^m \\ & + j_3\nabla\cdot\mu\, h_l\mu_k + j_4\mathcal{R}\mu_k h_l, \\ {}_L B_k = {} & -(1/\rho)[(j_3\nabla\cdot\mu + j_4\mathcal{R})h_k + j_1\mu^m{}_{;k}h_m \\ & + j_2\mu_{k;m}h^m + \tfrac{1}{2}\bar{\gamma}_2 e_{km}h^m - 2\mathcal{P}\mu_k] \\ t^{(\mu)}{}_{kl} = {} & 2[\beta_0\nabla\cdot\mu + \beta_3\mathcal{R} + j_3\mathcal{U}]\delta_{kl} \\ & + 2(\bar{\beta}_3\nabla\cdot\mu + \beta_9\mathcal{R} + j_4\mathcal{U})h_k h_l \\ & + 2(j_1 + j_2)\mu_{(k}h_{l)} + 2(\beta_1 + \beta_2)\mu_{(k;l)} \\ & + 2(\beta_5 + \beta_7)h_m\mu^m{}_{;(l}h_{k)} \\ & + 2(\beta_6 + \beta_8)h_{(l}\mu_{k);m}h^m, \end{aligned} \tag{9.7}$$

where we have set

$$b_0 = b_1 = \bar{\gamma}_0 = \bar{\gamma}_1 = \bar{\gamma}_3 = \bar{\gamma}_4 = \bar{\gamma}_5 = 0 \tag{9.8}$$

since, for saturation, the strain energy (9.6) does not depend on the magnitude of μ [23] but only on its direction (equivalently, its components). Sixteen material coefficients appearing in (9.7) and listed below remain to be experimentally determined. They are

$$\bar{\alpha}_i, \quad i = 0, 1, \cdots, 4, \quad j_i, \quad i = 1, \cdots, 4,$$
$$\beta_0, \quad \beta_1 + \beta_2, \quad \bar{\beta}_3, \quad \beta_5 + \beta_7, \quad \beta_6 + \beta_8, \quad \beta_9,$$
$$\bar{\gamma}_2. \tag{9.9}$$

If we set $\mathbf{h} \equiv \mathbf{0}$, we recover Eq. (8.12) given for the isotropic case.

The following model may be more suitable for computations. We may reasonably assume that the "exchange" terms (of quantum mechanical origin) are of little importance in t_{kl} and ${}_L B_k$ but retain their importance in $t^{(\mu)}{}_{kl}$. If we set $|{}_L\mathbf{B}| = O(|\mu|)$, we can neglect terms of the kind $|\mu|\cdot|\mathbf{e}|, |\mu|\cdot|\mathbf{e}|, \cdots$. We thus write the approximate constitutive equations for *incompressible linear transversely isotropic elastic magnetically saturated solids:*

$$t_{kl} = (-p + \bar{\alpha}_0 e + \tfrac{1}{2}\bar{\alpha}_2 \mathcal{E})\delta_{kl} + (\tfrac{1}{2}\bar{\alpha}_2 e + \bar{\alpha}_4 \mathcal{E}) h_k h_l + \bar{\alpha}_1 e_{kl} + \bar{\alpha}_3 e^m{}_{(l} h_{k)} h_m + \tfrac{1}{2}\bar{\gamma}_2 \mu_k h_l,$$
$${}_L B_k = (2\mathcal{P}/\rho)\mu_k, \tag{9.10}$$
$t^{(\mu)}{}_{kl}$, given by the third of Eqs. (9.7).

In this case, the rotation of the spin is affected by the Maxwell's magnetic field and the magnetization field. Second order gradients of the magnetization will appear in the velocity of rotation of μ.

10. CENTROSYMMETRIC CUBIC MATERIALS

This paragraph is intended for applying to ferromagnetic materials of cubic structure. We refer to Sec. 4 for the general features of the invariance under magnetic point groups and to Refs. 8, 10, and 24.

As an example we consider a cubic crystal of the magnetic class $\underline{m}3\underline{m} \in m_3$, the generators of which are

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}. \tag{10.1}$$

This material is centrosymmetric; Eqs. (7.7) and (7.8) are therefore valid. Thus we consider the following expansion of the strain energy function:

$$\Sigma = d^{kl} e_{kl} + \tfrac{1}{2} d^{klmn} e_{kl} e_{mn} + \tfrac{1}{2}\lambda^{klmn} e_{kl}\mu_m\mu_n + \tfrac{1}{2} a^{klmn}\mu_{l;k}\mu_{n;m} + \tfrac{1}{2} b^{mn}\mu_m\mu_n. \tag{10.2}$$

We consider only the *saturated* case which is of importance for ferromagnetic materials. Neglecting terms of order higher than that of μ^2 in t_{kl} and $|\mu|$ in ${}_L\mathbf{B}$, we get the approximate constitutive equations

$$t_{kl} = d_{kl} + d_{klmn} e^{mn} + b_{ml}\mu^m\mu_k + \tfrac{1}{2}\lambda_{klmn}\mu^m\mu^n,$$
$${}_L B_k = -(1/\rho)(b_{mk}\mu^m - 2\mathcal{P}\mu_k), \tag{10.3}$$
$$t^{(\mu)}{}_{kl} = 2a_{(lk)mn}\mu^{n;m}.$$

For the considered symmetry, the material coefficients appearing in (10.2) assume the form (see Sirotin[25])

$$b^{mn} = b_0\delta^{mn}, \quad d^{kl} = d_0\delta^{kl},$$
$$d^{klmn} = d\delta^{klmn} + d_{12}\delta^{kl}\delta^{mn} + d_{44}(\delta^{km}\delta^{ln}\delta^{kn}\delta^{lm}),$$
$$\lambda^{klmn} = \lambda\delta^{klmn} + \lambda_{12}\delta^{kl}\delta^{mn} + \lambda_{44}(\delta^{km}\delta^{ln} + \delta^{kn}\delta^{lm}),$$
$$a^{(lk)mn} = a\delta^{lkmn} + a_{12}\delta^{lk}\delta^{mn} + a_{44}(\delta^{lm}\delta^{kn} + \delta^{ln}\delta^{km}), \tag{10.4}$$

in which the symbol δ^{klmn} is equal to one if $k = l = m = n$ and zero otherwise. Upon carrying (10.4) into (10.2) and assuming the existence of a stress free state ($d_0 \equiv 0$), we can write the strain energy function in the form

$$\Sigma = \tfrac{1}{2} b_0\mu_s^2 + \tfrac{1}{2}\left[d\sum_i (e_{\underline{ii}})^2 + d_{12}e^2 + 2d_{44}\,\mathrm{tr}\,\mathbf{e}^2\right] + \tfrac{1}{2}\left[\lambda\sum_i (e_{\underline{ii}}\mu_{\underline{i}}\mu_{\underline{i}}) + \lambda_{12}e\mu_s^2 + 2\lambda_{44}\,\mathrm{tr}(\dot{e}:\mu\otimes\mu)\right] + \tfrac{1}{2}\left[a\sum_i (\mu_{\underline{i},\underline{i}})^2 + a_{12}(\mathrm{tr}\,\mathfrak{M})^2 + a_{44}(\mathrm{tr}\,\mathfrak{M}^2 + \mathrm{tr}\,\mathfrak{M}\,\mathfrak{M}^T)\right], \tag{10.5}$$

where we have made use of the notations (8.5)–(8.6). This equation is written for rectangular coordinates. For instance, we have

$$\sum_i (e_{\underline{ii}})^2 = e_{11}{}^2 + e_{22}{}^2 + e_{33}{}^2,$$
$$\sum_i (e_{\underline{ii}}\mu_{\underline{i}}\mu_{\underline{i}}) = e_{11}\mu_1^2 + e_{22}\mu_2^2 + e_{33}\mu_3^2. \tag{10.6}$$

Since for saturation Σ does not depend on the magnitude of μ we set

$$b_0 = \lambda_{12} = 0, \tag{10.7}$$

and we are left with 8 material coefficients, namely $d, d_{12}, d_{44}, \lambda, \lambda_{44}, a, a_{12}$ and a_{44}. For instance, in rectangular coordinates, Eq. (10.3) yields for an *incompressible* solid:

$$t_{11} = -p + d_{12}e + (d + 2d_{44})e_{11} + (\lambda/2 + \lambda_{44})\mu_1{}^2,$$
$$t_{12} = 2d_{44}e_{12} + \lambda_{44}\mu_1\mu_2,$$
$${}_L B_i = (2\mathcal{P}/\rho)\mu_i, \tag{10.8}$$
$$t^{(\mu)}{}_{11} = 2(a + a_{44} + a_{12})\mu_{1,1} + 2a_{12}(\mu_{2,2} + \mu_{3,3}),$$
$$t^{(\mu)}{}_{12} = a_{44}(\mu_{1,2} + \mu_{2,1}),$$
etc.,

while the term which represents the magnetostrictive energy in (10.5) is

$$\Sigma_{\text{m.s.e.}} = (\lambda_{44} + \lambda/2)(e_{11}\mu_1{}^2 + e_{22}\mu_2{}^2 + e_{33}\mu_3{}^2) + 2\lambda_{44}(e_{12}\mu_1\mu_2 + e_{23}\mu_2\mu_3 + e_{31}\mu_3\mu_1). \tag{10.9}$$

This form is similar to that of Landau and Lifshitz.[26] It shows that only two material coefficients are needed to describe the magnetostrictive effect for a centrosymmetric cubic crystal (cf. Ref. 19). Finally, we note that, as in the case of fully isotropic materials, the local field ${}_L\mathbf{B}$ is of no importance for such a structured medium. In fact, it is known that the mag-

netic anisotropy energy is of fourth order in cubic crystal[26] and therefore is relatively small. In the present work, its effects cannot be felt since the expansion of Σ has been cut at the third order.

This last class of materials is particularly important for it is in the realm of ferromagnetic materials that the saturated media have been studied experimentally for some decades[19,27]. The theoretical considerations are shown to possess physical reality. It is therefore expected that, in future articles, motions of special interest and special effects predicted by the theory will be examined more accurately for both exact nonlinear and approximate linear theories.

11. PROSPECTS

The classical (three-dimensional) theory of magnetized deformable materials presented in Part I and herein provides an insight into the *continuum* behavior of the interaction between deformable matter and the magnetization field. It takes into account in a macroscopical way two features of quantum mechanical origin: a repartition of electronic spins throughout the body and the interaction of neighboring spins. If we disregard these average effects, the theory still gives a basis for the nonlinear treatment of the magnetostrictive effect. The theory has been developed in the frame of quasimagnetostatics.

Several likely applications of the theory exist. We mention three such classes below:

(i) Nonlinear hemitropic and isotropic magnetically saturated media: The study of nonlinear deformations due to high intensity of applied magnetic fields can be made, thus casting light onto rich nonlinear effects. For instance, one may ask the legitimate questions: Is there any Kelvin-type effect? Is there a possibility to study universal motions in order to give a strong theoretical support to the laboratory determination of the material coefficients? All answers are certainly beyond the scope of the present articles.

(ii) Ferromagnetic media: The theory seems adequate to tackle on the one hand piezomagnetism and magnetostriction and, on the other, a phenomenological approach to the problem of *magnetic* domains. We have provided the necessary tools for a study of centrosymmetric cubic materials in the frame of an approximate theory. For example, it would be of interest to study the structure[28] of a wall separating two magnetic domains and the velocity of propagation of magnetic walls in the light of the present development.

(iii) Rubber bonded barium ferrite composite materials: We may reasonably assume that the theory is fitted for studying such materials. The "spin rotation" equation obtained here above would prove capable of describing the dynamical behavior of the small magnets embedded in the matrix of rubber-like material, upon action of an applied magnetic field. Although only approximate constitutive equations are given, it is expected that fully nonlinear ones will be constructed once minimal function bases for transverse isotropy are established. The study of large deformations such as bending or torsion superposed to an applied magnetic field of given direction should lead to interesting results.

Finally, we note that some categories of problems can be examined such as that of the superposition of infinitesimal deformations to a finite deformation field and the superposition of a small dynamical magnetic field upon a large static magnetic field.

The limitations of the theory are quite obvious: Only quasimagnetostatics and nondissipative processes (except for Sec. 8 in Part I) have been considered, thus excluding the electric field and the polarization and their related effects e.g., magnetoelectropolarizability, piezoelectricity···and the *strong* dynamical behaviors, e.g., waves. Furthermore, we note that the inclusion of the notions of temperature, conductivity, and resistance would have led to numerous other effects.

To remedy some of these limitations, we give in a forthcoming article a special relativistic theory of nonlinear elastic elastic solids that exhibit a repartition of electronic spins, an electromagnetic field acting upon the body. Current, charges, electric and polarization fields are no longer ignored and the treatment is fully dynamical due to the synthesis of space and time as one entity. It is thus expected to give a sound basis for the approximate theory developed above.

We close the present article by noting three prospects of interest:

(a) The theory developed above is concerned with a classical mechanical behavior of materials. Mechanical couple stresses and micromorphic[29] (more restrictly micropolar) deformation fields have been excluded. It seems that some magnetized materials may possess a plastic behavior, hence the possible introduction of "mechanical directors" (not to be confused with the "magnetization director" or spin introduced above). Other materials such as the aforementioned rubber-like material are indeed composites and therefore manifest a more involved mechanical behavior. More naturally, certain liquid crystals are known to possess both microscopic orientations and magnetic dipole moments. A synthesis of the present work and of the concept of micropolar medium should provide still further rich grounds for the exploration of new physical phenomena.

(b) The considered materials have been selected among "*simple*" materials. The treatment of more involved materials such as those gifted of hereditary characteristics, requires the study of functional constitutive equations. Approaches "à la" Eringen[30] or "à la" Noll[31] if the statement of balance laws constitutes the starting point, or "à la" Edelen[32] if use is made of a *non-local variational principle* to start with, may be envisaged.

(c) Finally we mention the attempt of Eringen and Kafadar[33] to develop a micromorphic theory of magnetism in matter. It is expected to explain micromagnetism phenomena, but it seems to be a task to define the energetic concepts without which the construction of a constitutive theory (or of a variational principle) remains incomplete. With the construction of the energetic concepts it will be possible to predict the behavior of materials exhibiting extra degrees of freedom of mechanical and electromagnetical origins. This theory would give an insight into the behavior of matter at an intermediate level in the midst of two realms, microphysics and continuum physics, though the formalism of the latter would be used throughout.

* This work was supported by the National Science Foundation.

1 G. A. Maugin and A. C. Eringen, J. Math. Phys. **13**, 143 (1972).
2 N. F. Jordan and A. C. Eringen, Intern. J. Eng. Sci. **2**, 59 (1964).
3 R. C. Dixon and A. C. Eringen, Intern. J. Eng. Sci. **3**, 359 (1965).
4 R. A. Grot and A. C. Eringen, Intern. J. Eng. Sci. **4**, 611 (1966).
5 A. C. Eringen, *Mechanics of Continua* (J. Wiley, New York, 1967), p. 149.
6 C. Truesdell and W. Noll, *Hanbuch der Physik*. Vol. 3/3, edited by S. Flügge (Springer-Verlag, Berlin, 1965), p. 52.
7 H. Weyl, *The Classical Groups* (Princeton U. P., New Jersey, 1946), Chap. 2, Art. 9.
8 R. R. Birss, Rept. Progr. Phys. **26**, 307 (1963).
9 I. S. Zheludev, Kristallografiya **5**, 376 (1960) [Sov. Phys. Crystallography **5**, 328 (1960)].
10 S. Bhagavantam, *Crystal Symmetry and Physical Properties* (Academic, New York, 1966), p. 78; see also Ref. 8.
11 A. S. Borovik-Romanov, Zh. Eksp. Teor. Fiz. **36**, 1954 (1959) [Sov. Phys. JETP **9**, 1390 (1959)].
12 G. F. Smith, Arch. Ratl. Mech. Anal. **18**, 282 (1965).
13 C. B. Kafadar and A. C. Eringen, Intern. J. Eng. Sci. **9**, 271 (1971).
14 A. J. M. Spencer, in *Continuum Physics*, Vol. I, edited by A. C. Eringen (Academic, New York, 1971), Chap. 3.
15 C. F. Smith, (a) Intern. J. Eng. Sci. **8**, 1 (1970); (b) Arch. Ratl. Mech. Anal. **36**, 161 (1970); C. C. Wang, Arch. Ratl. Mech. Anal. **36**, 166 (1970).
16 Provided the fact that the magnetization amplitude can be considered as constant throughout the body, a reasonable hypothesis but not quantitatively known on account of the lack of informations concerning these manufactured materials.
17 The process of linearization could be carried out directly from the use of Eqs. (2.10)–(2.12) by setting

$$x^k{}_{,K} \approx g^k{}_K + U^k{}_{:K}, \quad E_{KL} \approx e_{kl} g^k{}_K g^l{}_L,$$
$$\Pi_K \approx g^l{}_K \mu_l, \quad D_{AK} \approx g^l{}_K g^m{}_A \mu_{l;m},$$

where **U** is the displacement.
18 E. S. Suhubi, Intern. J. Eng. Sci. **7**, 993 (1969).
19 N. Akulov, Z. Physik **100**, 197 (1936).
20 A. C. Eringen, "Micropolar elasticity", in *Fracture*, Vol. 2, edited by H. Liebowitz (Academic, New York, 1968), p. 621.
21 J. L. Ericksen and R. S. Rivlin, J. Ratl. Mech. Anal. **3**, 281 (1954).
22 G. F. Smith and R. S. Rivlin, Quart. J. Appl. Math. **15**, 308 (1957).
23 The coefficients of the terms containing $\mathfrak{U}$ must be set equal to zero since, with (9.3), $\mathfrak{U}$ is none other than μ^2 in the reference configuration.
24 W. P. Mason, *Crystal Physics of Interaction Processes* (Academic, New York, 1966).
25 Yu. I. Sirotin, Dokl. Akad. Nauk SSSR **133**, 321 (1960) [Sov. Phys. Dokl. **5**, 774 (1961)] Kristallografia **5**, 157 (1960); **6**, 331 (1961); [Sov. Phys. Crystallography **5**, 157 (1960); **6**, 263 (1961)].
26 Though they express this energy in terms of stresses rather than strains, cf. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon, Oxford, 1960), Sec. 38.
27 M. Bozorth, J. Appl. Phys. **8**, 575 (1937).
28 We grant to "structure" the meaning given in "shock wave structure."
29 A. C. Eringen, Intern. J. Eng. Sci. **8**, 819 (1970).
30 Cf. Ref. 5, Chap. 5.
31 W. Noll, Arch. Ratl. Mech. Anal. **2**, 197 (1958).
32 D. G. B. Edelen, Intern. J. Eng. Sci. **7**, 269, 287, 373, 391, 401, 677, 843 (1969).
33 A. C. Eringen and C. B. Kafadar, J. Math. Phys. **11**, 1984 (1970).

Procedures in Quantum Mechanics without Von Neumann's Projection Axiom*

Paul A. Benioff
Chemistry Division, Argonne National Laboratory, Argonne, Illinois
(Received 31 January 1972; Revised Manuscript Received 10 April 1972)

The results of previous work are generalized to include procedures whose measurement operations correspond to expectations as defined by Davies. For such procedures Von Neumann's projection axiom is not in general applicable. Finite and infinite sequences of measurements and transformations as well as finite and infinite decision procedures are considered. It is shown that with each such procedure there is associated in a unique manner a probability operator measure.

I. INTRODUCTION

In an earlier paper,[1] hereafter denoted as I, processes consisting of finite or infinite sequences of observations separated by transformations (such as rotations, translations, etc.) were considered. The process Q was required to be such that there exist a Hilbert space $\mathcal{H}$ such that each observation in Q corresponded to a discrete self-adjoint operator in $B(\mathcal{H})$, the algebra of all bounded linear operators on $\mathcal{H}$. Also each transformation in Q, considered as a map with domain and range in the set of all state preparation procedures, was taken to correspond to a map: $S(\mathcal{H}) \to S(\mathcal{H})$ which was implementable by an isometry in $B(\mathcal{H})$. $S(\mathcal{H})$ is the set of all states on $\mathcal{H}$. Furthermore, Von Neumann's projection axiom[2] was required to be applicable.

The main result of I was that, with each finite or infinite process Q which satisfied these requirements, there is uniquely associated a probability operator measure $O^Q : \Sigma^Q \to B(\mathcal{H})$. Σ^Q is a σ-field of Borel subsets of Ω^Q, the set of all possible outcome sequences of Q, and $B(\mathcal{H})$ is the algebra of all bounded linear operators on $\mathcal{H}$.

In another paper[3] hereafter called II, this and other results of I were extended to include finite and infinite decision procedures. That is, for each j, the choice of operations in the procedure for the jth step of any path could depend on the operations and outcomes of previous observations. The decision procedures were required to satisfy the same restrictions as were imposed in I. That is, each observation in each procedure Q corresponded to a discrete self-adjoint operator in $B(\mathcal{H})$ and each transformation was implementable by means of an isometry in $B(\mathcal{H})$. Also Q was to be such that Von Neumann's projection axiom was applicable. Also, each path p of Q was required to be such that the length of every (proper) initial segment of p is finite.

Now as Margenau has pointed out[4] there are many measurement procedures which do not satisfy Von Neumann's projection axiom. Furthermore, if one wishes to consider observables with continuous spectra without replacing them by "coarse grained" observables which are discrete, then the projection axiom fails. That is, for discrete observables the mapping $\rho \to \rho' = \Sigma_i P_{x_i} \rho P_{x_i}$, where $\{P_{x_i} | i = 1, 2 \cdots\}$ is a complete set of eigenprojectors for a discrete observable A exists. However, if A is continuous, then no mapping of the form $\rho \to \rho'$ given above exists.[5] More generally, the map $\rho \to \rho'$ as defined above is a conditional expectation as defined by Nakamura and Umegaki,[6] and these exist only for discrete observables.

For these reasons it seems worthwhile to consider procedures which contain measurements for which Von Neumann's projection axiom is not applicable. In this work, these procedures will be treated by the methods of Davies[7] and Lewis.[5] The mathematical objects one works with in these methods are expectations and their dual—instruments. The basic correspon-

* This work was supported by the National Science Foundation.

1 G. A. Maugin and A. C. Eringen, J. Math. Phys. **13**, 143 (1972).
2 N. F. Jordan and A. C. Eringen, Intern. J. Eng. Sci. **2**, 59 (1964).
3 R. C. Dixon and A. C. Eringen, Intern. J. Eng. Sci. **3**, 359 (1965).
4 R. A. Grot and A. C. Eringen, Intern. J. Eng. Sci. **4**, 611 (1966).
5 A. C. Eringen, *Mechanics of Continua* (J. Wiley, New York, 1967), p. 149.
6 C. Truesdell and W. Noll, *Hanbuch der Physik*. Vol. 3/3, edited by S. Flügge (Springer-Verlag, Berlin, 1965), p. 52.
7 H. Weyl, *The Classical Groups* (Princeton U. P., New Jersey, 1946), Chap. 2, Art. 9.
8 R. R. Birss, Rept. Progr. Phys. **26**, 307 (1963).
9 I. S. Zheludev, Kristallografiya **5**, 376 (1960) [Sov. Phys. Crystallography **5**, 328 (1960)].
10 S. Bhagavantam, *Crystal Symmetry and Physical Properties* (Academic, New York, 1966), p. 78; see also Ref. 8.
11 A. S. Borovik-Romanov, Zh. Eksp. Teor. Fiz. **36**, 1954 (1959) [Sov. Phys. JETP **9**, 1390 (1959)].
12 G. F. Smith, Arch. Ratl. Mech. Anal. **18**, 282 (1965).
13 C. B. Kafadar and A. C. Eringen, Intern. J. Eng. Sci. **9**, 271 (1971).
14 A. J. M. Spencer, in *Continuum Physics*, Vol. I, edited by A. C. Eringen (Academic, New York, 1971), Chap. 3.
15 C. F. Smith, (a) Intern. J. Eng. Sci. **8**, 1 (1970); (b) Arch. Ratl. Mech. Anal. **36**, 161 (1970); C. C. Wang, Arch. Ratl. Mech. Anal. **36**, 166 (1970).
16 Provided the fact that the magnetization amplitude can be considered as constant throughout the body, a reasonable hypothesis but not quantitatively known on account of the lack of informations concerning these manufactured materials.
17 The process of linearization could be carried out directly from the use of Eqs. (2.10)–(2.12) by setting

$$x^k{}_{,K} \approx g^k{}_K + U^k{}_{:K}, \quad E_{KL} \approx e_{kl} g^k{}_K g^l{}_L,$$
$$\Pi_K \approx g^l{}_K \mu_l, \quad D_{AK} \approx g^l{}_K g^m{}_A \mu_{l;m},$$

where **U** is the displacement.
18 E. S. Suhubi, Intern. J. Eng. Sci. **7**, 993 (1969).
19 N. Akulov, Z. Physik **100**, 197 (1936).
20 A. C. Eringen, "Micropolar elasticity", in *Fracture*, Vol. 2, edited by H. Liebowitz (Academic, New York, 1968), p. 621.
21 J. L. Ericksen and R. S. Rivlin, J. Ratl. Mech. Anal. **3**, 281 (1954).
22 G. F. Smith and R. S. Rivlin, Quart. J. Appl. Math. **15**, 308 (1957).
23 The coefficients of the terms containing $\mathfrak{U}$ must be set equal to zero since, with (9.3), $\mathfrak{U}$ is none other than μ^2 in the reference configuration.
24 W. P. Mason, *Crystal Physics of Interaction Processes* (Academic, New York, 1966).
25 Yu. I. Sirotin, Dokl. Akad. Nauk SSSR **133**, 321 (1960) [Sov. Phys. Dokl. **5**, 774 (1961)] Kristallografia **5**, 157 (1960); **6**, 331 (1961); [Sov. Phys. Crystallography **5**, 157 (1960); **6**, 263 (1961)].
26 Though they express this energy in terms of stresses rather than strains, cf. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon, Oxford, 1960), Sec. 38.
27 M. Bozorth, J. Appl. Phys. **8**, 575 (1937).
28 We grant to "structure" the meaning given in "shock wave structure."
29 A. C. Eringen, Intern. J. Eng. Sci. **8**, 819 (1970).
30 Cf. Ref. 5, Chap. 5.
31 W. Noll, Arch. Ratl. Mech. Anal. **2**, 197 (1958).
32 D. G. B. Edelen, Intern. J. Eng. Sci. **7**, 269, 287, 373, 391, 401, 677, 843 (1969).
33 A. C. Eringen and C. B. Kafadar, J. Math. Phys. **11**, 1984 (1970).

Procedures in Quantum Mechanics without Von Neumann's Projection Axiom*

Paul A. Benioff
Chemistry Division, Argonne National Laboratory, Argonne, Illinois
(Received 31 January 1972; Revised Manuscript Received 10 April 1972)

The results of previous work are generalized to include procedures whose measurement operations correspond to expectations as defined by Davies. For such procedures Von Neumann's projection axiom is not in general applicable. Finite and infinite sequences of measurements and transformations as well as finite and infinite decision procedures are considered. It is shown that with each such procedure there is associated in a unique manner a probability operator measure.

I. INTRODUCTION

In an earlier paper,[1] hereafter denoted as I, processes consisting of finite or infinite sequences of observations separated by transformations (such as rotations, translations, etc.) were considered. The process Q was required to be such that there exist a Hilbert space $\mathcal{H}$ such that each observation in Q corresponded to a discrete self-adjoint operator in $B(\mathcal{H})$, the algebra of all bounded linear operators on $\mathcal{H}$. Also each transformation in Q, considered as a map with domain and range in the set of all state preparation procedures, was taken to correspond to a map: $S(\mathcal{H}) \to S(\mathcal{H})$ which was implementable by an isometry in $B(\mathcal{H})$. $S(\mathcal{H})$ is the set of all states on $\mathcal{H}$. Furthermore, Von Neumann's projection axiom[2] was required to be applicable.

The main result of I was that, with each finite or infinite process Q which satisfied these requirements, there is uniquely associated a probability operator measure $O^Q : \Sigma^Q \to B(\mathcal{H})$. Σ^Q is a σ-field of Borel subsets of Ω^Q, the set of all possible outcome sequences of Q, and $B(\mathcal{H})$ is the algebra of all bounded linear operators on $\mathcal{H}$.

In another paper[3] hereafter called II, this and other results of I were extended to include finite and infinite decision procedures. That is, for each j, the choice of operations in the procedure for the jth step of any path could depend on the operations and outcomes of previous observations. The decision procedures were required to satisfy the same restrictions as were imposed in I. That is, each observation in each procedure Q corresponded to a discrete self-adjoint operator in $B(\mathcal{H})$ and each transformation was implementable by means of an isometry in $B(\mathcal{H})$. Also Q was to be such that Von Neumann's projection axiom was applicable. Also, each path p of Q was required to be such that the length of every (proper) initial segment of p is finite.

Now as Margenau has pointed out[4] there are many measurement procedures which do not satisfy Von Neumann's projection axiom. Furthermore, if one wishes to consider observables with continuous spectra without replacing them by "coarse grained" observables which are discrete, then the projection axiom fails. That is, for discrete observables the mapping $\rho \to \rho' = \Sigma_i P_{x_i} \rho P_{x_i}$, where $\{P_{x_i} | i = 1, 2 \cdots\}$ is a complete set of eigenprojectors for a discrete observable A exists. However, if A is continuous, then no mapping of the form $\rho \to \rho'$ given above exists.[5] More generally, the map $\rho \to \rho'$ as defined above is a conditional expectation as defined by Nakamura and Umegaki,[6] and these exist only for discrete observables.

For these reasons it seems worthwhile to consider procedures which contain measurements for which Von Neumann's projection axiom is not applicable. In this work, these procedures will be treated by the methods of Davies[7] and Lewis.[5] The mathematical objects one works with in these methods are expectations and their dual—instruments. The basic correspon-

dence assumption is that each procedure Q must be such that each measurement operation in Q corresponds to an expectation (or an instrument). However, before working with procedures, we must first define expectations and instruments and give some easy properties. This is done in Sec. II. It is noted, among other things, that the use of expectations and instruments requires a generalization of the definition of an observable to be a probability operator measure rather than the more specialized spectral measure.

In Sec. III sequences of measurements are considered in which each measurement corresponds to an expectation. (Rather than carry both expectations and instruments through this work, which would be unnecessary, we will work almost exclusively with expectations.) Furthermore the value space need not be discrete. Thus in this section the requirement, present in I[1] and II,[3] that the observables be discrete is dropped.

The main result obtained is that with each finite or infinite sequence Q of measurements there is uniquely associated an observable (or probability operator measure) O^Q. In the case of finite sequences Q, there is also a unique expectation associated with the process, and the observable O^Q is merely a part of the expectation.

Section IV extends this result to finite and infinite decision procedures. Here the requirement that the value space of each measurement in Q be discrete is reimposed. The reason is that there are some mathematical problems in the nondiscrete case related to how one expresses the dependence of the choice of the expectation in step j in a path on the prior outcomes. In any case it is doubtful if procedures with uncountably many lines leading out of a vertex of the associated tree need be considered.

The Appendix gives some mathematical results necessary to carry out the constructions of this paper. The Kolmogorov[8] extension theorem, extended in I to include probability operator measures on Cartesian products of the real line, is extended further here to Cartesian products of arbitrary, complete separable metric spaces. Davies and Lewis[7] theorem on the combination of instruments is given for the combination of expectations. It is noted that these theorems can be shown to apply also to σ-compact, locally compact Hausdorff spaces.

II. EXPECTATIONS AND INSTRUMENTS

Expectations are defined as follows: Let X be a set (of outcomes of a measurement procedures) and $\mathfrak{B}(X)$ a set of Borel subsets of X and $B(\mathcal{H})$ the algebra of bounded linear operators on some Hilbert space $\mathcal{H}$. An expectation $\mathcal{E}$ is a map, $\mathcal{E}: \mathfrak{B}(X) \times B(\mathcal{H}) \to B(\mathcal{H})$ such that [$\mathbf{0}$ and $\mathbf{1}$ are the zero and identity operators of $B(\mathcal{H})$]:

(a) If $B \geq \mathbf{0}$ then $\mathcal{E}(E, B) \geq \mathbf{0}$ for each $E \in \mathfrak{B}(X)$.
(b) $\mathcal{E}(X, \mathbf{1}) = \mathbf{1}$.
(c) For each $B \in B(\mathcal{H})$, $\mathcal{E}(-, B)$ is strongly countably additive.
(d) For each complex number α, set E, and operator B, $\mathcal{E}(E, \alpha B) = \alpha \mathcal{E}(E, B)$.
(e) $\mathcal{E}(E, B + B') = \mathcal{E}(E, B) + \mathcal{E}(E, B')$.
(f) Let A_α be a monotone net of operators in $B(\mathcal{H})$ with $A = \text{s-lim}_\alpha A_\alpha$. Then for each E, $\mathcal{E}(E, A_\alpha)$ is a monotone net with $\mathcal{E}(E, A) = \text{s-lim}_\alpha \mathcal{E}(E, A_\alpha)$.

This definition is the same as that given by Davies.[7] The replacement of weak by strong convergence in properties (c) and (f) has no effect since weak and strong countable additivity are equivalent,[9] and in (f) weak and strong convergence are equivalent.[10]

A simple example of an expectation is given by

$$\mathcal{E}(E, B) = \sum_{x \in E} P_x B P_x \tag{1}$$

for each $E \subset X$ with X countable and B in $B(\mathcal{H})$. Equation (1) gives the well-known form of the expectation associated with any measurement procedure whose outcome set X is at most countably infinite and which satisfies Von Neumann's projection axiom.

Let S be the set of states on $\mathcal{H}$. Then an instrument is defined[5,7] to be a map $\mathcal{J}: \mathfrak{B}(X) \times S \to S$ such that (i) $\mathcal{J}(E, \rho) \geq \mathbf{0}$ if $\rho \geq \mathbf{0}$ for all $\rho \in S$; (ii) $\mathcal{J}(-, \rho)$ is strongly σ-additive; and (iii) $\mathrm{Tr}\mathcal{J}(X, \rho) = \mathrm{Tr}\rho$ for all $\rho \in S$.

Basically an expectation can be regarded as a conditioning on operators[5] just as an instrument is a conditioning on states.[5] That is, let $\mathcal{E}_M$ and $\mathcal{J}_M$ be the respective expectations and instrument associated with some measurement procedure M. Then $\mathcal{E}_M(E, B)$ is the conditioning of B which corresponds to the prior carrying out of M and finding the outcome in E. Similarly $\mathcal{J}_M(E, \rho)$ is the conditioning of ρ which corresponds to the carrying out M on a system in state ρ and finding the outcome in E.

The duality of $\mathcal{J}_M$ and $\mathcal{E}_M$ is evident from the fact that

$$\mathrm{Tr}(\rho \mathcal{E}_M(E, B)) = \mathrm{Tr}(\mathcal{J}_M(E, \rho) B) \tag{2}$$

for each ρ in S, E in $\mathfrak{B}(X)$ and B in $B(\mathcal{H})$. If M corresponds to a discrete observable and Von Neumann's projection axiom holds, then $\mathcal{J}_M$ is given by

$$\mathcal{J}_M(E, \rho) = \sum_{x \in E} P_x \rho P_x \tag{3}$$

for each $E \subset X$ and $\rho \in S$ and $\mathcal{E}_M$ is given by Eq. (1).

The association of expectations and instruments to measurement procedures also results in the generalization of the concept of an observable. In the usual quantum mechanics, a bounded observable corresponds to a self-adjoint operator in $B(\mathcal{H})$ or, equivalently, through the spectral theorem, to a spectral measure on $\mathfrak{B}(R)$, the set of Borel subsets of the real line. Here an observable corresponds[5,7] to a probability operator measure on $\mathfrak{B}(X)$. (A spectral measure is a probability operator measure whose range set is a set of mutually commuting projection operators. In general, the operators in the range set of a probability operator measure do not have to be projection operators nor do they have to commute.) Here, if M is a measurement procedure with associated expectation $\mathcal{E}_M$, then $\mathcal{E}_M(-, \mathbf{1})$ is the observable which M measures.[5]

Although the use of expectations and instruments and the consequent generalization of the definition of observables may appear counter intuitive, they do have some important advantages. One is that there are potentially at least, many measurement procedures

which do not satisfy the projection axiom, but can be described by expectations and observables in the general sense of Davies and Lewis.[5] Of course, there may be many measurement procedures which cannot be associated with expectations either. However, the increased generalization given by expectations and instruments does bring a potentially larger class of measurement procedures into the domain of the theory.

Another advantage is that, as will be seen, the compounding of measurement procedures[7] each described by an expectation (and observable) gives a procedure which is described by an expectation. Thus an observable is associated with the compound procedures, and it also corresponds to a measurement procedure. This does not hold in the usual quantum mechanical interpretation in which Von Neumann's projection axiom holds. There the compound procedure of measuring observable $A_1 = \sum_{y\in Y} yP_y$ followed by a measurement of observable $A_2 = \sum_{z\in Z} zP_z$, where A_1 and A_2 do not commute, has associated with it a map $\mathcal{E}(-, 1)$ defined by

$$\mathcal{E}(E, \mathbf{1}) = \sum_{\langle j, z\rangle \in E} P_y P_z P_y .$$

For each $E\subset Y \times Z$. But $\mathcal{E}(-, \mathbf{1})$ so defined, is not a spectral measure and thus the compound procedure does not correspond to any observable under the usual interpretation.

III. SEQUENCES OF MEASUREMENTS

In this and the next sections only those sequences Q of measurements will be considered which are such that each measurement operation in Q corresponds to exactly one expectation. Any Q for which this is true will be said to satisfy the *correspondence assumption* as there corresponds to each such Q a unique sequence of expectations. Clearly this is a weaker restriction than was used in I and II where the projection axiom was required to hold.

Also, in this and the next section, the simplifying restriction to processes with no transformations separating the measurements will be made. This is an inessential simplification made simply to conserve on notation. It will be shown later that the results obtained extend easily to processes with transformations separating the measurements.

Let Q denote a process consisting of an infinite sequence of measurement operations and $\{\mathcal{E}_j^Q \mid j = 0, 1, \cdots\}$ the corresponding infinite sequence of expectations. That is, for each j, under the correspondence assumption, $\mathcal{E}_j^Q$ is the unique expectation associated with the measurement operation $Q(j)$. Q is also assumed to be such that for each j, the outcome space X_j is a complete separable metric space. This is a weak restriction since it includes the real line R or any countable set such as the integers with discrete topology.[11] Also, if each X_j is a complete separable metric space, so is the Cartesian product $\otimes_{j=0}^{n-1} X_j$ for each n.[12]

Now for each n, let Q_n denote the first n steps of Q, $\{\mathcal{E}_j^Q \mid j = 0, 1, \cdots, n-1\}$ the corresponding sequence of expectations $X^n = \otimes_{j=0}^{n-1} X_j$ the Cartesian product of the outcome spaces and $\mathfrak{B}(X^n)$ the Borel system of subsets of X^n. By Theorem 2 in the Appendix there is associated with Q_n a unique expectation $\mathcal{E}^{Q_n}$: $\mathfrak{B}(X^n) \times B(\mathfrak{K}) \to B(\mathfrak{K})$, such that for each set E in $\mathfrak{B}(X^n)$ of the form

$$E = E_0 \times E_1 \times \cdots \times E_{n-1} \tag{4}$$

with $E_j \in \mathfrak{B}(X^j)$, one has

$$\mathcal{E}^{Q_n}(E, B) = \mathcal{E}_0^Q(E_0, \mathcal{E}_1^Q(E_1, \ldots, \mathcal{E}_{n-1}^Q(E_{n-1}, B)\cdots)) \tag{5}$$

For each B in $B(\mathfrak{K})$. By the same theorem one also has that for each set E of the form $E = F \times G$ with $F\in\mathfrak{B}(X^l)$ and $G\in\mathfrak{B}(X^{n-l})$ with $X^{n-l} = X_l \times \cdots \times X_{n-1}$,

$$\mathcal{E}^{Q_n}(E, B) = \mathcal{E}^{Q_l}(F, \mathcal{E}^{Q_{n-l}}(G, B)) . \tag{6}$$

Thus one sees that by means of the correspondence assumption, with each finite process Q_n whose outcome spaces satisfy the restriction given, there is associated a unique expectation $\mathcal{E}^{Q_n}$ defined on $\mathfrak{B}(X^n) \times B(\mathfrak{K})$ which satisfies Eq. (4). In particular, the unique probability operator measure associated with the process Q_n is given by $\mathcal{E}^{Q_n}(-, \mathbf{1})$.

If one assumes that to each probability operator measure there corresponds a unique observable (the converse is already assumed in this section), then the above shows that a unique observable is associated with the process Q_n, i.e., the observable corresponding to $\mathcal{E}^{Q_n}(-, \mathbf{1})$.

In order to better understand this description, one has the following: If each outcome space X_j is countable, then the above results give

$$\mathcal{E}^{Q_n}(E, B) = \sum_{\varphi_n \in E} \mathcal{E}_0^Q(\{\varphi_n(0)\}, \mathcal{E}_1^Q(\{\varphi_n(1)\}, \ldots, \mathcal{E}_{n-1}^Q(\{\varphi_n(n-1)\}, B)\ldots) \tag{7}$$

for each $E\in\mathfrak{B}(X^n)$ and $B\in B(\mathfrak{K})$. φ_n denotes an element of X^n and $\{\varphi_n(j)\}$ is the subset of X_j containing the single element $\varphi_n(j)$. If, furthermore, Von Neumann's projection axiom[2] holds, then repeated use of Eq. (1) in (7) gives

$$\mathcal{E}^{Q_n}(E, B) = \sum_{\varphi_n \in E} P_{\varphi_n(0)}^{Q(0)} P_{\varphi_n(1)}^{Q(1)} \cdots P_{\varphi_n(n-1)}^{Q(n-1)} \times B P_{\varphi_n(n-1)}^{Q(n-1)} \cdots P_{\varphi_n(1)}^{Q(1)} P_{\varphi_n(0)}^{Q(0)}, \tag{8}$$

which is a well-known result.

These results show that with each finite sequence of measurements one can, under the correspondence assumption, associate a unique expectation defined on the product space X^n. The question now arises whether such an association is possible for infinite sequences.

To this end, let $X^\omega = \otimes_{j=0}^{\infty} X_j$ be the set of all possible infinite sequences of outcomes for the process Q. Define $\mathfrak{F}$ to be the field of all Borel cylinder subsets of X^ω and $\mathfrak{B}(X^\omega)$, the minimal σ field over $\mathfrak{F}$. That is, each $E\in\mathfrak{F}$ has the form $E = F \times X_n \times X_{n+1} \times \cdots$ with $F\in\mathfrak{B}(X^n)$ for some n. The question now becomes "under the correspondence assumption does there exist a unique expectation $\mathcal{E}^Q : \mathfrak{B}(X^\omega) \times B(\mathfrak{K}) \to B(\mathfrak{K})$ such that for each $E\in\mathfrak{F}$ with Borel base F in $\mathfrak{B}(X^n)$ and $B\in B(\mathfrak{K})$,

$$\mathcal{E}^{Q}(E,B) = \text{s-lim}_m \mathcal{E}_n^{Q_n}(F, \mathcal{E}_n^{Q}(X_n, \ldots, \mathcal{E}_{n+m-1}^{Q}(X_{n+m-1}, B)\ldots)), \tag{9}$$

where s-lim means convergence in the strong operator topology?"

It can be shown by means of examples that for some sequences of expectations the limit of Eq. (9) exists and for others it does not exist. However, one has the following result.

Set $B = \mathbf{1}$ in Eq. (9). Then from $\mathcal{E}_j^{Q}(X_j, \mathbf{1}) = \mathbf{1}$ for each j, one has that the sequence of probability operator measures $\{\mathcal{E}^{Q_n}(-, \mathbf{1}) | n = 1, 2\cdots\}$ is consistent, and the limit of Eq. (9) exists trivially and equals $\mathcal{E}^{Q_n}(F, \mathbf{1})$. By Theorem 1 of the Appendix there exists a unique probability operator measure $\mathcal{E}^{Q}(-, \mathbf{1})$: $\mathfrak{B}(X^\omega) \to B(\mathfrak{H})$ such that, for each $E \in \mathfrak{F}$ with base F in $\mathfrak{B}(X^n)$ for some n,

$$\mathcal{E}^{Q}(E, \mathbf{1}) = \mathcal{E}^{Q_n}(F, \mathbf{1}). \tag{10}$$

Thus one has the result that (under the correspondence assumption) with each infinite process Q there is associated a unique observable or probability operator measure $\mathcal{E}^{Q}(-, \mathbf{1})$ which satisfies Eq. (10) and which contains the statistical properties of Q. This follows from the fact that for each E in $\mathfrak{B}(X^\omega)$ and each state ρ, $\text{Tr}(\rho \mathcal{E}^{Q}(E, \mathbf{1}))$ is the probability that carrying out Q on a system in state ρ yields an outcome sequence in E.

IV. DECISION PROCEDURES

In the above, the methods of Davies[7] and Lewis[5] have been used to treat finite and infinite sequences of measurements. Here these methods are extended to include decision procedures. However, in order to avoid many mathematical problems, the restriction that the outcome space of each measurement procedure be at most countably infinite is reimposed. This is a quite minor restriction since it is doubtful if one needs to consider decision procedures whose associated trees have vertices with uncountably many lines leading out.

Let Q denote an infinite path decision procedure and τ_Q the associated tree. That is, Q is such that all paths in τ_Q are infinitely long. It is also required that each path in Q be such that each initial segment be of finite length. The necessary definitions and properties of trees and their association with decision procedures are given in II, and the reader is referred there for details. We continue to suppress inclusion of transformations or, more generally, of any operation (other than the identity observable) representable as an isometry on the underlying Hilbert space. (The inclusion of procedures with one or more finite paths will be discussed later on.)

For each n let Q_n denote the first n steps of Q. That is, Q_n is obtained from Q by cutting each path of Q (or τ_Q) between the nth and $(n+1)$th steps and discarding the infinite terminal segments.

Under the correspondence assumption each measurement in Q corresponds to an expectation with an associated set of outcomes. For each $n = 1, 2, \ldots$, define S_n to be the set of all possible outcome sequences φ_n $(= \varphi_n(0), \ldots, \varphi_n(n-1))$ of length n associated with carrying out the first n steps of Q. That is, for each n

$$\begin{aligned} S_1 &= X_\Phi, \\ S_{n+1} &= \bigcup_{\varphi_n \in S_n} \varphi_n {}^* X_{\varphi_n}, \end{aligned} \tag{11}$$

where $\varphi_n {}^* X_{\varphi_n}$ is the set of all sequences φ_{n+1} of length $n+1$ such that $\varphi_{n+1}(j) = \varphi_n(j)$ for $j = 0, 1, \ldots, n-1$ and $\varphi_{n+1}(n) \in X_{\varphi_n}$.

To motivate this definition, one notes that the first measurement in Q, denoted by $Q(\phi)$, has the (discrete) outcome space X_Φ, corresponds to an expectation $\mathcal{E}_\Phi^{Q}$, and is step 0 of Q. Φ denotes the empty set. For each $n = 1, 2\cdots$ and each $\varphi_n \in S_n$, let $\mathcal{E}_{\varphi_n}^{Q}$ be the expectation corresponding to the measurement operation $Q(\varphi_n)$ and X_{φ_n} the (discrete) outcome space of $Q(\varphi_n)$. $Q(\varphi_n)$ denotes the measurement operation, Q assigns to step n of any path in τ_Q which corresponds to observing the outcome sequence φ_n in the first n steps of Q. [Step number n is the $(n+1)$th step.]

Let Σ^n denote the set of all subsets of S_n. For each n one defines a mapping $O^{Q_n}: \Sigma^n \to \mathfrak{B}(\mathfrak{H})$ as follows: For each $\varphi_n \in S_n$

$$O^{Q_n}(\{\varphi_n\}) = \mathcal{E}_{\varphi_{n,0}}^{Q}(\{\varphi_n(0)\}, \mathcal{E}_{\varphi_{n,1}}^{Q}(\{\varphi_n(1)\}, \ldots, \mathcal{E}_{\varphi_{n,n-1}}^{Q}(\{\varphi_n(n-1)\}, \mathbf{1})\cdots), \tag{12}$$

where $\varphi_{n,j}$ denotes the first j elements of φ_n and $\varphi_{n,0} = \Phi$. For each $E \subset S_n$,

$$O^{Q_n}(E) = \sum_{\varphi_n \subset E} O^{Q_n}(\{\varphi_n\}), \tag{13}$$

where strong convergence is implied.

It must now be shown that O^{Q_n} is a probability operator measure. To this end, one first shows that the strong limit implied by Eq. (13) exists. Consider $O^{Q_n}(S_n)$. This can be written as

$$O^{Q_n}(S_n) = \sum_{\varphi_1 \in S_1} \sum_{\substack{\varphi_2 \in S_2 \\ \varphi_{2,1} = \varphi_1}} \cdots \sum_{\substack{\varphi_n \in S_1 \\ \varphi_{n,n-1} = \varphi_{n-1}}} \mathcal{E}_{\varphi_{n,0}}^{Q}(\{\varphi_n(0)\}, \mathcal{E}_{\varphi_{n,1}}^{Q}(\{\varphi_n(1)\}, \ldots, \mathcal{E}_{\varphi_{n,n-1}}^{Q}(\{\varphi_n(n-1)\}, \mathbf{1}))\cdots), \tag{14}$$

where $\varphi_{n,j}$ denotes the first j elements of φ_n. By repeated use of properties (a)–(c) and (f) in the definition of expectations, one has for the last sum in Eq. (14),

$$\sum_{\substack{\varphi_n\in S_n\\ \varphi_{n,n-1}=\varphi_{n-1}}} \mathcal{E}^{Q}_{\varphi_{n,0}}(\{\varphi_n(0)\}, \mathcal{E}^{Q}_{\varphi_{n,1}}(\{\varphi_n(1)\}, \ldots, \mathcal{E}^{Q}_{\varphi_{n,n-1}}(\{\varphi_n(n-1),1\})\cdots)$$

$$= \mathcal{E}^{Q}_{\varphi_{n,0}}(\{\varphi_n(0)\}, \mathcal{E}^{Q}_{\varphi_{n,1}}(\{\varphi_n(1)\}, \ldots, \sum_{\substack{\varphi_n\in S_n\\ \varphi_{n,n-1}=\varphi_{n-1}}} \mathcal{E}^{Q}_{\varphi_{n,n-1}}(\{\varphi_n(n-1)\},1))\ldots)$$

$$= \mathcal{E}^{Q}_{\varphi_{n,0}}(\{\varphi_n(0)\}, \mathcal{E}^{Q}_{\varphi_{n,1}}(\{\varphi_n(1)\}, \ldots, \mathcal{E}^{Q}_{\varphi_{n,n-2}}(\{\varphi_n(n-2)\},1))\ldots). \tag{15}$$

The second equality of Eq. (15) follows directly from properties(b) and (c) in the definition of expectations. The first equality follows from the fact that by property (a) the internal sum in the middle part of Eq. (15), which can be written as the limit of a sequence of partial sums, is the limit of a monotone sequence of positive operators. Property (f) allows one to shift the sum stepwise to the left.

Substitution of Eq. (15) into Eq. (14) and the use of $\mathcal{E}^{Q}_{\varphi_{n,j}} \equiv \mathcal{E}^{Q}_{\varphi_{n-1,j}}$ in each term of the sum of Eq. (14) ($\varphi_{n,j} = \varphi_{n-1,j}$ for $j = 0, 1, \ldots n-2$) and repetition of this process over and over gives finally

$$O^{Q_n}(S^n) = \sum_{\varphi_1\in S_1} \mathcal{E}_{\varphi_{1,0}}(\{\varphi_1(0)\},1) = \mathcal{E}_{\varphi_{1,0}}(S_1,1) = 1 \tag{16}$$

as $\varphi_{1,0} = \Phi$ is independent of φ_1.

From this one has that $O^{Q_n}(E)$ exists for each $E\subset S^n$. This follows from the fact that $O^{Q_n}(\{\varphi_n\}) \geq \mathbf{0}$ for each $\varphi_n\in S_n$ and that $O^{Q_n}(E)$ is the limit of a nondecreasing sequence of partial sums bounded from above[11] by $O^{Q_n}(S^n) = \mathbf{1}$.

Finally the strong countable additivity of O^{Q_n} follows from the fact that if an infinite sum of positive operators exists, then the limit is independent of the ordering of the terms in the sum. Thus for any countable pairwise disjoint sequence $\{E_j | j = 1, 2\cdots\}$ of subsets of S_n with $E = \cup_m E_m$,

$$O^{Q_n}(E) = \sum_{\varphi_m\in E} O^{Q_n}(\{\varphi_n\}) = \sum_m \sum_{\varphi_n\in E_m} O^{Q_n}(\{\varphi_n\}) = \sum_m O^{Q_n}(E_m), \tag{17}$$

and O^{Q_n} is a probability operator measure.

Thus one sees that for each decision procedure Q and each n the correspondence assumption associates to the decision procedure Q_n consisting of the first n steps of Q, a unique observable or probability operator measure O^{Q_n}. Clearly O^{Q_n} satisfies Eqs. (12) and (13) and describes the statistical properties of Q_n. The latter follows from the fact that the probability that carrying out Q_n on a system in state ρ gives an outcome sequence in E is given by $\mathrm{Tr}(\rho\, O^{Q_n}(E))$.

In order to use these results to assign an observable O^Q to Q, one defines X_n by $X_n = \cup_{\varphi_n\in S_n} X_{\varphi_n}$. Since each X_{φ_n} is countable, it can be assigned a suitable metric to make it a complete separable metric space.[11] Also since S_n is countable, X_n is a complete separable metric space.[13]

Let $\mathcal{B}(X)$ and $\mathcal{B}(X^n)$ be the σ-fields of Borel subsets of X and $X^n, = X_0 \times X_1 \times \cdots \times X_{n-1}$, respectively. Let X^ω denote the set of all infinite sequences of X: $\mathfrak{F}$ the field of all Borel cylinder subsets of X^ω, and $\mathcal{B}(X^\omega)$ the minimal σ-field over $\mathfrak{F}$. For each n, define Ω^n to be the subset of X^ω defined by $\Omega^n = S_n \times X_{n+1} \times X_{n+2} \times \cdots$. Since S_n is at most countably infinite and each X_{φ_n} is a Borel subset of X_n, $S_n\in\mathcal{B}(X^n)$ and in fact, each subset of S_n is a Borel subset of X^n. Clearly $\Omega^n\in\mathfrak{F}$ for each n.

Define Ω^Q by $\Omega^Q = \cap_n \Omega^n$. Ω^Q is the set of all outcome sequences of Q. Also $\Omega^Q\in\mathcal{B}(X^\omega)$ as $\mathcal{B}(X^\omega)$ is closed under countable intersections. Let $\mathfrak{F}^Q$ be the ring of all cylinder subsets of Ω^Q with bases in S_n for some n. That is, each E in $\mathfrak{F}^Q$ has the structure

$$E = (F \times X_{n+1} \times X_{n+2} \times \cdots) \cap \Omega^Q \tag{18}$$

for some $F\subset S_n$ and some n. Let Σ^Q be the minimal σ-ring over $\mathfrak{F}^Q$. Clearly Σ^Q is a sub σ-ring of $\mathcal{B}(X^\omega)$.

In order to show that there is a unique probability operator measure O^Q associated with Q, one proceeds as in II. For each n one defines $O^{Q'_n}$ on $\mathcal{B}(X^n)$ by

$$O^{Q'_n}(B) = O^{Q_n}(B\cap S_n) \tag{19}$$

for each $B\in\mathcal{B}(X^n)$ where O^{Q_n} is defined by Eqs. (12) and (13). Clearly $O^{Q'_n}$ is well defined and is a probability operator measure on $\mathcal{B}(X^n)$.

One must first show that the $O^{Q'_n}$ are consistent on the $\mathcal{B}(X^n)$ for $n = 1, 2, \cdots$. First assume that the O^{Q_n} are consistent on Σ^n, and let $m > n$. The for each set A in $\mathcal{B}(X^m)$ of the form $A = D \times X^{m-n}$ with D in $\mathcal{B}(X^n)$ one has from Eq. (19), with $E = A\cap S_m$ and $F = D\cap S_n$,

$$O^{Q'_m}(A) = O^{Q_m}(E) = O^{Q_n}(F) = O^{Q'_n}(D). \tag{20}$$

It remains to show that the O^{Q_n} are consistent. To this end, let E be a subset of S_m given by

$$E = \bigcup_{\varphi_n\in F} \bigcup_{\substack{\varphi_m\in S_m\\ \varphi_{m,n}=\varphi_n}} \{\varphi_m\}$$

for some subset F of S_n. [Note that by Eq. (11), E and F are such that there are sets A and D defined as for Eq. (20) such that $E = A\cap S_m$ and $F = D\cap S_n$.] From Eqs. (12) and (13) one has

$$O^{Q_m}(E) = \sum_{\varphi_n\in F} \sum_{\substack{\varphi_{n+1}\in S_{n+1}\\ \varphi_{n+1,n}=\varphi_n}} \sum_{\substack{\varphi_m\in S_m\\ \varphi_{m,n}=\varphi_n}} O^{Q_m}(\{\varphi_m\})$$

$$= \sum_{\varphi_n\in F} \sum_{\substack{\varphi_{n+1}\in S_{n+1}\\ \varphi_{n+1,n}=\varphi_n}} \cdots \sum_{\substack{\varphi_m\in S_m\\ \varphi_{m,n}=\varphi_n}} \mathcal{E}^{Q}_{\varphi_{m,0}}(\{\varphi_m(0)\}, \ldots,$$

$$\mathcal{E}^{Q}_{\varphi_{m,n+1}}(\{\varphi_m(n+1)\}, \ldots, \mathcal{E}^{Q}_{\varphi_m}(\{\varphi_m(m-1)\},1))\cdots).$$

By the use of the same techniques that were used to obtain Eq. (16), the above expression gives

$$O^{Q_m}(E) = \sum_{n\in F} \mathcal{E}^{Q}_{\varphi_{n,0}}(\{\varphi_n(0)\}, \ldots, \mathcal{E}^{Q}_{\varphi_n}(\{\varphi_n(n-1)\}, \mathbf{1}))\ldots) = O^{Q_n}(F), \quad (21)$$

and thus the O^{Q_n} are consistent.

Since the $O^{Q'_n}$ are consistent, Theorem 1 of the Appendix gives the result that there exists a unique probability operator measure $O^{Q'}$ on $\mathfrak{B}(X^\omega)$ such that for each set B in $\mathfrak{F}$ with base D in $\mathfrak{B}(X^n)$

$$O^{Q'}(B) = O^{Q'_n}(D). \quad (22)$$

By Eq. (18) and the fact that Ω^Q is in $\mathfrak{B}(X^\omega)$, $\mathfrak{F}^Q$ and Σ^Q are a subring and sub σ-ring of $\mathfrak{F}$ and $\mathfrak{B}(X^\omega)$, respectively. Thus one can define a map $O^Q: \Sigma^Q \to \mathfrak{B}(\mathfrak{K})$ as the restriction of $O^{Q'}$ to Σ^Q. That is,

$$O^Q(E) = O^{Q'}(E) \quad (23)$$

for each E in Σ^Q. Since $O^{Q'}$ is strongly continuous from above[1] on $\mathfrak{B}(X^\omega)$ and $O^{Q'}(\Omega^n) = \mathbf{1}$ (Eq. (19)] and the Ω^n are nonincreasing, one has that $O^Q(\Omega^Q) = O^{Q'}(\Omega^Q) = \text{s-lim}_n O^{Q'}(\Omega_n) = \mathbf{1}$. Thus O^Q is a probability operator measure. Also for each $E\in\mathfrak{F}^Q$ with base $F\subset S_n$ for some n, Eqs. (22) and (23) give

$$O^Q(E) = O^{Q_n}(F), \quad (24)$$

The uniqueness of O^Q follows from the uniqueness of $O^{Q'}$ and the fact that O^Q as given by Eq. (23) is well defined.

Thus one has the result that, under the correspondence assumption, with each infinite decision procedure Q consisting of measurement procedures with countable outcome spaces only, there is associated a unique probability operator measure or observable O^Q, which contains the statistical properties of the process.

This result extends easily to finite path decision procedures as well as those with both finite and infinite paths. To see this, the methods of II, used for a similar extension, are followed. Let Q be a decision procedure with one or more finite paths, and let Q' be the infinite path procedure obtained by adding to each finite path of Q an infinite sequence of repetitions of the measurement procedure whose only outcome is 1 and whose corresponding expectation $\mathcal{E}: \{\Phi, \{1\}\} \times B(\mathfrak{K}) \to B(\mathfrak{K})$ is defined by $\mathcal{E}(\Phi, B) = \mathbf{0}$, and

$$\mathcal{E}(\{1\}, B) = B \quad (25)$$

for each B in $\mathfrak{K}$. This corresponds to adding an infinite sequence of repetitions of the measurements of the identity observable and extends each finite outcome sequence of Q by adding on an infinite sequence of 1's to give the corresponding outcome sequence of Q'.

It first must be shown that the operator associated with each finite path of Q by the appropriate expectation string is also the same operator associated with the extended path of Q'. To this end, let p be a finite path in Q with corresponding outcome sequence φ^p, and let p' be the extension of p in Q' and $\varphi^{p'}$ the corresponding outcome sequence. Let φ^p be of length n and $\varphi^{p'}_m$ be the initial segment of $\varphi^{p'}$ of length m. By construction, for all j, if $0 \leq j < n$ $\varphi^p(j) = \varphi^{p'}(j)$ and if $j \geq n$, $\varphi^{p'}(j) = 1$.

For each $m > n$ one has from Eqs. (12) and (25)

$$O^{Q'_m}(\{\varphi^{p'}_m\}) = \mathcal{E}^{Q'}_{\varphi^{p'}_{m,0}}(\{\varphi^{p'}_m(0)\}, \ldots, \mathcal{E}^{Q'}_{\varphi^{p'}_{m,n-1}}(\{\varphi^{p'}_m(n-1)\}, \mathcal{E}^{Q'}_{\varphi^{p'}_{m,n}}(\{\varphi^{p'}_m(n)\}, \ldots, \mathcal{E}^{Q'}_{\varphi^{p'}_m}(\{\varphi^{p'}_m(m-1)\}, \mathbf{1}))\cdots) = \mathcal{E}^{Q'}_{\varphi^{p'}_{m,0}}(\{\varphi^{p'}_m(0)\}, \ldots, \mathcal{E}^{Q'}_{\varphi^{p'}_{m,n-1}}(\{\varphi^{p'}_m(n-1)\}, \mathbf{1}), \quad (26)$$

where the last equality arises from repeated use of Eq. (25), and $\varphi^{p'}_{m,j}$ denotes the first j elements of $\varphi^{p'}$ (or of $\varphi^{p'}_m$). By construction, for each $j \leq n-1$, $\mathcal{E}^{Q'}_{\varphi^{p'}_j} = \mathcal{E}^{Q}_{\varphi^{p}_j}$, and thus one obtains

$$O^{Q'_m}(\{\varphi^{p'}_m\}) = \mathcal{E}^{Q}_{\varphi^{p}_0}(\{\varphi^{p}(0)\}, \ldots, \mathcal{E}^{Q}_{\varphi^{p}_n}(\{\varphi^{p}_n(n-1)\}, \mathbf{1}), \quad (27)$$

where the right-hand operator is just the operator one would assign to path p of Q in a direct construction. Furthermore, since the right-hand side of Eq. (26) is independent of m if $m > n$ [$\varphi^{p'}_{m,j} = \varphi^{p'}_j$ and $\varphi^{p'}_m(j) = \varphi^{p'}(j)$ for $j = 0, 1, \ldots, n-1$], one has

$$\lim_m O^{Q'_m}(\{\varphi^{p'}_m\}) = \mathcal{E}^{Q}_{\varphi^{p}_0}(\{\varphi^{p}(0)\}, \cdots, \mathcal{E}^{Q}_{\varphi^{p}}(\varphi^{p}(n-1), \mathbf{1}))\ldots). \quad (28)$$

From these results, the probability operator measure $O^{Q'}$ uniquely associated with Q' can be used to associate a probability operator measure O^Q with Q which satisfies Eq. (12). One first defines the σ-field Σ^Q as follows: Let F^Q and I^Q be the respective sets of all finite and infinite outcome sequences of Q. Let Σ^Q be the set of all sets of the form $E_1 \cup E_2$ where $E_1 \subset F^Q$ and $E_2 = B \cap I^Q$ for some Borel subset B of X^ω.

To define $O^Q: \Sigma^Q \to B(\mathfrak{K})$, one first lets E be any set in Σ^Q which contains infinite sequences only. By construction the corresponding paths in Q and Q' are identical, and one sets $O^Q(E) = O^{Q'}(E)$.

Next let φ^p be any finite outcome sequence corresponding to a finite path p of Q with length n, and let $\varphi^{p'}$ and p' be the corresponding outcome sequence and path of Q'. Define $O^Q(\{\varphi^p\})$ by

$$O^Q(\{\varphi^p\}) = O^{Q'}(\{\varphi^{p'}\}), \quad (29)$$

and let $E^{p'}_{\varphi^p_m}$ be the set of all outcome sequences of Q' which have φ^p_m as an initial segment. Clearly by construction $E^{p'}_{\varphi^p_m} = \{\varphi^{p'}\}$ and $O^{Q'}(E^{p'}_{\varphi^p_m}) = O^{Q'_m}(\{\varphi^{p'}_m\})$ for all $m > n$. Since $O^{Q'}$ is a probability operator measure it is continuous from above,[1] and one has that $O^{Q'}(\{\varphi^{p'}\}) = \lim_m O^{Q'_m}(\{\varphi^{p'}_m\})$. Equations (28) and (29) then give

$$O^Q(\{\varphi^p\} = \mathcal{E}^{Q}_{\varphi^{p}_0}(\{\varphi^{p}(0)\}, \ldots, \mathcal{E}^{Q}_{\varphi^{p}}(\varphi^{p}(n-1), \mathbf{1}))\ldots), \quad (30)$$

which is just what Eq. (12) would give for a direct construction.

For any set E containing finite outcome sequences of Q only, set

$$O^Q(E) = \sum_{\varphi\in E} O^Q(\{\varphi\}) \quad (31)$$

with $O^Q(\{\varphi\})$ given by Eq. (29). One has $O^Q(E) = O^{Q'}(F)$, where F is the set of outcome sequences of Q' corresponding to those in E. Finally let $E = E_1 \cup E_2$ with E_1 and E_2 containing finite and infinite sequences only. In this case define $O^Q(E_1 \cup E_2)$ by

$$O^Q_{E_1 \cup E_2} = O^Q_{E_1} + O^Q_{E_2} = O^{Q'}_{E_1} + O^{Q'}_{E_2}$$

with $O^Q_{E_1}$ and $O^Q_{E_2}$ given as above.

Since $O^{Q'}$ is a probability operator measure so is O^Q, and one has the desired result that O^Q, constructed through $O^{Q'}$, is the probability operator measure associated with the decision procedure Q. O^Q is unique since Q' is the unique extension of Q which allows Eq. (30) to be satisfied, and $O^{Q'}$ is unique.

So far the procedures Q were assumed to consist of measurement procedures only without any intervening transformations. It is easy to see that the results obtained here hold without this restriction. To see this, let the procedure Q contain transformations (implementable by isometries in the common Hilbert space of the process) in between the measurement procedures. Let φ be any infinite outcome sequence of Q. Then by repeated use of Eq. (2) which gives $\mathrm{Tr}(\mathcal{J}(E, V\rho V^\dagger)B) = \mathrm{Tr}(\rho V^\dagger \mathcal{E}(E,B)V)$, one has the result that the right-hand sides of Eqs. (5) and (12) are replaced by

$$V_0^{Q\dagger}\mathcal{E}_0^Q(E_0, V_1^{Q\dagger}\mathcal{E}_1^Q(E_1, \ldots, V_{n-1}^{Q\dagger}\mathcal{E}_{n-1}^Q(E_{n-1}, B) \times V_{n-1}^Q)\ldots)V_1^Q)V_0^Q$$

and

$$V_{\varphi_{n,0}}^{Q\dagger}\mathcal{E}_{\varphi_{n,0}}^Q(\{\varphi_n(0)\}, V_{\varphi_{n,1}}^{Q\dagger}\mathcal{E}_{\varphi_{n,1}}^Q(\{\varphi_n(1)\}, \ldots, V_{\varphi_{n,n-1}}^{Q\dagger}\mathcal{E}_{\varphi_{n,n-1}}^Q(\{\varphi_n(n-1)\}, \mathbf{1})V_{\varphi_{n,n-1}}^Q)\cdots)V_{\varphi_{n,1}}^Q)V_{\varphi_{n,0}}^Q,$$

respectively, where V_j^Q and $V_{\varphi_{n,j}}^Q$ are the isometries corresponding to the appropriate transformations in Q. Note that, in the decision procedures, the possible dependence of the transformation on previous outcomes is taken care of by the index "$\varphi_{n,j}$" on V^Q. Also the above expressions reduce to the right-hand sides of Eqs. (5) and (12) if V_j and $V_{\varphi_{n,j}}$ correspond to the identity for each $j = 0, 1, \ldots, n-1$.

By Theorem 3 of the Appendix, if $\mathcal{E}$ is an expectation so is $V^\dagger \mathcal{E} V$ for any isometry V. Thus the association of probability operator measures (or observables) with procedures applies also to infinite sequences of measurements (Sec. III) with intervening transformations and to decision procedures which contain transformations; for one uses the above replacements in Eqs. (5) and (12) and proceeds exactly as before, using Theorem 3 when necessary, to obtain the desired association.

ACKNOWLEDGMENT

The author thanks the referee for several valuable comments which included an example for which the limit of Eq. (9) does not exist.

APPENDIX

A. Operator Valued Measures

Let Σ be a σ-ring of subsets of some set Ω and $B(\mathcal{H})$ the algebra of all bounded linear operators on a Hilbert space $\mathcal{H}$. A map $O: \Sigma \to B(\mathcal{H})$ is an operator-valued measure if $O(\Phi) = \mathbf{0}$ and O is strongly countably additive. O is self-adjoint or positive if, for each E in Σ, $O(E)$ is a self-adjoint or positive operator, respectively. If O is positive, Σ is a σ field, and $O(\Omega) = \mathbf{1}$, then O is a probability operator measure. Each probability operator measure O and each state ρ define a scalar probability measure P_ρ by $P_\rho(E) = \mathrm{Tr}(\rho O(E))$ for each E in Σ.

An operator-valued measure has the following easy properties: O is finitely additive. If $E \subset F$, then

$$O(F) = O(E) + O(F - E). \tag{A1}$$

If further O is positive, then

$$O(F) \geq O(E). \tag{A2}$$

If O is finitely additive, positive, and strongly continuous from above at Φ or from below, then O is strongly countably additive.[1,14] Weak countable additivity is equivalent to strong countable additivity, but not to uniform countable additivity.[9]

If Σ is the minimal σ ring over a ring S of subsets of Ω and O' is a bounded positive operator measure defined on S, then, as Berberian[14] has shown, there is a unique extension O of O' onto Σ such that $O = O'$ on S.

An operator-valued measure O is bounded if there exists a constant M such that

$$\|O(E)\| < M \tag{A3}$$

for each E in Σ. O is decomposable if there exist four positive operator measures O_1, O_2, O_3, O_4 such that

$$O = O_1 - O_2 + iO_3 - iO_4. \tag{A4}$$

It will be seen that there are many decomposable O's.

Let X be a Hausdorff space and $\mathfrak{B}(X)$ the σ field of subsets of X generated by the open subsets. A bounded positive operator measure on $\mathfrak{B}(X)$ is regular if for each $E \in \mathfrak{B}(X)$

$$O(E) = \sup\{O(C) \mid C \subset E \text{ and } C \text{ compact}\},$$
$$O(E) = \inf\{O(V) \mid V \supset E \text{ and } V \text{ open}\}.$$

If X is a complete separable metric space, it follows from a result of Parthasarathy[12] that very bounded positive operator measure on $\mathfrak{B}(X)$ is regular.

For our purposes the most important property of operator measures is the extension of such measures onto infinite product spaces from finite product spaces. One has the following theorem:

Theorem 1: Let $\{X_j \mid j = 1, 2\cdots\}$ be a sequence of complete separable metric spaces, $X^n = X_1 \times X_2 \times \cdots X_n$ and $\mathfrak{B}(X^n)$ the σ field of Borel subsets of X^n, and let $\{O_n \mid n = 1, 2\cdots\}$ be a consistent sequence of positive operator measures with (X^n) the domain of O_n. Let $X^\omega = X_1 \times X_2 \times \cdots$ and $\mathfrak{F}$ be the field of all Borel cylinder subsets of X^ω and $\mathfrak{B}(\mathfrak{F})$ the minimal σ field over $\mathfrak{F}$.

Then there exists a unique positive operator measure O on $\mathfrak{B}(\mathfrak{F})$ such that for each $E \in \mathfrak{F}$ with base $F \in \mathfrak{B}(X^n)$,

$$O(E) = O_n(F). \tag{A5}$$

Proof: If one can show that $\{O_n | n = 1, 2\cdots\}$ defines a unique positive operator measure O' on $\mathfrak{F}$, then, by Berberian's extension theorem,[14] one has the desired result. To this end, one must show that O', defined on $\mathfrak{F}$ by Eq. (A5) with $O'(E)$ replacing $O(E)$, is finitely additive and strongly continuous at Φ from which it follows that O' is strongly countably additive.

To prove strong continuity at Φ, it is sufficient to prove the converse. That is, let $\{E_l | l = 1, 2, \dots\}$ be a nonincreasing sequence of sets in $\mathfrak{F}$ such that there exists an $\epsilon > 0$ and a ψ in $\mathfrak{K}$ such that $\|O'E_l)\psi\| > \epsilon$ for each l. We have to prove $\lim_l E_l$ is not empty.

Let n_l and F_l be the base index and base for each $E_l[E_l = F_l \times X_{n_l+1} \times \cdots$ with F_l in $\mathfrak{B}(X^{n_l})]$. Since each O_n is regular on $\mathfrak{B}(X^n)$, for each $\delta > 0$ there is a compact set C_l in $\mathfrak{B}(X^{n_l})$ with $C_l \subset F_l$ such that $\|O_{n_l}(F_l - C_l)\psi\| \delta < /2^{l+1}$.

From here on the proof will not be given as it is an exact repetition of the proof given elsewhere for the real line.[1,8,15] One notes that since each X^n is a metric space each compact subset of X^n is sequentially compact and closed.[16] QED

B. Expectations

The definition of an expectation $\mathcal{E}$ is given at the beginning of Sec. II of the main text and will be referred to often.

Many properties of expectations are simple consequences of the definition. One has from properties (e) and (c) of the definition

$$\mathcal{E}(E, \mathbf{0}) = \mathbf{0} = \mathcal{E}(\phi, B) \tag{A6}$$

for any E in $\mathfrak{B}(X)$ and B in $B(\mathfrak{K})$. Also $\mathcal{E}(E, B)$ is self-adjoint if and only if B is. To see this, let B be self-adjoint and set

$$B = B_+ - (B_+ - B), \tag{A7}$$

where B_+ is the positive operator given by

$$B_+ = \int_0^\infty r d\mathcal{S}_r^B, \tag{A8}$$

where $\mathcal{S}^B$ is the spectral measure of B and $\mathcal{S}_r^B = \mathcal{S}^B((-\infty, r])$. Since $B_+ - B$ is also a positive operator, properties (a), (d), and (f) give that $\mathcal{E}(E, B)$ is self-adjoint. If B is not self-adjoint, then one can write $B = B_1 + iB_2$ with B_1, B_2 self-adjoint and repeat the above to show that $\mathcal{E}(E, B)$ is not self-adjoint. Another immediate property is that

$$B \geq B' \quad \text{implies} \quad \mathcal{E}(E, B) \geq \mathcal{E}(E, B')) \tag{A9}$$

for all E and self-adjoint B and B'. This follows from (a), (d), and (e), and $B - B' \geq 0$. Also, one notes that for each B, $\mathcal{E}(-, B)$ is a decomposable operator measure. This follows from setting $B = B_a + iB_b$ and using Eqs. (A7) and (A8) to decompose B_a and B_b further to get $B = B_1 - B_2 + iB_3 - iB_4$ where B_1, B_2, B_3, and B_4 are positive operators. Properties (a), (d), and (f) give the desired result that $\mathcal{E}(E, B) = \mathcal{E}(E, B_1) - \mathcal{E}(E, B_2) + i\mathcal{E}(E, B_3) - i\mathcal{E}(E, B_4)$ for each E in $\mathfrak{B}(X)$. This answers the question for such measures raised in conjunction with Eq. (A4). Thus one sees that $\mathcal{E}$ is a family of operator measures indexed by the operators in $B(\mathfrak{K})$ and which satisfies properties (a), (b) and (d)–(f) in the definition.

Finally one has that

$$\mathcal{E}(E, \mathbf{1}) = \mathbf{0} \quad \text{implies} \quad \mathcal{E}(E, B) = \mathbf{0} \tag{A10}$$

for each $B \in B(\mathfrak{K})$. Since any operator can be decomposed into the sum of four positive operators [Eqs. (A7) and (A8)] by properties (d) and (f), it is sufficient to prove this for positive operators. Let B be a positive operator in $B(\mathfrak{K})$ and consider $B' = B/\|B\|$. Since $\mathbf{0} \leq B' \leq \mathbf{1}$, Eq. (A9) and $\mathcal{E}(E, \mathbf{1}) = \mathbf{0}$ imply that $\mathcal{E}(E, B') = \mathbf{0}$ and by (d) $\mathcal{E}(E, B) = \mathbf{0}$.

The next theorem refers to the composition of a finite number of expectations.

Theorem 2 (Davies and Lewis): Let $X_1, X_2, \dots X_n$ each be a complete separable metric space with $\mathfrak{B}(X_1), \dots, \mathfrak{B}(X_n)$, the respective σ-fields of Borel subsets of $X_1, \dots, X_n$. Let $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_n$ be expectations defined on $\mathfrak{B}(X_1) \times B(\mathfrak{K}), \dots, \mathfrak{B}(X_n) \times B(\mathfrak{K})$, respectively. Then there exists a unique expectation $\mathcal{E}: \mathfrak{B}(X^n) \times B(\mathfrak{K}) \to B(\mathfrak{K})$, where $\mathfrak{B}(X^n)$ is the system of Borel subsets of $X^n = X_1 \times \cdots \times X_n$, such that for each rectangle $E = E_1 \times \cdots \times E_n$ with E_j in $\mathfrak{B}(X_j)$ for $j = 1, 2, \dots, n$,

$$\mathcal{E}(E, B) = \mathcal{E}_1(E_1, \mathcal{E}_2(E_2, \dots, \mathcal{E}_n(E_n, B))\cdots) \tag{A11}$$

for each B in $B(\mathfrak{K})$.

Proof: Davies and Lewis[5] have shown that, for $n = 2$, the combination of two instruments is an instrument. Clearly this holds for any finite n. The theorem then follows from the fact that expectations and instruments are uniquely related by a 1–1 corespondence[7] through Eq. (2) of the text. QED

Theorems 1 and 2 can be generalised to the case in which the $X_1, \dots, X_n$ are σ-compact, locally compact Hausdorff spaces. In this case one defines Baire and Borel expectations and regularity in an obvious way. The one can prove that every Baire expectation is regular, and every Baire expectation extends to a unique regular Borel expectation $\mathcal{E}'$ such that $\mathcal{E}' = \mathcal{E}$ on the Baire sets.[17] Theorem 2, just proved, then applies with minor changes to a sequence of Baire expectations, and one can combine these results to show that a sequence $\mathcal{E}_1, \dots, \mathcal{E}_n$ of regular Borel expectations on $\mathrm{Bo}(X_1) \times B(\mathfrak{K}), \dots, \mathrm{Bo}(X_n) \times B(\mathfrak{K})$ generates a unique regular Borel expectation on $\mathrm{Bo}(X^n) \times B(\mathfrak{K})$.[17] [$\mathrm{Bo}(X)$ = minimal σ ring over the set of compact subsets of X.]

The following simple result is needed in the text.

Theorem 3: If V is an isometry on a Hilbert space $\mathfrak{K}$ and $\mathcal{E}$ is an expectation on $\mathfrak{B}(\Omega) \times B(\mathfrak{K})$ for any set Ω, then $V^\dagger \mathcal{E} V$ is an expectation.

Proof: Properties (a), (b), (d), and (f) of the definitions of an expectation are obvious. Property (e) follows from

$$\left\| V^\dagger \mathcal{E}(E, B)V - \sum_{j=1}^{m} V^\dagger \mathcal{E}(E_j B)V \ \psi \right\|$$

$$\leq \|V^\dagger\| \cdot \left\| \left(\mathcal{E}(E,B) - \sum_{j=1}^{m} \mathcal{E}(E_j,B) \right) V\psi \right\| \to 0$$

as $m \to \infty$,

and property (f) follows from

$$\|V^\dagger \mathcal{E}(E,B)V - V^\dagger \mathcal{E}(E,B_\alpha)V)\psi\| < \|V^\dagger\| \cdot \|(\mathcal{E}(E,B) - \mathcal{E}(E,B_\alpha))V\psi\| \to 0$$

as α increases. QED

---

* Work performed under the auspices of the U.S. Atomic Energy Commission.

1 P.A. Benioff, J. Math. Phys. **13**, 231 (1972).
2 J. Von Neumann, *Mathematical Foundations of Quantum Mechanics*, Translated by R. T. Beyer (Princeton U.P., Princeton, N.J., 1955), pp. 335-51.
3 P. A. Benioff, J. Math. Phys., **13**, 908 (1972).
4 H. Margenau, Phil. Sci. **30**, 1 (1963).
5 E. B. Davies and J. T. Lewis, Commun. Math. Phys. **17**, 239 (1970).
6 M. Nakamura and H. Umegaki, Math. Japan. **7**, 151 (1962).
7 E. B. Davies, J. Functional Anal. **6**, 318 (1970).
8 A. N. Kolmogorov, *Foundations of Probability Theory* (Chelsea, New York, 1956), 2nd English ed., pp. 29-33.
9 I. Dobrakov, Czech. Math. J. **20**, 511 (1970).
10 S. K. Berberian, *Notes on Spectral Theory* (Nostrand, Reinhold, New York, 1966), pp. 6, 7.
11 J. L. Kelley, *General Topology* (Nostrand, Reinhold, New York, 1955), p. 146; J. Dugundji, *Topology* (Allyn and Bacon, Boston, 1966), pp. 181, 183, 294.
12 K. R. Parthasarathy, *Probabilities on Metric Spaces* (Academic, New York, 1967), Chaps. II, V.
13 G. W. Mackey, Trans. Amer. Math. Soc. **85**, 134 (1957).
14 Reference 10, p. 15.
15 M. Loève, *Probability Theory* (Nostrand, Reinhold, New York, 1963), 3rd ed., pp. 92-4.
16 N. Dunford and J. T. Schwartz, *Linear Operators* (Interscience, New York, 1957), Vol. I p. 21.
17 P. A. Benioff, unpublished work.

Ward-Takahashi Relations in Massive Yang-Mills Theory*†

T. A. Barnebey‡

Department of Physics, University of California, Los Angeles, California 90024

(Received 24 February 1972)

The equivalence of massive Yang-Mills theory in the simplest gauge to a theory of vector electrodynamics is reviewed. The resultant Feynman rules and $SU(2)$ invariance are used to derive Ward-Takahashi relations among the corrected vertices which are valid to all orders. One of these is employed to find a new Ward identity. A brief discussion of divergence problems is given.

1. INTRODUCTION

Investigations of massive Yang-Mills theory[1-4] have generally been formal and/or complicated by efforts to include space-time dependent gauge transformations. It is the purpose of this paper to adopt a more pedestrian approach, and examine the theory in the simplest gauge (called the vector gauge[5]) by standard perturbative methods.

The ultimate goal of previous papers as well as the present work is to determine whether massive Yang-Mills theory is renormalizable. Since the key to the renormalizability of spinor electrodynamics is the existence of Ward relations, it would seem that the first step toward the goal is to look for them in this case also. This step is successfully carried out here.

Since the title of this work is similar to that of a paper by Veltman,[3] perhaps the main distinctions between the two should be explicitly stated. Broadly speaking, they differ in approach (this paper does not employ the spurious scalars introduced by Veltman) and in the generality of the results (the Ward-Takahashi relations derived here are valid off, as well as on, the mass shell).

The development begins with Sec. 2, where the Yang-Mills Lagrangian is interpreted as representing a theory of vector electrodynamics in which the "photon" has the same mass as the charged particles. In Sec. 3, it is shown that nonlocal gauge invariance may be used to rotate Wightman functions corresponding to different processes into one another; in this way the apparent asymmetry introduced by labeling one gauge field component "neutral" and the others "charged" is removed.

The Feynman rules for the perturbation expansion are derived in Sec. 4, and Ward-Takahashi relations among the bare propagators and vertices are obtained in Sec. 5.

The similarity of the massive Yang-Mills theory to a theory of vector electrodynamics provides motivations for the theorems concerning general tree diagrams which are presented in Sec. 6. Section 7 contains derivations of Ward-Takahashi relations between total propagators and vertices which are based on these theorems. It is also shown there that although the distinction between neutral and charged fields is central to the theorems, the generality of the Ward-Takahashi relations is not impaired, since, essentially, gauge invariance allows a neutral field to be rotated into a charged one.

In Secs. 8 and 9 one of the Ward-Takahashi relations is applied to find a Ward identity between the renormalization constants of the formally renormalized theory.

Section 10 concludes the paper with discussions of the field ordering implied by the theorems and the question of divergences.

2. FORMALISM

The Lagrangian of the massive Yang-Mills field in the vector gauge is

$$\mathcal{L} = -\tfrac{1}{4} D^a_{\mu\nu} D^{\mu\nu}_a + \tfrac{1}{2} m^2 D^a_\mu D^\mu_a ,$$

where

$$D^a_{\mu\nu} \equiv \partial_\mu D^a_\nu - \partial_\nu D^a_\mu - g C_{abc} D^b_\mu D^c_\nu ,$$

$a, b, c, \cdots$ label the $SU(n)$ components of the vector fields and the C's are structure constants.

$$\leq \|V^\dagger\| \cdot \left\| \left(\mathcal{E}(E,B) - \sum_{j=1}^{m} \mathcal{E}(E_j,B) \right) V\psi \right\| \to 0$$

as $m \to \infty$,

and property (f) follows from

$$\|V^\dagger \mathcal{E}(E,B)V - V^\dagger \mathcal{E}(E,B_\alpha)V)\psi\| < \|V^\dagger\| \cdot \|(\mathcal{E}(E,B) - \mathcal{E}(E,B_\alpha))V\psi\| \to 0$$

as α increases. QED

---

* Work performed under the auspices of the U.S. Atomic Energy Commission.

1 P.A. Benioff, J. Math. Phys. **13**, 231 (1972).
2 J. Von Neumann, *Mathematical Foundations of Quantum Mechanics*, Translated by R. T. Beyer (Princeton U.P., Princeton, N.J., 1955), pp. 335-51.
3 P. A. Benioff, J. Math. Phys., **13**, 908 (1972).
4 H. Margenau, Phil. Sci. **30**, 1 (1963).
5 E. B. Davies and J. T. Lewis, Commun. Math. Phys. **17**, 239 (1970).
6 M. Nakamura and H. Umegaki, Math. Japan. **7**, 151 (1962).
7 E. B. Davies, J. Functional Anal. **6**, 318 (1970).
8 A. N. Kolmogorov, *Foundations of Probability Theory* (Chelsea, New York, 1956), 2nd English ed., pp. 29-33.
9 I. Dobrakov, Czech. Math. J. **20**, 511 (1970).
10 S. K. Berberian, *Notes on Spectral Theory* (Nostrand, Reinhold, New York, 1966), pp. 6, 7.
11 J. L. Kelley, *General Topology* (Nostrand, Reinhold, New York, 1955), p. 146; J. Dugundji, *Topology* (Allyn and Bacon, Boston, 1966), pp. 181, 183, 294.
12 K. R. Parthasarathy, *Probabilities on Metric Spaces* (Academic, New York, 1967), Chaps. II, V.
13 G. W. Mackey, Trans. Amer. Math. Soc. **85**, 134 (1957).
14 Reference 10, p. 15.
15 M. Loève, *Probability Theory* (Nostrand, Reinhold, New York, 1963), 3rd ed., pp. 92-4.
16 N. Dunford and J. T. Schwartz, *Linear Operators* (Interscience, New York, 1957), Vol. I p. 21.
17 P. A. Benioff, unpublished work.

Ward-Takahashi Relations in Massive Yang-Mills Theory*†

T. A. Barnebey‡

Department of Physics, University of California, Los Angeles, California 90024

(Received 24 February 1972)

The equivalence of massive Yang-Mills theory in the simplest gauge to a theory of vector electrodynamics is reviewed. The resultant Feynman rules and $SU(2)$ invariance are used to derive Ward-Takahashi relations among the corrected vertices which are valid to all orders. One of these is employed to find a new Ward identity. A brief discussion of divergence problems is given.

1. INTRODUCTION

Investigations of massive Yang-Mills theory[1-4] have generally been formal and/or complicated by efforts to include space-time dependent gauge transformations. It is the purpose of this paper to adopt a more pedestrian approach, and examine the theory in the simplest gauge (called the vector gauge[5]) by standard perturbative methods.

The ultimate goal of previous papers as well as the present work is to determine whether massive Yang-Mills theory is renormalizable. Since the key to the renormalizability of spinor electrodynamics is the existence of Ward relations, it would seem that the first step toward the goal is to look for them in this case also. This step is successfully carried out here.

Since the title of this work is similar to that of a paper by Veltman,[3] perhaps the main distinctions between the two should be explicitly stated. Broadly speaking, they differ in approach (this paper does not employ the spurious scalars introduced by Veltman) and in the generality of the results (the Ward-Takahashi relations derived here are valid off, as well as on, the mass shell).

The development begins with Sec. 2, where the Yang-Mills Lagrangian is interpreted as representing a theory of vector electrodynamics in which the "photon" has the same mass as the charged particles. In Sec. 3, it is shown that nonlocal gauge invariance may be used to rotate Wightman functions corresponding to different processes into one another; in this way the apparent asymmetry introduced by labeling one gauge field component "neutral" and the others "charged" is removed.

The Feynman rules for the perturbation expansion are derived in Sec. 4, and Ward-Takahashi relations among the bare propagators and vertices are obtained in Sec. 5.

The similarity of the massive Yang-Mills theory to a theory of vector electrodynamics provides motivations for the theorems concerning general tree diagrams which are presented in Sec. 6. Section 7 contains derivations of Ward-Takahashi relations between total propagators and vertices which are based on these theorems. It is also shown there that although the distinction between neutral and charged fields is central to the theorems, the generality of the Ward-Takahashi relations is not impaired, since, essentially, gauge invariance allows a neutral field to be rotated into a charged one.

In Secs. 8 and 9 one of the Ward-Takahashi relations is applied to find a Ward identity between the renormalization constants of the formally renormalized theory.

Section 10 concludes the paper with discussions of the field ordering implied by the theorems and the question of divergences.

2. FORMALISM

The Lagrangian of the massive Yang-Mills field in the vector gauge is

$$\mathcal{L} = -\tfrac{1}{4} D^a_{\mu\nu} D^{\mu\nu}_a + \tfrac{1}{2} m^2 D^a_\mu D^\mu_a ,$$

where

$$D^a_{\mu\nu} \equiv \partial_\mu D^a_\nu - \partial_\nu D^a_\mu - g C_{abc} D^b_\mu D^c_\nu ,$$

$a, b, c, \cdots$ label the $SU(n)$ components of the vector fields and the C's are structure constants.

$\mathcal{L}$ is invariant under the transformations

$$D_\mu^{a'}(x) = R_{ab} D_\mu^b(x) \tag{2.1}$$

if R_{ab} satisfies

$$\partial_\mu R_{ab} = 0, \quad R_{ac}R^T_{cb} = \delta_{ab}$$

and

$$R_{ad}R_{be}R_{cf}C_{def} = C_{abc}\,.$$

This implies that there are conserved currents in the theory, which in turn suggests that there are Ward relations.

In the remainder of this paper, only the theory corresponding to the group $SU(2)$ will be studied. There are then three independent vector fields, and if the following identifications are made,

$$A_\mu \equiv D^3_\mu(x),$$
$$B_\mu \equiv (1/\sqrt{2})(D^1_\mu - iD^2_\mu), \quad B^+_\mu \equiv (1/\sqrt{2})(D^1_\mu + iD^2_\mu),$$

the Lagrangian becomes

$$\mathcal{L} = -\tfrac{1}{4}a_{\mu\nu}a^{\mu\nu} + \tfrac{1}{2}m^2A_\mu A^\mu - \tfrac{1}{2}b_{\mu\nu}b^{+\mu\nu} + m^2B_\mu B^{+\mu}, \tag{2.2}$$

where

$$a_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu - ig(B^+_\mu B_\nu - B_\mu B^+_\nu),$$
$$b_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu + ig(A_\mu B_\nu - A_\nu B_\mu).$$

$\mathcal{L}$ now appears to describe a theory of vector electrodynamics, in which the neutral and charged particles have the same mass. Hence the case $SU(2)$ is convenient in that it is particularly amenable to familiar interpretation.[6]

3. $SU(2)$ INVARIANCE

The invariance of the theory under constant rotations in $SU(2)$ space means that the various physical processes are not all independent. The relations between them are most easily displayed in terms of Wightman functions.

In order to illustrate the method of deriving these relations, consider

$$\langle 0|A_\mu(x)A_\nu(y)|0\rangle\,.$$

If $\mathcal{R}$ is the unitary Hilbert space operator corresponding to a rotation R_{ab} in Eq. (2.1), it follows from $SU(2)$ invariance of the vacuum that

$$\langle 0|A_\mu(x)A_\nu(y)|0\rangle = \langle 0|\mathcal{R}^{-1}A_\mu(x)\mathcal{R}\mathcal{R}^{-1}A_\nu(y)\mathcal{R}|0\rangle.$$

Now choose $\mathcal{R} = \mathcal{R}_M$ with

$$\mathcal{R}_M^{-1}\begin{pmatrix}B_\mu\\B^+_\mu\\A_\mu\end{pmatrix}\mathcal{R}_M = \begin{pmatrix}i/2 & -i/2 & -i/\sqrt{2}\\ i/2 & -i/2 & i/\sqrt{2}\\ i/\sqrt{2} & i/\sqrt{2} & 0\end{pmatrix}\begin{pmatrix}B_\mu\\B^+_\mu\\A_\mu\end{pmatrix}. \tag{3.1}$$

Then

$$\langle 0|A_\mu(x)A_\nu(y)|0\rangle = \tfrac{1}{2}\langle 0|[B_\mu(x)B_\nu(y) + B^+_\mu(x)B^+_\nu(y) + B_\mu(x)B^+_\nu(y) + B^+_\mu(x)B_\nu(y)]|0\rangle\,. \tag{3.2}$$

Another $\mathcal{R}$ operator may be used to reduce the expression on the right; this time $\mathcal{R}$ is chosen to be $\mathcal{R}_C$, where

$$\mathcal{R}_C^{-1}\begin{pmatrix}B_\mu\\B^+_\mu\\A_\mu\end{pmatrix}\mathcal{R}_C = \begin{pmatrix}0 & e^{-i\varphi} & 0\\ e^{i\varphi} & 0 & 0\\ 0 & 0 & -1\end{pmatrix}\begin{pmatrix}B_\mu\\B^+_\mu\\A_\mu\end{pmatrix}, \tag{3.3}$$

and ϕ is an arbitrary real number. This rotation, of course, corresponds to charge conjugation in electrodynamics. When invariance under this operation is invoked, it is inferred that

$$\langle 0|B_\mu(x)B_\nu(y)|0\rangle = \langle 0|B^+_\mu(x)B^+_\nu(y)|0\rangle = 0.$$

In general, any Wightman function containing an unequal number of B and B^+ fields vanishes. It may also be shown that

$$\langle 0|B^+_\mu(x)B_\nu(y)|0\rangle = \langle 0|B_\mu(x)B^+_\nu(y)|0\rangle.$$

Hence, Eq. (3.2) becomes

$$\langle 0|A_\mu(x)A_\nu(y)|0\rangle = \langle 0|B_\mu(x)B^+_\nu(y)|0\rangle. \tag{3.4}$$

By proceeding in similar fashion, using invariance under transformations which mix charged and neutral fields, such as $\mathcal{R}_M$, and invariance under charge conjugation $\mathcal{R}_C$, it is possible to relate other Wightman functions involving equal numbers of different types of fields to each other. Some additional examples are given below:

$$\langle 0|B^+_\mu(x)B_\nu(y)A_\rho(z)|0\rangle = -\langle 0|B_\mu(x)B^+_\nu(y)A_\rho(z)|0\rangle, \tag{3.5}$$

$$\langle 0|B^+_\mu(x)B_\nu(y)A_\rho(z)|0\rangle = -\langle 0|B^+_\mu(x)A_\nu(y)B_\rho(z)|0\rangle, \tag{3.6}$$

$$\langle 0|A_\mu(x)A_\nu(y)B_\rho(z)B^+_\sigma(w)|0\rangle = \langle 0|B_\mu(x)B^+_\nu(y)A_\rho(z)A_\sigma(w)|0\rangle, \tag{3.7}$$

$$\langle 0|B^+_\mu(x)B^+_\nu(y)B_\rho(z)B_\sigma(w)|0\rangle = \langle 0|A_\mu(x)A_\nu(y)A_\rho(z)A_\sigma(w)|0\rangle - \langle 0|A_\mu(x)A_\nu(y)B^+_\rho(z)B_\sigma(w)|0\rangle. \tag{3.8}$$

These equations are valid to all orders in the perturbation expansion and, therefore, in each order separately. They will be referred to again in Sec. 7, after some interaction representation results have been discussed.

4. FEYNMAN RULES

A. The Renormalized Lagrangian

The Feynman rules of any theory of vector electrodynamics are not trivial to derive, because the time components of the vector fields must be treated as dependent variables. However, Nakamura[7] and Tzou,[8] building on the work of Lee and Yang,[9] have shown that in the present case the noncovariant parts of the propagators may be dropped and the Feynman rules again become simple if the interaction Hamiltonian is taken to be the negative of the interaction Lagrangian.[10] The rules found in those papers are extended to include renormalizations in this section, and are employed in the remaining sections.

It is not known whether the massive Yang–Mills theory is renormalizable; but it is still possible to formally allow for proper normalization of physical quantities by introducing renormalization constants into the Lagrangian. This is accomplished with the following replacements:

$$m \to Z^{1/2} m, \quad g \to Z_1 Z_2^{-3/2} g, \quad D_a^\mu \to Z_2^{1/2} D_a^\mu .$$

Notice all fields are assumed to be renormalized by the same constant; this must be the case if $SU(2)$ symmetry is to remain unbroken by the renormalization procedure. The Lagrangian, Eq. (2.2), now becomes

$$\mathcal{L} = -\tfrac{1}{4} Z_2 A_{\mu\nu} A^{\mu\nu} + \tfrac{1}{2} Z Z_2 m^2 A_\mu A^\mu - \tfrac{1}{2} Z_2 B_{\mu\nu} B^{+\mu\nu} + Z Z_2 m^2 B_\mu B^{+\mu} \quad (4.1)$$

with

$$A_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - i Z_1 Z_2^{-1} g (B_\mu^+ B_\nu - B_\mu B_\nu^+),$$
$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu + i Z_1 Z_2^{-1} g (A_\mu B_\nu - A_\nu B_\mu).$$

The constants m and g now appearing in $\mathcal{L}$ are the observed mass and coupling constant.

The free field Lagrangian $\mathcal{L}_F$ is taken to be

$$\mathcal{L}_F = -\tfrac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) + \tfrac{1}{2} m^2 A_\mu A^\mu - \tfrac{1}{2}(\partial_\mu B_\nu - \partial_\nu B_\mu)(\partial^\mu B^{+\nu} - \partial^\nu B^{+\mu}) + m^2 B_\mu B^{+\mu}, \quad (4.2)$$

and the interaction Lagrangian $\mathcal{L}_I$ is defined by

$$\mathcal{L} = \mathcal{L}_F + \mathcal{L}_I,$$

so

$$\begin{aligned}\mathcal{L}_I = {} & Z_2(Z-1)\tfrac{1}{2} m^2 A_\mu A^\mu + Z_2(Z-1) m^2 B_\mu B^{+\mu} \\ & + (Z_2 - 1)[-\tfrac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu) \\ & \times (\partial^\mu A^\nu - \partial^\nu A^\mu) + \tfrac{1}{2} m^2 A_\mu A^\mu] \\ & + (Z_2 - 1)[-\tfrac{1}{2}(\partial_\mu B_\nu - \partial_\nu B_\mu) \\ & \times (\partial^\mu B^{+\nu} - \partial^\nu B^{+\mu}) + m^2 B_\mu B^{+\mu}] \\ & + i Z_1 g[\partial_\mu A_\nu (B^{+\mu} B^\nu - B^\mu B^{+\nu}) \\ & + \partial_\mu B_\nu (A^\mu B^{+\nu} - B^{+\mu} A^\nu) + \partial_\mu B_\nu^+ (B^\mu A^\nu - B^\mu A^\nu)] \\ & + Z_1^2 Z_2^{-1} g^2 [\tfrac{1}{2}(B_\mu^+ B_\nu B^{+\mu} B^\nu - B_\mu^+ B_\nu^+ B^\mu B^\nu) \\ & - (A_\mu B^{+\mu} A_\nu B^\nu - A_\mu A^\mu B_\nu^+ B^\nu)].\end{aligned}$$

B. Vertices

If broken lines represent neutral particles and solid directed lines represent charged particles, the arrow indicating the flow direction of positive charge, then

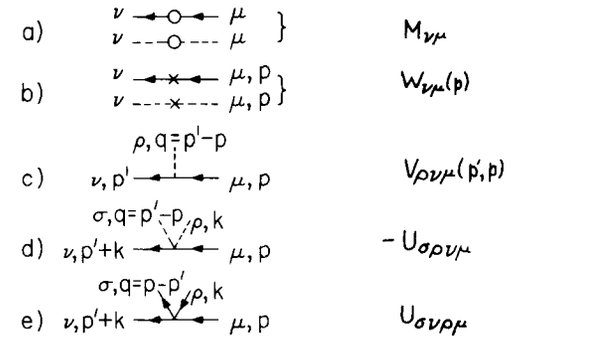


FIG. 1. The vertices associated with $\mathcal{L}_I$. The arrows on the charged lines indicate the flow direction of positive charge.

the vertices derived from $\mathcal{L}_I$ are those shown in Fig. 1. The symbols appearing there are defined by

$$M_{\nu\mu} = i(2\pi)^4 (Z-1) Z_2 m^2 g_{\nu\mu}, \quad (4.3)$$

$$W_{\nu\mu}(p) = -i(2\pi)^4 (Z_2 - 1)[g_{\nu\mu}(p^2 - m^2) - p_\nu p_\mu], \quad (4.4)$$

$$V_{\rho\nu\mu}(p',p) = i(2\pi)^4 Z_1 g[(p'+p)_\rho g_{\nu\mu} + (p'-2p)_\nu g_{\mu\rho} + (p-2p')_\mu g_{\rho\nu}], \quad (4.5)$$

$$U_{\sigma\rho\nu\mu} = i(2\pi)^4 Z_1^2 Z_2^{-1} g^2 [2 g_{\sigma\rho} g_{\nu\mu} - g_{\sigma\nu} g_{\mu\rho} - g_{\sigma\mu} g_{\rho\nu}]. \quad (4.6)$$

The following symmetry properties of the bare vertex functions are easily proved from their definitions:

$$\begin{aligned} V_{\rho\nu\mu}(p',p) &= -V_{\rho\nu\mu}(-p',-p) = V_{\rho\mu\nu}(p,p') \\ &= -V_{\mu\nu\rho}(p',p'-p) \\ &= -V_{\nu\rho\mu}(p-p',p), \end{aligned} \quad (4.7a)$$

$$V_{\rho\nu\mu}(p',p) + V_{\nu\mu\rho}(p,p) + V_{\mu\rho\nu}(p',p') = 0, \quad (4.7b)$$

$$U_{\sigma\rho\nu\mu} = U_{\nu\mu\sigma\rho} = U_{\rho\sigma\nu\mu} = U_{\sigma\rho\mu\nu}, \quad (4.7c)$$

$$U_{\sigma\rho\nu\mu} + U_{\sigma\nu\mu\rho} + U_{\sigma\mu\rho\nu} = 0. \quad (4.7d)$$

C. Propagators

That the charged and neutral particles have the same propagator is implied by Eq. (3.4). If their common free field propagator is denoted by $S^{\nu\mu}(p)$, then

$$S^{\nu\mu}(p) = -\frac{i}{(2\pi)^4} \frac{g^{\nu\mu} - (p^\nu p^\mu / m^2)}{p^2 - m^2 + i\epsilon}. \quad (4.8)$$

Its inverse $S_{\nu\mu}^{-1}(p)$, defined by

$$S_{\nu\rho}^{-1}(p) S^{\rho\mu}(p) = \delta_\nu^\mu,$$

is

$$S_{\nu\mu}^{-1}(p) = i(2\pi)^4 [g_{\nu\mu}(p^2 - m^2) - p_\nu p_\mu]. \quad (4.9)$$

Note that the bilinear vertex function $W_{\nu\mu}$, Eq. (4.4), may be written as

$$W_{\nu\mu}(p) = -(Z_2 - 1) S_{\nu\mu}^{-1}(p). \quad (4.10)$$

When amplitudes are computed, the free propagator will be modified by diagrams involving vertices. The effect of the bilinear vertices (Figs. 1a and 1b) can be calculated in closed form, so they need not be explicitly included on internal lines. To do this it is only necessary to realize that any two points in a diagram connected in lowest order by $S^{\nu\mu}(p)$ are connected to all orders by $\Delta^{\nu\mu}(p)$ defined as

$$\Delta = S + SWS + SWSWS + \cdots .$$

(The space–time indices have been suppressed for simplicity; matrix products are implied.) When Eq. (4.10) is used, this becomes

FIG. 2. The unrenormalized charged (a) and neutral (b) particle propagators.

$$\Delta = [1 - (Z_2 - 1) + (Z_2 - 1)^2 - \cdots]S$$

or

$$\Delta^{\nu\mu}(p) = Z_2^{-1}S^{\nu\mu}(p). \tag{4.11}$$

Similarly, two points connected by $\Delta^{\nu\mu}$ in one diagram will be effectively joined by $D^{\nu\mu}$ in a complete amplitude, where

$$D = \Delta + \Delta M\Delta + \Delta M\Delta M\Delta + \cdots .$$

With the definition of $M_{\nu\mu}$, Eq. (4.3), and of $\Delta^{\nu\mu}$, Eq. (4.11), it may be shown that

$$D^{\nu\mu}(p) = -\frac{i}{(2\pi)^4} Z_2^{-1} \frac{g^{\nu\mu} - (p^\nu p^\mu / Zm^2)}{p^2 - Zm^2 + i\epsilon}. \tag{4.12}$$

Clearly, $D^{\nu\mu}(p)$ is an unrenormalized propagator which includes all contributions from the bilinear vertex counterterms. Its graphical representations are given in Fig. 2.

The relationship of $S^{\nu\mu}$ to $D^{\nu\mu}$ is most directly expressible in terms of their inverses. The inverse of $D^{\nu\mu}$ turns out to be

$$D^{-1}_{\nu\mu} = i(2\pi)^4 Z_2[g_{\nu\mu}(p^2 - Zm^2) - p_\nu p_\mu], \tag{4.13}$$

and the desired relation is

$$D^{-1}_{\nu\mu}(p) = S^{-1}_{\nu\mu}(p) - M_{\nu\mu} - W_{\nu\mu}(p). \tag{4.14}$$

D. Amplitudes

The elements introduced above are to be assembled into diagrams and associated amplitudes according to rules following from the standard Dyson–Wick procedure, with one notable addition. It turns out to be expedient to compute and include contractions of two free fields with the same space–time argument, instead of defining them to be zero. The reason is not uncovered until later in the development, so the detailed discussion of this point is postponed until Sec. 10A. For now, it is simply noted that diagrams resulting from the new rule are understood to be included in any calculation described in the remaining sections.

5. THE BARE WARD RELATIONS

Before going on to the consideration of more general diagrams, some simple relations involving the unrenormalized propagator and the bare vertices are presented here.

From the definition of $V_{\rho\nu\mu}(p', p)$, Eq. (4.5), and of $D^{-1}_{\nu\mu}(p)$, Eq. (4.13), it follows that

$$(p' - p)^\rho V_{\rho\nu\mu}(p', p) = gZ_1Z_2^{-1}[D^{-1}_{\nu\mu}(p') - D^{-1}_{\nu\mu}(p)] \tag{5.1}$$

or

$$(p' - p)^\gamma D^{\nu\beta}(p') V_{\gamma\beta\alpha}(p', p) D^{\alpha\mu}(p) = -gZ_1Z_2^{-1}[D^{\nu\mu}(p') - D^{\nu\mu}(p)]. \tag{5.2a}$$

Similarly,

$$D^{\nu\beta}(p') V_{\gamma\beta\alpha}(p', p) D^{\gamma\mu}(p' - p) p^\alpha = gZ_1Z_2^{-1}[D^{\nu\mu}(p') - D^{\nu\mu}(p' - p)] \tag{5.2b}$$

and

$$p'^\beta D^{\nu\gamma}(p' - p) V_{\gamma\beta\alpha}(p', p) D^{\alpha\mu}(p) = -gZ_1Z_2^{-1}[D^{\nu\mu}(p' - p) - D^{\nu\mu}(p)]. \tag{5.2c}$$

These last equations are Ward–Takahashi relations (WTR). The first, (5.2a), is directly analogous to the one found in spinor electrodynamics.[11] The other two, Eqs. (5.2b) and (5.2c), exist as a consequence of the extra symmetry in the Yang–Mills theory.

A WTR for $U_{\sigma\rho\nu\mu}$ follows trivially from the definitions of $U_{\sigma\rho\nu\mu}$ and $V_{\rho\nu\mu}(q, q)$, Eqs. (4)–(6). It is

$$q^\sigma U_{\sigma\rho\nu\mu} = gZ_1Z_2^{-1} V_{\rho\nu\mu}(q, q).$$

Using the definition of $V_{\rho\nu\mu}(p', p)$, this may be written in a more complicated, but equivalent, form that will be needed in Sec. 7D2: namely

$$(p' - p)^\sigma U_{\sigma\rho\nu\mu} = gZ_1Z_2^{-1}[V_{\rho\nu\mu}(p' + k, p') - V_{\rho\nu\mu}(p + k, p)]. \tag{5.3}$$

Other WTR's may be inferred from the symmetries of $U_{\sigma\rho\nu\mu}$ Eq. (4.7c).

Ward relations are derived from the WTR's above by allowing p to approach p'. For example,

$$D^{\nu\beta}(p) V_{\rho\beta\alpha}(p, p) D^{\alpha\mu}(p) = -gZ_1Z_2^{-1} \frac{\partial}{\partial p^\rho} D^{\nu\mu}(p) \tag{5.4}$$

and

$$U_{\sigma\rho\nu\mu} = gZ_1Z_2^{-1} \frac{\partial}{\partial p^\sigma} V_{\rho\nu\mu}(p, p). \tag{5.5}$$

These equations have the usual physical interpretation: Differentiation of a diagram with respect to a momentum has the effect of inserting into that diagram an external neutral line with zero momentum. Equation (5.4) shows the new line is added to a propagator. Equation (5.5) shows the new line is added at a V vertex, changing it to a U vertex.

It is interesting to note that all the results in this section which involve $D^{\nu\mu}$ remain valid if $D^{\nu\mu}$ is replaced by $\Delta^{\nu\mu}$. Generally speaking, the bare WTR's are the same for any nonzero value of the mass appearing in the propagators.

6. THE GENERALIZED LEE THEOREM

Some facts about certain important sets of diagrams will now be established. They are applied later to generalize the WTR's of Sec. 5 and to prove a Ward identity. The theorem and corollaries presented here are direct extensions of analogous theorems proved by Lee in connection with another theory of vector electrodynamics.[12]

Definition 1: Let $G_{\nu\mu}(c; p)$ be the algebraic expression corresponding to an arbitrary diagram of the general form shown in Fig. 3.[13] The main charged particle line of such diagrams shall, henceforth, be known as the trunk. c specifies the particular configuration of the diagram; that is, the number of the various types of vertices attached to the trunk and the order in which they are arranged. All configurations considered have at least one vertex. p is the momentum entering the trunk; the momentum leaving

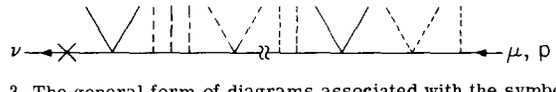


FIG. 3. The general form of diagrams associated with the symbols $G_{\nu\mu}(c; p)$. The momenta and indices of the branches attached to the trunk are suppressed; this information is formally contained in the index c.

the opposite end of the trunk is, of course, p plus the sum of the net momenta entering through the "branches" at each vertex. Whereas the trilinear and quadrilinear vertices of Figs. 1c–e may appear anywhere along the trunk, the bilinear vertices, Figs. 1a and 1b can only occur at the trunk ends in diagrams comprising scattering amplitudes because their appearance elsewhere has been taken into account by using $D^{\nu\mu}(p)$ as the unrenormalized propagator (see Sec. 4C).

Definition 2: $G_{\rho\nu\mu}(c;p,q)$ is defined to be the sum of all distinct diagrams derivable by attaching one additional neutral line, with index ρ and momentum q, to the trunk of the configuration c.

To illustrate these definitions and obtain some information needed for the Theorem, consider this

Lemma: Denote by $G_{\nu\mu}(1;p)$ any one of the five charged line vertices shown in Fig. 1. These are the simplest configurations. Then

$$q^\rho G_{\rho\nu\mu}(1;p,q) = gZ_1Z_2^{-1}[D^{-1}_{\mu\alpha}(p')D^{\alpha\beta}(p'-q)G_{\beta\mu}(1;p) - G_{\nu\alpha}(1;p+q)D^{\alpha\beta}(p+q)D^{-1}_{\beta\mu}(p)], \quad (6.1)$$

where p' is the momentum leaving the trunks of the summands of $G_{\rho\nu\mu}(1;p,q)$.

Proof: The five $G_{\nu\mu}(1)$'s together with the associated $G_{\rho\nu\mu}(1)$'s prescribed by Def. 2 are given in Fig. 4. The method of proof is straightforward calculation. Only the most complicated case, 1c, shall be worked out here.

When the Feynman rules of Sec. 4 are used to translate the diagrams for $G_{\rho\nu\mu}(1c;p,q)$ into a momentum space expression, the result is

$$G_{\rho\nu\mu}(1c;p,q) = V_{\sigma\nu\alpha}(p',p+q)D^{\alpha\beta}(p+q)V_{\rho\beta\mu}(p+q,p) + V_{\rho\nu\alpha}(p',p+k)D^{\alpha\beta}(p+k)V_{\sigma\beta\mu}(p+k,p) - U_{\sigma\rho\nu\mu}.$$

The bare WTR's of Section V then imply

$$q^\rho G_{\rho\nu\mu}(1c;p,q) = gZ_1Z_2^{-1}V_{\sigma\nu\alpha}(p',p+q)D^{\alpha\beta}(p+q) \times [D^{-1}_{\beta\mu}(p+q) - D^{-1}_{\beta\mu}(p)] + gZ_1Z_2^{-1}[D^{-1}_{\nu\alpha}(p') - D^{-1}_{\nu\alpha}(p+k)]D^{\alpha\beta}(p+k)V_{\sigma\beta\mu}(p+k,p) - gZ_1Z_2^{-1}V_{\sigma\nu\mu}(q,q),$$

which reduces to

$$q^\rho G_{\rho\nu\mu}(1c;p,q) = gZ_1Z_2^{-1}[D^{-1}_{\nu\alpha}(p')D^{\alpha\beta}(p+k)V_{\sigma\beta\mu}(p+k,p) - V_{\sigma\nu\alpha}(p',p+q)D^{\alpha\beta}(p+q)D^{-1}_{\beta\mu}(p)]$$

or

$$q^\rho G_{\rho\nu\mu}(1c;p,q) = gZ_1Z_2^{-1}[D^{-1}_{\nu\alpha}(p')D^{\alpha\beta}(p'-q)G_{\beta\mu}(1c;p) - G_{\nu\alpha}(1c;p+q)D^{\alpha\beta}(p+q)D^{-1}_{\beta\mu}(p)]. \quad (6.2)$$

In the reduction, the relation

$$V_{\sigma\nu\alpha}(p+q+k,p+q) - V_{\sigma\nu\mu}(p+k,p) - V_{\sigma\nu\mu}(q,q) = 0$$

was used. It follows directly from the linearity of $V_{\rho\nu\mu}(p',p)$ in its two variables [see Eq. (4.5)].

With the development leading to Eq. (6.2), the Lemma has been proved for case 1c. Similar calculations establish it for the remaining cases. QED

Theorem (the generalized Lee theorem): for *any* configuration c,

$$q^\rho G_{\rho\nu\mu}(c;p,q) = gZ_1Z_2^{-1}[D^{-1}_{\nu\alpha}(p')D^{\alpha\beta}(p'-q)G_{\beta\mu}(c,p) - G_{\nu\alpha}(c;p+q)D^{\alpha\beta}(p+q)D^{-1}_{\beta\mu}(p)], \quad (6.3)$$

where p' is the momentum leaving the trunks of the summands of $G_{\rho\nu\mu}(c;p,q)$.

Proof: Each possible configuration may be classified according to which of the five types of vertices (see Fig. 1) terminates its trunk. The Theorem must then be proved for each class. Details are outlined below for the class terminating with the vertex of Fig. 1e; this is a case not included in Lee's work.[12]

If $G_{\nu\mu}(c;p)$ is within the class of configurations under consideration, it may be written as

$$G_{\nu\mu}(c;p) = U_{\sigma_2\nu\sigma_1\alpha}D^{\alpha\beta}(p'')G_{\beta\mu}(c';p),$$

where p'' is the momentum leaving the trunk of $G_{\beta\mu}(c';p)$, and σ_1, σ_2 are the indices of the charged line "branches" which distinguish c from c'. In words, c is derived from another configuration c' by attaching the appropriate vertex. Definition 2 leads to this equation

$$G_{\rho\nu\mu}(c;p,q) = V_{\rho\nu\alpha}(p',p'-q)D^{\alpha\beta}(p'-q) \times U_{\sigma_2\beta\sigma_1\gamma}D^{\gamma\delta}(p'')G_{\delta\mu}(c';p) + U_{\sigma_2\nu\sigma_1\alpha}D^{\alpha\beta}(p''+q)G_{\rho\beta\mu}(c';p,q).$$

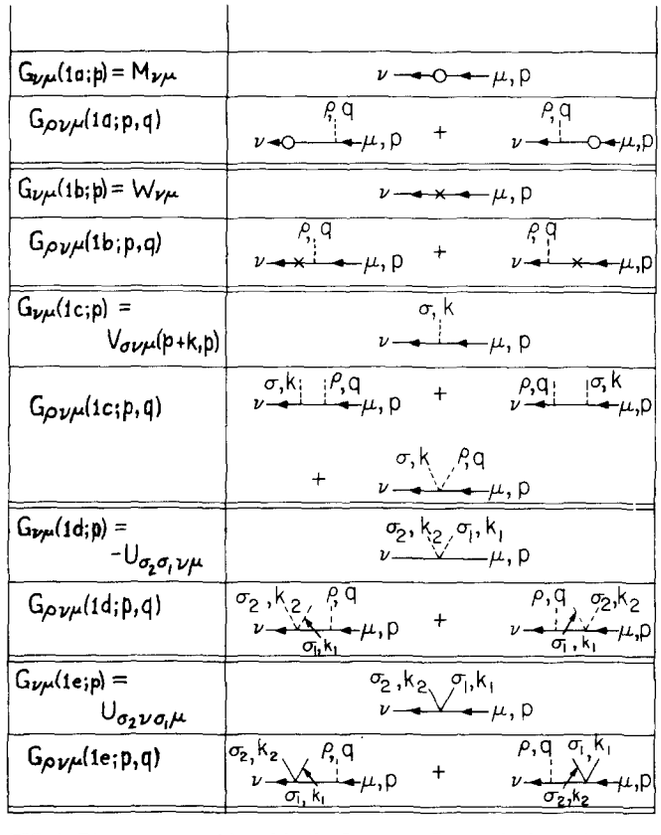


FIG. 4. Diagrams considered in the Lemma. For each elementary configuration $G_{\nu\mu}(1;p)$ the appropriate contributions to the three index symbol $G_{\rho\nu\mu}(1;p,q)$ are given. The latter are found in accordance with Def. 2.

If the theorem is assumed valid for $G_{\rho\beta\mu}(c';p,q)$, then a little computation, making use of the bare WTR's in Sec. 5, shows it to be also true for $G_{\rho\nu\mu}(c;p,q)$.

With parallel arguments it may be shown that adding a vertex of any type to a configuration that satisfies the Theorem produces another configuration which also satisfies it. But the previous Lemma, Eq. (6.1), established the Theorem for each elementary configuration. Therefore it is true in all cases by induction.

QED

Definition 3: The quantity $F_{\rho\nu\mu}(c;p,q)$ is defined by

$$F_{\rho\nu\mu}(c;p,q) = G_{\rho\nu\mu}(c;p,q) - V_{\rho\nu\alpha}(p',p'-q)D^{\alpha\beta}(p'-q)G_{\beta\mu}(c;p) - G_{\nu\alpha}(c;p+q)D^{\alpha\beta}(p+q)V_{\rho\beta\mu}(p+q,p),$$

i.e., $F_{\rho\nu\mu}$ is formed in the same way as $G_{\rho\nu\mu}$ except that contributions from the two diagrams with the neutral line attached to an end of the trunk of $G_{\nu\mu}$ are omitted.

Corollary 1:

$$q^\rho F_{\rho\nu\mu}(c;p,q) = gZ_1Z_2^{-1}[G_{\nu\mu}(c;p) - G_{\nu\mu}(c;p+q)]. \tag{6.4}$$

Proof: This relation follows from Eq. (6.3), Def. 3, and the bare WTR's after a straightforward calculation.

Definition 4: Define $G_\rho(c;p,q)$ by

$$G_\rho(c;p,q) = D^{\mu\nu}(p)[F_{\rho\nu\mu}(c;p,q) + V_{\rho\nu\alpha}(p,p-q)D^{\alpha\beta}(p-q)G_{\beta\mu}(c;p)].$$

Graphically, $G_\rho(c)$ is the sum of all distinct diagrams found by attaching one neutral line to $G_{\nu\mu}(c)$ after its trunk has been closed into a loop and before the internal loop momentum p has been integrated over.

Corollary 2:

$$q^\rho G_\rho(c;p,q) = gZ_1Z_2^{-1}[D^{\mu\nu}(p-q)G_{\nu\mu}(c;p) - D^{\mu\nu}(p)G_{\nu\mu}(c;p+q)]$$

and

$$q^\rho \int dp\, G_\rho(c;p,q) = 0. \tag{6.5}$$

Proof: The first equation may be inferred directly from Def. 4, Eq. (6.4) and the bare WTR's of Sec. 5. The second is found when a displacement of the integration variable p is made.[14]

7. GENERAL WARD–TAKAHASHI RELATIONS

A. Total Amplitudes

The corollaries of the preceding section were proved by singling out the neutral lines for special consideration. It will be shown here that the $SU(2)$ invariance of the theory of "vector electrodynamics" being investigated implies additional relations in which charged lines play the role previously assigned only to neutrals.

As a prelude, note that scattering amplitudes have this form:

$$\frac{\epsilon^\sigma_{\lambda_N}(p_N)}{(2\pi)^{3/2}} \cdots \frac{\epsilon^\nu_{\lambda_2}(p_2)}{(2\pi)^{3/2}} \frac{\epsilon^\mu_{\lambda_1}(p_1)}{(2\pi)^{3/2}} \times T_{\sigma\ldots\nu\mu}(p_N,\ldots,p_2,p_1)\,\delta(\Sigma p_{\text{in}} - \Sigma p_{\text{out}}). \tag{7.1}$$

Each factor $(2\pi)^{-3/2}\,\epsilon^\mu_\lambda(p)$ represents an incoming or outgoing particle, $\epsilon^\mu_\lambda(p), \lambda = 1,2,3$, being its appropriate polarization vector.

The T functions are calculated by summing all diagrams contributing to the process in question. They may also be expressed in terms of time ordered Wightman functions, and the explicit *LSZ* formulas are given in Appendix A. The symmetries of the T functions corresponding to various processes are derived from the latter formulation by applying relations of the type given in Sec. 3.

The results of Sec. 3 directly imply the symmetry properties of total amplitudes only. But certain important partial sums of diagrams, such as self-energies and corrected vertices, also have symmetries that can be used to derive more general WTR's than are allowed by the corollaries of Sec. 6 alone. Specific examples are presented in the remainder of this section.

B. The Total Propagator

As a consequence of Eq. (3.4), the total propagator $S'^{\nu\mu}(p)$ is common to both kinds of particle. It is given by

$$S' = D + D\Pi D + D\Pi D\Pi D + \cdots,$$
$$S' = D(1 - \Pi D)^{-1}, \tag{7.2}$$

or, alternatively,

$$S'^{-1} = D^{-1} - \Pi, \tag{7.3}$$

where $\Pi_{\nu\mu}(p)$ is the sum of all proper "self-energy" diagrams associated with either the charged or neutral particles. $\Pi_{\nu\mu}(p)$ includes no explicit bilinear vertices since the unrenormalized propagator $D^{\nu\mu}$ has been used throughout, and it already contains the contributions of the counterterms (see Sec. 4C).

Interpreted as the neutral particle self-energy, $\Pi_{\nu\mu}$ is the sum of all proper diagrams with two neutral external lines. The summands may be constructed as follows: (1) Write down all proper diagrams with one external line—a neutral line labeled by μ, and call these the primary diagrams; (2) attach one more neutral line, labeled by ν, in all possible ways to the primaries to find secondary diagrams; (3) sum the secondary diagrams to obtain $\Pi_{\nu\mu}$. All the charged lines in the primary diagrams must be closed into loops to conserve charge, so the second step creates the quantities $G_\nu(c;q,p)$ of Def. 4 as pieces of the secondary diagrams. Furthermore, since all internal loop momenta, here denoted q, are to be integrated over in accordance with Feynman's rules, the self-energy contributions are seen to be expressible in terms of the type of integral appearing in Corollary 2, Eq. (6.5). The three steps outlined above which lead to this conclusion are illustrated with a particular example in Appendix C.

From the considerations in the previous paragraph, it follows that Eq. (6.5) may be applied to all the terms constituting the function $\Pi_{\nu\mu}(p)$; consequently

$$p^{\nu}\Pi_{\nu\mu}(p) = 0. \tag{7.4}$$

It is concluded that $\Pi_{\nu\mu}$ must have the form

$$\Pi_{\nu\mu}(p) = (g_{\nu\mu}p^2 - p_{\nu}p_{\mu})F(p^2). \tag{7.5}$$

C. The Total Trilinear Vertex

With the $SU(2)$ relations (3.5) and (3.6), space–time reflection invariance and the LSZ formulas of Appendix A, it may be shown that the T function associated with the total amplitude for the process in Fig. 5[15] has the exchange symmetries (4.7a) previously found for $V_{\rho\nu\mu}(p',p)$.

Denote the total trilinear vertex by $V'_{\rho\nu\mu}(p',p)$. It differs from the T function by terms involving external line corrections. But the total external charged line correction is the same as that of the external neutral lines, because both depend only upon $S'^{\nu\mu}$, $\Pi_{\nu\mu}$ and the bilinear vertices of Figs. 1a and 1b.[16] So it might be expected that external line corrections do not affect symmetry properties, and, in fact, it is easily proved that $V'_{\rho\nu\mu}$ has the same exchange properties as the T function and $V_{\rho\nu\mu}$. Specifically

$$\begin{aligned} V'_{\rho\nu\mu}(p',p) &= -V'_{\rho\nu\mu}(-p',-p) = V'_{\rho\mu\nu}(p,p') \\ &= -V'_{\mu\nu\rho}(p',p'-p) = -V'_{\nu\rho\mu}(p-p',p). \end{aligned} \tag{7.6}$$

The total vertex is given by

$$V'_{\rho\nu\mu}(p',p) = V_{\rho\nu\mu}(p',p) + v_{\rho\nu\mu}(p',p). \tag{7.7}$$

where $v_{\rho\nu\mu}(p',p)$ is the total correction to the bare vertex. Generally speaking, corrections to the bare vertex are generated by attaching a neutral line in all possible ways to the charged particle self-energy diagrams. The total correction is the sum of all the ensuing diagrams except those with the form shown in Fig. 6. The latter are excluded, because they are external line corrections rather than contributions to the vertex. Hence the relation of $\Pi_{\nu\mu}$ to $v_{\rho\nu\mu}$ is just such that Corollary 1, Eq. (6.4), leads to

$$(p'-p)^{\rho}v_{\rho\nu\mu}(p',p) = gZ_1Z_2^{-1}[\Pi_{\nu\mu}(p) - \Pi_{\nu\mu}(p')]. \tag{7.8}$$

And this, together with Eq. (7.3) and the bare WTR (5.1), implies

$$(p'-p)^{\rho}V'_{\rho\nu\mu}(p',p) = gZ_1Z_2^{-1}[S'^{-1}_{\nu\mu}(p') - S'^{-1}_{\nu\mu}(p)]. \tag{7.9}$$

Two more WTR's may be inferred with the aid of Eqs. (7.6).

The conclusions of this subsection may be summarized as follows: In the bare WTR's (5.2), make the replacements

$$V_{\rho\nu\mu} \to V'_{\rho\nu\mu} \quad \text{and} \quad D^{\nu\mu} \to S'^{\nu\mu};$$

the resulting equations are correct WTR's for the total trilinear vertex.

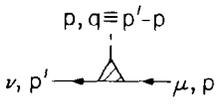


FIG. 5. The form of the diagrams contributing to the three particle T function and $V'_{\rho\nu\mu}(p'p)$.

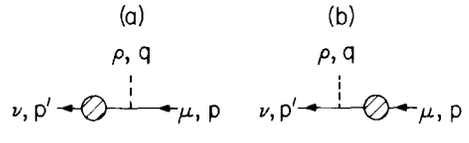


FIG. 6. Diagrams omitted from the corrected vertex $V'_{\rho\nu\mu}(p'p)$.

D. The Total Quadrilinear Vertices

1. Four Neutral External Lines

Let $\bar{U}_{\sigma\rho\nu\mu}(k,p'-k,p)$ be the sum of all proper diagrams of the type shown in Fig. 7. All momenta have been directed inward, so $\bar{U}_{\sigma\rho\nu\mu}$ is clearly symmetric under the interchange of any two external lines. There is of course no bare contribution to this vertex.

Since all charged lines in the summands of $\bar{U}_{\sigma\rho\nu\mu}$ must form closed loops, it follows from Corollary 2, Eq. (6.5), by an argument paralleling the one used in the case of $\Pi_{\nu\mu}$ (Sec. 7B) that

$$\begin{aligned} p^{\mu}\bar{U}_{\sigma\rho\nu\mu}(k,p'-k,p) &= (p'-k)^{\nu}\bar{U}_{\sigma\rho\nu\mu}(k,p'-k,p) \\ &= k^{\rho}\bar{U}_{\sigma\rho\nu\mu}(k,p'-k,p) = q^{\sigma}\bar{U}_{\sigma\rho\nu\mu}(k,p'-k,p) = 0. \end{aligned} \tag{7.10}$$

These last relations are directly analogous to the similar result for light–light scattering in spinor electrodynamics. As in that case, they mean the constant term in a Taylor expansion of $\bar{U}_{\sigma\rho\nu\mu}$ vanishes.

However, it is not clear in the present instance that the remaining terms are finite, as they are in spinor electrodynamics.

2. Two Neutral-Two Charged External Lines

Consider the process in Fig. 8. Obviously, the total amplitude and its associated T function are invariant under exchange of the two neutral lines. Other exchange symmetries are found from $SU(2)$ relations such as (3.7).

Let $t'_{\sigma\rho\nu\mu}(k;p'+k,p)$ be the sum of all diagrams contributing to the process except those involving external line corrections. As in the trilinear vertex case (Sec. 7C), it may be demonstrated that $t'_{\sigma\rho\nu\mu}$ has the same symmetries as the T function. However, $t'_{\sigma\rho\nu\mu}$ is not the appropriate quadrilinear vertex, because it still includes diagrams of the type shown in Fig. 9. If the sum of these contributions is denoted $\delta t'_{\sigma\rho\nu\mu}(k;p'+k,p)$, then the Feynman rules yield

$$\begin{aligned} \delta t'_{\sigma\rho\nu\mu}(k;p'+k,p) &= V'_{\sigma\nu\mu}(p'+k,p+k)S'^{\beta\alpha}(p+k)V'_{\rho\alpha\mu}(p+k,p) \\ &+ V'_{\rho\nu\beta}(p'+k,p+q)S'^{\beta\alpha}(p+q)V'_{\sigma\alpha\mu}(p+q,p), \end{aligned} \tag{7.11}$$

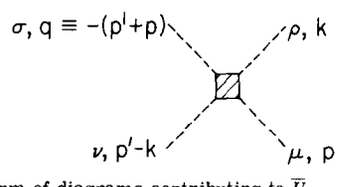


FIG. 7. The form of diagrams contributing to $\bar{U}_{\sigma\rho\nu\mu}(k,p'-k,p)$. All neutral particles are assumed to be incoming.

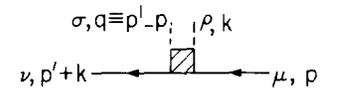


FIG. 8. The form of diagrams contributing to $t'_{\sigma\rho\nu\mu}(k;p'+k,p)$ and $U'_{\sigma\rho\nu\mu}(k;p'+k,p)$.

and the total corrected vertex $U'_{\sigma\rho\nu\mu}(k;p'+k,p)$ is defined by

$$U'_{\sigma\rho\nu\mu}(k;p'+k,p) = t'_{\sigma\rho\nu\mu}(k;p'+k,p) - \delta t'_{\sigma\rho\nu\mu}(k;p'+k,p). \quad (7.12)$$

The symmetries of $V'_{\rho\nu\mu}$, Eqs. (7.6), may be used to prove $\delta t'_{\sigma\rho\nu\mu}$ behaves under exchanges in the same way as the T function and $t'_{\sigma\rho\nu\mu}$, so $U'_{\sigma\rho\nu\mu}$ also has the same properties. These well-advertised symmetries are now finally revealed in terms of $U'_{\sigma\rho\nu\mu}$:

$$U'_{\sigma\rho\nu\mu}(k;p'+k,p) = U'_{\sigma\rho\mu\nu}(k;-p,-p'-k) = U'_{\rho\sigma\nu\mu}(q;p'+k,p) = U'_{\nu\mu\sigma\rho}(p;-q,k). \quad (7.13)$$

The vertex corrections to the bare process may be constructed from the diagrams for the quantity $v_{\rho\nu\mu}(p',p)$ introduced in Eq. (7.7) by attaching an extra neutral line at all possible points, except along the incoming or outgoing charged lines (the latter contribute to $\delta t'_{\sigma\rho\nu\mu}$). Hence Corollary 1 is again applicable, and together with the bare WTR (5.3), yields

$$q^\sigma U'_{\sigma\rho\nu\mu}(k;p'+k,p) = gZ_1Z_2^{-1}[V'_{\rho\nu\mu}(p+k,p) - V'_{\rho\nu\mu}(p'+k,p')]. \quad (7.14)$$

Additional WTR's are implied by the symmetry equations (7.13).

3. Four Charged External Lines

$t''_{\sigma\rho\nu\mu}(k;p'+k,p)$ is defined to be the sum of all diagrams with the form shown in Fig. 10, except those with external line corrections. Since it turns out that symmetry relations between such quantities are the same as those among the total amplitudes and their associated T functions, it may be inferred from the $SU(2)$ equation (3.8) that

$$t''_{\sigma\rho\nu\mu}(k;p'+k,p) = \bar{U}_{\sigma\rho\nu\mu}(-k,p'+k,p) - t'_{\sigma\rho\nu\mu}(-k;-p'-k,p). \quad (7.15)$$

With the help of Eq. (7.12), this becomes

$$t''_{\sigma\rho\nu\mu}(k;p'+k,p) + \delta t'_{\sigma\rho\nu\mu}(-k;-p'-k,p) = \bar{U}_{\sigma\rho\nu\mu}(-k,p'-k,p) - U'_{\sigma\rho\nu\mu}(-k;-p'-k,p). \quad (7.16)$$

But Eqs. (7.6) and (7.11) may be utilized to find

$$\begin{aligned}\delta t'_{\sigma\rho\nu\mu}(-k;-p'-k,p) &= -V'_{\beta\sigma\nu}(p'+p,p'+k)S'^{\beta\alpha}(p-k)V'_{\alpha\rho\mu}(k,p) \\ &\quad - V'_{\beta\rho\nu}(k,p'+k)S'^{\beta\alpha}(p')V'_{\alpha\sigma\mu}(p'+p,p) \\ &\equiv -\delta t''_{\sigma\rho\nu\mu}(k;p'+k,p),\end{aligned}$$

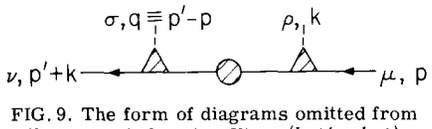


FIG. 9. The form of diagrams omitted from the corrected vertex $U'_{\sigma\rho\nu\mu}(k;p'+k,p)$.

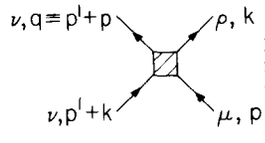


FIG. 10 The form of diagrams contributing to $t''_{\sigma\rho\nu\mu}(k;p'+k,p)$ and $U''_{\sigma\rho\nu\mu}(k;p'+k,p)$.

and $\delta t''_{\sigma\rho\nu\mu}$ is seen upon inspection of the above expressions to be the sum of all the diagrams with the form of Fig. 11. These are precisely the contributions that must be removed from $t''_{\sigma\rho\nu\mu}$ to arrive at the corrected vertex $U''_{\sigma\rho\nu\mu}$ for the process in question. More exactly,

$$U''_{\sigma\rho\nu\mu}(k;p'+k,p) = t''_{\sigma\rho\nu\mu}(k;p'+k,p) - \delta t''_{\sigma\rho\nu\mu}(k;p'+k,p).$$

Equation (7.16) then reads

$$U''_{\sigma\rho\nu\mu}(k;p'+k,p) = \bar{U}_{\sigma\rho\nu\mu}(-k,p'+k,p) - U'_{\sigma\rho\nu\mu}(-k;-p'-k,p). \quad (7.17)$$

The exchange symmetries of $U''_{\sigma\rho\nu\mu}$ may be found from the relation above or inferred from the LSZ formula for the total amplitude (Appendix A). They are the following:

$$U''_{\sigma\rho\nu\mu}(k;p'+k,p) = U''_{\sigma\rho\mu\nu}(k;p,p'+k) = U''_{\rho\sigma\nu\mu}(q;p'+k,p) = U''_{\nu\mu\sigma\rho}(p;q,k). \quad (7.18)$$

When the WTR's for $\bar{U}_{\sigma\rho\nu\mu}$, Eq. (7.10), and $U'_{\sigma\rho\nu\mu}$, Eq. (7.14), are employed in conjunction with Eq. (7.17), the result is

$$q^\sigma U''_{\sigma\rho\nu\mu}(k;p'+k,p) = -q^\sigma U'_{\sigma\rho\nu\mu}(-k;-p'-k,p)$$

or

$$q^\sigma U''_{\sigma\rho\nu\mu}(k;p'+k,p) = gZ_1Z_2^{-1}[V'_{\rho\nu\mu}(p-k,p) + V'_{\rho\nu\mu}(p'+k,p')]. \quad (7.19)$$

Once again, other WTR's may be derived by taking advantage of the exchange symmetries (7.18).

E. Concluding Remarks

Two of the relations found in this section, (7.9) and (7.14), are similar to WTR's appearing in Lee's work,[12] although in the present instance more bare vertices have been included. Equation (7.19) and the WTR's obtained by applying the exchange symmetries (7.6), (7.13), and (7.18) are new, however, and indigenous to the massive Yang-Mills theory.

Incidentally, WTR's similar to those presented above may be derived, in principle, without specializing to the $SU(2)$ case or identifying the fields with charged and neutral particles. A generalized Lee theorem similar to the one in Sec. 6 has been proved by the author for the $SU(n)$ massive Yang-Mills theory. However, additional terms appear on the right-hand

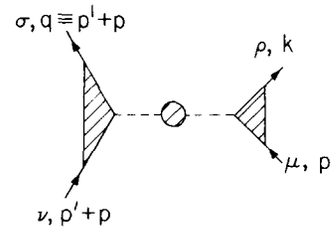


FIG. 11. The form of diagrams omitted from the correct vertex $U''_{\sigma\rho\nu\mu}(k;p'+k,p)$.

side of the analog of Eq. (6.3), and these complicate the WTR's among the total vertices. Therefore, only the simple, easily interpreted $SU(2)$ case has been described here.

The method employed and the WTR's derived above constitute the major results of this work. The remainder of the paper consists essentially of an application and a discussion of some related problems.

8. FORMAL RENORMALIZATION

A. Introduction

As an application of a general Ward–Takahashi relation, a Ward identity shall be derived in the next section. Here, some preliminaries are presented; it is necessary to develop the conditions which formally determine the renormalization constants.

The total propagator and vertices satisfying Ward–Takahashi relations were discussed in Sec. 7. They already contain renormalization counter terms, and they become the renormalized propagator and vertices when specific values are chosen for the constants Z, Z_1, and Z_2. The conditions used to evaluate renormalization constants are always partly conventional, and this is expecially true in the present case, since the physical system described by the massive Yang–Mills theory has not been specified. For definiteness, physical conditions shall be imposed which are analogous to the ones appearing in spinor electrodynamics.

Once the conditions are postulated, the theory is formally renormalized, although it is not yet known whether this procedure removes the divergences.

A new expression for the propagator will prove useful in what follows. Note that if Eq. (4.14) is used in Eq. (7.3), the result is

$$S'^{-1} = S^{-1} - \Pi', \tag{8.1}$$

where

$$\Pi'_{\nu\mu}(p) \equiv \Pi_{\nu\mu}(p) + M_{\nu\mu} + W_{\nu\mu}(p). \tag{8.2}$$

A little manipulation produces this relation:

$$S' = S + S'\Pi'S. \tag{8.3}$$

B. The Mass Condition

If $S'^{\nu\mu}(p)$ is to be the appropriate renormalized propagator, it must have a pole at the physical mass squared, that is at m^2, with residue one.[17] More exactly, the mass condition shall be expressed by

$$S'^{\nu\rho}(p)S^{-1}_{\rho\mu}(p)\epsilon_\lambda^{\ \mu}(p) \xrightarrow[p^2\to m^2]{} \epsilon_\lambda^{\ \nu}(p); \tag{8.4}$$

it may presumably be satisfied by a particular choice of the renormalization constants appearing in $S'^{\nu\mu}(p)$. Equation (8.3) is then seen to imply

$$S'^{\nu\rho}(p)\Pi'_{\rho\mu}(p)\epsilon_\lambda^{\ \mu}(p) \xrightarrow[p^2\to m^2]{} 0. \tag{8.5}$$

C. The Wavefunction Normalization Condition

When the LSZ formalism is utilized to compute the S matrix $\mathcal{S}$ for the process in which one particle comes in with momentum p and polarization λ and goes out with momentum p' and polarization λ', the result is

$$\begin{aligned}\mathcal{S}(p',\lambda';p,\lambda) &= 2p_0\delta_{\lambda\lambda'}\delta(\mathbf{p}'-\mathbf{p}) \\ &\quad + [\epsilon^\nu_{\lambda'}(p')/(2\pi)^{3/2}]S^{-1}_{\nu\sigma}(p')S'^{\sigma\rho}(p)S^{-1}_{\rho\mu}(p) \\ &\quad \times [\epsilon^\mu_\lambda(p)/(2\pi)^{3/2}]\delta(p'-p).\end{aligned}$$

But with proper normalization and the conventions of Appendix A,

$$\mathcal{S}(p',\lambda';p,\lambda) = 2p_0\delta_{\lambda\lambda'}\delta(\mathbf{p}'-\mathbf{p}) \tag{8.6}$$

must hold for physical states. Hence the condition

$$\epsilon^\nu_{\lambda'}(p)S^{-1}_{\nu\sigma}(p)S'^{\sigma\rho}(p)S^{-1}_{\rho\mu}(p)\epsilon_\lambda^{\ \mu}(p) \xrightarrow[p^2\to m^2]{} 0.$$

When Eqs. (8.3) and (8.4) are taken into account, this becomes

$$\epsilon^\nu_{\lambda'}(p)\Pi'_{\nu\mu}(p)\epsilon^\mu_\lambda(p) \xrightarrow[p^2\to m^2]{} 0. \tag{8.7}$$

D. The Charge Condition

Consider the matrix element

$$\begin{aligned}&{}_{\text{out}}\langle p',\lambda'|a^+_{\text{in}}(q,\lambda'')|p,\lambda\rangle_{\text{in}} \\ &\quad \equiv \frac{\epsilon^\rho_{\lambda''}(q)}{(2\pi)^{3/2}}\ \frac{\epsilon^\nu_{\lambda'}(p')}{(2\pi)^{3/2}}\ \frac{\epsilon^\mu_\lambda(p)}{(2\pi)^{3/2}}\ T_{\rho\nu\mu}(p',p)\,\delta(p'-p-q),\end{aligned}$$

where $|p,\lambda\rangle_{\text{in}}$ and $|p',\lambda'\rangle_{\text{out}}$ are single-charged particle in and out states, respectively, and $a^+_{\text{in}}(q,\lambda'')$ creates an incoming neutral particle. In Appendix B it is shown that when the neutral particle operator is contracted out and the motion equation of the neutral field is used, this matrix element becomes

$$\begin{aligned}&-i\int dx e^{-iqx}\epsilon^\rho_{\lambda''}(q)\,{}_{\text{out}}\langle p,\lambda|[\tilde{J}_\rho(x) - (Z-1) \\ &\qquad \times m^2 A_\rho(x)]|p,\lambda\rangle_{\text{in}} \\ &\quad = \epsilon^\rho_{\lambda''}(q)\ \frac{\epsilon^\nu_{\lambda'}(p')}{(2\pi)^{3/2}}\ T_{\rho\nu\mu}(p',p)\ \frac{\epsilon^\mu_\lambda(p)}{(2\pi)^{3/2}}\ \delta(p'-p-q),\end{aligned} \tag{8.8}$$

where $\tilde{J}_\rho$ differs from the total "electromagnetic" current only by an additional divergence. Therefore

$$\int d\mathbf{x}\tilde{J}_0(x) = Q,$$

Q being the total "electromagnetic" charge operator. When the limit $q \to 0$ is taken in Eq. (8.8), and the equation

$$\sum_{\lambda=1}^{3}\epsilon^\rho_\lambda(0)\epsilon^\sigma_\lambda(0) = -g^{\rho\sigma}$$

(see Appendix A) is used to remove the $\epsilon_{\lambda''}{}^\rho(0)$ factors, the following relation is discovered:

$$\begin{aligned}&-i2\pi\delta^1(0)\,{}_{\text{out}}\langle p',\lambda'|Q|p,\lambda\rangle_{\text{in}} \\ &\quad = -i(Z-1)\int dx\,{}_{\text{out}}\langle p',\lambda'|A_0(x)|p,\lambda\rangle_{\text{in}} \\ &\qquad + [\epsilon^\nu_{\lambda'}(p')/(2\pi)^{3/2}]\,T_{0\nu\mu}(p,p)[\epsilon^\mu_\lambda(p)/(2\pi)^{3/2}] \\ &\qquad \times \delta^1(0)\delta(\mathbf{p}'-\mathbf{p}),\end{aligned}$$

in which the one-dimensional delta function has been denoted by $\delta^1(0)$. The matrix element of $A_0(x)$ appearing above may be expressed in terms of $T_{0\nu\mu}$ by reversing the contraction procedure; the result is

$$\int dx\, {}_{\text{out}}\langle p', \lambda' | A_0(x) | p, \lambda \rangle_{\text{in}} = (2\pi)^4 S^{0\rho}(0) \times [\epsilon^{\nu}_{\lambda'}(p)/(2\pi)^{3/2}] T_{\rho\nu\mu}(p,p) [\epsilon^{\mu}_{\lambda}(p)/(2\pi)^{3/2}] \times \delta^1(0)\, \delta(\mathbf{p}' - \mathbf{p})$$
$$= (i/m^2)[\epsilon^{\nu}_{\lambda'}(p)/(2\pi)^{3/2}]\, T_{0\nu\mu}(p,p) [\epsilon^{\mu}_{\lambda}(p)/(2\pi)^{3/2}] \times \delta^1(0)\delta(\mathbf{p}' - \mathbf{p}).$$

Therefore the matrix element of the charge operator becomes

$$-i2\pi_{\text{out}}\langle p', \lambda' | Q | p, \lambda \rangle = Z[\epsilon^{\nu}_{\lambda'}(p)/(2\pi)^{3/2}]\, T_{0\nu\mu}(p,p) \times [\epsilon^{\mu}_{\lambda}(p)/(2\pi)^{3/2}] \delta(\mathbf{p}' - \mathbf{p}).$$

Now by analogy with spinor electrodynamics, $|p, \lambda\rangle_{\text{in}}$ is assumed to be an eigenstate of Q; specifically,

$$Q | p, \lambda \rangle = g | p, \lambda \rangle. \tag{8.9}$$

But by the definition of the S matrix and Eq. (8.6),

$${}_{\text{out}}\langle p', \lambda' | p, \lambda \rangle_{\text{in}} \equiv S(p', \lambda'; p, \lambda) = 2p_0 \delta_{\lambda'\lambda} \delta(\mathbf{p}' - \mathbf{p}).$$

So, finally, the condition

$$\epsilon^{\nu}_{\lambda'}(p)\, T_{0\nu\mu}(p,p) \epsilon^{\mu}_{\lambda}(p) \xrightarrow[p^2 \to m^2]{} -iZ^{-1} g (2\pi)^4 2p_0 \delta_{\lambda'\lambda} \tag{8.10}$$

is reached.

Equations (8.5), (8.7), and (8.10) are the conditions from which the three constants Z, Z_1, and Z_2 may in principle be determined.

9. A WARD IDENTITY

A. Derivation

The function $T_{\rho\nu\mu}(p', p)$ introduced in the previous section is the sum of all diagrams with the general form shown in Fig. 5. However, some of these diagrams do not contribute to the expression on the left side of condition (8.10). For example, consider Fig. 6b. The sum of all diagrams of this type is, in accordance with the Feynman rules,

$$V_{\rho\nu\beta}(p', p) S'^{\beta\alpha}(p) \Pi'_{\alpha\mu}(p), \tag{9.1}$$

and the mass condition (8.5) implies

$$\epsilon^{\nu}_{\lambda'}(p) V_{\rho\nu\beta}(p,p) S'^{\beta\alpha}(p) \Pi'_{\alpha\mu}(p) \epsilon^{\mu}_{\lambda}(p) \xrightarrow[p^2 \to m^2]{} 0,$$

so that Fig. 6b does, in fact, not enter into condition (8.10). Similarly, it is seen that all diagrams with corrections to the external charged lines may be neglected.

The same is not true of corrections to the external neutral line, because the neutral momentum is not on the mass shell. These contributions must be evaluated and included. They have the form shown in Fig. 12, and their sum is

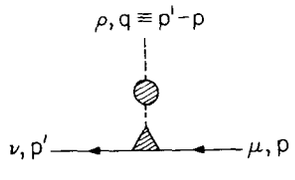


FIG. 12. The form of the external neutral line correlations to the trilinear vertex $V'_{\rho\nu\mu}(p', p)$.

$$\Pi'_{\rho\beta}(q) S'^{\beta\alpha}(q) V'_{\alpha\nu\mu}(p', p).$$

As a consequence of the above considerations and the definition of $V'_{\rho\nu\mu}$ as the complete vertex function, it is possible to write

$$\epsilon^{\nu}_{\lambda'}(p) T_{0\nu\mu}(p,p) \epsilon^{\mu}_{\lambda}(p) = \epsilon^{\nu}_{\lambda'}(p)[V'_{0\nu\mu}(p,p) + \Pi'_{0\beta}(0) S'^{\beta\alpha}(0) V'_{\alpha\nu\mu}(p,p)] \epsilon^{\mu}_{\lambda}(p)$$

if $p^2 = m^2$. But the expression for $S'^{\nu\mu}$, Eq. (8.1), and the definition of $\Pi'_{\nu\mu}$, Eq. (8.2), together with the form of $\Pi_{\nu\mu}$, Eq. (7.5), may be utilized to find[18]

$$S'^{\beta\alpha}(0) = \frac{i}{(2\pi)^4} \frac{(ZZ_2)^{-1}}{m^2} g^{\beta\alpha},$$

and

$$\Pi'_{\beta\alpha}(0) = i\,(2\pi)^4\,(ZZ_2 - 1)\, m^2 g_{\beta\alpha}.$$

Therefore

$$\epsilon^{\nu}_{\lambda'}(p) T_{0\nu\mu}(p,p) \epsilon^{\mu}_{\lambda}(p) = (ZZ_2)^{-1} \epsilon^{\nu}_{\lambda'}(p) V'_{0\nu\mu}(p,p) \epsilon^{\mu}_{\lambda}(p). \tag{9.2}$$

At this point, the WTR for $V'_{\rho\nu\mu}$ is applied. In Eq. (7.9) let p approach p'; the result is

$$V'_{\rho\nu\mu}(p,p) = g Z_1 Z_2^{-1} \frac{\partial}{\partial p^{\rho}} S'^{-1}_{\nu\mu}(p). \tag{9.3}$$

From conditions (8.4) and (8.5) and Eq. (8.1), it is inferred that

$$S'^{-1} = S^{-1} + O[(p^2 - m^2)^2],$$

so

$$\frac{\partial}{\partial p^{\rho}} S'^{-1}_{\nu\mu}(p) = \frac{\partial}{\partial p^{\rho}} S^{-1}_{\nu\mu}(p) + O(p^2 - m^2).$$

When this is inserted into Eq. (9.3) and it in turn is used in (9.2), the relation

$$\epsilon_{\lambda'^{\nu}}(p) T_{0\nu\mu}(p,p) \epsilon_{\lambda}{}^{\mu}(p) = -i\, g (2\pi)^4 Z^{-1} Z_1 Z_2^{-2}\, 2p_0 \delta_{\lambda'\lambda}$$

is found at $p^2 = m^2$. The last step is to use this formula with condition (8.10) and arrive at the following Ward identity:

$$Z_1 Z_2^{-2} = 1. \tag{9.4}$$

B. General Comments

The standard Ward identity of spinor electrodynamics states that Z_1 is the same as Z_2; Eq. (9.4) is a different relation. The difference has its origin in the fact that, unlike the photon, the neutral particle in the present theory has a finite mass, so the charge condition imposed here, Eq. (8.10), involves an off mass shell amplitude. Such is not the case in the analogous spinor electrodynamics derivation, because a $q = 0$ photon is still on its mass shell.

The discrepancy between the usual Ward identity and the result above prompts a discussion of the distinctions between the present development (which is adopted from Nishijima[11]) and the usual textbook presentation of renormalization constants. The two attitudes may be characterized as follows: (a) In the standard approach, the constants Z_1 and Z_2 are primarily vertex and propagator renormalizers; their values are chosen by requiring that divergences be removed from the theory; (b) according to the method utilized in

this paper, physical quantities, i.e., the coupling constant and field operators, are renormalized, and the renormalization constants are evaluated with physical conditions such as equations (8.4), (8.6) and (8.9).

Whichever way the Z's are defined, the Ward identity then follows from the appropriate WTR. Approach (b) has been used here because it is not yet known how, or even whether, divergences may be removed from the massive Yang–Mills theory by simple vertex and propagator renormalizations, so the Z's cannot be evaluated in accordance with method (a). Such considerations do not prevent the formal imposition of physical conditions, however, which is the essential step in method (b).

If Z_1 and Z_2 had been assumed to be vertex and propagator renormalization constants, as in (a), the question of evaluation having been temporarily ignored, the standard Ward identity would have been recovered. In this sense, (a) and (b) lead to different Ward identities in massive Yang–Mills theory, in contradistinction with the spinor electrodynamics case, where both methods yield the same result. The difference is due, of course, to the nonvanishing mass of the neutral particle.

The new Ward identity has a simple meaning in connection with the relation of Q to the I-spin generator T_3. The three conserved charges Q_a, $a = 1, 2, 3$ ($Q_3 = Q$), corresponding to the currents which act as sources of the D_a^{μ} fields have commutators that may be computed from the canonical rules associated with Lagrangian (4.1) (recall that the charged and neutral fields are related to the D_a^{μ} fields as described in Sec. 2). These commutators turn out to be

$$(Q_a, Q_b) = i\, Z_1 Z_2^{-2}\, g\, C_{abc} Q_c.$$

Since the generators of I-spin transformations T_a have the commutators

$$(T_a, T_b) = i\, C_{abc} T_c,$$

it is possible to put

$$Q_a = Z_1 Z_2^{-2}\, g\, T_a,$$

or, because of the Ward Identity, Eq. (9.4),

$$Q_a = g\, T_a.$$

In particular,

$$Q = g\, T_3, \tag{9.5}$$

which is analogous to the usual relation between the charge and I-spin operators in an $SU(2)$ theory without baryons.[19]

It follows that assumption (8.9) is equivalent to

$$T_3 |p, \lambda\rangle_{\text{in}} = 1 |p, \lambda\rangle_{\text{in}},$$

i.e., the "positively charged" particle state is the $T_3 = 1$ member of an I-spin multiplet. Similarly, the "neutral" and "negatively charged" free particle states may be shown to be eigenvectors with $T_3 = 0$ and $T_3 = -1$, respectively, so they complete the $T = 1$ multiplet.

Equation (9.5) and its consequences are not unexpected; they merely show that the analogy with spinor electrodynamics, introduced to motivate the renormalization conditions, is maintained throughout these formal manipulations.

10. DISCUSSION

A. Field Ordering

Equation (7.4) must, of course, be true in each order of the perturbation series for $\Pi_{\nu\mu}$. This has an interesting consequence in lowest order. The lowest-order neutal self-energy contribution must be constructed from the diagram in Fig. 13 according to Def. 4 in Sec. 6; otherwise, Corollary 2 is inapplicable and Eq. (7.4) is not satisfied. The resulting diagrams are shown in Fig. 14. Diagrams of the type in Fig. 14b shall be called leafs.

Leaf diagrams do not appear in the usual Dyson–Wick expansion of the S matrix, because contractions of two fields at the same space–time point are not usually allowed. But their presence seems desirable, since it simplifies the self-energy expression. El-Ghabaty *et al.*[20] have also noticed, in another context, that the inclusion of leaf diagrams leads to simplifications.

The modification of the traditional procedure necessary to produce these diagrams is minimal. One must only allow the time ordering operation in the interaction representation expansion of the S matrix to order the fields within each interaction Hamiltonian as well as the Hamiltonians themselves. The order of products of fields having the same time argument is then defined to be the limit of the time-ordered product of the fields at different times as the times become equal.

This prescription, in principle, also produces the diagram in Fig. 13 as part of the amplitudes in addition to the leaf diagrams; but a simple calculation shows that the expression for this diagram vanishes. The analogous diagram appearing in spinor electrodynamics when the new ordering rule is applied may be similarly disposed of, so that well tested theory remains unaltered.

B. Divergences

It is not yet known whether the renormalization conditions of Sec. 8 remove the infinities from the theory. Although determinations of degrees of divergence by direct power counts seem to preclude this possibility, several papers have shown that power counting is misleading,[8,3,21] so the subject is far from closed.

In many investigations of renormalizability, propagators less divergent than $D^{\nu\mu}(p)$, Eq. (4.12), are produced by introducing scalar fields in the Lagrangian.[2-4] However, some of the single loop results thus found[2,3]

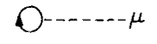


FIG. 13. Primary diagram from which lowest order contributions to $\Pi_{\nu\mu}(p)$ are constructed according to Def. 4.

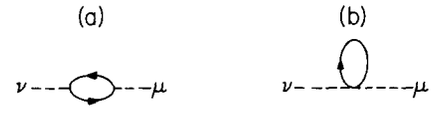


FIG. 14. The lowest-order contributions to $\Pi_{\nu\mu}(p)$ (the charged and the neutral particle self-energy).

may be reproduced in a less sophisticated way with Lagrangian (4.1), by writing $D^{\nu\mu}$ as the sum of a more convergent propagator, for instance $\tilde{D}^{\nu\mu}$, where

$$\tilde{D}^{\nu\mu}(p) = -\frac{i}{(2\pi)^4} Z_2^{-1} \frac{g^{\nu\mu} - (p^\nu p^\mu/p^2)}{p^2 - Zm^2 + i\epsilon}$$

and a correction $\delta D^{\nu\mu}$. Then, when the external states are physical, it is found that while there are extra expressions in the amplitudes due to the $\delta D^{\nu\mu}$s, no term is more divergent (by power count) than the self-energy in spinor electrodynamics. The only inputs needed to show this are the bare WTR's in Sec. 5. Therefore it is a matter of taste whether the extra terms are interpreted as loop diagrams for scalar particles or the consequences of an attempt to rewrite things such that the correct degrees of divergence are manifest. In fact, Wong[22] has reproduced a special case of Veltman's single loop Generalized Ward identity[3] without recourse to scalar fields in the Lagrangian. It is not clear that the agreement is maintained for multiple loop diagrams; but evidently spurious scalars might be unnecessary elements in any proof or disproof of renormalizability.

More recently, 't Hooft[4] has developed a massive gauge theory based on the Yang–Mills Lagrangian, which appears to be renormalizable. However, the modifications introduced are sufficiently drastic that no conclusion can be drawn about the renormalizability of the simple Yang–Mills theory discussed here.

The developments in this paper are expected to be useful in proving the theory much less divergent than it appears. For example, it may be shown from the Theorem and Corollary 2 in Sec. 6 that the $p^\nu p^\mu/m^2$ terms in the neutral particle propagators can be dropped when physical amplitudes are computed. Although this fact does not completely explain the low orders of divergence found by Veltman[3] or by Glashow,[21] it illustrates one way WTR's reduce the number of infinities.

That the WTR's found here are valid off the mass shell could prove important. For instance, Eq. (7.4) means the $p^\nu p^\mu/m^2$ terms of any propagator joined to a self-energy diagram may be omitted. Hence internal lines with self-energy bubbles do not contribute as many powers of the momentum as would otherwise be expected. Furthermore, by an actual calculation of the second order self-energy, the author has found that even off the mass shell it is only as divergent as its spinor electrodynamics analogue. From these observations, it is inferred that certain double loops, those which consist of loops on the internal lines of other loops, are also no more divergent than their spinor electrodynamics counterparts.

The results mentioned in the last two paragraphs shall be discussed in more detail in a future paper, along with other matters related to the divergence question.

ACKNOWLEDGMENTS

The author would like to thank Robert J. Finkelstein for suggesting this work and for the abundant and valuable advice offered while it was in progress and during the preparation of the manuscript. Thanks are also extended to John O. Mouton and Joel S. Kvitky for many interesting discussions about the material presented here and related topics.

APPENDIX A: CONVENTIONS–LSZ FORMULAS

The free field expansions corresponding to $\mathcal{L}_F$, Eq. (4.2), are

$$A^\mu(x) = (2\pi)^{-3/2} \int dk\ \theta(k_0)\delta(k^2 - m^2)\Sigma_\lambda \epsilon_\lambda{}^\mu(k) \times [a(k,\lambda)e^{-ikx} + a^+(k,\lambda)e^{ikx}],$$

$$B^\mu(x) = (2\pi)^{-3/2} \int dk\ \theta(k_0)\delta(k^2 - m^2)\Sigma_\lambda \epsilon_\lambda{}^\mu(k) \times [b(k,\lambda)e^{-ikx} + d^+(k,\lambda)e^{ikx}],$$

where the three vectors $\epsilon_\lambda{}^\mu(k)$, $\lambda = 1, 2, 3$, satisfy

$$k_\mu \epsilon_\lambda^\mu(k) = 0, \qquad \text{when } k^2 = m^2,$$

$$\Sigma_\lambda \epsilon_\lambda^\mu(k)\epsilon_\lambda^\nu(k) = -[g^{\mu\nu} - (k^\mu k^\nu/m^2)],$$

$$g_{\mu\nu}\epsilon_\lambda^\mu(k)\epsilon_{\lambda'}^\nu(k) = -\delta_{\lambda\lambda'},$$

and the creation, annihilation operators have these nonzero commutators

$$(a(k,\lambda), a^+(k',\lambda')) = (b(k,\lambda), b^+(k',\lambda')) = (d(k,\lambda), d^+(k',\lambda')) = 2k_0\delta_{\lambda\lambda'}\delta(\mathbf{k} - \mathbf{k}').$$

With the standard LSZ procedure, the S matrix elements, minus contributions from processes that do not involve true scattering, are found to have the form

$$\int dx(-i)[\epsilon_{\lambda_1}^\mu(p_1)/(2\pi)^{3/2}]e^{ip_1x}\kappa_{\mu\alpha}(x)\cdots \times \int dy(-i)[\epsilon_{\lambda_2}^\nu(p_2)/(2\pi)^{3/2}]e^{ip_2y}\kappa_{\nu\beta}(y)\cdots \times \int dz(-i)[\epsilon_{\lambda_3}^\mu(p_3)/(2\pi)^{3/2}]e^{ip_3z}\kappa_{\rho\gamma}(z)\cdots \times \langle 0|T(B^\alpha(x)\cdots B^{+\beta}(y)\cdots A^\gamma(z)\cdots)|0\rangle,$$

where

$$\kappa_{\mu\nu}(x) \equiv g_{\mu\nu}(\Box_x + m^2) - \frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu}. \tag{A1}$$

In and out states correspond to the factors in the above expression according to Table I.

A reference to Eq. (4.9) and integrations by parts lead to this alternative form of the S-matrix elements:

$$\frac{\epsilon_{\nu_1}^\mu(p_1)}{(2\pi)^{3/2}} S_{\mu\alpha}^{-1}(p_1)\cdots\frac{\epsilon_{\lambda_2}^\nu(p_2)}{(2\pi)^{3/2}} S_{\nu\beta}^{-1}(p_2)\cdots\frac{\epsilon_{\lambda_3}^\rho(p_3)}{(2\pi)^{3/2}}S_{\rho\lambda}^{-1}(p_3)\cdots \times \int\frac{dx}{(2\pi)^4}e^{ip_1x}\cdots\int\frac{dy}{(2\pi)^4}e^{ip_2y}\cdots\int\frac{dz}{(2\pi)^4}e^{ip_3z}\cdots \times \langle 0|T(B^\alpha(x)\cdots B^{+\beta}(y)\cdots A^\gamma(z)\cdots)|0\rangle.$$

The T functions introduced in Sec. 7 are now seen to be given by

$$T_{\mu\cdots\nu\cdots\rho\cdots}(p_1\cdots p_2\cdots p_3\cdots)\delta(\Sigma p_{\text{in}} - \Sigma p_{\text{out}}) = S_{\mu\alpha}^{-1}(p_1)\cdots S_{\nu\beta}^{-1}(p_2)\cdots S_{\rho\gamma}^{-1}(p_3)\cdots \times \int\frac{dx}{(2\pi)^4}e^{ip_1x}\cdots\int\frac{dy}{(2\pi)^4}e^{ip_2y}\cdots\int\frac{dz}{(2\pi)^4}e^{ip_3z}\cdots \times \langle 0|T(B^\alpha(x)\cdots B^{+\beta}(y)\cdots A^\gamma(z)\cdots)|0\rangle.$$

TABLE I.

	out state	in state
b particle (charge $+g$) with momentum p	$e^{ipx}B^{\mu}(x)$	$e^{-ipx}B^{+\mu}(x)$
d particle (charge $-g$) with momentum p	$e^{ipx}B^{+\mu}(x)$	$e^{-ipx}B^{\mu}(x)$
a particle (neutral) with momentum p	$e^{ipx}A^{\mu}(x)$	$e^{-ipx}A^{\mu}(x)$

For reference in the text, it is noted here that the momemtum space propagator $S'^{\nu\mu}(p)$ is expressed in terms of time-ordered Wightman functions by

$$S'^{\nu\mu}(p)\delta(p'-p) = \int \frac{dx}{(2\pi)^4}\int \frac{dy}{(2\pi)^4} e^{ip'x}e^{ipy} \times \langle 0|T(B^{\nu}(x)B^{+\mu}(y))|0\rangle = \int \frac{dx}{(2\pi)^4}\int \frac{dy}{(2\pi)^4} e^{ip'x}e^{ipy} \times \langle 0|T(A^{\nu}(x)A^{\mu}(y)|0\rangle.$$

APPENDIX B: MATRIX ELEMENTS OF THE CURRENT

Suppose just one neutral particle in operator is contracted out of an arbitrary amplitude. According to Appendix A, the result is

$$\langle 2|a^{\dagger}_{\rm in}(p,\lambda)|1\rangle = -i\int dx[\epsilon_{\lambda}{}^{\mu}(p)/(2\pi)^{3/2}]e^{-ipx}\kappa_{\mu\nu}(x) \times \langle 2|A^{\nu}(x)|1\rangle.$$

The remaining in and out operators are contained in the state vectors $|1\rangle$ and $|2\rangle$.

Since the definition of $\kappa_{\mu\nu}(x)$, Eq. (A1), implies

$$\kappa_{\mu\nu}(x)A^{\nu}(x) \equiv \partial^{\nu}A_{\nu\mu} + m^2A_{\mu} + iZ_1Z_2^{-1}g\partial^{\nu}(B^{+}_{\nu}B_{\mu} - B^{+}_{\mu}B_{\nu}),$$

the field equations derived from Lagrangian (4.1) may be used to find

$$\kappa_{\mu\nu}(x)A^{\nu}(x) = J_{\mu} - (Z-1)m^2A_{\mu} + iZ_1Z_2^{-1}g\partial^{\nu}(B^{+}_{\nu}B_{\mu} - B^{+}_{\mu}B_{\nu})$$

or

$$\kappa_{\mu\nu}(x)A^{\nu}(x) = \tilde{J}_{\mu} - (Z-1)m^2A_{\mu}, \tag{B1}$$

where J_{μ} is the "electromagnetic" current given by

$$J^{\nu} = iZ_1Z_2^{-1}g(B^{+}_{\mu}B^{\mu\nu} - B_{\mu}B^{+\mu\nu})$$

and

$$\tilde{J}_{\mu} \equiv J_{\mu} + iZ_1Z_2^{-1}g\partial^{\nu}(B^{+}_{\nu}B_{\mu} - B^{+}_{\mu}B_{\nu}). \tag{B2}$$

When Eq. (B1) is applied, the matrix element above is transformed into

$$\langle 2|a^{\dagger}_{\rm in}(p,\lambda)|1\rangle = -i\int dx[\epsilon_{\lambda}{}^{\mu}(p)/(2\pi)^{3/2}]e^{-ipx} \times \langle 2[\tilde{J}_{\mu} - (Z-1)A_{\mu}]|1\rangle. \tag{B3}$$

APPENDIX C: CONSTRUCTION OF NEUTRAL SELF-ENERGY DIAGRAMS

Consider the primary diagram in Fig. 15. Two of the configurations from which it may be constructed are shown in Fig. 16.

In this appendix, the configurations in Figs. 16a and 16b shall be denoted by

$$G_{\sigma_2\sigma_1;\beta\alpha}(q_1) \quad (= U_{\sigma_2\beta\sigma_1\alpha}) \tag{C1a}$$

and

$$G_{\sigma_2\sigma_1\mu;\beta\alpha}(q_2), \tag{C1b}$$

respectively, i.e., the all inclusive symbol c introduced in Sec. 6 shall be replaced here by explicit indices corresponding to the relevant external line branches. These additional indices appear on the left of the semicolons in the expressions (C1a) and (C1b). For the purpose of Appendix C it is not necessary to display all the external momenta.

The Feynman rules summarized in Sec. 4 may be used to find a momentum space function associated with Fig. 15. In terms of the expressions defined above, the result may be written in either of two equivalent ways: omitting numerical factors, they are the following:

$$\int dq_1\int dq_2\, D^{\alpha\beta}(q_1)G_{\sigma_2\sigma_1;\beta\alpha}(q_1)D^{\sigma_1\delta}(q_2+p) \times V_{\mu\delta\gamma}(q_2+p,q_2)D^{\gamma\sigma_2}(q_2) \tag{C2a}$$

and

$$\int dq_1\int dq_2\, D^{\sigma_1\sigma_2}(q_1)G_{\sigma_2\sigma_1\mu;\beta\alpha}(q_2)D^{\alpha\beta}(q_2). \tag{C2b}$$

When step two of the prescription given in Sec. 7B is followed, the secondary diagrams obtained from Fig. 15 are seen to be those shown in Fig. 17.

There is only one way to attach a neutral line to the left hand loop in Fig. 15, and Fig. 17a shows the resulting diagram. According to Def. 4 in Sec. 6 and the definition of $G_{\sigma_2\sigma_1;\beta\alpha}(q_1)$ above, the appropriate

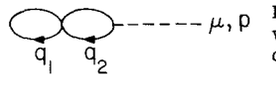


FIG. 15. The primary diagram from which self-energy contributions are constructed in Appendix C.

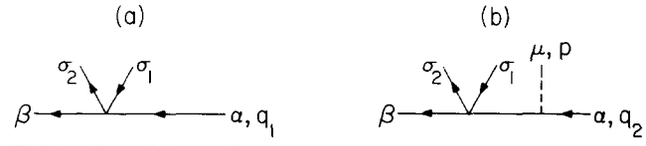


FIG. 16. Two of the configurations comprising the primary diagram in Fig. 15. In the text they are denoted by (a) $G_{\sigma_2\sigma_1;\beta\alpha}(q_1)$ and (b) $G_{\sigma_2\sigma_1\mu;\beta\alpha}(q_2)$.

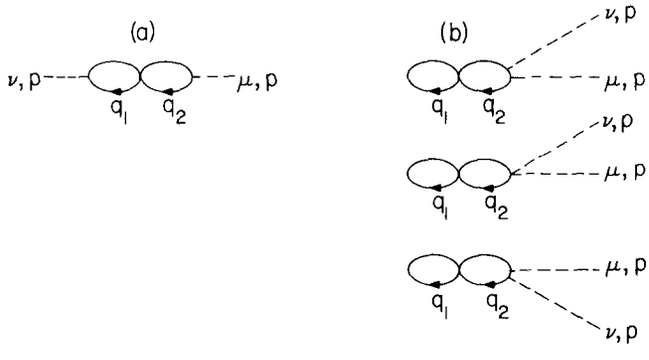


FIG. 17 The secondary diagrams derived from Fig. 15 in accordance with the construction procedure illustrated in Appendix C. See expressions (C3) and (C4) for the relevant momentum space functions.

symbol for the left-hand loop in Fig. 17a is $G_{\sigma_2\sigma_1;\nu}(q_1, p)$. Therefore, interpretation (C2a) of Fig. 15 implies that a correct expression for the diagram in Fig. 17a is

$$\Sigma^{(1)}_{\nu\mu}(p) \equiv \int dq_1 \int dq_2\, G_{\sigma_2\sigma_1;\nu}(q_1, p) D^{\sigma_1\delta}(q_2 + p) \times V_{\mu\delta\gamma}(q_2 + p, q_2) D^{\gamma\sigma_2}(q_2). \quad \text{(C3)}$$

Similarly, Def. 4 in Sec. 6, together with the definition of $G_{\sigma_2\sigma_1\mu;\beta\alpha}(q_2)$ above and interpretation (C2b) of Fig. 15, implies that the three diagrams in Fig. 17b sum to

$$\Sigma^{(2)}_{\nu\mu}(p) \equiv \int dq_1 \int dq_2 D^{\sigma_1\sigma_2}(q_1) G_{\sigma_2\sigma_1\mu;\nu}(q_2, p). \quad \text{(C4)}$$

Corollary 3, Eq. (6.6), may be applied to expressions (C3) and (C4) separately, with these results:

$$p^\nu \Sigma^{(1)}_{\nu\mu} = p^\nu \Sigma^{(2)}_{\nu\mu} = 0.$$

Hence, if $\Sigma_{\nu\mu}(p)$ denotes the sum of all the self-energy contributions shown in Fig. 17, then

$$\Sigma_{\nu\mu}(p) = \Sigma^{(1)}_{\nu\mu}(p) + \Sigma^{(2)}_{\nu\mu}(p)$$

and

$$p^\nu \Sigma_{\nu\mu}(p) = 0. \quad \text{(C5)}$$

Since all neutral particle self-energy contributions may be constructed along the lines illustrated here, Eq. (C5) is also valid if $\Sigma_{\nu\mu}$ is replaced by $\Pi_{\nu\mu}$ as indicated in the text.

---

* Supported in part by the National Science Foundation.

† Submitted in partial satisfaction of the requirements for the Ph.D. degree, Department of Physics, University of California, Los Angeles.

‡ Present address: 1711 Purdue Avenue, No. 9, Los Angeles, California 90025.

1 A. Salam and J. Strathdee, "On Equivalent Formulations of Massive Vector Field Theories," Trieste preprint IC/70/8 (1970).
2 D. G. Boulware, Ann. Phys. (N. Y.) **56**, 140 (1970).
3 M. Veltman, "Generalized Ward Identities and Yang-Mills Fields," Cern preprint Ref. TH. 1147-CERN (1970).
4 G. 't. Hooft, "Renormalizable Lagrangians for Massive Yang-Mills Fields," Print-71-2144 Utrecht.
5 R. Finkelstein and L. Staunton, Ann. Phys. (N.Y.) **54**, 97 (1969).
6 The classical consequences of this interpretation in the massless case have already been studied by R. Finkelstein, Rev. Mod. Phys. **36**, 632 (1964).
7 M. Nakamura, Progr. Theoret. Phys. **33**, 279 (1965).
8 K. H. Tzou, Nuovo Cimento **33**, 286 (1964).
9 T. D. Lee and C. N. Yang, Phys, Rev. **128**, 885 (1962).
10 This result has been extended to the general $SU(n)$ massive Yang-Mills theory by R. J. Finkelstein, J. S. Kvitky, and J. O. Mouton, UCLA preprint 71/TEP/36 (1971).
11 See, for instance, K. Nishijima, *Fields and Particles* (Benjamin, New York, 1969), p. 240.
12 T. D. Lee, Phys. Rev. **128**, 899 (1962), Appendix A.
13 Notice that in this paper the usual conventions relating diagrams to S-matrix contributions have been slightly modified. External lines do not correspond to factors describing in and out states, but rather to the indices to which these factors are attached. Hence the expressions associated with the diagrams for a given process must be multiplied by in and out factors in order to compute scattering amplitudes. Formula (7.1) displays the precise form.
14 Strictly speaking, this displacement is not always mathematically permissible, because some G_ρ contain divergent integrals. A suitable regularization procedure can presumably be found to rigorize such steps as the displacement in Corollary 3, if the consequences of the present work turn out to justify the effort.
15 The cross-hatched shapes in the figures of this paper are meant to signify any diagram or sum of diagrams with the appropriate lines leading in and out.
16 Expression (9.1), in which definition (8.2) is used, contains the exact formula.
17 See, for instance, S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Harper and Row, New York, 1962), p. 636.
18 It has been tacitly assumed here that the function $F(q^2)$ in Eq. (7.5) is not singular as $q \to 0$. Lowest order calculations tend to corroborate this assumption.
19 See, for example, S. Gasiorowicz, *Elementary Particle Physics* (Wiley, New York, 1966), p. 252.
20 S. L. El-Ghabaty, S. Gupta, W. Weihoffen, Phys. Rev. D **2**, 1130 (1971).
21 S. L. Glashow and J. Iliopoulos, Phys. Rev. D **3**, 1043 (1971).
22 S. K. Wong, Phys. Rev. D **3**, 945 (1971).

Quantum Dynamics of Higher-Derivative Fields

F. Riewe and A. E. S. Green
Department of Physics and Astronomy, University of Florida, Gainesville, Florida 32601
(Received 14 February 1971)

The generalized quantum field theory which follows from Lagrangians containing arbitrarily high-order derivatives is formulated in an indefinite metric space. Particular attention is given to conservation laws and canonical commutation relations. The Heisenberg equations of motion are derived.

1. INTRODUCTION

Because of problems encountered by quantum field theory associated with short distances, several theories have been proposed which are nonlocal over small space-time regions. Particularly promising are those similar to the generalized electrodynamics of Podolsky[1] and Bopp[2] and the generalized meson field theory of Green,[3] in which the nonlocality results from the presence of higher-order derivatives in the field equations. In the present work we will concentrate on the mathematical properties of higher-derivative nonlocal theories. The formalism we develop can be shown to have a physical interpretation in terms of the form factors of particles.[4]

Higher-order field equations of the type we shall discuss were first introduced in order to remove certain inconsistencies which arose in the traditional treatments of classical electrodynamics and in the early attempts to develop a quantum electrodynamics. When the use of propagator cutoffs and infinite-renormalization techniques produced a successful quantum electrodynamics, interest in higher-derivative field theories temporarily subsided.

However, if the same renormalization procedures are applied to other problems, such as those involving strongly interacting particles, they cannot always remove the infinities and divergences which beset many of the theories. For this reason a number of more recent works[5] have made essential use of higher-derivative fields or the resulting regularization[6] or indefinite-metric-space methods. It is therefore use-

symbol for the left-hand loop in Fig. 17a is $G_{\sigma_2\sigma_1;\nu}(q_1, p)$. Therefore, interpretation (C2a) of Fig. 15 implies that a correct expression for the diagram in Fig. 17a is

$$\Sigma^{(1)}_{\nu\mu}(p) \equiv \int dq_1 \int dq_2\, G_{\sigma_2\sigma_1;\nu}(q_1, p) D^{\sigma_1\delta}(q_2 + p) \times V_{\mu\delta\gamma}(q_2 + p, q_2) D^{\gamma\sigma_2}(q_2). \quad \text{(C3)}$$

Similarly, Def. 4 in Sec. 6, together with the definition of $G_{\sigma_2\sigma_1\mu;\beta\alpha}(q_2)$ above and interpretation (C2b) of Fig. 15, implies that the three diagrams in Fig. 17b sum to

$$\Sigma^{(2)}_{\nu\mu}(p) \equiv \int dq_1 \int dq_2 D^{\sigma_1\sigma_2}(q_1) G_{\sigma_2\sigma_1\mu;\nu}(q_2, p). \quad \text{(C4)}$$

Corollary 3, Eq. (6.6), may be applied to expressions (C3) and (C4) separately, with these results:

$$p^\nu \Sigma^{(1)}_{\nu\mu} = p^\nu \Sigma^{(2)}_{\nu\mu} = 0.$$

Hence, if $\Sigma_{\nu\mu}(p)$ denotes the sum of all the self-energy contributions shown in Fig. 17, then

$$\Sigma_{\nu\mu}(p) = \Sigma^{(1)}_{\nu\mu}(p) + \Sigma^{(2)}_{\nu\mu}(p)$$

and

$$p^\nu \Sigma_{\nu\mu}(p) = 0. \quad \text{(C5)}$$

Since all neutral particle self-energy contributions may be constructed along the lines illustrated here, Eq. (C5) is also valid if $\Sigma_{\nu\mu}$ is replaced by $\Pi_{\nu\mu}$ as indicated in the text.

---

* Supported in part by the National Science Foundation.

† Submitted in partial satisfaction of the requirements for the Ph.D. degree, Department of Physics, University of California, Los Angeles.

‡ Present address: 1711 Purdue Avenue, No. 9, Los Angeles, California 90025.

1 A. Salam and J. Strathdee, "On Equivalent Formulations of Massive Vector Field Theories," Trieste preprint IC/70/8 (1970).
2 D. G. Boulware, Ann. Phys. (N. Y.) **56**, 140 (1970).
3 M. Veltman, "Generalized Ward Identities and Yang-Mills Fields," Cern preprint Ref. TH. 1147-CERN (1970).
4 G. 't. Hooft, "Renormalizable Lagrangians for Massive Yang-Mills Fields," Print-71-2144 Utrecht.
5 R. Finkelstein and L. Staunton, Ann. Phys. (N.Y.) **54**, 97 (1969).
6 The classical consequences of this interpretation in the massless case have already been studied by R. Finkelstein, Rev. Mod. Phys. **36**, 632 (1964).
7 M. Nakamura, Progr. Theoret. Phys. **33**, 279 (1965).
8 K. H. Tzou, Nuovo Cimento **33**, 286 (1964).
9 T. D. Lee and C. N. Yang, Phys, Rev. **128**, 885 (1962).
10 This result has been extended to the general $SU(n)$ massive Yang-Mills theory by R. J. Finkelstein, J. S. Kvitky, and J. O. Mouton, UCLA preprint 71/TEP/36 (1971).
11 See, for instance, K. Nishijima, *Fields and Particles* (Benjamin, New York, 1969), p. 240.
12 T. D. Lee, Phys. Rev. **128**, 899 (1962), Appendix A.
13 Notice that in this paper the usual conventions relating diagrams to S-matrix contributions have been slightly modified. External lines do not correspond to factors describing in and out states, but rather to the indices to which these factors are attached. Hence the expressions associated with the diagrams for a given process must be multiplied by in and out factors in order to compute scattering amplitudes. Formula (7.1) displays the precise form.
14 Strictly speaking, this displacement is not always mathematically permissible, because some G_ρ contain divergent integrals. A suitable regularization procedure can presumably be found to rigorize such steps as the displacement in Corollary 3, if the consequences of the present work turn out to justify the effort.
15 The cross-hatched shapes in the figures of this paper are meant to signify any diagram or sum of diagrams with the appropriate lines leading in and out.
16 Expression (9.1), in which definition (8.2) is used, contains the exact formula.
17 See, for instance, S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Harper and Row, New York, 1962), p. 636.
18 It has been tacitly assumed here that the function $F(q^2)$ in Eq. (7.5) is not singular as $q \to 0$. Lowest order calculations tend to corroborate this assumption.
19 See, for example, S. Gasiorowicz, *Elementary Particle Physics* (Wiley, New York, 1966), p. 252.
20 S. L. El-Ghabaty, S. Gupta, W. Weihoffen, Phys. Rev. D **2**, 1130 (1971).
21 S. L. Glashow and J. Iliopoulos, Phys. Rev. D **3**, 1043 (1971).
22 S. K. Wong, Phys. Rev. D **3**, 945 (1971).

Quantum Dynamics of Higher-Derivative Fields

F. Riewe and A. E. S. Green
Department of Physics and Astronomy, University of Florida, Gainesville, Florida 32601
(Received 14 February 1971)

The generalized quantum field theory which follows from Lagrangians containing arbitrarily high-order derivatives is formulated in an indefinite metric space. Particular attention is given to conservation laws and canonical commutation relations. The Heisenberg equations of motion are derived.

1. INTRODUCTION

Because of problems encountered by quantum field theory associated with short distances, several theories have been proposed which are nonlocal over small space-time regions. Particularly promising are those similar to the generalized electrodynamics of Podolsky[1] and Bopp[2] and the generalized meson field theory of Green,[3] in which the nonlocality results from the presence of higher-order derivatives in the field equations. In the present work we will concentrate on the mathematical properties of higher-derivative nonlocal theories. The formalism we develop can be shown to have a physical interpretation in terms of the form factors of particles.[4]

Higher-order field equations of the type we shall discuss were first introduced in order to remove certain inconsistencies which arose in the traditional treatments of classical electrodynamics and in the early attempts to develop a quantum electrodynamics. When the use of propagator cutoffs and infinite-renormalization techniques produced a successful quantum electrodynamics, interest in higher-derivative field theories temporarily subsided.

However, if the same renormalization procedures are applied to other problems, such as those involving strongly interacting particles, they cannot always remove the infinities and divergences which beset many of the theories. For this reason a number of more recent works[5] have made essential use of higher-derivative fields or the resulting regularization[6] or indefinite-metric-space methods. It is therefore use-

ful to develop a rigorous mathematical framework for use in such applications.

The prototype equation for the generalized field theory we shall discuss is the field equation of Podolsky's electrodynamics[1]:

$$(1-\Box/m^2)\Box A_\mu = 0, \tag{1.1}$$

where we use the notation

$$\Box = \partial_\mu \partial^\mu = \nabla^2 - \left(\frac{\partial}{\partial t}\right)^2, \qquad \hbar = c = 1.$$

This equation was also proposed by Bopp[2] and appears in later treatments[7] of the electrodynamics suggested by Landé.[8] The classical theory of generalized electrodynamics was quantized by Podolsky and Kikuchi,[9] while Montgomery[10] and Green[11] investigated the consequences of the theory.[12]

The generalized meson field equation

$$(1-\Box/m_2^2)(\Box - m_1^2)\phi_\mu = 0 \tag{1.2}$$

was proposed by Green,[3] who later extended it to the more general form[13,14]

$$\left[\prod_{r=2}^{N}\left(1-\frac{\Box}{m_r^2}\right)\right](\Box - m_1^2)\phi_\mu = 0. \tag{1.3}$$

It is the mathematical treatment of equations of this type which is the major concern of the present investigation.

The mathematics used to treat these higher-derivative field equations is based on the generalized classical mechanics of Ostrogradsky.[15] The formalism was extended to include continuous-field equations by the above authors and also by Chang[16] and de Wet.[17]

The development of generalized field theory which we will present in Sec. 3 will be more general than these early treatments in a number of ways. First, the formalism will be valid in indefinite metric spaces (discussed in Sec. 2), in which the norm of a vector can be positive, negative, or zero. There will be a detailed treatment of such topics as the derivation of conservation laws, the relation between spin and statistics, and the derivation of the generalized Heisenberg equations of motion. In Sec. 4 the formalism will be compared with similar discussions, particularly those by Misra[18] and by Barut and Mullen.[19]

2. INDEFINITE METRIC SPACES

When higher-order field equations are quantized, it is often convenient, and sometimes necessary, to use an indefinite metric state space. For this reason we briefly review those mathematical properties of indefinite metric spaces which we will need later. Additional material and further references are given by Pandit[20] and Nagy[21] and by several of the authors in Ref. 5.

The use of an indefinite metric state space in quantum field theory was suggested by Dirac[22] and discussed in detail by Pauli.[23] In such a space the norm of a function ψ is given by

$$||\psi|| = \int \psi^* \eta \psi d^3x, \tag{2.1}$$

or, in simpler notation,

$$||\psi|| = \langle\psi|\eta|\psi\rangle. \tag{2.2}$$

The metric operator η is chosen to be Hermitian,

$$\langle\psi|\eta|\chi\rangle^* = \langle\chi|\eta|\psi\rangle, \tag{2.3}$$

where the asterisk (*) denotes complex conjugation, and the inverse η^{-1} is assumed to exist. Here the vectors $\langle\psi|$ and $|\chi\rangle$ have the usual complex scalar product and thus the basis vectors of the space can be chosen to be orthonormal:

$$\langle i|j\rangle = \delta_{ij}. \tag{2.4}$$

The eigenvalues of η are not required to be positive, so that the norm of a vector can be positive, negative, or zero.

We will find it more convenient to work with the indefinite metric basis vectors $(i|$ and $|j)$, which are usually chosen to be related to $\langle i|$ and $|j\rangle$ by

$$(i|j) = \langle i|\eta|j\rangle. \tag{2.5}$$

We will make the more specific choice

$$(i| = \langle i|\eta, \qquad |j) = |j\rangle. \tag{2.6}$$

In an indefinite metric space the adjoint (or pseudo-Hermitian conjugate) A^* of an operator A is defined by

$$(\psi|A|\chi)^* = (\chi|A^*|\psi). \tag{2.7}$$

It is generally not the same as the Hermitian conjugate $A^\dagger$, defined by

$$\langle\psi|A|\chi\rangle^* = \langle\chi|A^\dagger|\psi\rangle, \tag{2.8}$$

since the two are related by

$$A^* = \eta^{-1}A^\dagger\eta. \tag{2.9}$$

An operator is said to be self-adjoint, or pseudo-Hermitian, if it satisfies

$$H = H^*, \tag{2.10}$$

or

$$H = \eta^{-1}H^\dagger\eta. \tag{2.11}$$

The adjoint has the property that if

$$A|a) = a|a), \tag{2.12}$$

then

$$(a|A^* = (a|a^*. \tag{2.13}$$

Thus, if H is a pseudo-Hermitian operator with eigenvalues h and h' corresponding to eigenvectors $|h)$ and $|h')$, then we have

$$(h^* - h')(h|h') = 0. \tag{2.14}$$

In many cases of interest we will find that although $(h|h)$ can be either positive or negative, it will not be zero. It then follows that all eigenvalues of H are real and eigenvectors corresponding to unequal eigenvalues are orthogonal. Therefore, pseudo-Hermitian operators with no zero-norm eigenvectors have the familiar properties possessed by Hermitian operators in a positive definite space.

If the pseudo-Hermitian operator H has eigenvectors of norm zero, the corresponding eigenvalues can be complex. However, the expectation value, defined by

$$\langle H\rangle = (\psi|H|\psi), \tag{2.15}$$

is real, since

$$\langle H\rangle^* = (\psi|H|\psi)^* = (\psi|H|\psi) = \langle H\rangle. \tag{2.16}$$

If $|h)$ is an eigenstate of H with complex eigenvalue, then we must have $(h|h) = 0$. Hence the expectation value of H with respect to a state with a complex eigenvalue is always zero.

In an indefinite metric space the transformations of greatest interest are pseudo-unitarity:

$$\begin{aligned} A' &= UAU^*, \\ |\psi') &= U|\psi), \end{aligned} \tag{2.17}$$

where the pseudo-unitary operator U satisfies

$$U^*U = UU^* = 1 \tag{2.18}$$

or

$$\eta^{-1}U^\dagger\eta U = U\eta^{-1}U^\dagger\eta = 1. \tag{2.19}$$

Probability is conserved under pseudo-unitary transformations since eigenvalues and scalar products preserve their values.

With this brief review we now proceed to discuss higher-derivative field equations.

3. QUANTUM DYNAMICS

Quantum field theory involves both the vector-space formalism of the preceding section and the equations of motion and commutation relations for specific dynamical systems. We shall now develop the mathematical formalism of higher-derivative fields based on the usual quantum dynamics of local fields.

The use of higher-order-derivative Lagrangian densities provides theories which differ mathematically from the quantized theory of local fields in two respects. First, the higher derivatives add mathematical complications for both classical and quantized fields. Second, such theories sometimes, but not always, require the use of an indefinite metric state space for quantization. Because of these differences, we must redevelop the basic formalism of quantum dynamics.

Perhaps the best discussion of the quantum dynamics of local fields is found in the papers of Schwinger.[24,25] Prior to the work of Schwinger the usual procedure was to develop a classical field theory and then quantize it by introducing commutation (or anticommutation) relations. Schwinger, however, used quantized fields throughout and derived the commutation relations, rather than assuming them. This procedure is especially important in nonlocal field theory since many of the problems of interpretation center around the commutation relations. Since we can see precisely what assumptions are being made, we will know to what extent the theory can be modified, if necessary.

Although there have been a number of discussions of higher-order-Lagrangian field theories, only a few have followed Schwinger's approach. In particular, Misra[18] extended Schwinger's work to include higher derivatives, and Barut and Mullen[19] assumed an indefinite metric space and also derived the generalized form of Hamilton's canonical equations. For reasons discussed in Sec. 4, their treatment differs from ours except for the initial definitions. However, their papers will serve as a useful guide since certain parts, notably those corresponding to our Eqs. (3.10)–(3.14), require little or no modification. Wherever practical, we shall use their notation, which is an extension of that used by Schwinger. In contrast to the work of Barut and Mullen, our results will be consistent with those of the authors discussed in Sec. 1, including the more recent works.[5]

Schwinger bases his development on the basic postulate that the operator δW which generates infinitesimal transformations is obtained by variation of the action integral

$$W_{12} = \int_{\sigma_2}^{\sigma_1} d^4x\mathcal{L}, \tag{3.1}$$

where σ_1 and σ_2 are spacelike space–time surfaces and $\mathcal{L}$ is the Lagrangian density operator, or Lagrange function. Schwinger assumes that the space is positive definite and also that $\mathcal{L}$ is a function only of a field operator $\phi(x)$ and its 4-gradient $\phi_\mu = \partial_\mu\phi(x) = \partial\phi/\partial x^\mu$. We shall relax both of these restrictions by allowing the possibility of an indefinite metric space and also by allowing $\mathcal{L}$ to depend on derivatives of arbitrarily high order, designated by

$$\phi_{\mu(n)} = \partial_{\mu(n)}\phi(x) = \partial_{\mu_1}\partial_{\mu_2}\cdots\partial_{\mu_n}\phi(k). \tag{3.2}$$

The vectors of the indefinite metric space describing the states of the system are

$$(\xi_1'\sigma_1| \quad \text{and} \quad |\xi_1'\sigma_1), \tag{3.3}$$

where ξ_1' represents the eigenvalues of a complete set of commuting operators ξ on the space–time surface σ_1. We are using the notation of the previous section, so that the rounded brackets indicate an indefinite metric.

The variation of a vector is defined by

$$\delta|\xi_1'\sigma_1) = |\bar{\xi}_1'\bar{\sigma}_1) - |\xi_1'\sigma_1), \tag{3.4}$$

where $|\bar{\xi}_1'\bar{\sigma}_1)$ differs from $|\xi_1'\sigma_1)$ by an infinitesimal pseudo-unitary transformation

$$|\bar{\xi}_1'\bar{\sigma}_1) = U(\sigma_1)|\xi_1'\sigma_1) = [1 - iF(\sigma_1)]|\xi_1'\sigma_1), \tag{3.5}$$

in which $F(\sigma_1)$ is self-adjoint, or pseudo-Hermitian. Then

$$\delta|\xi_1'\sigma_1) = -iF(\sigma_1)|\xi_1'\sigma_1) \tag{3.6}$$

and

$$\delta(\xi_1'\sigma_1| = (\xi_1'\sigma_1|iF(\sigma_1), \tag{3.7}$$

so that

$$\delta(\xi_1'\sigma_1|\xi_2'\sigma_2) = i(\xi_1'\sigma_1|[F(\sigma_1) - F(\sigma_2)]|\xi_2'\sigma_2). \tag{3.8}$$

Our fundamental assumption is that $F(\sigma_1) - F(\sigma_2)$ is the variation of the operator W_{12} defined in Eq. (3.1). Thus the variational principle is

$$\begin{aligned} \delta(\xi_1'\sigma_1|\xi_2'\sigma_2) &= i(\xi_1'\sigma_1|[F(\sigma_1) - F(\sigma_2)]|\xi_2'\sigma_2) \\ &= i(\xi_1'\sigma_1|\delta W_{12}|\xi_2'\sigma_2) \\ &= i(\xi_1'\sigma_1|\delta\int_{\sigma_2}^{\sigma_1} d^4x\mathcal{L}|\xi_2'\sigma_2). \end{aligned} \tag{3.9}$$

Since F is self-adjoint, $\mathcal{L}$ must be also. This result differs from that of Barut and Mullen[19] for reasons we will discuss in Sec. 4.

Next we carry out the variation of the action integral. We break the variation into two parts:

$$\delta W_{12} = \int_{\sigma_2}^{\sigma_1} d^4x \delta_0 \mathcal{L} + \left(\int_{\sigma_1} - \int_{\sigma_2}\right) d\sigma_\mu \delta x^\mu \mathcal{L}. \tag{3.10}$$

The first term is the variation $\delta_0 W_{12}$, which leaves the value of $\mathcal{L}$ fixed on the boundary surfaces, while the second results from a displacement δx^μ of the boundary surfaces. By defining

$$L^{\mu(n)} = \frac{1}{n!} \sum_{\text{permutations of } \mu(n)} \frac{\partial \mathcal{L}}{\partial \phi_{\mu(n)}}, \tag{3.11}$$

the variation of $\mathcal{L}(\phi, \phi_\mu, \ldots, \phi_{\mu(N)})$ may be written

$$\begin{aligned}\delta_0 \mathcal{L} &= \sum_{n=0}^{N} \frac{\partial \mathcal{L}}{\partial \phi_{\mu(n)}} \delta_0 \phi_{\mu(n)} \\ &= \sum_{n=0}^{N} (-1)^n \partial_{\mu(n)} L^{\mu(n)} \delta_0 \phi \\ &\quad + \partial_\mu \left(\sum_{n=0}^{N-1} \pi^{\mu\nu(n)} \partial_{\nu(n)} \delta_0 \phi \right),\end{aligned} \tag{3.12}$$

where

$$\pi^{\mu(n)} = \sum_{m=0}^{N-n} (-1)^m \partial_{\lambda(m)} L^{\mu(n)\lambda(m)}. \tag{3.13}$$

In the expression for $\delta_0 \mathcal{L}$, the order of the operators cannot be changed. However, we will follow Schwinger[24] by assuming that identical contributions are obtained from terms which differ only in the position of $\delta_0 \phi$. By substituting Eq. (3.12) into Eq. (3.10) and using Gauss' theorem, we see that if δW_{12} vanishes for variations in which $\phi(x)$ is held fixed on the boundary surfaces, the Euler–Lagrange equations result:

$$\sum_{n=0}^{N} (-1)^n \partial_{\mu(n)} L^{\mu(n)} = 0. \tag{3.14}$$

Also the generating function is found to be

$$F(\sigma) = \int d\sigma_\mu \left(\mathcal{L} \delta x^\mu + \sum_{n=0}^{N-1} \pi^{\mu\tau(n)} \partial_{\tau(n)} \delta_0 \phi \right). \tag{3.15}$$

The variation $\delta \phi_{\tau(n)}$ may be written as the sum of two terms: $\delta_0 \phi_{\tau(n)}$, which is the variation at a fixed point, and $\phi_{\nu\tau(n)} \delta x^\nu$, which corresponds to an infinitesimal Lorentz transformation

$$\delta x^\mu = a^\mu - \epsilon^{\mu\nu} x_\nu, \tag{3.16}$$

in which

$$\epsilon^{\mu\nu} = -\epsilon^{\nu\mu} = \partial^\mu \delta x^\nu. \tag{3.17}$$

If $\phi(x)$ has some additional transformation properties, such as vector or spinor character, we symbolize these by the component $\phi^a(x)$. The variation then has an extra term:

$$\delta \phi^a_{\tau(n)} = \delta_0 \phi^a_{\tau(n)} + \phi^a_{\nu\tau(n)} \delta x^\nu + \tfrac{1}{2} \epsilon^{\alpha\beta} S^a_{\alpha\beta b} \phi^b_{\tau(n)}. \tag{3.18}$$

For a scalar field $\phi(x)$, we have $S^a_{\alpha\beta b} = 0$. For simplicity we shall suppress the indices a and b on $\phi(x)$ and $S_{\alpha\beta}$.

Using

$$\delta_0 \phi_{\tau(n)} = \partial_{\tau(n)} \delta_0 \phi, \tag{3.19}$$

the generating function may now be written

$$\begin{aligned}F(\sigma) = \int d\sigma_\mu \Big(&\mathcal{L} \delta x^\mu + \sum_{n=0}^{N-1} \pi^{\mu\tau(n)} (\delta \phi_{\tau(n)} \\ &- \phi_{\nu\tau(n)} \delta x^\nu - \tfrac{1}{2} \epsilon^{\alpha\beta} S_{\alpha\beta} \phi_{\tau(n)}) \Big).\end{aligned} \tag{3.20}$$

Define the quantity

$$\begin{aligned}f^{\mu\lambda\nu}_{(n)} = \tfrac{1}{2} [&\pi^{\mu\tau(n)} S^{\lambda\nu} \phi_{\tau(n)} + \pi^{\nu\tau(n)} S^{\lambda\mu} \phi_{\tau(n)} \\ &+ \pi^{\lambda\tau(n)} S^{\nu\mu} \phi_{\tau(n)}].\end{aligned} \tag{3.21}$$

From

$$f^{\mu\lambda\nu}_{(n)} = -f^{\lambda\mu\nu}_{(n)} \tag{3.22}$$

we have

$$\begin{aligned}\epsilon_{\lambda\nu} f^{\mu\lambda\nu}_{(n)} &= \tfrac{1}{2} \epsilon_{\lambda\nu} (f^{\mu\lambda\nu}_{(n)} - f^{\lambda\mu\nu}_{(n)}) = \tfrac{1}{2} \pi^{\mu\tau(n)} S^{\lambda\nu} \phi_{\tau(n)} \epsilon_{\lambda\nu} \\ &= \partial_\lambda (f^{\mu\lambda\nu}_{(n)} \delta x_\nu) + \partial_\lambda f^{\lambda\mu\nu}_{(n)} \delta x_\nu.\end{aligned} \tag{3.23}$$

Since, for a sufficiently large surface,

$$\int d\sigma_\mu \partial_\lambda (f^{\mu\lambda\nu}_{(n)} \delta x_\nu) = 0, \tag{3.24}$$

$F(\sigma)$ can be written as

$$F(\sigma) = \int d\sigma_\mu \left(\sum_{n=0}^{N-1} \pi^{\mu\tau(n)} \delta \phi_{\tau(n)} + T^{\mu\nu} \delta x_\nu \right), \tag{3.25}$$

where

$$T^{\mu\nu} = \mathcal{L} g^{\mu\nu} - \sum_{n=0}^{N-1} (\pi^{\mu\tau(n)} \partial^\nu \phi_{\tau(n)} + \partial_\lambda f^{\lambda\mu\nu}_{(n)}). \tag{3.26}$$

The important operator $T^{\mu\nu}$ is the stress tensor operator, or energy–momentum tensor operator. From it we can find the Hamiltonian density operator

$$\mathcal{H} = T^{44} = \sum_{n=0}^{N-1} (\pi^{4\tau(n)} \dot{\phi}_{\tau(n)} + \partial_\lambda f^{\lambda 44}_{(n)}) - \mathcal{L}. \tag{3.27}$$

Conservation laws may be found corresponding to invariance principles. If the Lagrangian density operator $\mathcal{L}$ is symmetric with respect to some transformation, then for variations which involve only that transformation we have

$$\delta W_{12} = F(\sigma_1) - F(\sigma_2) = 0. \tag{3.28}$$

For example, if $\mathcal{L}$ is Lorentz invariant, then variations involving only Lorentz transformations can be written in the form

$$\delta x_\nu = a_\nu - \epsilon_{\nu\lambda} x^\lambda \tag{3.29}$$

and

$$\delta \phi_{\tau(n)} = 0. \tag{3.30}$$

For this case the operator $F(\sigma)$ is

$$F_{\delta x}(\sigma) = a_\mu P^\mu(\sigma) + \tfrac{1}{2} \epsilon_{\mu\nu} J^{\mu\nu}(\sigma), \tag{3.31}$$

where

$$P^\nu(\sigma) = \int d\sigma_\mu T^{\mu\nu} \tag{3.32}$$

and

$$J^{\mu\nu} = \int d\sigma_\lambda M^{\lambda\mu\nu}, \tag{3.33}$$

$$M^{\lambda\mu\nu} = T^{\lambda\nu} x^\mu - T^{\lambda\mu} x^\nu. \tag{3.34}$$

Since $\delta W_{12} = 0$, we have

$$P^\nu(\sigma_1) - P^\nu(\sigma_2) = 0 \tag{3.35}$$

and

$$J^{\mu\nu}(\sigma_1) - J^{\mu\nu}(\sigma_2) = 0, \tag{3.36}$$

which are the conservation laws for momentum and angular momentum. The differential forms of these integral conservation laws are

$$\partial_\mu T^{\mu\nu} = 0 \tag{3.37}$$

and

$$\partial_\lambda M^{\lambda\mu\nu} = 0, \tag{3.38}$$

which imply that the stress tensor is symmetric:

$$T^{\mu\nu} - T^{\nu\mu} = \partial_\lambda M^{\lambda\mu\nu} = 0. \tag{3.39}$$

Conservation laws also result if $\mathfrak{L}$ is symmetric with respect to some modification of the field. For example, the law of conservation of charge is obtained if $\mathfrak{L}$ is symmetric under the phase transformation $\phi \to e^{iq\delta\lambda}\phi$; i.e., $\delta\phi = -iq\delta\lambda\phi$. The generating operator is then

$$F_{\delta\lambda}(\sigma) = -iq \int d\sigma_\mu \sum_{n=0}^{N-1} \pi^{\mu\tau(n)}\phi_{\tau(n)}\delta\lambda = Q(\sigma)\delta\lambda, \tag{3.40}$$

where

$$Q(\sigma) = \int d\sigma_\mu j^\mu \tag{3.41}$$

and

$$j^\mu = -iq \sum_{n=0}^{N-1} \pi^{\mu\tau(n)}\phi_{\tau(n)}. \tag{3.42}$$

Thus phase invariance of $\mathfrak{L}$ implies the conservation law for the charge operator:

$$Q(\sigma_1) - Q(\sigma_2) = 0. \tag{3.43}$$

It should be clear from the above that conservation laws and their derivations are essentially the same for a higher-order-derivative field theory as for a local field theory.

It should be noted that if two Lagrangian density operators differ only by a 4-divergence,

$$\bar{\mathfrak{L}} = \mathfrak{L} + \partial_\nu f^\nu, \tag{3.44}$$

then the same Euler-Lagrange equations result, since the action integral operators will differ only by a surface integral:

$$\bar{W}_{12} = W_{12} + \left(\int_{\sigma_1} - \int_{\sigma_2}\right) d\sigma_\nu f^\nu. \tag{3.45}$$

We shall make use of this ambiguity in the choice of $\mathfrak{L}$ later.

Next we derive the commutation relations for the field quantities. To simplify notation, we define

$$d\sigma_\mu = n_\mu d\sigma, \tag{3.46}$$

where n_μ is a unit vector. We shall abbreviate $n_\mu \pi^{\mu\tau(n)}$ as $\pi^{\tau(n)}$ and write

$$\pi^n\phi_n \equiv \sum_{n=0}^{N-1} \pi^{\tau(n)}\phi_{\tau(n)}. \tag{3.47}$$

Thus for variations $\delta x_\nu = 0$, which hold the coordinates fixed, the generating operator may be written

$$\begin{aligned} F_{\delta\phi}(\sigma) &= \int d\sigma_\mu \sum_{n=0}^{N-1} \pi^{\mu\tau(n)}\delta\phi_{\tau(n)} \\ &\equiv \int d\sigma \pi^n \delta\phi_n. \end{aligned} \tag{3.48}$$

Another form for the generating operator is obtained if $\mathfrak{L}$ is altered by adding the 4-divergence

$$f^\nu = -\pi^{\nu n}\phi_n. \tag{3.49}$$

We then have

$$\delta \int d\sigma_\nu f^\nu = -\delta \int d\sigma \pi^n \phi_n = -\int d\sigma(\pi^n \delta\phi_n + \delta\pi^n\phi_n), \tag{3.50}$$

so that

$$\bar{F}(\sigma) \equiv F_{\delta\pi}(\sigma) = -\int d\sigma\phi\pi^n\phi_n. \tag{3.51}$$

In certain cases, it is possible that some of the $\pi^{\tau(n)}$ will be identically zero. These are called "variables of constraint," since their Euler–Lagrange equations have the form of equations of constraint.

To obtain the commutation relation for the unconstrained variables, we consider the operator $G(\phi, \pi)$ and its variation $\delta_\phi G$:

$$(\xi'\sigma|\delta_\phi G|\xi''\sigma) = \delta(\xi'\sigma|G|\xi''\sigma). \tag{3.52}$$

Since the variation δ acts only on the commuting operator set ξ on σ, we have

$$\begin{aligned} \delta(\xi'\sigma|G|\xi''\sigma) &= [\delta(\xi'\sigma|]G|\xi''\sigma) + (\xi'\sigma|G[\delta|\xi''\sigma)] \\ &= -i(\xi'\sigma|[G, F_{\delta\phi}]|\xi''\sigma), \end{aligned} \tag{3.53}$$

so that

$$[G, F_{\delta\phi}] = i\delta_\phi G. \tag{3.54}$$

Similarly we find

$$[G, F_{\delta\pi}] = i\delta_\pi G. \tag{3.55}$$

The special cases $G = \phi_{\mu(k)}$ and $G = \pi^{\mu(k)}$ yield

$$\begin{aligned} &\left[\phi_{\mu(k)}, \int d\sigma\pi^n\delta\phi_n\right] = i\delta\phi_{\mu(k)}, \\ &\left[\pi^{\mu(k)}, \int d\sigma\pi^n\delta\phi_n\right] = 0, \\ &\left[\int d\sigma\delta\pi^n\phi_n, \pi^{\mu(k)}\right] = i\delta\pi^{\mu(k)}, \\ &\left[\int d\sigma\delta\pi^n\phi_n, \phi_{\mu(k)}\right] = 0. \end{aligned} \tag{3.56}$$

In order to make the correct choice of commutation versus anticommutation relations, we must examine the connection between spin and statistics. Following Schwinger,[24] we note that for integral spin $\mathfrak{L}$ is invariant under time reversal, while the sign of $\mathfrak{L}$ is reversed under time reversal for half-odd-integral spin. Recall that in the derivation of the Euler–Lagrange equations it was postulated that terms which differed only in the position of $\delta_0\phi$ would give identical contributions. Such terms occur with the same sign for integral spin and with opposite signs for half-odd-integral spin. Thus for the first case $\delta_0\phi$ must commute with its neighbors while in the second case it must anticommute. Thus we have

$$\begin{aligned} &\int d\sigma'[\phi_{\mu(k)}(x), \pi^n(x')]_\pm \delta\phi_n(x') = i\delta\phi_{\mu(k)}(x), \\ &\int d\sigma'[\pi^{\mu(k)}(x), \pi^n(x')]_\pm \delta\phi_n(x') = 0, \\ &\int d\sigma'[\delta\pi^n(x')][\phi_n(x'), \pi^{\mu(k)}(x)]_\pm = i\delta\pi^{\mu(k)}(x), \\ &\int d\sigma'[\delta\pi^n(x')][\phi_n(x'), \phi^{\mu(k)}(x)]_\pm = 0, \end{aligned} \tag{3.57}$$

where

$$[A,B]_{\pm} = AB \pm BA. \tag{3.58}$$

We make the additional assumption that $\phi_{\mu(k)}$ and $\pi^{\nu\tau(n)}$ commute (anticommute) unless $k = n$. Then, for x and x' on the same spacelike surface σ, we have the fundamental commutation relations

$$n_\nu[\phi_{\mu(k)}(x), \pi^{\nu\tau(n)}(x')]_{\pm} = in_\nu \delta^{\nu\tau(n)}_{\mu(k)} \delta^3(\mathbf{x}-\mathbf{x}'),$$
$$[\phi_{\mu(k)}(x), \phi_{\tau(n)}(x')]_{\pm} = n_\lambda n_\nu [\pi^{\lambda\mu(k)}(x), \pi^{\nu\tau(n)}(x')]_{\pm} = 0. \tag{3.59}$$

Here

$$\delta^{\tau(n)}_{\mu(k)} = \begin{cases} 1 & \text{if } \tau(n) \text{ is a permutation of } \mu(k) \\ 0 & \text{otherwise} \end{cases} \tag{3.60}$$

and $\delta^3(\mathbf{x}-\mathbf{x}')$ is the three-dimensional delta function. From the above we see that $\phi_{\mu(k)}$ and $n_\nu \pi^{\nu\tau(k)}$ are canonical conjugates and also that the usual relationship between spin and statistics holds. The commutation relations for the variables of constraint can be found by writing them as explicit functions of the canonical variables.

Finally we shall find Heisenberg's equation of motion. Noting that the last term in the expression for $T^{\mu\nu}$ is a 4-divergence, whose integral vanishes, we see that P^ν is given by

$$P^\nu = -\int d\sigma \pi^n \partial^\nu \phi_n + \int d\sigma^\nu \mathcal{L}. \tag{3.61}$$

The variation, either δ_ϕ or δ_π, of the second term gives

$$\begin{aligned} \delta \int d\sigma^\nu \mathcal{L} &= \int d\sigma^\nu \partial_\mu \left(\sum_{n=0}^{N-1} \pi^{\mu\tau(n)} \partial_{\tau(n)} \delta\phi \right) \\ &= \int d\sigma \partial^\nu \pi^n \delta\phi_n, \end{aligned} \tag{3.62}$$

where we have used the expression for $\delta_0 \mathcal{L}$, Eq. (3.12). Thus

$$\delta P^\nu = \int d\sigma (\partial^\nu \pi^n \delta\phi_n - \delta\pi^n \partial^\nu \phi_n). \tag{3.63}$$

By substituting P^ν for G and using the above expression in Eqs. (3.59) and (3.60), we find the Heisenberg canonical equations of motion:

$$\begin{aligned} \partial^\nu \phi_{\mu(n)} &= i[\phi_{\mu(n)}, P^\nu], \\ n_\mu \partial^\nu \pi^{\mu\tau(n)} &= i n_\mu [\pi^{\mu\tau(n)}, P^\nu]. \end{aligned} \tag{3.64}$$

For most practical applications it will be convenient to choose n_μ to point in the 4-direction so that

$$\partial^\nu \pi^{4\tau(n)} = i[\pi^{4\tau(n)}, P^\nu], \tag{3.65}$$

and similarly for previous equations. The only one of the Heisenberg equations which is usually needed is the simplest one:

$$i\dot{\phi} = [\phi, H], \tag{3.66}$$

where $H = P^4$.

4. DISCUSSION

We have presented a treatment of higher-derivative fields which follows the quantum dynamics of Schwinger.[24] Application of the formalism and its physical interpretation will be given in another paper.[4] As noted previously, our results are consistent with the work of most authors, but differ from those of Misra[18] and of Barut and Mullen.[19] In this section we discuss what we feel are the advantages of our approach.

The present investigation extends the work of Schwinger to include indefinite metric spaces and higher-derivative Lagrangians. In both cases the method of extension is different from that of Barut and Mullen. In our notation, the indefinite-metric-space equation $A|\alpha) = |\beta)$ and its dual, $(\alpha|A^* = (\beta|$, can also be written as $A|\alpha\rangle = |\beta\rangle$ and $\langle\alpha|\eta A^* = \langle\beta|\eta$ by using Eq. (2.6). In place of this last expression, Barut and Mullen choose to write $\langle\alpha|A^*\eta = \langle\beta|\eta$ (our notation) which is apparently inconsistent. (Note that in *their* notation the adjoint of A is $A^\dagger$, while we follow the more common convention of designating it as A^*.)

As a result their variational principle takes the form

$$\delta W_{12} = \eta^{-1} F(\sigma_1)\eta - F(\sigma_2). \tag{4.1}$$

In order to obtain the expression in our Eq. (3.9), they assume that $\mathcal{L}$ and F are "metrically invariant," i.e., they commute with the metric. However, from Eq. (2.11) we see that metrical invariance of a self-adjoint operator requires the operator to be Hermitian. Since Lagrangians which require an indefinite metric are not generally Hermitian,[4] the concept of metrical invariance appears to be of limited usefulness. In any case, the problems of misplaced metrics appear to be avoided in the present work.

Barut and Mullen[19] also treat the higher-derivative fields differently. As a result they obtain different expressions for the fundamental operators $T^{\mu\nu}$ and $F(\sigma)$ and for any result involving $T^{\mu\nu}$, $F(\sigma)$, or $\pi^{\mu(n)}$. Their procedure follows that of Misra[18] and is based on the assumption that a derivative of ϕ can be varied independent of ϕ itself, but that since ϕ is fixed on the boundary surface, derivatives of ϕ tangent to the surface are not independent of ϕ. Because of this they rewrite all 4-gradients in terms of tangential derivatives, making their formalism more cumbersome with no apparent advantages. Their basic formulas therefore differ from those of the other authors discussed in Sec. 1.

We conclude that our treatment has the advantage of being a direct extension of the work of Schwinger, while avoiding the problems encountered by Barut and Mullen involving metrical invariance and tangential derivatives. We now have all the mathematical detail necessary for most applications. Specific examples and their physical interpretation will be given in another paper.[4]

ACKNOWLEDGMENT

Early phases of this work were supported in part by the U.S. Air Force Office of Scientific Research Grant No. AFOSR-68-1397.

---

[1] B. Podolsky, Phys. Rev. **62**, 68 (1942).
[2] F. Bopp, Ann. Physik **38**, 345 (1940).
[3] A. E. S. Green, Phys. Rev. **73**, 26 (1948).
[4] F. Riewe and A. E. S. Green, J. Math. Phys. **13**, 1374 (1972), following paper.
[5] Recent works which make fundamental use of higher-derivative

fields, regularization, or indefinite metric spaces include the following: M. E. Arons, M. Y. Han, and E. C. G. Sudarshan, Phys. Rev. **137**, B1085 (1965); W. Heisenberg, *Introduction to the Unified Field Theory of Elementary Particles* (Interscience, New York, 1966); J. Tarski, J. Math. Phys. **7**, 560 (1966); D. A. Dubin and J. Tarski, *ibid.* 574 (1966); T. G. Pavlopoulos, Phys. Rev. **159**, 1106 (1967); T. D. Lee and G. C. Wick, Nucl. Phys. **B9**, 209 (1969); **B10**, 1 (1969); Phys. Rev. D **2**, 1033 (1970); D **3**, 1046 (1971); H. P. Dürr and E. Seiler, Nuovo Cimento **A66**, 734 (1970); A. M. Gleeson and E. C. G. Sudarshan, Phys. Rev. D **1**, 474 (1970); A. M. Gleeson, R. J. Moore, H. Rechenberg, and E. C. G. Sudarshan, *ibid.* **4**, 2242 (1971); P. N. Dobson, Jr., J. Math. Phys. **12**, 1207 (1971); R. Casalbuoni, R. Gatto, and G. Longhi, Phys. Rev. D **3**, 1499 (1971); N. Nakanishi, *ibid.*, 811, 1343 (1971); J. Gomatam, *ibid.*, 1292 (1971); J. Schwinger, *ibid.*, 1967 (1971); R. H. Tucker and C. L. Hammer, *ibid.*, 2448 (1971); D. L. Nordstrom, *ibid.* **4**, 1611 (1971).

[6] W. Pauli and F. Villars, Rev. Mod. Phys. **21**, 434 (1949).

[7] A. Landé and L. H. Thomas, Phys. Rev. **65**, 175 (1944).

[8] A. Landé, Phys. Rev. **60**, 121 (1941).

[9] B. Podolsky and C. Kikuchi, Phys. Rev. **65**, 228 (1944); **67**, 184 (1945).

[10] D. J. Montgomery, Phys. Rev. **69**, 117 (1946).

[11] A. E. S. Green, Phys. Rev. **72**, 628 (1947).

[12] B. Podolsky and P. Schwed, Rev. Mod. Phys. **20**, 40 (1948).

[13] A. E. S. Green, Phys. Rev. **73**, 519 (1948); **75**, 1926 (1949).

[14] For a review, see A. E. S. Green and T. Sawada, Rev. Mod. Phys. **39**, 594 (1967).

[15] M. Ostrogradsky, Mem. Acad. St. Petersburg **6**, 385 (1850); E. T. Whittaker, *Treatise on the Analytical Dynamics of Particles and Rigid Bodies* (Cambridge U.P., Cambridge, 1937), 4th ed., p. 265.

[16] T. S. Chang, Proc. Cambridge Phil. Soc. **42**, 132 (1946); **44**, 76 (1948).

[17] J. S. de Wet, Proc. Cambridge Phil. Soc. **44**, 546 (1948), Proc. Roy. Soc. (London) **A195**, 365 (1949).

[18] S. P. Misra, Indian J. Phys. **33**, 461 (1959).

[19] A. O. Barut and G. H. Mullen, Ann. Phys. (N.Y.) **20**, 203 (1962); A. O. Barut, *Electrodynamics and Classical Theory of Fields and Particles* (Macmillan, New York, 1964).

[20] L. K. Pandit, Suppl. Nuovo Cimento **11**, 157 (1959).

[21] K. L. Nagy, Suppl. Nuovo Cimento **17**, 924 (1960); *State Vector Spaces with Indefinite Metric in Quantum Field Theory* (Noordhoff, Groningen, The Netherlands, 1966).

[22] P. A. M. Dirac, Proc. Roy. Soc. (London) **A180**, 1 (1942).

[23] W. Pauli, Rev. Mod. Phys. **15**, 175 (1943).

[24] J. Schwinger, Phys. Rev. **82**, 914 (1951).

[25] J. Schwinger, Phys. Rev. **91**, 713 (1953).

Physical Interpretation of Higher-Derivative Field Theories

F. Riewe and A. E. S. Green
University of Florida, Gainesville, Florida 32601
(Received 14 February 1972)

Polynomial-type field equations are shown to have a realistic physical interpretation in terms of particle form factors, both for classical fields and for "dipole-regularized" quantized fields. Form factors arising from such field equations are found to give a reasonable description of the electromagnetic structure of the proton.

1. INTRODUCTION

The basic mathematical formalism required by higher-derivative, or "generalized," quantum field theories appears to be straightforward and free from inconsistencies.[1] However, when particular higher-derivative field equations are treated in detail, nonphysical results often appear.[2-4] Usually the difficulties relate to negative-energy states and the corresponding lack of conservation of probability. It is our purpose to show that a generalized field theory can be constructed which is mathematically consistent and which has a realistic physical interpretation in terms of the form factors of interacting particles.

Our discussion will center around the Nth-order polynomial meson-field equation introduced by Green.[5,6] We begin by reviewing the procedure for the quantization of generalized free-field equations for the cases in which the roots of the polynomial are real and unequal, real and equal, or complex-conjugate pairs. We then present a physical interpretation involving particle form factors which eliminates certain difficulties from classical generalized electrodynamics. It is shown that the higher-derivative field equation for the interaction of point particles is mathematically equivalent to the usual equations describing particle interactions characterized by form factors. This correspondence is shown to hold for scalar source particles, in agreement with Ueda and Green,[7] and is extended to the case of spin-$\frac{1}{2}$ particles.

The only quantized higher-derivative theories which are generally believed to avoid nonphysical results are those involving "ghost particles."[8] We will demonstrate that the ghost particles can be considered to be the intermediate particles characterizing the form factors of particle interactions, in contrast to the usual assumption that they are physically unobservable. We then test the validity of this interpretation by comparing the form factors derived from physically admissible generalized field theories with experiment. We show that the simplest charged-particle form factors consistent with generalized field theory give a realistic description of the electromagnetic structure of the proton. Hence the form-factor interpretation of generalized field theory is both mathematically consistent and physically reasonable.

2. FREE-FIELD EQUATIONS

A. General Formalism

The polynomial free-field equation

$$\left[\prod_{r=2}^{N}\left(1-\frac{\Box}{m_r^2}\right)\right](\Box - m_1^2)\phi(x) = 0, \tag{2.1}$$

introduced by Green,[5,6] has been frequently discussed for the case in which the polynomial roots m_r (often called "regulator masses" when $r > 1$) are real and unequal.[9] Here we are using $\Box = \nabla^2 - \partial^2/\partial t^2$ with $\hbar = c = 1$. Certain more recent "regularized" models, such as those of Lee[10] and Heisenberg,[11] involve m's which are complex conjugates or are equal in pairs. Quantization often requires the use of an indefinite metric space, as discussed by Nagy.[3,4]

We will now give a brief discussion in which the formalism used by Green is extended to include equal or complex regulator masses. Our basic purpose in this section is to present the mathematical treatment of polynomial free-field equations and to show the negative-energy difficulties encountered by the physical interpretation in terms of auxiliary fields. For general applications our treatment should be more convenient than other discussions[3,4] involving specific models.

The Fourier solution of Eq. (2.1) for unequal roots can be expressed as a sum of auxiliary fields,

fields, regularization, or indefinite metric spaces include the following: M. E. Arons, M. Y. Han, and E. C. G. Sudarshan, Phys. Rev. **137**, B1085 (1965); W. Heisenberg, *Introduction to the Unified Field Theory of Elementary Particles* (Interscience, New York, 1966); J. Tarski, J. Math. Phys. **7**, 560 (1966); D. A. Dubin and J. Tarski, *ibid.* 574 (1966); T. G. Pavlopoulos, Phys. Rev. **159**, 1106 (1967); T. D. Lee and G. C. Wick, Nucl. Phys. **B9**, 209 (1969); **B10**, 1 (1969); Phys. Rev. D **2**, 1033 (1970); D **3**, 1046 (1971); H. P. Dürr and E. Seiler, Nuovo Cimento **A66**, 734 (1970); A. M. Gleeson and E. C. G. Sudarshan, Phys. Rev. D **1**, 474 (1970); A. M. Gleeson, R. J. Moore, H. Rechenberg, and E. C. G. Sudarshan, *ibid.* **4**, 2242 (1971); P. N. Dobson, Jr., J. Math. Phys. **12**, 1207 (1971); R. Casalbuoni, R. Gatto, and G. Longhi, Phys. Rev. D **3**, 1499 (1971); N. Nakanishi, *ibid.*, 811, 1343 (1971); J. Gomatam, *ibid.*, 1292 (1971); J. Schwinger, *ibid.*, 1967 (1971); R. H. Tucker and C. L. Hammer, *ibid.*, 2448 (1971); D. L. Nordstrom, *ibid.* **4**, 1611 (1971).

[6] W. Pauli and F. Villars, Rev. Mod. Phys. **21**, 434 (1949).

[7] A. Landé and L. H. Thomas, Phys. Rev. **65**, 175 (1944).

[8] A. Landé, Phys. Rev. **60**, 121 (1941).

[9] B. Podolsky and C. Kikuchi, Phys. Rev. **65**, 228 (1944); **67**, 184 (1945).

[10] D. J. Montgomery, Phys. Rev. **69**, 117 (1946).

[11] A. E. S. Green, Phys. Rev. **72**, 628 (1947).

[12] B. Podolsky and P. Schwed, Rev. Mod. Phys. **20**, 40 (1948).

[13] A. E. S. Green, Phys. Rev. **73**, 519 (1948); **75**, 1926 (1949).

[14] For a review, see A. E. S. Green and T. Sawada, Rev. Mod. Phys. **39**, 594 (1967).

[15] M. Ostrogradsky, Mem. Acad. St. Petersburg **6**, 385 (1850); E. T. Whittaker, *Treatise on the Analytical Dynamics of Particles and Rigid Bodies* (Cambridge U.P., Cambridge, 1937), 4th ed., p. 265.

[16] T. S. Chang, Proc. Cambridge Phil. Soc. **42**, 132 (1946); **44**, 76 (1948).

[17] J. S. de Wet, Proc. Cambridge Phil. Soc. **44**, 546 (1948), Proc. Roy. Soc. (London) **A195**, 365 (1949).

[18] S. P. Misra, Indian J. Phys. **33**, 461 (1959).

[19] A. O. Barut and G. H. Mullen, Ann. Phys. (N.Y.) **20**, 203 (1962); A. O. Barut, *Electrodynamics and Classical Theory of Fields and Particles* (Macmillan, New York, 1964).

[20] L. K. Pandit, Suppl. Nuovo Cimento **11**, 157 (1959).

[21] K. L. Nagy, Suppl. Nuovo Cimento **17**, 924 (1960); *State Vector Spaces with Indefinite Metric in Quantum Field Theory* (Noordhoff, Groningen, The Netherlands, 1966).

[22] P. A. M. Dirac, Proc. Roy. Soc. (London) **A180**, 1 (1942).

[23] W. Pauli, Rev. Mod. Phys. **15**, 175 (1943).

[24] J. Schwinger, Phys. Rev. **82**, 914 (1951).

[25] J. Schwinger, Phys. Rev. **91**, 713 (1953).

Physical Interpretation of Higher-Derivative Field Theories

F. Riewe and A. E. S. Green
University of Florida, Gainesville, Florida 32601
(Received 14 February 1972)

Polynomial-type field equations are shown to have a realistic physical interpretation in terms of particle form factors, both for classical fields and for "dipole-regularized" quantized fields. Form factors arising from such field equations are found to give a reasonable description of the electromagnetic structure of the proton.

1. INTRODUCTION

The basic mathematical formalism required by higher-derivative, or "generalized," quantum field theories appears to be straightforward and free from inconsistencies.[1] However, when particular higher-derivative field equations are treated in detail, nonphysical results often appear.[2-4] Usually the difficulties relate to negative-energy states and the corresponding lack of conservation of probability. It is our purpose to show that a generalized field theory can be constructed which is mathematically consistent and which has a realistic physical interpretation in terms of the form factors of interacting particles.

Our discussion will center around the Nth-order polynomial meson-field equation introduced by Green.[5,6] We begin by reviewing the procedure for the quantization of generalized free-field equations for the cases in which the roots of the polynomial are real and unequal, real and equal, or complex-conjugate pairs. We then present a physical interpretation involving particle form factors which eliminates certain difficulties from classical generalized electrodynamics. It is shown that the higher-derivative field equation for the interaction of point particles is mathematically equivalent to the usual equations describing particle interactions characterized by form factors. This correspondence is shown to hold for scalar source particles, in agreement with Ueda and Green,[7] and is extended to the case of spin-$\frac{1}{2}$ particles.

The only quantized higher-derivative theories which are generally believed to avoid nonphysical results are those involving "ghost particles."[8] We will demonstrate that the ghost particles can be considered to be the intermediate particles characterizing the form factors of particle interactions, in contrast to the usual assumption that they are physically unobservable. We then test the validity of this interpretation by comparing the form factors derived from physically admissible generalized field theories with experiment. We show that the simplest charged-particle form factors consistent with generalized field theory give a realistic description of the electromagnetic structure of the proton. Hence the form-factor interpretation of generalized field theory is both mathematically consistent and physically reasonable.

2. FREE-FIELD EQUATIONS

A. General Formalism

The polynomial free-field equation

$$\left[\prod_{r=2}^{N}\left(1-\frac{\Box}{m_r^2}\right)\right](\Box - m_1^2)\phi(x) = 0, \tag{2.1}$$

introduced by Green,[5,6] has been frequently discussed for the case in which the polynomial roots m_r (often called "regulator masses" when $r > 1$) are real and unequal.[9] Here we are using $\Box = \nabla^2 - \partial^2/\partial t^2$ with $\hbar = c = 1$. Certain more recent "regularized" models, such as those of Lee[10] and Heisenberg,[11] involve m's which are complex conjugates or are equal in pairs. Quantization often requires the use of an indefinite metric space, as discussed by Nagy.[3,4]

We will now give a brief discussion in which the formalism used by Green is extended to include equal or complex regulator masses. Our basic purpose in this section is to present the mathematical treatment of polynomial free-field equations and to show the negative-energy difficulties encountered by the physical interpretation in terms of auxiliary fields. For general applications our treatment should be more convenient than other discussions[3,4] involving specific models.

The Fourier solution of Eq. (2.1) for unequal roots can be expressed as a sum of auxiliary fields,

$$\phi(x) = \sum_{r=1}^{N} \phi_r(x), \tag{2.2}$$

where

$$\phi_r(x) = \frac{1}{(2\pi)^3} \int d^3k [\phi_r(\mathbf{k}) e^{ik_r x} + \bar{\phi}_r(\mathbf{k}) e^{-ik_r x}],$$
$$k_r = (\mathbf{k}, \omega_r), \quad \omega_r = (\mathbf{k}^2 + m_r^2)^{1/2}. \tag{2.3}$$

Here k_r and x are 4-vectors with the scalar product $k_r x = \mathbf{k}\cdot\mathbf{r} - \omega_r t$. This solution is identical to that given by Green[6] except that we have not specified the relationship between the field components $\phi_r(\mathbf{k})$ and $\bar{\phi}_r(\mathbf{k})$. We will see that the relationship depends on whether the regulator mass m_r is real or complex and on whether quantization is carried out in an indefinite metric space or in the usual positive definite Hilbert space.

The field equation can be derived from the Lagrangian density

$$\mathcal{L} = \tfrac{1}{2}\phi(x) \left[\prod_{r=2}^{N} \left(1 - \frac{\Box}{m_r^2}\right)\right] (\Box - m_1^2)\phi(x) \tag{2.4}$$

(or from any $\mathcal{L}$ differing from this by a 4-divergence). Paralleling the procedure used by Green to find the Hamiltonian H, the expression for the Hamiltonian density can be obtained from $\mathcal{L}$,[1] integrated, and the Fourier solution for $\phi(x)$ inserted to obtain

$$H = \sum_{r=1}^{N} H_r, \tag{2.5}$$

where

$$H_r = \frac{1}{(2\pi)^3} \gamma_r \int d^3k \omega_r^2 (\bar{\phi}_r \phi_r + \phi_r \bar{\phi}_r) \tag{2.6}$$

and

$$\gamma_r = \frac{m_1^2}{m_r^2} \prod_{s=1}^{N'} \left(1 - \frac{m_r^2}{m_s^2}\right). \tag{2.7}$$

The prime (′) on the product symbol indicates that the factor with $r = s$ is omitted. The Heisenberg equation of motion[1] leads to the commutation relations

$$[\phi_r(\mathbf{k}), \bar{\phi}_s(\mathbf{k}')] = (2\gamma_r \omega_r)^{-1} \delta_{rs}\, \delta(\mathbf{k} - \mathbf{k}'). \tag{2.8}$$

The operators $\phi_r(x)$ are often interpreted as representing individual free fields having energies determined by H_r. In this "auxiliary-field" interpretation the fields $\phi_r(x)$ are given the alternative definition

$$\phi_r(x) = \left[\prod_{s=1}^{N'} \frac{\Box - m_s^2}{m_r^2 - m_s^2}\right] \phi(x), \tag{2.9}$$

which satisfies Eq. (2.2). The fields then obey the equations

$$(\Box - m_r^2)\phi_r(x) = 0, \tag{2.10}$$

which have Fourier solutions given by Eq. (2.3). Thus the two definitions are equivalent.

Because the fields $\phi_r(x)$ obey the set of apparently independent equations (2.10), they are often viewed as actually being a set of independent fields.[2] Unfortunately, when calculations are based on this assumption, they yield unphysical results, such as negative-energy states. In Sec. 3 we will show that for classical fields the equations are not independent, so that the difficulties do not occur. Similarly, quantized fields discussed in Sec. 2C have relationships among the ϕ_r when the m_r are complex conjugates or are equal in pairs. In Sec. 4 we show that these quantized fields have a realistic physical interpretation.

B. Real-Root Polynomial Equations

For the special case in which all of the m's in Eq. (2.1) are real and unequal, $\bar{\phi}_r(\mathbf{k})$ must be the adjoint $\phi_r^*(\mathbf{k})$ of $\phi_r(\mathbf{k})$ in order that $\phi(x)$ be self-adjoint. [Note that if the space used for quantization has an indefinite metric, the Hermitian conjugate $\phi_r^\dagger(\mathbf{k})$ will not be the same as the adjoint $\phi_r^*(\mathbf{k})$.] In particular, if the masses are numbered such that $m_{r+1} > m_r$, then the γ's alternate in sign: $\gamma_r = (-1)^{r+1}|\gamma_r|$. Thus for even r, the component $\phi_r(\mathbf{k})$ obeys "wrong-sign" commutation relations:

$$[\phi_r(\mathbf{k}), \phi_r^*(\mathbf{k}')] = -|2\gamma_r \omega_r|^{-1} \delta(\mathbf{k} - \mathbf{k}'). \tag{2.11}$$

For odd r the normal commutation relations (without the minus sign) obtain.

There are two fundamentally different methods which attempt to overcome the problem of the wrong-sign relations: the formalism originally used by Green[6] and the use of an indefinite metric space suggested by Matthews.[12] We will discuss each in turn. In the first method one defines the annihilation operators

$$a_r(\mathbf{k}) = \begin{cases} |2\gamma_r \omega_r|^{1/2} \phi_r(\mathbf{k}), & \text{odd } r, \\ |2\gamma_r \omega_r|^{1/2} \phi_r^*(\mathbf{k}), & \text{even } r, \end{cases} \tag{2.12}$$

In this case $\phi(x)$ can be taken to be Hermitian in a positive metric space so that $\phi_r^*(\mathbf{k}) = \phi_r^\dagger(\mathbf{k})$.

If we make the usual simplifying choice to work with a single value of $\mathbf{k}$ in a space with a countable number of degrees of freedom, the commutation relations take the form

$$[a_r, a_s^\dagger] = \delta_{rs}, \tag{2.13}$$

with all other commutators equal to zero. The Hilbert-space states are defined by

$$a_r|0\rangle = 0, \quad \langle 0|0\rangle = 1,$$
$$|n_r\rangle = (n_r!)^{-1/2} (a_r^\dagger)^{n_r} |0\rangle,$$
$$|n_r n_s \cdots n_t\rangle = |n_r\rangle |n_s\rangle \cdots |n_t\rangle. \tag{2.14}$$

The quantity n_1 is the number of "actual" particles, while n_r, $r > 1$, gives the number of "auxiliary," or "regulator," particles. The Hamiltonian is

$$H = \tfrac{1}{2} \sum_{r=1}^{N} (-1)^{r+1} \omega_r (a_r a_r^\dagger + a_r^\dagger a_r). \tag{2.15}$$

We now see why there are difficulties of interpretation for even the free-particle field equations. Because the ω's are positive, the expectation value of the Hamiltonian is not positive definite. Thus the field can exist in states of negative energy. For this reason the theory is often considered to be unacceptable.

Just as the quantization of the electromagnetic field proposed by Gupta[13] and Bleuler[14] introduces an indefinite metric space to overcome negative-energy

difficulties, it has been suggested[12] that an indefinite metric space be used in the quantization of higher-derivative field equations. In such a space, the annihilation and creation operators are

$$d_r(\mathbf{k}) = (2\gamma_r\omega_r)^{1/2}\phi_r(\mathbf{k}),$$
$$d_r^\dagger(\mathbf{k}) = (2\gamma_r\omega_r)^{1/2}\phi_r^*(\mathbf{k}), \tag{2.16}$$

which satisfy

$$[d_r, d_s^\dagger] = \delta_{rs}. \tag{2.17}$$

Since the Hamiltonian

$$H = \tfrac{1}{2}\sum_r \omega_r(d_r^\dagger d_r + d_r d_r^\dagger) \tag{2.18}$$

appears to have a positive definite form, it might seem that there are no difficulties associated with it.

However, from Eq. (2.16) we see that the operation denoted by an asterisk (*) cannot be the same as Hermitian conjugation since $(2\gamma_r\omega_r)^{1/2}$ is imaginary for even r, so that the field operator $\phi(x)$ is not Hermitian, and neither is the Langrangian from which the field equation is derived. Nevertheless, an indefinite metric vector space can be constructed in which $\phi(x)$ is self-adjoint.

The space in which $\phi(x)$ is self-adjoint is defined by

$$d_r|0) = 0, \quad (0|0) = 1,$$
$$|n_r) = (n_r!)^{-1/2}(d_r^*)^{n_r}|0), \tag{2.19}$$
$$|n_r n_s \cdots n_t) = |n_r)|n_s)\cdots|n_t).$$

In such a space the expectation value of the Hamiltonian with respect to the state $|n_i)$ is

$$\langle H\rangle = (-1)^{n_i\delta_i}n_i\omega_i, \tag{2.20}$$

where

$$\delta_i = \begin{cases} 0, & \text{odd } i, \\ 1, & \text{even } i. \end{cases} \tag{2.21}$$

Unfortunately, the energy $\langle H\rangle$ is still indefinite. In fact, the factor $(-1)^{n_i\delta_i}$, which alternates in sign as more particles are added to the state, makes the result appear even less physical than before. The use of an indefinite metric does not lead to a physically reasonable expression for $\langle H\rangle$ in this case.

To date, there does not appear to be a quantization of higher-derivative field equations with real, unequal masses which completely eliminates the problems of negative-energy states as long as the fields $\phi_r(x)$ are considered to be independent of one another. However, work with the Lee model[10] and the Heisenberg unified field theory[11] has shown that quantization is more successful if there is a relationship between the masses, such as equality or complex conjugation. We next discuss the free-field equations for such cases.

C. Complex- and Dipole-Root Polynomial Equations

If some of the m's in Eq. (2.1) are complex, then the self-adjoint Fourier solution $\phi(x)$ is of the form

$$\phi(x) = \sum_{r=1}^{N}\phi_r(x),$$
$$\phi_r(x) = \frac{1}{(2\pi)^3}\int d^3k\{\phi_r(\mathbf{k})e^{ik_r x} + [\phi_r(\mathbf{k})e^{ik_r x}]^*\},$$
$$k_r = (\mathbf{k}, \omega_r), \quad \omega_r = (\mathbf{k}^2 + m_r^2)^{1/2}. \tag{2.22}$$

For complex m_r the sign of ω_r must be chosen by convention.

When one of the polynomial roots is complex, its complex conjugate must also be a root for the energy to be real. If the two complex-conjugate roots are denoted by m_1 and m_2, the Hamiltonian is then

$$H = \frac{1}{(2\pi)^3}\int d^3k(\gamma_1\omega_1^2(\phi_1\phi_2^* + \phi_2^*\phi_1) + \gamma_2\omega_2^2(\phi_2\phi_1^* + \phi_1^*\phi_2)] + \sum_{r=3}^{N}H_r. \tag{2.23}$$

The commutation relations for $\phi_1(\mathbf{k}), \phi_2(\mathbf{k})$ are

$$[\phi_i(\mathbf{k}), \phi_j^*(\mathbf{k}')] = (2\gamma_i\omega_i)^{-1}(1 - \delta_{ij})\delta(\mathbf{k} - \mathbf{k}'), \tag{2.24}$$

where i and j take on only the values 1 and 2.

Creation and annihilation operators are defined as before by Eq. (2.16) in an indefinite metric space characterized by Eq. (2.19). The energy eigenvalues corresponding to the state $|n_1 n_2)$ are complex, so that[1] $(n_1 n_2|H|n_1 n_2) = 0$. Hence the fields $\phi_1(x)$ and $\phi_2(x)$ do not contribute to the energy and the problem of negative energies does not arise.

The "dipole" field equation, which is the limiting case with $m_1 = m_2$, requires special attention since the Fourier solution has the unusual form

$$\phi(x) = \frac{1}{(2\pi)^3}\int d^3k[\phi_1 + \omega t\phi_2)e^{ikx} + (\phi_1^* + \omega t\phi_2^*)e^{-ikx}] + \sum_{r=3}^{N}\phi_r,$$
$$k = (\mathbf{k}, \omega), \quad \omega = (\mathbf{k}^2 + m_1^2)^{1/2} = (\mathbf{k}^2 + m_2^2). \tag{2.25}$$

However, the Hamiltonian and commutation relations are similar to the complex-conjugate case. Although the state space has vectors of zero norm,[3,4] a consistent treatment is possible and the expectation value of H with respect to the state $|n_1 n_2)$ remains zero.

Both the complex-conjugate and the dipole free-field equations avoid the nonphysical results associated with negative energies. However, it can be shown that when certain higher-order interactions are considered, the complex-conjugate case allows transitions to negative-energy states and is therefore unsatisfactory.[15] Nevertheless, "dipole-regularized" equations, or "dipole-ghost" equations, of the type

$$\left[\prod_{j=2}^{M}\left(1 - \frac{\Box}{m_j^2}\right)^2\right](\Box - m_1)\phi(x) = 0 \tag{2.26}$$

appear to give meaningful results,[16] particularly when applied to the models of Lee[10] and Heisenberg.[11] We now turn to the question of the physical interpretation of such equations.

3. PHYSICAL INTERPRETATION OF CLASSICAL FIELDS

A. Nonrelativistic Spinless Source Fields

Since the existence of auxiliary fields and ghost particles is required by many field-theoretical models, it is important to understand their physical significance. In this section we will consider the case of the generalized classical electromagnetic field and present a physically consistent interpretation in terms of particle form factors. In Sec. 4 the feasibility of extending the interpretation to quantized fields is investigated. We will find that the problems of interpretation encountered with classical free fields will not occur if the source of the field is taken into consideration.

First consider the case of a classical source particle characterized by a charge density $\rho(\mathbf{r})$. It is assumed that the particle is radially symmetric and nonrotating and also that it can be treated nonrelativistically. Spin and relativistic effects will be discussed later.

As usual, the form factor of the particle is defined to be the Fourier transform of the charge density:

$$F(\mathbf{k}^2) = \frac{1}{q}\int d^3r\, e^{i\mathbf{k}\cdot\mathbf{r}}\rho(r), \tag{3.1}$$

where q is the total charge of the particle. Since the particle is radially symmetric, the form factor is a function only of $\mathbf{k}^2$.

For our purposes, form factors will be useful because of their connection with generalized field theory. We will show that applying generalized electrodynamics to classical point particles yields results identical to those obtained by applying the usual electrodynamics to particles having form factors. The present work is based on that of Ueda and Green,[7] and extends it to apply to particles with spin $\frac{1}{2}$.

We will consider a generalization of electrodynamics in which the 4-vector potential $A = (A, \phi)$ obeys the classical field equation

$$f(\Box)\Box A_\mu = -4\pi j_\mu, \tag{3.2}$$

where j_μ is the 4-vector current density $j = (\mathbf{j}, \rho)$ which is the source of the field. For the special case $f(\Box) = (1 - \Box/\Lambda^2)$ the theory is the same as that of Podolsky[17] and Bopp.[18] Our physical interpretation will, however, be different from that of the above authors.

To determine the relationship between generalized electrodynamics and form factors, consider a generalized field with a point-charge source. We will work in the rest frame of the particle, so that $\dot{\phi} = 0$. The generalization of Poisson's equation for a point charge q located at the origin is then

$$f(\nabla^2)\nabla^2\phi(r) = -4\pi q\,\delta(\mathbf{r}). \tag{3.3}$$

Defining $\phi(\mathbf{k})$ by

$$\phi(\mathbf{r}) = \frac{1}{(2\pi)^3}\int d^3k\phi(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{3.4}$$

we find that the Fourier transform of Poisson's equation is

$$f(-\mathbf{k}^2)(-\mathbf{k}^2)\phi(\mathbf{k}) = -4\pi q \tag{3.5}$$

or

$$(-\mathbf{k}^2)\phi(\mathbf{k}) = -4\pi q F(\mathbf{k}^2), \tag{3.6}$$

where

$$F(\mathbf{k}^2) = 1/f(-\mathbf{k}^2). \tag{3.7}$$

Taking the reverse Fourier transform yields

$$\nabla^2\phi(\mathbf{r}) = -4\pi\rho(\mathbf{r}), \tag{3.8}$$

where

$$\rho(\mathbf{r}) = \int d^3k q F(\mathbf{k}^2)e^{i\mathbf{k}\cdot\mathbf{r}}; \tag{3.9}$$

i.e., $F(\mathbf{k}^2)$ is the form factor corresponding to the density $\rho(\mathbf{r})$. Thus the static potential of a point charge as calculated from generalized electrodynamics is the same as the potential which satisfies the usual Poisson equation for Maxwell's theory of electromagnetism, provided that the Maxwell particle has the form factor $f^{-1}(-\mathbf{k}^2)$.

In the generalized classical electrodynamics we have thus far considered, there are no nonphysical results since the generalized Poisson equation,

$$f(\nabla^2)\nabla^2\phi(\mathbf{r}) = -4\pi q\delta(\mathbf{r}), \tag{3.10}$$

is mathematically equivalent to the usual equation,

$$\nabla^2\phi(\mathbf{r}) = -4\pi\rho(\mathbf{r}). \tag{3.11}$$

Given either of the equations, the other can be calculated from it. Since $\phi(\mathbf{r})$ can be calculated from the second of these equations, there can be nothing unphysical about it, even though it is also a solution of the first. However, for free fields, the situation changes.

A free field is a solution to the field equation in a region where $j_\mu = 0$. The two free-field wave equations corresponding to the pair of equivalent Poisson equations are

$$f(\Box)\Box A_\mu = 0 \tag{3.12}$$

and

$$\Box A_\mu = 0. \tag{3.13}$$

Unfortunately, these two are *not* equivalent. The first has more solutions; given only the second, it is impossible to reconstruct the first. While Eq. (3.13) is the usual form of the wave equation, some authors feel that the classical auxiliary-particle interpretation of Eq. (3.12) is not physically reasonable.[2] If $f(\Box)$ is an Nth-order polynomial in $\Box$ with nonzero roots, then according to Eq. (2.10) the generalized free-field equation (3.12) is equivalent to

$$\Box A_\mu^{(1)} = 0 \tag{3.14}$$

plus the set of N equations

$$(\Box - m_r^2)A_\mu^{(r)} = 0. \tag{3.15}$$

If these equations are thought to be independent, their physical interpretation appears obvious. The first is the usual wave equation for the electromagnetic potential, or, equivalently, the equation for a zero-mass

particle, while the others are the wave equations of particles of mass m_γ. Both $A_\mu^{(1)}$ and $A_\mu^{(\gamma)}$ are solutions to the generalized wave equation so that the general solution for A_μ is a linear combination of $A_\mu^{(1)}$ and $A_\mu^{(\gamma)}$. As pointed out by Feynman,[19] this violates physical experience since one particle will travel at the speed of light while the others must move more slowly. A burst of light obeying a generalized wave equation would gradually separate into wavefronts moving at different velocities. Thus it would seem that generalized electrodynamics does not offer a reasonable description of free fields, even classically.

For classical fields a solution to the dilemma is to reexamine the concept of "free fields." If the usual Maxwell's equations are solved for the case that the source j_μ is a moving point particle, the results are the well-known Liénard–Wiechert fields. The expressions for the fields contain two types of terms: the "near field" terms, which fall off like R^{-2} where R is the distance at the retarded time, and the "radiation" terms, which fall off like R^{-1} and vanish when the particle is not accelerated. Although the near field can be thought of as moving with the particle, the radiation field propagates outward at the speed of light and does not depend on the future position of the particle. After its emission it is free of the particle and hence is often called a "free field."

If the source is an extended particle with a charge density other than a δ function, then at any point in space and time all of the components of j_μ will not generally be zero, so that the radiation field is part of the solution of $\Box A_\mu = -4\pi j_\mu$, while the homogeneous equation $\Box A_\mu = 0$ is not a valid field equation. For the same reason, we will take

$$f(\Box)\Box A_\mu = -4\pi j_\mu, \tag{3.16}$$

where j_μ is nonzero, as the equation from which to determine the generalized radiation field, or free field.

In analogy to the equivalence of Eqs. (3.10) and (3.11) there is a function j'_μ such that

$$\Box A_\mu = -4\pi j'_\mu \tag{3.17}$$

has the same solution as Eq. (3.16). Thus the Liénard–Wiechert radiation-field solution of Eq. (3.17), which, of course, is physically reasonable, is also the "free-field" solution of the generalized equation (3.16). Hence by treating free fields as radiation fields the paradoxes associated with classical auxiliary fields can be avoided.

B. Relativistic Spin-$\frac{1}{2}$ Source Particles

We next show that the same relationship between generalized fields and form factors is valid for relativistic source particles with spin $\frac{1}{2}$. For the present we will consider the Dirac equation as a "classical" wave equation in the sense that there will be no attempt to quantize it by introducing commutation relations. The discussion will be illuminating, not only because the form factor interpretation is being extended to the case of spin $\frac{1}{2}$, but also because we will employ an entirely different formalism from that of the preceding discussion.

For spin-$\frac{1}{2}$ particles like the proton, which are known to have form factors and anomalous magnetic moments, Foldy[20] has suggested the following equation:

$$\left(\not\partial + m - i\sum_{n=0}^{\infty}(\epsilon_n\Box^n\not A + \mu_n\Box^n\sigma^{\mu\nu}F_{\mu\nu})\right)\psi = 0. \tag{3.18}$$

The constants ϵ_n and μ_n characterize the charge and magnetic-moment distributions of the particle. In particular, ϵ_0 is the particle's charge and μ_0 is its anomalous magnetic moment. $F_{\mu\nu}$ is the field tensor

$$F_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu \tag{3.19}$$

and

$$\sigma^{\mu\nu} = \tfrac{1}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu). \tag{3.20}$$

The equation was chosen to be the most general extension of the Dirac equation to describe a particle whose structure can be characterized by electric and magnetic moments, provided the additional terms are Lorentz covariant, gauge invariant, linear in the potentials, and do not vanish for zero particle momentum. It can be shown[21] that ϵ_n and μ_n completely characterize the form factors of spin-$\frac{1}{2}$ particles.

We will now show that the Foldy–Dirac equation (3.18), combined with the usual electromagnetic field equations, is mathematically equivalent to the interaction of a Dirac particle with a generalized electromagnetic field. The magnetic moment terms can be rewritten in a more convenient form: $\sigma^{\mu\nu}F_{\mu\nu} = 2\partial^\mu A_\mu - 2\not\partial\not A$. We will work in the Lorentz gauge, so that

$$\sigma^{\mu\nu}F_{\mu\nu} = -2\not\partial\not A. \tag{3.21}$$

The Foldy–Dirac equation can now be rewritten:

$$\left(\not\partial + m - i\sum_{n=0}^{\infty}[\epsilon_n\not\partial^{2n}\not A + \mu_n\not\partial^{2n}(-2\not\partial\not A)]\right)\psi = 0, \tag{3.22}$$

or, by regrouping terms and choosing new constants α_m,

$$\left(\not\partial + m - i\sum_{m=0}^{\infty}\alpha_m\not\partial^m\not A\right)\psi = 0. \tag{3.23}$$

Let $\bar A_\mu(x)$ be a potential for which the spin-$\frac{1}{2}$ particle obeys the point-particle equation

$$(\not\partial - ie\not{\bar A} + m)\psi = 0. \tag{3.24}$$

It follows that

$$\epsilon_0\not{\bar A} = \sum_{m=0}^{\infty}\alpha_m\not\partial^m\not A. \tag{3.25}$$

We then find

$$\epsilon_0\Box\not{\bar A} = -4\pi\sum_{m=0}^{\infty}\alpha_m\not\partial^m j. \tag{3.26}$$

The Fourier transform is

$$\epsilon_0(-k^2)\not{\bar A}(k) = -4\pi\sum_{m=0}^{\infty}\alpha_m(i\not k)^m\not j(k). \tag{3.27}$$

Since the matrix $\not k$ has the inverse $(\not k)^{-1} = \not k/k^2$, it is possible to determine the series expansion of the inverse of an arbitrary function of $\not k$, so that $\not{\bar A}(k)$ obeys the equation

$$f(i\not k)(-k^2)\not{\bar A}(k) = -4\pi\not j(k), \tag{3.28}$$

where

$$f(i\not k) = \epsilon_0\left(\sum_{m=0}^{\infty}\alpha_m(i\not k)^m\right)^{-1}. \tag{3.29}$$

The inverse Fourier transform gives

$$f(\partial)\Box\vec{A}(x) = -4\pi j(x). \tag{3.30}$$

We have thus shown that the Foldy–Dirac equation combined with the usual equations of electrodynamics is mathematically equivalent to the Dirac equation (3.24) if the electromagnetic potential is taken to obey the generalized wave equation (3.30). Hence the form-factor interpretation can equally well be applied to source particles with spin $\frac{1}{2}$. No problems appear classically as long as the electromagnetic field is always considered to have an extended source. We next consider the effects of quantization.

4. PHYSICAL INTERPRETATION OF QUANTIZED FIELDS

The classical form-factor interpretation of the previous section carries over directly into quantum theory. In quantum field theory the form factor $F(k^2)$ is a vertex operator, which describes the coupling between fields. For example, the vertex diagram for interacting scalar fields given in Fig. 1(a) represents a process with amplitude $(k^2 - m^2)^{-1}F(k^2)$. In theoretical calculations of actual form factors[22] the lowest-order contributions come from intermediate states which are particle resonances. Figure 1(b) gives such a form factor, characterized by a single intermediate particle with mass λ. The amplitude for this diagram is the same as for Fig. 1(a), so that the two processes can be considered to be equivalent.

A third process having the same amplitude is illustrated in Fig. 1(c). In this case there is a point vertex and a generalized field which obeys

$$(1 - \Box/\lambda^2)(\Box - m^2) = 0. \tag{4.1}$$

We see that just as in the classical case the usual interaction of a field with a particle having a form factor is equivalent to a generalized field interacting with a point particle. The auxiliary particles, or ghost particles, of mass λ in Fig. 1(c) correspond to the intermediate particles in Fig. 1(b). For the general case of a polynomial field the same correspondence applies. For example, a field with two regulator masses λ_1 and λ_2 corresponds to a form factor $F(k^2) = (1 + k^2/\lambda_1^2)^{-1}(1 + k^2/\lambda_2^2)^{-1}$. In this case the intermediate particle λ_2 is itself coupled to the source field via a second intermediate particle λ_1.

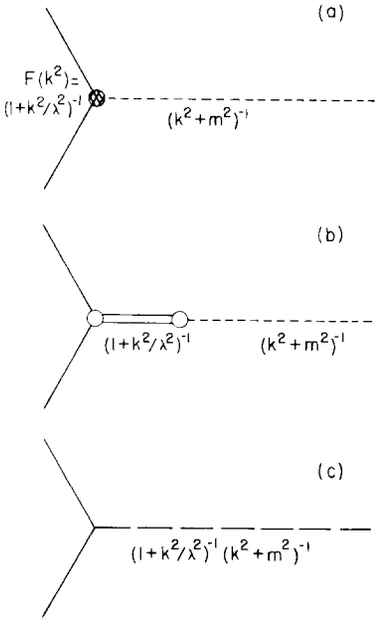


FIG. 1. Three vertex diagrams, each representing a process with amplitude $(1 + k^2/\lambda^2)^{-1}(k^2 + m^2)^{-1}$. (a) Vertex with monopole form factor. (b) Vertex characterized by an intermediate particle of mass λ. (c) Point vertex with a generalized field having a "ghost particle" of mass λ.

For higher-order diagrams the interpretation can still be applied. However, as discussed in Sec. 2 certain calculations involving such diagrams lead to nonphysical results unless the generalized field obeys the dipole-regularized field equation (2.26). Thus from Eq. (3.7) the only admissible form factors must be of the type

$$F_N(k^2) = \left[\prod_{j=1}^{N}\left(1 + \frac{k^2}{\Lambda_j^2}\right)^2\right]^{-1}. \tag{4.2}$$

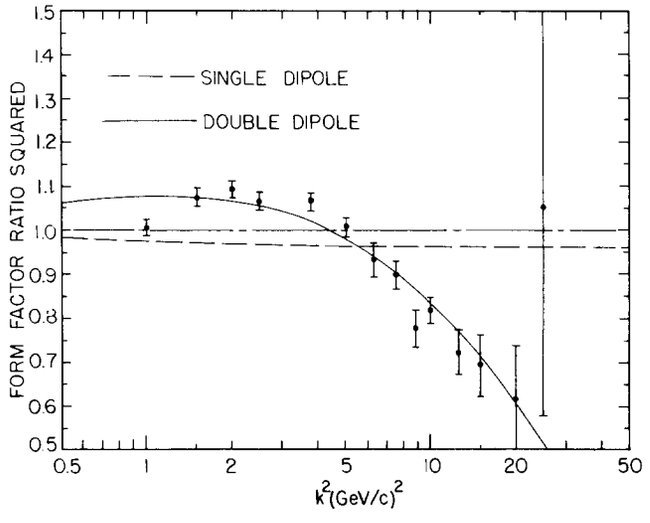


FIG. 2. Form factors consistent with generalized field theory. Dashed line: single dipole with $\Lambda = 0.702$ (GeV/c)^2. Solid line: double dipole with $\Lambda_1 = 0.744$ (GeV/c)^2 and $\Lambda_2 = 107.7$ (GeV/c)^2.

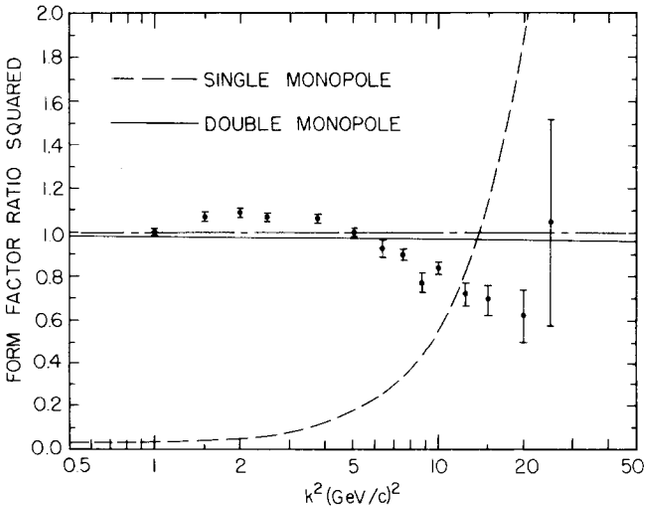


FIG. 3. Form factors not consistent with generalized field theory. Dashed line: single monopole with $\Lambda^2 = 0.033$ (GeV/c)^2. Solid line: double monopole with $\Lambda_1^2 = 0.701$ (GeV/c)^2 and $\Lambda_2 = 0.705$ (GeV/c)^2.

The result that a form factor consistent with generalized field theory must be a product of dipoles provides a test of the theory. Since the only charged particle with an experimentally well-known form factor is the proton, it will be used as the basis of our test.

Following the procedure of Green and Ueda,[23] we will compare the square of the theoretical form factor with the experimental scattering cross sections, using SLAC[24] electron–proton scattering data.

The simplest dipole form factor is the single dipole

$$F_1(k^2) = (1 + k^2/\Lambda^2)^{-2}, \tag{4.3}$$

which has one adjustable parameter. F_1 was first discovered empirically and is known as the Hofstadter–Wilson dipole form factor. Hofstadter[25] and Wilson[26] choose $\Lambda^2 = 0.71(\mathrm{GeV}/c)^2$.

The dipole form factor is often used as the standard against which others can be compared. Thus, in Fig. 2, we plot the ratios of the data and theoretical calculations to the dipole form factor. This method clearly shows any discrepancies between theory and experiment. In the figure we show the dipole form factor both for $\Lambda^2 = 0.710(\mathrm{GeV}/c)^2$ and also for $\Lambda^2 = 0.702(\mathrm{GeV}/c)^2$, which gives a best fit to the more recent data. As a measure of the accuracy of the agreement with experiment we use the usual quantity χ^2. The readjusted dipole gives $\chi^2 = 6.96$. The double dipole form factor,

$$F_2(k^2) = (1 + k^2/\Lambda_1^2)^{-2}(1 + k^2/\Lambda_2^2)^{-2}, \tag{4.4}$$

gives a much better fit, shown in Fig. 2. The value of χ^2 in this case is 1.28.

We thus find that a satisfactory fit to the data can be obtained with only two adjustable parameters, if the form factor is derived from generalized field theory. In contrast, the form factor which is a product of monopole terms,

$$F = \prod_{j=1}^{N} \left(1 + \frac{k^2}{\Lambda_j^2}\right)^{-1}, \tag{4.5}$$

was found by Green and Ueda[23] to require three parameters in order to obtain significant improvement over the one-parameter dipole form factor. For purposes of comparison with the dipole case, the best fits to the data for monopole-type form factors for $N = 1$ and $N = 2$ are given in Fig. 3. The $N = 1$ case gives a very poor fit, with $\chi^2 = 346.8$. For $N = 2$, the best fit essentially duplicates the (one-parameter) Hofstadter–Wilson dipole form factor, again with $\chi^2 = 6.96$.

A similar situation occurs if λ_1 and λ_2 are chosen to be a pair of complex conjugates, in which case the best fit requires λ_1 and λ_2 to be approximately equal. It therefore appears that the requirement that the proton electromagnetic form factor be consistent with generalized field theory results in a very realistic functional form.

We conclude that higher-derivative quantum field theories have a realistic physical interpretation in terms of form factors, in which the "ghost particles" associated with generalized fields represent intermediate states contributing to the form factor.

ACKNOWLEDGMENT

Early phases of this work were supported in part by the U.S. Air Force Office of Scientific Research Grant No. AFOSR-68-1397.

1 F. Riewe and A. E. S. Green, J. Math. Phys. **13**, 1368 (1972), preceding paper.
2 A. Pais and G. E. Uhlenbeck, Phys. Rev. **79**, 145 (1950).
3 K. L. Nagy, Suppl. Nuovo Cimento **17**, 924 (1960).
4 K. L. Nagy, *State Vector Spaces with Indefinite Metric in Quantum Theory* (Noordhoff, Groningen, Netherlands, 1966).
5 A. E. S. Green, Phys. Rev. **73**, 26, 519 (1948).
6 A. E. S. Green, Phys. Rev. **75**, 1926 (1949).
7 T. Ueda and A. E. S. Green, Nucl. Phys. **B10**, 289 (1969). A similar interpretation, in terms of "partly dressed" particles, is given by A. O. Barut, P. Cordero, and G. C. Ghirardi, Phys. Rev. **182**, 1844 (1969).
8 K. L. Nagy, Ref. 4, Chap. 3.
9 For a review see A. E. S. Green and T. Sawada, Rev. Mod. Phys. **39**, 594 (1967).
10 T. D. Lee, Phys. Rev. **95**, 1329 (1954); W. Heisenberg, Nucl. Phys. **4**, 532 (1957).
11 W. Heisenberg, Rev. Mod. Phys. **29**, 269 (1957); W. Heisenberg, *Introduction to the Unified Field Theory of Elementary Particles* (Interscience, New York, 1966).
12 P. T. Matthews, Proc. Cambridge Phil. Soc. **45**, 441 (1949).
13 S. N. Gupta, Proc. Phys. Soc. (London) **A63**, 691 (1950).
14 K. Bleuler, Helv. Phys. Acta **23**, 567 (1950).
15 K. L. Nagy, Ref. 4, pp. 84–85; A. M. Gleeson and E. C. G. Sudarshan, Phys. Rev. D **1**, 474 (1970).
16 K. L. Nagy, Ref. 4, p. 84. Note that the objections raised by N. Nakanishi, Phys. Rev. D **3**, 811, 1343 (1971), are answered by T. D. Lee and G. C. Wick, Phys. Rev. D **3**, 1046 (1971).
17 B. Podolsky, Phys. Rev. **62**, 68 (1942).
18 F. Bopp, Ann. Physik **38**, 345 (1940).
19 R. P. Feynman, Phys. Rev. **74**, 939 (1948).
20 L. I. Foldy, Phys. Rev. **87**, 688 (1952).
21 See, for example, G. Salzman, Phys. Rev. **105**, 1076 (1957); K. W. McVoy and L. Van Hove, *ibid.* **125**, 1034 (1962).
22 S. D. Drell and F. Zachariasen, *Electromagnetic Structure of Nucleons* (Oxford U.P., Oxford, 1961).
23 A. E. S. Green and T. Ueda, Phys. Rev. Letters **21**, 1499 (1968); **22**, 119 (1969).
24 D. H. Coward, H. DeStaebler, R. A. Early, J. Litt, A. Minten, L. W. Mo, W. K. H. Panofsky, R. E. Taylor, M. Breidenbach, J. I. Friedman, H. W. Kendall, P. N. Kirk, B. C. Barish, J. Mar, and J. Pine, Phys. Rev. Letters **20**, 292 (1968).
25 T. Janssens, R. Hofstadter, E. B. Hughes, and M. R. Yearian, Phys. Rev. **142**, 922 (1966).
26 M. Goitein, J. R. Dunning, and R. Wilson, Phys. Rev. Letters **18**, 1018 (1967).

Scattering by Two Charged Centers*

Ming Chiang Li
Department of Physics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
(Received 7 February 1972)

The scattering of a charged particle by two fixed charged centers is discussed. The scattering potential is long-range and spheroidal. It is pointed out that the general method for handling the short-range spheroidal potential is not directly applicable here. The integral differential equation for the Coulomb spheroidal phase shift is given in the text. The behavior of the phase shift is discussed in Born's approximation. A method for solving for the radial spheroidal Coulomb wavefunction is also given.

1. INTRODUCTION

Nonspherical, particularly spheroidal, scattering gives rise to interesting problems. It was introduced quite early in the classical scattering theory of sound and electromagnetic waves. The study of acoustic scattering[1] by circular disks and apertures, which are spheroidal in form, was of importance for the so-called Rayleigh-disk method of measuring sound intensity. Furthermore, a great deal of work has been done on electromagnetic scattering by spheroids, disks, and apertures. For example, Meixner and Andrajewski and Flammer[2] treated the problem of the scattering of plane electromagnetic waves by a perfectly conducting circular disk. The effect of circular disks and apertures on the radiation from electric and magnetic dipoles was investigated by Meixner and Flammer.[3] In addition, the problem of the scattering of plane electromagnetic waves incident on a perfectly conducting prolate spheroid has been studied by Schultz, Siegel, and collaborators.[4]

Spheroidal scattering is interesting not only from the point of view of classical physics, but also from that of quantum mechanics. It was shown in the early years of quantum mechanics by Burrau[5] that the Coulomb potential produced by a pair of fixed charges (of whatever relative sign) is spheroidally symmetric. Hence the scattering of an electron by the fixed charge pair is spheroidal. The scattering of slow electrons by diatomic molecules has been considered to be spheroidal by Stier and by Fisk.[6] (See Naghara, Takayanagi and Hara[7] for extensions of Stier and Fisk.)

The importance of spheroidal potentials can also be traced to the physics of elementary particles. At present, elementary particles are not always treated as mathematical particles. In fact the elementary particles are neither elementary, nor particles if one employs a classical description. Rather, they are treated as a composite system with some structure. This view has been incorporated in the quark,[8] parton,[9] and droplet models.[10] Among them, the one most directly related to the spheroidal scattering is the droplet model, in which the high energy scattering process is treated as a wave passing through a Lorentz-contracted optical medium. The medium has either a disc or a pancake shape, both of which are spheroidal. In other words, even in high energy physics spheroidal scattering plays a significant role.

Despite its importance, spheroidal scattering is not as well studied as is spherical scattering. Investigations to date have been mostly on the scattering from discontinuous spheroidal boundaries and on the numerical calculation for comparison with observed experimental phenomena. Thus, any further study on spheroidal scattering should be of considerable interest in physics.

Scattering by a short-range spheroidal potential has been our main concern in an earlier paper.[11] The present work can be considered as a continuation of the above-mentioned, and is devoted to the long-range Coulomb potential. In Sec. 2 we point out that the Coulomb potential produced by a pair of two equal charged centers is spheroidal. A general spheroidal phase shift expression for short range potential is given; it is a straightforward generalization of the one in the previous paper, in which the incident wave is limited to the z-direction of the coordinate system. This expression is shown to be unsuitable for the Coulomb case. In Sec. 3 the spheroidal Coulomb phase shift is discussed and is expressed in term of an integro-differential equation. In Sec. 4 we deal with a Born approximation for the spheroidal Coulomb phase shift. It is shown that spheroidal and spherical high order Coulomb phase shifts are the same. In Sec. 5 we present a method solving the integro-differential equation for the radial wavefunction and the validity of the solution is verified.

2. POTENTIAL

A pair of equal fixed point charges are located on z axis with coordinates $\mathbf{r}_a = (0, 0, d/2)$ and $\mathbf{r}_b = (0, 0, -d/2)$, where d is the separation between the two charges. The potential V at distance $\mathbf{r}$ has the form

$$V = Q/|\mathbf{r} - \mathbf{r}_a| + Q/|\mathbf{r} - \mathbf{r}_b|. \tag{2.1}$$

Q is the charge on each point. The above potential, although not spherically symmetric, does possess another type of symmetry. This can be seen through the introduction of the prolate spheroidal coordinate system,[12] in which the distance $\mathbf{r} = (x, y, z)$ is expressed as

$$\begin{aligned} x &= (d/2)[(1-\eta^2)(\xi^2-1)]^{1/2}\cos\phi, \\ y &= (d/2)[(1-\eta^2)(\xi^2-1)]^{1/2}\sin\phi, \\ z &= (d/2)\eta\xi, \end{aligned} \tag{2.2}$$

with $1 \leq \xi < \infty$, $-1 \leq \eta \leq 1$, $0 \leq \phi \leq 2\pi$. The separation d between the two fixed charges is now called the interfocal distance for such a spheroidal coordinate system. Parameters ξ, η, ϕ are referred to as the prolate spheroidal coordinates. Under the limit,

$$d \to 0, \quad \tfrac{1}{2}d\xi \to r, \quad \text{and} \quad \eta \to \cos\theta, \tag{2.3}$$

the spheroidal coordinate system is reduced to a spherical one, where r and θ are the spherical coordinates. Potential V in Eq. (2.1) in terms of prolate spheroidal coordinates can be expressed as

$$V = (4Q/d)[\xi/(\xi^2 - \eta^2)]. \tag{2.4}$$

This is a potential with prolate spheroidal symmetry.

The Schrödinger equation for describing a charged particle scattered by the potential in Eq. (2.4) has the form

$$-(\hbar^2/2\mu)\nabla^2\psi + qV\psi = (\hbar^2k^2/2\mu)\psi, \tag{2.5}$$

where μ is the mass, q the charge, and k the momentum of the incident particle. Equation (2.5) may be expressed simply in the prolate spheroidal coordinates as

$$\left[\frac{\partial}{\partial\eta}(1-\eta^2)\frac{\partial}{\partial\eta} + \frac{\partial}{\partial\xi}(\xi^2-1)\frac{\partial}{\partial\xi} + \left(\frac{1}{1-\eta^2} - \frac{1}{\xi^2-1}\right)\frac{\partial^2}{\partial\phi^2} + (\tfrac{1}{2}kd)^2(\xi^2-\eta^2) - \frac{2\mu dQq}{\hbar^2}\xi\right]\psi = 0. \tag{2.6}$$

The above equation is separable, and can be solved through the spheroidal phase shift analysis method. This method has been extensively discussed in the classical scattering theories of sound and electromagnetic waves, and was used in the scattering of electrons by diatomic molecules.[6,7] The general method for describing scattering by a short-range oblate spheroidal potential can be found in a more recent paper.[11] In terms of the spheroidal phase shift analysis, the scattering amplitude may be expressed as

$$f_k(\theta,\phi) = \frac{1}{ik}\sum_n\sum_m \frac{(2-\delta_{0m})}{N_{mn}(c)} S_{mn}(c,\eta') \times S_{mn}(c,\eta)(e^{2i\delta_{mn}} - 1)\cos m(\phi-\phi'), \tag{2.7}$$

where

$$c = \tfrac{1}{2}kd, \qquad \theta = \cos^{-1}\eta. \tag{2.8}$$

During scattering, the incident particle is in the direction specified by polar angle $\theta' \equiv \cos^{-1}\eta'$ and azimuthal angle ϕ'. $N_{mn}(c)$ are the normalization constants of the prolate spheroidal angle functions $S_{mn}(c,\eta)$:

$$\int_{-1}^{1} S_{mn}(c,\eta)S_{mn'}(c,\eta)\,d\eta = \begin{cases} 0, & \text{for } n \neq n', \\ N_{mn}(c), & \text{for } n = n. \end{cases} \tag{2.9}$$

The problem now is the determination of spheroidal phase shifts δ_{mn}. It is the same problem as that encountered in the spherical phase shift analysis, which can be considered as a special case of the present analysis under the limit $d \to 0$. In the spherical analysis, the Coulomb potential is a particular one, and is considered separately as in contrast with short-range potentials.[13] This is also true for the spheroidal Coulomb potential; its peculiarity will be briefly explained as follows.

For a general short range spheroidal potential, one can obtain an integral equation[11] for the spheroidal phase shift δ_{mn}

$$e^{i\delta_{mn}}\sin\delta_{mn} = -c\int_1^\infty d\xi\, R^{(1)}_{mn}(c,\xi)\tilde{U}(\xi)\tilde{T}_{mn}(c,\xi), \tag{2.10}$$

where $R^{(1)}_{mn}(c,\xi)$ is the spheroidal radial function, the spheroidal short-range potential $\tilde{V}(\xi,\eta,\phi)$ is expressed as

$$(\mu d^2/2\hbar^2)(\xi^2-\eta^2)\tilde{V}(\xi,\eta,\phi) \equiv \tilde{U}(\xi), \tag{2.11}$$

$\tilde{T}_{mn}(c,\xi)$ is defined through the expansion of the scattered wave $\psi_k^{(+)}(\mathbf{r})$

$$\psi_k^{(+)}(\mathbf{r}) = 2\sum_n\sum_m i^n \frac{(2-\delta_{0m})}{N_{mn}(c)} S_{mn}(c,\eta') \times S_{mn}(c,\eta)\cos m(\phi-\phi')\,\tilde{T}_{mn}(c,\xi) \tag{2.12}$$

and has the asymptotic expression

$$\tilde{T}_{mn}(c,\xi) \underset{\xi\to\infty}{\longrightarrow} \frac{1}{c\xi}e^{i\delta_{mn}}\cos[c\xi - \tfrac{1}{2}(n+1)\pi + \delta_{mn}]. \tag{2.13}$$

The Coulomb potential in Eq. (2.4) is long range in nature, and causes the integral in Eq. (2.10) to be logarithmically divergent. On account of this, new formulations should be sought to describe the long-range spheroidal Coulomb phase shift.

3. COULOMB PHASE SHIFTS

In this section we would like to derive an explicit expression for the spheroidal phase shifts from potential V in Eq. (2.4). We shall begin with Schrödinger's equation (2.6). By the usual separation procedure, its solutions can be obtained in the form of the Lamé products:

$$\psi_{mn} = T_{mn}(c,\xi)S_{mn}(c,\eta)e^{im\phi}, \tag{3.1}$$

where $S_{mn}(c,\eta)$ is the prolate spheroidal angle function as used in Eq. (2.7). Function $T_{mn}(c,\xi)$ satisfies the ordinary differential equation

$$\frac{d}{d\xi}\left((\xi^2-1)\frac{d}{d\xi}T_{mn}(c,\xi)\right) - \left(\lambda_{mn}(c) - c^2\xi^2 + Ad\xi + \frac{m^2}{\xi^2-1}\right)T_{mn}(c,\xi) = 0, \tag{3.2}$$

where

$$A = 2\mu Qq/\hbar^2. \tag{3.3}$$

Constant m is an integer, which comes from the single value requirement of the wavefunction. Constant $\lambda_{mn}(c)$ is the eigenvalue of function $S_{mn}(c,\eta)$ and can be expressed as

$$\lambda_{mn}(c) = \sum_k l^{mn}_{2k}c^{2k}. \tag{3.4}$$

The first two coefficients are found to be

$$l^{mn}_0 = n(n+1), \tag{3.5}$$

$$l^{mn}_2 = \tfrac{1}{2}\left(1 - \frac{(2m-1)(2m+1)}{(2n-1)(2n+3)}\right). \tag{3.6}$$

By substituting $c\xi = y$ and $T_{mn}(c,\xi) = [(y^2-c^2)/y^2]^{m/2} \times Y_{mn}(y)$, Eq. (3.2) becomes as

$$(y^2-c^2)\frac{d^2}{dy^2}Y_{mn}(y) + \left(2y + 2m\frac{c^2}{y}\right)\frac{d}{dy}Y_{mn}(y) - \left(\lambda_{mn}(c) - y^2 + \frac{2A}{k}y + \frac{m(m+1)c^2}{y^2}\right)Y_{mn}(y) = 0. \tag{3.7}$$

Under the limit $d \to 0$, for which two fixed point charges coincide with each other, Eq. (3.7) is reduced to the form

$$y^2\frac{d^2}{dy^2}H_n(y) + 2y\frac{d}{dy}H_n(y) - [n(n+1) - y^2 + 2\bar{\eta}y]H_n(y) = 0, \tag{3.8}$$

where we have used

$$\lim_{d\to 0} Y_{mn}(y) = H_n(y) \tag{3.9}$$

and

$$\bar{\eta} = A/k. \tag{3.10}$$

Equation (3.8) is a radial Coulomb wave equation. Its general solutions are well known[14]:

$$H_n(y) = c_1 y^{-1} F_n(\bar{\eta}, y) + c_2 y^{-1} G_n(\bar{\eta}, y), \tag{3.11}$$

where c_1 and c_2 are arbitrary constants. $F_n(\bar{\eta}, y)$ is the regular Coulomb wavefunction and $G_n(\bar{\eta}, y)$ is the irregular (logarithmic) Coulomb wavefunction. From recurrence relations of Coulomb wavefunctions one obtains the following relations:

$$\frac{dH_n(y)}{dy} = \frac{1}{2n+1}\{(n^2+\bar{\eta}^2)^{1/2} H_{n-1}(y) - [(n+1)^2+\bar{\eta}^2]^{1/2} H_{n+1}(y)\}, \tag{3.12}$$

$$y^{-1}H_n(y) = \frac{1}{n(n+1)(2n+1)}\{n[(n+1)^2+\bar{\eta}^2]^{1/2} H_{n+1}(y) + (n+1)(n^2+\bar{\eta}^2)^{1/2} H_{n-1}(y) - (2n+1)\bar{\eta} H_n(y)\}. \tag{3.13}$$

The radiation Green's function $\tilde{G}_n(y,y')$ of the Coulomb wave equation (3.8) satisfies the inhomogeneous equation

$$y^2 \frac{d^2}{dy^2}\tilde{G}_n(y,y') + 2y\frac{d}{dy}\tilde{G}_n(y,y') - [n(n+1) - y^2 + 2\bar{\eta}y]\tilde{G}_n(y,y') = \delta(y-y'), \tag{3.14}$$

the regularity requirement at $y = 0$, and the radiation condition

$$\lim_{y\to\infty} \tilde{G}_n(y,y') = \text{const} \times (1/y)\exp[i(y - \bar{\eta}\ln 2y)]. \tag{3.15}$$

Following the routine construction procedure for Green's function, we find

$$\tilde{G}_n(y,y') = \frac{-i}{yy'}\begin{cases} F_n(\bar{\eta},y)[F_n(\bar{\eta},y') - iG_n(\bar{\eta},y')], & y < y', \\ F_n(\bar{\eta},y')[F_n(\bar{\eta},y) - iG_n(\bar{\eta},y)], & y' < y. \end{cases} \tag{3.16}$$

Because of the long-range nature of the Coulomb force, the radial function $T_{mn}(c,\xi)$ has the asymptotic form

$$T_{mn}(c,\xi) \xrightarrow[c\xi\to\infty]{} (1/c\xi)e^{i\delta_{mn}}\cos[c\xi - \bar{\eta}\ln(2c\xi) - \tfrac{1}{2}(n+1)\pi + \delta_{mn}], \tag{3.17}$$

where δ_{mn} is the spheroidal Coulomb phase shift. The asymptotic form for function $Y_{mn}(y)$ follows directly:

$$Y_{mn}(y) \xrightarrow[y\to\infty]{} (1/y)e^{i\delta_{mn}}\cos[y - \bar{\eta}\ln(2y) - \tfrac{1}{2}(n+1)\pi + \delta_{mn}]. \tag{3.18}$$

By using the Green's function $\tilde{G}_n(y,y')$ in Eq. (3.16), the solution for Eq. (3.7), which is regular at $y = c$ and satisfies asymptotic condition (3.18), has the form

$$Y_{mn}(y) = \frac{1}{y}e^{i\sigma_n}F_n(\bar{\eta},y) + \int_c^\infty dy'\tilde{G}_n(y,y')U_{mn}(y')Y_{mn}(y'), \tag{3.19}$$

where

$$U_{mn}(y')Y_{mn}(y') = c^2\frac{d^2}{dy'^2}Y_{mn}(y') - \frac{2mc^2}{y'}\frac{d}{dy'}Y_{mn}(y') + \left(\lambda_{mn}(c) - n(n+1) + \frac{m(m+1)c^2}{y'^2}\right)Y_{mn}(y'),$$
$$\sigma_n = \arg\Gamma(n+1+i\bar{\eta}). \tag{3.20}$$

σ_n is the spherical Coulomb phase shift. The factor $e^{i\sigma_n}$ in Eq. (3.19) is used to guarantee the proper asymptotic behavior of function $Y_{mn}(y)$ in the limit $c \to 0$. By utilizing the asymptotic forms of Coulomb wavefunctions[14]

$$F_n(\bar{\eta},y) \xrightarrow[y\to\infty]{} \cos[y - \bar{\eta}\ln(2y) - \tfrac{1}{2}(n+1)\pi + \sigma_n],$$
$$G_n(\bar{\eta},y) \xrightarrow[y\to\infty]{} -\sin[y - \bar{\eta}\ln(2y) - \tfrac{1}{2}(n+1)\pi + \sigma_n], \tag{3.21}$$

the following equation can be obtained from Eq. (3.19) in the limit $y \to \infty$:

$$e^{i\delta_{mn}}\cos[y - \bar{\eta}\ln(2y) - \tfrac{1}{2}(n+1)\pi + \delta_{mn}] = e^{i\sigma_n}\cos[y - \bar{\eta}\ln(2y) - \tfrac{1}{2}(n+1)\pi + \sigma_n] - i\exp i[y - \bar{\eta}\ln(2y) - \tfrac{1}{2}(n+1)\pi + \sigma_n] \times \int_c^\infty dy'\frac{F_n(\bar{\eta},y')}{y'}U_{mn}(y')Y_{mn}(y'). \tag{3.22}$$

By equating the coefficients in front of the exponential $\exp i[y - \bar{\eta}\ln(2y) - \tfrac{1}{2}(n+1)\pi]$ at Eq. (3.22), we arrive at a simple formula

$$e^{i\delta_{mn}}\sin(\delta_{mn} - \sigma_n) = -\int_c^\infty dy\,\frac{F_n(\bar{\eta},y)}{y}U_{mn}(y)Y_{mn}(y). \tag{3.23}$$

It is easy to see from Eq. (3.23) that in the limit $c \to 0$ the spheroidal Coulomb phase shifts δ_{mn} reduces to the spherical Coulomb phase shifts σ_n.

4. BORN APPROXIMATION

The Born approximation for spheroidal Coulomb phase shifts is discussed here. In the approximation, we make the following substitution:

$$Y_{mn}(y) \to (1/y)e^{i\sigma_n}F_n(\bar{\eta},y). \tag{4.1}$$

Namely, only the zero-order approximation is made to the radial function $Y_{mn}(y)$ in the integro-differential equation (3.19). From Eq. (3.23), one arrives at an equation for the Born phase shift δ^B_{mn}:

$$e^{i(\delta^B_{mn}-\sigma_n)}\sin(\delta^B_{mn} - \sigma_n) = -\int_c^\infty dy\left(\frac{F_n(\bar{\eta},y)}{y}\right)^2 U_{mn}(y). \tag{4.2}$$

By using Eqs. (3.8), (3.12), and (3.13), one obtains the expression

$$e^{i(\delta^B_{mn}-\sigma_n)}\sin(\delta^B_{mn} - \sigma_n) = -c^2\int_c^\infty\frac{dy}{y^2}F_n(\bar{\eta},y)[I_1(y)F_{n+1}(\bar{\eta},y) + I_2(y)F_n(\bar{\eta},y)], \tag{4.3}$$

where

$$I_1(y) = \frac{2(m+1)[(n+1)^2 + \bar{\eta}^2]^{1/2}}{(n+1)y},$$

$$I_2(y) = \frac{\lambda_{mn}(c) - n(n+1)}{c^2} - 1 + \frac{2\bar{\eta}}{y}\frac{(n-m)}{(n+1)} + \frac{(n-m)(n-m-1)}{y^2}. \quad (4.4)$$

The Coulomb wavefunction $F_n(\bar{\eta}, y)$ is bounded. This boundness makes the improper integral (4.3) defined. If, instead, one uses the short-range integral equation (2.10) to evaluate spheroidal Coulomb phase shift, then, under the Born approximation of Eq. (2.12) with Coulomb potential Eq. (2.4), one would have to deal with an integral

$$\int_c^\infty d\xi\ \xi[R_{mn}^{(1)}(c,\xi)]^2. \quad (4.5)$$

The asymptotic form of the radial function[12]

$$R_{mn}^{(1)}(c,\xi) \underset{\xi\to\infty}{\longrightarrow} (1/c\xi)\sin[c\xi - \tfrac{1}{2}(n+1)\pi] \quad (4.6)$$

makes integral (4.5) logarithmic divergent. The reason for such a difference between Eqs. (2.10) and (3.23) is not difficult to understand. The divergence in Eq. (2.10), in case of the Coulomb potential, is due to an improper handling of the asymptotic form of radial function $\tilde{T}_{mn}(c,\xi)$ as expressed in Eq. (2.13). In the expression of the long-range contribution $\bar{\eta}\ln(2\xi)$ is neglected, and which is infinite at $\xi\to\infty$. In deriving Eq. (3.23), such a long-range contribution is first taken into consideration, and all short-range dependence, although involving differentiation, is treated as an effective perturbed "potential" $U_{mn}(y)$. In this way the logarithmic divergence is avoided. The latter method is quite similar to the one for treating a modified Coulomb potential in the spherical phase shift analysis.

Some estimates on the behavior of Coulomb spheroidal phase shifts δ_{mn} can be made through Eq. (4.3). It is difficult to obtain an analytic expression for the integral in Eq. (4.3), instead an approximation will be used in the discussion. The integral involved is expressed in terms of the regular Coulomb wavefunction $F_n(\bar{\eta}, y)$ and the inverse powers of the integration variable y. For large order n, in different geometric regions, $F_n(\bar{\eta}, y)$ has respective approximated forms. We shall divide such regions as $y \ll n$, $y \sim n$, and $y \gg n$. These forms in the lowest-order approximation are as follows

$$F_n(\bar{\eta}, y) \sim (e^{-\pi\bar{\eta}/2}/\sqrt{2}\,e)(ey/2n)^{n+1}, \quad \text{for } y \ll n, \quad (4.7)$$

$$F_n(\bar{\eta}, y) \sim \tfrac{1}{2}\sqrt{2\pi y}\, e^{-\pi\bar{\eta}/2} J_{n+1/2}(y), \quad \text{for } y \sim n, \quad (4.8)$$

$$F_n(\bar{\eta}, y) \sim \sin[y - \bar{\eta}\ln(2y) - n\pi/2 + \sigma_n], \quad \text{for } y \gg n. \quad (4.9)$$

The function $F_n(\bar{\eta}, y)$ in region $y \ll n$ decreases quite rapidly according to the power of n. In region $y \gg n$, the function $F_n(\bar{\eta}, y)$ oscillates according to its asymptotic form Eq. (3.21). From the behavior of the Bessel function $J_{n+1/2}(y)$, we can obtain a detailed description[14] of the function $F_n(\bar{\eta}, y)$ in the transition region $y \sim n$:

$$F_n(\bar{\eta}, y) \sim \left(\frac{\pi y}{2}\right)^{1/2} e^{-\pi\bar{\eta}/2}\left(\frac{2}{3n}\right)^{1/3}\left(\frac{3^{-1/3}}{\Gamma(\frac{2}{3})} + \frac{2^{1/3}\xi_0}{\Gamma(\frac{1}{3})}\right), \quad \text{for } y \text{ near } n, \quad (4.10)$$

$$F_n(\bar{\eta}, y) \sim \left(\frac{\xi_1}{4(1-\xi_1^2)^{1/2}}\right)^{1/2} \exp\{-\tfrac{1}{2}\pi\bar{\eta} - (n+\tfrac{1}{2})\xi_2\} \quad \text{for } y < n, \text{ or } n < y, \quad (4.11)$$

where

$$\xi_0 = (2y - 2n - 1)/2n^{1/3},$$
$$\xi_1 = 2y/(2n+1), \quad (4.12)$$
$$\xi_2 = \int_{\xi_2}^1 \frac{(1-t^2)^{1/2}}{t}\,dt.$$

For y near to n, $F_n(\bar{\eta}, y)$ is smooth and does not vary too much. For $y < n$, ξ_2 is positive real and monotonically increases with the decrease of variable y. Then $F_n(\bar{\eta}, y)$ decreases with y to the region $y \ll n$. For $y > n$, ξ_2 is purely imaginary and its magnitude increases with y. This means that $F_n(\bar{\eta}, y)$ starts to oscillate toward its asymptotic form in region $y \gg n$.

Based on these properties of the Coulomb wavefunction $F_n(\bar{\eta}, y)$, one can estimate the integration in Eq. (4.3) for the large order of n. Due to the smallness of $F_n(\bar{\eta}, y)$ in region $y \ll n$, the contribution from this region is not important to the integration. Then the contribution mainly comes from regions $y \sim n$ and $y \gg n$ for which the approximations (4.8) and (4.9) are applicable. Now the estimation for integration in Eq. (4.3) may be expressed as

$$\begin{aligned} e^{i(\delta^B_{mn} - \sigma_n)}\sin(\delta^B_{mn} - \sigma_n) \sim{}& -\frac{\pi c^2}{2} e^{-\pi\bar{\eta}} \int_c^\infty \frac{dy}{y} J_{n+1/2}(y) \\ &\times [I_1(y) J_{n+3/2}(y) + I_2(y) J_{n+1/2}(y)] \\ &+ c^2 \int_{n_u}^\infty \frac{dy}{y^2}\{I_2(y)[\sin^2(y - n\pi/2) \\ &- \sin^2(y - \bar{\eta}\ln(2y) - n\pi/2 + \sigma_n)] \\ &- \tfrac{1}{2} I_1(y)[\sin(2y - n\pi) \\ &- \sin(2y - 2\bar{\eta}\ln(2y) - n\pi + 2\sigma_n)]\} \end{aligned} \quad (4.13)$$

where n_u is the boundary between the regions $y \sim n$ and $y \gg n$. In arriving at Eq. (4.13) we have used the smallness of Bessel function in region $y \ll n$ and its asymptotic form. Estimation of the various integrations is not too difficult and yields

$$e^{i(\delta^B_{mn} - \sigma_n)}\sin(\delta^B_{mn} - \sigma_n) \sim 0(n^{-1}), \quad \text{for large } n; \quad (4.14)$$

however, the spherical Coulomb phase shift σ_n in Eq. (3.20) has the asymptotic form

$$\sigma_n \sim \bar{\eta}\ln(n+1). \quad (4.15)$$

Equations (3.14) and (3.15) reveal that the spheroidal Coulomb phase shift only deviates from the spherical one at small orders, and at large orders they are the same.

5. RADIAL FUNCTION

A further investigation of the spheroidal Coulomb phase shift δ_{mn} as expressed in Eq. (3.23) leads to a study of the radial function $Y_{mn}(y)$. A formal expression for $Y_{mn}(y)$ can be obtained through Eq. (3.19) by

a direct iteration procedure. The resulting form is often called the Born expansion

$$\tilde{Y}_{mn}(y) \equiv y e^{-i\sigma_n} Y_{mn}(y) = F_n(\bar{\eta},y) + \sum_{i=1}^{\infty} \int_c^{\infty} dy' K_{mn}^i(y,y') F_n(\bar{\eta},y'), \quad (5.1)$$

where

$$K_{mn}^1(y,y') F_n(\bar{\eta},y') = y \tilde{G}_n(y,y') U_{mn}(y') y^{-1} F_n(\bar{\eta},y'), \quad (5.2)$$

$$\begin{aligned} K_{mn}^{i+1}(y,y') F_n(\bar{\eta},y') &= \int_c^{\infty} dy'' K_{mn}^1(y,y'') K_{mn}^i(y'',y') F_n(\bar{\eta},y') \\ &= \int_c^{\infty} dy_1 \cdots \int_c^{\infty} dy_i K_{mn}^1(y,y_1) \\ &\quad \times K_{mn}^1(y_1,y_2) \cdots K_{mn}^1(y_i,y') F_n(\bar{\eta},y') \\ &= \int_c^{\infty} dy_1 \cdots \int_c^{\infty} dy_i\, y \tilde{G}_n(y,y_1) \\ &\quad \times U_{mn}(y_1) \tilde{G}_n(y_1,y_2) \cdots U_{mn}(y_i) \tilde{G}_n(y_i,y') \\ &\quad \times U_{mn}(y') y'^{-1} F_n(\bar{\eta},y'). \end{aligned} \quad (5.3)$$

By using Eqs. (3.8), (3.12), and (3.13), we can express the factors in Eqs. (5.2) and (5.3) in the following forms:

$$U_{mn}(y) y^{-1} F_n(\bar{\eta},y) = c^2 y^{-1} [I_1(y) F_{n+1}(\bar{\eta},y) + I_2(y) F_n(\bar{\eta},y)], \quad (5.4)$$

$$\begin{aligned} U_{mn}(y) \tilde{G}_n(y,y') &= c^2 y^{-2} \delta(y-y') \\ &\quad - ic^2 y^{-1} [I_1(y) F_{n+1}(\bar{\eta},y) + I_2(y) F_n(\bar{\eta},y)] \\ &\quad \times y'^{-1} F_n(\bar{\eta},y') \\ &\quad - c^2 \begin{cases} y^{-1}[I_1(y) F_{n+1}(\bar{\eta},y) + I_2(y) F_n(\bar{\eta},y)] \\ \quad \times y'^{-1} G_n(\bar{\eta},y'), \ y < y', \\ y^{-1}[I_1(y) G_{n+1}(\bar{\eta},y) + I_2(y) G_n(\bar{\eta},y)] \\ \quad \times y'^{-1} F_n(\bar{\eta},y'), \ y' < y, \end{cases} \end{aligned} \quad (5.5)$$

where $I_1(y)$ and $I_2(y)$ are the functions given in Eq. (4.4). The above procedure eliminates the differentiation which appears in Eqs. (5.2) and (5.3). In Eq. (5.5) there is a δ function, which comes from the second derivative term of the "potential" $U_{mn}(y)$ with respect to the Green's function $\tilde{G}_n(y,y')$. The presence of such a term prevents the convergence of the formal Born expansion in Eq. (5.1). The reason is simple. The δ function contributes a set of terms in Eq. (5.1) of the form

$$c^{2i-2} \int_c^{\infty} dy_1\, y \tilde{G}_n(y,y_1) y_1^{-(2i-2)} U_{mn}(y_1) y_1^{-1} F_n(\bar{\eta},y_1). \quad (5.6)$$

The series formed by such terms does not converge for any constant c. This obscures the formal Born expansion in Eq. (5.1) for the radial function $Y_{mn}(y)$.

The obstacle is easy to circumvent through a slightly different approach. Let us return to the original radial integral equation (3.19). We observe that it is the second-order derivative of the "potential" $U_{mn}(y)$ which causes the Born expansion to diverge. To avoid this, we will consider a related integral equation, with the removal of the second order derivative, which arises from Eq. (3.19) by multiplying potential $U_{mn}(y)$:

$$Y_{mn}^u(y) = F_n^u(\bar{\eta},y) + (c^2/y^2) Y_{mn}^u(y) + \int_c^{\infty} dy' G_n^u(y,y') Y_{mn}^u(y'), \quad (5.7)$$

where

$$\begin{aligned} Y_{mn}^u(y) &= U_{mn}(y) Y_{mn}(y), \\ F_n^u(\bar{\eta},y) &= U_{mn}(y) y^{-1} e^{i\sigma_n} F_n(\bar{\eta},y), \\ G_{nm}^u(y,y') &= U_{mn}(y) \tilde{G}_n(y,y') - (c^2/y^2)\delta(y-y'). \end{aligned} \quad (5.8)$$

The regularity of the radial function $Y_{mn}(y)$ at point c yields

$$F_n^u(\bar{\eta},c) + \int_c^{\infty} dy' G_{nm}^u(c,y') Y_{mn}^u(y') = 0. \quad (5.9)$$

From Eqs. (5.7) and (5.9) we obtain

$$Y_{mn}^u(y) = f_n^u(\bar{\eta},y) + \int_c^{\infty} dy' K_{mn}(y,y') Y_{mn}^u(y'), \quad (5.10)$$

where

$$\begin{aligned} f_n^u(\eta,y) &= [y^2/(y^2-c^2)][F_n^u(\bar{\eta},y) - F_n^u(\bar{\eta},c)], \\ K_{mn}(y,y') &= [y^2/(y^2-c^2)][G_{nm}^u(y,y') - G_{nm}(c,y')]. \end{aligned} \quad (5.10')$$

Equation (5.10) is a proper Fredholm's integral equation. Its kernel is regular and contains neither δ function nor derivative. By the iteration procedure a Born expansion can be obtained to yield the solution

$$\begin{aligned} Y_{mn}^u(y) = f_n^u(\bar{\eta},y) &+ \sum_{i=1}^{\infty} \int_c^{\infty} dy_1 \cdots \int_c^{\infty} dy_i \\ &\times K_{nm}(y,y_1) K_{nm}(y_1,y_2) \cdots \\ &\times K_{nm}(y_{i-1},y_i) f_n^u(\bar{\eta},y_i). \end{aligned} \quad (5.11)$$

Now let us study the validity of the above equation. The regular Coulomb wavefunction is bounded, and goes to zero as y^{n+1} as $y \to 0$. These properties have been manifested in approximated forms of Eqs. (4.7), (4.8), and (4.9). The factor $U_{mn}(y) y^{-1} F_n(\bar{\eta},y)$ goes to zero at least like y^{-1} at $y \to \infty$ and cannot be more singular than y^{-2} at $y = 0$. Then, along with Eq. (5.10), we can find that the function $f_n^u(\bar{\eta},y)$ is bounded by

$$|f_n^u(\bar{\eta},y)| \leq M/y, \quad (5.12)$$

where M is a constant. The irregular Coulomb wavefunction $G_n(\bar{\eta},y)$ is unbounded only at the origin $y = 0$, and has the singular behavior

$$G_n(\bar{\eta},y) \sim y^{-n}, \quad \text{at } y \to 0. \quad (5.13)$$

Such an unboundedness does not yield any difficulty in estimating each of the terms in the Born expansion (5.11). The function $G_n(\bar{\eta},y)$ appears in the expansion through the Green's function $\tilde{G}_n(y,y')$ and is accompanied by the function $F_n(\bar{\eta},y)$, which has $(n+1)$ zero as shown in Eq. (4.7). In the combination the argument of the function $F_n(\bar{\eta},y)$ is always smaller than the argument for the function $G_n(\bar{\eta},y)$. In other words the singular behavior of the function $G_n(\bar{\eta},y)$ is suppressed. Then the factor $U_{mn}(y) G_{mn}^u(y,y')$ has the same type of singularity and asymptotic behavior as the factor $U_{mn}(y) y^{-1} F_n(\bar{\eta},y)$. Now it is easy to show that kernel $K_{mn}(y,y')$ in Eq. (5.10) has the bound

$$|K_{mn}(y,y')| \leq c^2 N/yy', \quad (5.14)$$

where N is a constant. From Eqs. (5.12) and (5.14) one can conclude without any difficulty that the Born expansion is uniformly convergent for

$$cN < 1. \tag{5.15}$$

The convergence verifies the validity of the Born expansion in Eq. (5.11).

From Eqs. (3.19), (5.8), and (5.11) the final form for the radial function $Y_{mn}(y)$ can be expressed as

$$Y_{mn}(y) = (1/y)e^{i\sigma_n} F_n(\bar{\eta}, y)$$
$$+ \int_c^\infty dy' \tilde{G}_n(y, y')[f_n^u(\bar{\eta}, y')$$
$$+ \sum_{i=1}^{\infty} \int_c^\infty dy_1 \cdots \int_c^\infty dy_i K_{nm}(y', y_1)$$
$$\times K_{nm}(y_1, y_2) \cdots K_{nm}(y_{i-1}, y_i) f_n^u(\bar{\eta}, y_i)]. \tag{5.16}$$

Since the series is uniformly convergent, Eq. (5.16) is a valid expression. As $c \to 0$, the integral part reduces to zero and the function $y^{-1}e^{i\sigma_n}F_n(\bar{\eta}, y)$ is the zeroth-order approximation for the radial function $Y_{mn}(y)$.

* Work supported in part by the National Science Foundation.

1 M. Kotani, Proc. Phys.-Math. Soc. Japan **15**, 30 (1933); C. J. Bouwkamp, J. Math. Phys. **26**, 79 (1947).
2 J. Meixner and W. Andrejewski, Ann. Phys. (N.Y.) **7**, 157 (1950); C. Flammer, J. Appl. Phys. **24**, 1218, 1224 (1953).
3 J. Meixner, Ann. Phys. (N.Y.) **12**, 227 (1953); C. Flammer, "Radiation from Electric and Magnetic Dipoles in the Presence of a Conducting Circular Disk," Technical Report No. 49 (Stanford Research Institute, 1955).
4 F. V. Schultz, "Scattering by a Prolate Spheroid," Report UMM-42 (University of Michigan, 1950); K. M. Siegel, B. H. Gere, I. Marx, and F. B. Sleator, "The Numerical Determination of the Radar Cross Section of a Prolate Spheroid," Report UMM-126 (University of Michigan, 1953).
5 O. Burrau, Kgl. Danske Videnskab. Selskab, Mat. Fys. Medd. **7**, 14 (1927).
6 H. Stier, Z. Physik **76**, 439 (1932); J. B. Fisk, Phys. Rev. **49**, 167 (1936).
7 S. Nagahara, J. Phys. Soc. Japan **8**, 165 (1953); **9**, 52 (1954); K. Takayanagi Progr. Theoret. Phys. (Kyoto) Suppl. **40**, 216 (1967) S. Hara, J. Phys. Soc. Japan **27**, 1009.
8 M. Gell-Mann, Phys. Letters **8**, 214 (1964).
9 R. P. Feynman, *High Energy Collisions*, edited by C. N. Yang *et al.* (Gordon and Breach, New York, 1969).
10 N. Byers and C. N. Yang, Phys. Rev. **142**, B 976 (1966); M. C. Li, *ibid.* D **1**, 326 (1970); S. P. Almeida and M. C. Li, Particles Nuclei **1**, 24 (1970).
11 M. C. Li, J. Math. Phys. **12**, 936 (1971).
12 C. Flammer, *Spheroidal Wave Functions* (Stanford U. P., Palo Alto, Calif., 1957); J. A. Stratton, P. M. Morse, L. J. Chu, J. D. C. Little, and F. J. Corbato, *Spheroidal Wave function* (Wiley, New York, 1956); P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).
13 L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968).
14 M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Natl. Bur. Stds., Washington, D.C., 1964); L. L. Yost, J. A. Wheeler, and G. Breit, Phys. Rev. **49**, 174 (1936); I. A. Stegun and M. Abramowitz *ibid.* **98**, 1851 (1955).

On the Decomposition of Direct Products of Irreducible Representations

Daniel B. Litvin
Department of Physics, University of British Columbia, Vancouver, Canada
(Received 27 March 1972)

A lemma concerning irreducible representations contained in the decomposition of a direct product of irreducible representations of simply reducible groups is generalized to arbitrary decomposable unitary and nonunitary groups.

I. INTRODUCTION

In the application of group theory in physics the problem very often arises of decomposing a direct product of two irreducible representation into a sum of irreducible parts. In the theory of solid state physics such a decomposition is required in defining selection rules in scattering processes in magnetic and nonmagnetic crystals.[1,2] A classical example of this is the addition of angular momentum in quantum mechanics. Wigner,[3] using a classification of irreducible representations given by Frobenius and Schur, proved a lemma concerning irreducible representations contained in the decomposition of a direct product of irreducible representations of simply reducible groups. The three-dimensional rotation group is a simply reducible group, and, for example, the fact that the addition of integer angular momenta does not contain half-integer momenta can be deduced directly from Wigner's lemma.

The purpose of this work is to generalize Wigner's lemma. We first review the Frobenius and Schur classification of irreducible representations and Wigner's lemma for simply reducible groups. This lemma is then generalized to arbitrary decomposable unitary and nonunitary groups.

II. SIMPLY REDUCIBLE GROUPS

Let Δ^k denote the kth irreducible representation, and u the elements of a unitary group G. Frobenius and Schur have shown that the irreducible representations of the group G can be classified into three cases[4]:

Case A: $\Delta^k(u)$ is equivalent to $\Delta^k(u)^*$ and potentially real, i.e., can be brought into real form.

Case B: $\Delta^k(u)$ is equivalent to $\Delta^k(u)^*$ and pseudoreal, i.e., can not be brought into real form.

Case C: $\Delta^k(u)$ is not equivalent to $\Delta^k(u)^*$.

For Cases A and B, $\Delta^k(u)$ is equivalent to $\Delta^k(u)^*$:

$$\Delta^k(u)^* = \beta_k^{-1}(u)\beta_k$$

and

$$\beta_k\beta_k^* = C_k E,$$

where $C_k = +1$ or -1 for Cases A and B, respectively.

A group is called simply reducible if[3]:

(1) Every element is equivalent to its reciprocal.

where N is a constant. From Eqs. (5.12) and (5.14) one can conclude without any difficulty that the Born expansion is uniformly convergent for

$$cN < 1. \tag{5.15}$$

The convergence verifies the validity of the Born expansion in Eq. (5.11).

From Eqs. (3.19), (5.8), and (5.11) the final form for the radial function $Y_{mn}(y)$ can be expressed as

$$Y_{mn}(y) = (1/y)e^{i\sigma_n} F_n(\bar{\eta}, y)$$
$$+ \int_c^\infty dy' \tilde{G}_n(y, y')[f_n^u(\bar{\eta}, y')$$
$$+ \sum_{i=1}^{\infty} \int_c^\infty dy_1 \cdots \int_c^\infty dy_i K_{nm}(y', y_1)$$
$$\times K_{nm}(y_1, y_2) \cdots K_{nm}(y_{i-1}, y_i) f_n^u(\bar{\eta}, y_i)]. \tag{5.16}$$

Since the series is uniformly convergent, Eq. (5.16) is a valid expression. As $c \to 0$, the integral part reduces to zero and the function $y^{-1}e^{i\sigma_n}F_n(\bar{\eta}, y)$ is the zeroth-order approximation for the radial function $Y_{mn}(y)$.

* Work supported in part by the National Science Foundation.

1 M. Kotani, Proc. Phys.-Math. Soc. Japan **15**, 30 (1933); C. J. Bouwkamp, J. Math. Phys. **26**, 79 (1947).
2 J. Meixner and W. Andrejewski, Ann. Phys. (N.Y.) **7**, 157 (1950); C. Flammer, J. Appl. Phys. **24**, 1218, 1224 (1953).
3 J. Meixner, Ann. Phys. (N.Y.) **12**, 227 (1953); C. Flammer, "Radiation from Electric and Magnetic Dipoles in the Presence of a Conducting Circular Disk," Technical Report No. 49 (Stanford Research Institute, 1955).
4 F. V. Schultz, "Scattering by a Prolate Spheroid," Report UMM-42 (University of Michigan, 1950); K. M. Siegel, B. H. Gere, I. Marx, and F. B. Sleator, "The Numerical Determination of the Radar Cross Section of a Prolate Spheroid," Report UMM-126 (University of Michigan, 1953).
5 O. Burrau, Kgl. Danske Videnskab. Selskab, Mat. Fys. Medd. **7**, 14 (1927).
6 H. Stier, Z. Physik **76**, 439 (1932); J. B. Fisk, Phys. Rev. **49**, 167 (1936).
7 S. Nagahara, J. Phys. Soc. Japan **8**, 165 (1953); **9**, 52 (1954); K. Takayanagi Progr. Theoret. Phys. (Kyoto) Suppl. **40**, 216 (1967) S. Hara, J. Phys. Soc. Japan **27**, 1009.
8 M. Gell-Mann, Phys. Letters **8**, 214 (1964).
9 R. P. Feynman, *High Energy Collisions*, edited by C. N. Yang *et al.* (Gordon and Breach, New York, 1969).
10 N. Byers and C. N. Yang, Phys. Rev. **142**, B 976 (1966); M. C. Li, *ibid.* D **1**, 326 (1970); S. P. Almeida and M. C. Li, Particles Nuclei **1**, 24 (1970).
11 M. C. Li, J. Math. Phys. **12**, 936 (1971).
12 C. Flammer, *Spheroidal Wave Functions* (Stanford U. P., Palo Alto, Calif., 1957); J. A. Stratton, P. M. Morse, L. J. Chu, J. D. C. Little, and F. J. Corbato, *Spheroidal Wave function* (Wiley, New York, 1956); P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).
13 L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968).
14 M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Natl. Bur. Stds., Washington, D.C., 1964); L. L. Yost, J. A. Wheeler, and G. Breit, Phys. Rev. **49**, 174 (1936); I. A. Stegun and M. Abramowitz *ibid.* **98**, 1851 (1955).

On the Decomposition of Direct Products of Irreducible Representations

Daniel B. Litvin
Department of Physics, University of British Columbia, Vancouver, Canada
(Received 27 March 1972)

A lemma concerning irreducible representations contained in the decomposition of a direct product of irreducible representations of simply reducible groups is generalized to arbitrary decomposable unitary and nonunitary groups.

I. INTRODUCTION

In the application of group theory in physics the problem very often arises of decomposing a direct product of two irreducible representation into a sum of irreducible parts. In the theory of solid state physics such a decomposition is required in defining selection rules in scattering processes in magnetic and nonmagnetic crystals.[1,2] A classical example of this is the addition of angular momentum in quantum mechanics. Wigner,[3] using a classification of irreducible representations given by Frobenius and Schur, proved a lemma concerning irreducible representations contained in the decomposition of a direct product of irreducible representations of simply reducible groups. The three-dimensional rotation group is a simply reducible group, and, for example, the fact that the addition of integer angular momenta does not contain half-integer momenta can be deduced directly from Wigner's lemma.

The purpose of this work is to generalize Wigner's lemma. We first review the Frobenius and Schur classification of irreducible representations and Wigner's lemma for simply reducible groups. This lemma is then generalized to arbitrary decomposable unitary and nonunitary groups.

II. SIMPLY REDUCIBLE GROUPS

Let Δ^k denote the kth irreducible representation, and u the elements of a unitary group G. Frobenius and Schur have shown that the irreducible representations of the group G can be classified into three cases[4]:

Case A: $\Delta^k(u)$ is equivalent to $\Delta^k(u)^*$ and potentially real, i.e., can be brought into real form.

Case B: $\Delta^k(u)$ is equivalent to $\Delta^k(u)^*$ and pseudoreal, i.e., can not be brought into real form.

Case C: $\Delta^k(u)$ is not equivalent to $\Delta^k(u)^*$.

For Cases A and B, $\Delta^k(u)$ is equivalent to $\Delta^k(u)^*$:

$$\Delta^k(u)^* = \beta_k^{-1}(u)\beta_k$$

and

$$\beta_k\beta_k^* = C_k E,$$

where $C_k = +1$ or -1 for Cases A and B, respectively.

A group is called simply reducible if[3]:

(1) Every element is equivalent to its reciprocal.

(2) The direct product of any two irreducible representations contains no representation more than once.

The first condition means that all irreducible representations of a simply reducible group are either Case A or Case B.

The following lemma for simply reducible groups has been proven by Wigner[3]:

Lemma 1: The direct product of two Case A or two Case B irreducible representations of a simply reducible group contains only Case A irreducible representations; the direct product of a Case A and Case B irreducible representation contains only Case B irreducible representations.

III. UNITARY GROUPS

For an arbitrary decomposable unitary group G we prove the following lemma:

Lemma 2: The direct product of two Case A or two Case B irreducible representations of an arbitrary decomposable unitary group G does not contain Case B irreducible representations; the direct product of a Case A and a Case B irreducible representation does not contain Case A irreducible representations.

The direct product, for example, of two Case A irreducible representations contains only Case A or Case C irreducible representations, each representation possibly more than once. For simply reducible groups Lemma 2 is identical to Lemma 1.

Proof of Lemma 2: We take the direct product $\Delta(u) = \Delta^i(u) \times \Delta^j(u)$, where Δ^i and Δ^j are either Case A or Case B irreducible representations, that is,

$$\Delta^i(u)^* = \beta_i^{-1}\Delta^i(u)\beta_i, \qquad \beta_i\beta_i^* = C_iE,$$
$$\Delta^j(u)^* = \beta_j^{-1}\Delta^j(u)\beta_j, \qquad \beta_j\beta_j^* = C_jE. \tag{1}$$

We show that if the decomposition of the direct product contains the irreducible representation Δ^k equivalent to Δ^{k*},

$$\Delta^k(u)^* = \beta_k^{-1}\Delta^k(u)\beta_k, \qquad \beta_k\beta_k^* = C_kE, \tag{2}$$

then $C_k = C_iC_j$.

The direct product $\Delta(u)$ is decomposed via a similarity transformation with a unitary matrix U:

$$\Delta_\gamma(u) = U^{-1}\Delta(u)U.$$

We assume that Δ_γ is in the following form

$$\Delta_\gamma = \begin{pmatrix} \Delta^k & & & & & \\ & \Delta^k & & & & \\ & & \ddots & & & \\ & & & \Delta^k & & \\ & & & & \Delta^p & \\ & & & & & \ddots \end{pmatrix},$$

where Δ^k appears n times and is assumed to be equivalent to Δ^{k*}, i.e., is either a Case A or Case B irreducible representation.

Using (1), we have

$$\Delta_\gamma(u)^* = U^{-1}\Delta(u)^*U^* = [U^{-1}(\beta_i \times \beta_j)U^*]^{-1}\Delta_\gamma(U)[U^{-1}(\beta_i \times \beta_j)U^*].$$

Denoting $U^{-1}(\beta_i \times \beta_j)U^*$ by β, we write the preceeding relation as

$$\Delta_\gamma(u)^* = \beta^{-1}\Delta_\gamma(u)\beta, \tag{3}$$

where $\beta\beta^* = C_iC_jE$.

β is subdivided into blocks corresponding in dimension to the irreducible representations appearing in Δ_γ:

$$\beta = \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots \\ \beta_{21} & \beta_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix}.$$

From (3) we have for $i,j = 1, 2, \ldots, n$

$$\Delta^k(u)^* = \beta_{ij}^{-1}\Delta^k(u)\beta_{ij}. \tag{4}$$

The β_{ij} for $i \leq n$ and $j > n$, and $j \leq n$ and $i > n$, are zero for they connect nonequivalent irreducible representations. β therefore is of the form

$$\beta = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1n} & & \\ \vdots & & \vdots & \bigcirc & \\ \beta_{n1} & \cdots & \beta_{nn} & & \\ & \bigcirc & & \beta_{n+1,n+1}\cdots & \\ & & & \vdots & \end{pmatrix}.$$

We consider now only the submatrix of β containing the matrices β_{ij}, $i,j = 1, 2, \ldots, n$, and denote this by $\bar{\beta}$. From the properties of β we have

$$\begin{pmatrix} \Delta^k(u)^* & & \\ & \ddots & \\ & & \Delta^k(u)^* \end{pmatrix} = \bar{\beta}^{-1} \begin{pmatrix} \Delta^k(u) & & \\ & \ddots & \\ & & \Delta^k(u) \end{pmatrix} \bar{\beta} \tag{5}$$

and

$$\bar{\beta}\bar{\beta}^* = C_iC_jE. \tag{6}$$

We will show that $\bar{\beta}$ can be transformed into the quasidiagonal form:

$$\begin{pmatrix} \alpha & & & \\ & \alpha & & \\ & & \ddots & \\ & & & \alpha \end{pmatrix}.$$

From (2) and (5) we then have $\Delta^{k*} = \alpha^{-1}\Delta^k\alpha$ and $\alpha\alpha^* = C_kE$, and from (6) that $\alpha\alpha^* = C_iC_jE$, thus giving $C_k = C_iC_j$ proving Lemma 2.

The matrix $\bar{\beta}$ of relation (5) is not unique. (5) will remain unchanged under any similarity transformation with a unitary matrix of the form $A \times E$, where E is of the same dimension as the irreducible representation Δ^k and A is an arbitrary unitary matrix of dimension n, the number of times Δ^k appears in (5).[5]

The matrix $\bar{\beta}$ can be replaced by

$$(A \times E)\bar{\beta} \tag{7}$$

without changing the form of relation (5).

We seek a matrix A that will put (7) in the required quasidiagonal form. To do this, we look at the structure of the matrix $\bar{\beta}$. From (4) for $i = j = 1$ and for a general i and j

$$\Delta^k(u)^* = \beta_{11}^{-1}\Delta^k(u)\beta_{11},$$
$$\Delta^k(u)^* = \beta_{ij}^{-1}\Delta^k(u)\beta_{ij}$$

from which we have

$$\beta_{ij}\beta_{11}^{-1}\Delta^k(u) = \Delta^k(u)\beta_{ij}\beta_{11}^{-1}$$

giving, by Schur's lemma,[6] $\beta_{ij} = \lambda_{ij}\beta_{11}$, where λ_{ij} is a constant. β can be written now as

$$\bar{\beta} = \begin{pmatrix} \lambda_{11}\beta_{11} & & \lambda_{1n}\beta_{11} \\ \vdots & & \vdots \\ \lambda_{n1}\beta_{11} & \cdots & \lambda_{nn}\beta_{11} \end{pmatrix} = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{n1} & & \lambda_{nn} \end{pmatrix} \times \beta_{11} \equiv \lambda \times \beta_{11}.$$

Since both $\bar{\beta}$ and β_{11} are unitary matrices, λ is also unitary. Finally, by choosing $A = \lambda^{-1}$, (7) takes on the required quasidiagonal form and the proof of Lemma 2 is complete.

IV. NONUNITARY GROUPS

A nonunitary group M contains elements half of which are unitary and half antiunitary. The unitary elements form an invariant subgroup G of index two, and we can write M as

$$M = G + Ga_0,$$

where a_0 is a fixed antiunitary element.

Corepresentations D^k of a nonunitary group M are constructed in one of three ways depending on the following classification of the irreducible representations Δ^k of the unitary subgroup G[7]:

Type I: $\Delta^k(u)$ is equivalent to $\Delta^k(a_0^{-1}ua_0)^*$, $\Delta^k(a_0^{-1}ua_0)^* = \beta_k^{-1}\Delta^k(u)\beta_k$ and $\beta_k\beta_k^* = \Delta^k(a_0^2)$.

Type II: $\Delta^k(u)$ is equivalent to $\Delta^k(a_0^{-1}ua_0)^*$, $\Delta^k(a_0^{-1}ua_0)^* = \beta_k^{-1}\Delta^k(u)\beta_k$ but $\beta_k\beta_k^* = -\Delta^k(a_0^2)$.

Type III: $\Delta^k(u)$ is not equivalent to $\Delta^k(a_0^{-1}ua_0)^*$.

The three types of corepresentations corresponding to the above classification of the irreducible representation of the unitary subgroup G are[7]

Type I: $D^k(u) = \Delta^k(u)$, $D^k(ua_0) = \Delta^k(u)\beta_k$.

Type II:

$$D^k(u) = \begin{pmatrix} \Delta^k(u) & \\ & \Delta^k(u) \end{pmatrix}, \quad D^k(ua_0) = \begin{pmatrix} & \Delta^k(u)\beta_k \\ -\Delta^k(u)\beta_k & \end{pmatrix}. \tag{8}$$

Type III:

$$D^k(u) = \begin{pmatrix} \Delta^k(u) & \\ & \Delta^k(a_0^{-1}ua_0)^* \end{pmatrix}, \quad D^k(ua_0) = \begin{pmatrix} & \Delta^k(ua_0^2) \\ \Delta^k(a_0^{-1}ua_0)^* & \end{pmatrix}.$$

The decomposition of direct products of two corepresentations of a nonunitary group M can be analyzed in terms of the decomposition of direct products of irreducible representations of the unitary subgroup G.

Let C_{ij}^k be the number of times the corepresentation D^k is contained in the direct product $D^i \times D^j$. C_{ij}^k is calculated from[8]

$$C_{ij}^k = \frac{\sum_u \chi(D^i(u))\,\chi(D^j(u))\,\chi(D^k(u))^*}{\sum_u \chi(D^k(u))\,\chi(D^k(u))^*}, \tag{9}$$

where $\chi(D^i(u))$ is the trace of $D^i(u)$. The number of times an irreducible representation Δ^k of the subgroup G of M is contained in the direct product $\Delta^i \times \Delta^j$ is denoted by d_{ij}^k and calculated from

$$d_{ij}^k = (l_k/n)\sum_u \chi(\Delta^i(u))\,\chi(\Delta^j(u))\,\chi(\Delta^k(u))^*, \tag{10}$$

where l_k is the dimension of Δ^k and n the order of the group G.

By using the explicit form of the corepresentations (8), the C_{ij}^k defined by (9) can be written in terms of the d_{ij}^k defined by (10). The explicit form of the relation depends on the type of the corepresentation D^i, D^j, and D^k. The relations between the C_{ij}^k and the d_{ij}^k, taken from Ref. 9, are listed in Table I.

TABLE I: The number of times the corepresentation D^k is contained in the direct product $D^i \times D^j$, denoted by C_{ij}^k, is given in terms of the d_{ij}^k, the number of times the irreducible representation Δ^k is contained in the direct product $\Delta^i \times \Delta^j$. Primed suffices, as in $d_{i'j}^k$, denote that the irreducible representation $\Delta^i(a_0^{-1}ua_0)^*$ replaces $\Delta^i(u)$ in the direct product.

D^i	D^j	D^k	C_{ij}^k
I	I	I	d_{ij}^k
I	I	II	$\frac{1}{2}d_{ij}^k$
I	I	III	d_{ij}^k
I	II	I	$2d_{ij}^k$
I	II	II	d_{ij}^k
I	II	III	$2d_{ij}^k$
I	III	I	$d_{ij}^k + d_{ij'}^k$
I	III	II	$\frac{1}{2}d_{ij}^k + \frac{1}{2}d_{ij'}^k$
I	III	III	$d_{ij}^k + d_{ij'}^k$
II	II	I	$4d_{ij}^k$
II	II	II	$2d_{ij}^k$
II	II	III	$4d_{ij}^k$
II	III	I	$2d_{ij}^k + 2d_{ij'}^k$
II	III	II	$d_{ij}^k + d_{ij'}^k$
II	III	III	$2d_{ij}^k + 2d_{ij'}^k$
III	III	I	$d_{ij}^k + d_{ij'}^k + d_{i'j}^k + d_{i'j'}^k$
III	III	II	$\frac{1}{2}d_{ij}^k + \frac{1}{2}d_{ij'}^k + \frac{1}{2}d_{i'j}^k + \frac{1}{2}d_{i'j'}^k$
III	III	III	$d_{ij}^k + d_{ij'}^k + d_{i'j}^k + d_{i'j'}^k$

We prove the following lemmas:

Lemma 3: The direct product of two Type I or two Type II irreducible representations of the sub-

group G of an arbitrary decomposable nonunitary group $M = G + Ga_0$ does not contain Type II irreducible representations; the direct product of a Type I and a Type II irreducible representation does not contain Type I irreducible representations.

Lemma 4: The direct product of two Type I or two Type II corepresentations of an arbitrary decomposable nonunitary group M does not contain Type II corepresentations; the direct product of a Type I and a Type II corepresentation does not contain Type I corepresentations.

Proof of Lemma 3: We take the direct product $\Delta = \Delta^i \times \Delta^j$, where Δ^i and Δ^j are each either Type I or Type II irreducible representations of the subgroup G of a nonunitary group M:

$$\Delta^i(a_0^{-1}ua_0)^* = \beta_i^{-1}\Delta^i(u)\beta_i, \quad \beta_i\beta_i^* = C_i\Delta^i(a_0^2),$$
$$\Delta^j(a_0^{-1}ua_0)^* = \beta_j^{-1}\Delta^j(u)\beta_j, \quad \beta_j\beta_j^* = C_j\Delta^j(a_0^2).$$

If the decomposition of the direct product contains the irreducible representation $\Delta^k(u)$ equivalent to $\Delta^k(a_0^{-1}ua_0)^*$, possibly more than once,

$$\Delta^k(a_0^{-1}ua_0)^* = \beta_k^{-1}\Delta^k(u)\beta_k, \quad \beta_k\beta_k^* = C_k\Delta^k(a_0^2),$$

then $C_k = C_iC_j$. The remainder of this proof is parallel to the proof of Lemma 2.

The proof of Lemma 4 follows immediately from Lemma 3 and Table I.

ACKNOWLEDGMENT

The author would like to thank Professor J. Zak for helpful discussions.

1 R. J. Elliott and R. Loudon, J. Phys. Chem. Solids **15**, 146 (1960).
2 A. P. Cracknell, Prog. Theoret. Phys. **38**, 1252 (1967).
3 E. P. Wigner, Amer. J. Math. **63**, 57 (1941).
4 M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1962), p. 139.
5 G. F. Koster, Phys. Rev. **109**, 227 (1958).
6 Reference 4, p. 98.
7 E. P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* (Academic , New York, 1959), Chap. 28.
8 G. F. Karavaev, Fiz. Tverd. Tela **6**, 3676 (1964) [Sov. Phys. Solid State **6**, 2943 (1965)].
9 C. J. Bradley and B. L. Davis, Rev. Mod. Phys. **40**, 359 (1968).

The Symmetric Group and the Gel'fand Basis of $U(3)$. Generalizations of the Dirac Identity

K. J. Lezuo
*Institut für Kernphysik, Universität Mainz, Mainz, Germany**
(Received 23 March 1972)

It is shown that the symmetrization of N particle states by means of the orthogonal units of the algebra of the symmetric group S_N yields the Gel'fand basis states of the irreducible representations of $U(3)$. The existence of generalizations of the Dirac identity is demonstrated, and a connection between the symmetrized two- and three-body exchange operators and the invariants of $U(3)$ is established.

INTRODUCTION

The study of the unitary group $U(3)$ and, more generally, of $U(n)$ is of great interest to present day physics. Most well known is the successful classification of the elementary particles according to the octet model as proposed by Gell-Mann and Ne'eman[1] in 1962. A physically different application of the theory of the unitary groups has been to the many particle system. In fact, a great deal of the development of the theory—associated with the names of Racah and Wigner[2]—has been done towards the goal of classifying the electronic states in the atom. More recently, the theory of the unitary groups has been used to obtain approximate solutions of the nuclear many-body problem.[3]

The study of the many-body system leads, in a rather natural manner, to consideration of the operations which permute the particles and, thus, to the introduction of the symmetric group S_N. The connection between the two groups $U(n)$ and S_N has been known since the work of Young and Frobenius around 1900. Later, recognizing the importance of the concepts for quantum mechanics, Weyl[4] continued research along these lines and laid the foundation for our present understanding of the subject. He formulated the concept of duality and gave it an expression in a number of theorems. These early investigations have been concerned with the irreducible representations and have used the characters as tools. It was only within the past decade that a systematic investigation of the basis states has been taken up, pursued mainly by Biedenharn[5,6] and also by Moshinsky[7,8] and their collaborators. Yet, the relevance of the symmetric group for the Gel'fand[9] states has been considered to a limited extent only. Moshinsky[10] showed that a certain class of Gel'fand states had a definite permutational symmetry, and Ciftan and Biedenharn[11] and Ciftan[12] used the concept of "hooks" (which originally has its proper meaning in the symmetric group) to construct the Gel'fand states of $U(4)$.

In the present paper we show that the duality between $U(n)$ and S_N can be extended to the individual basis states defined by the subgroup decomposition[6] $U(n) \supset U(n-1) \supset \cdots \supset U(1)$ on the one side and by an analogous chain on the other side. It will be shown that the Gel'fand states can be obtained by use of operations of S_N only, thus supplying a link to the understanding of the hook structure concept for the unitary groups. In addition to their transformation properties under the unitary groups, the Gel'fand states will be seen to transform like the basis states of the irreducible representations of S_N. The situation will be pictured by introducing a "combined Young–Weyl tableau."

As a first step we shall demonstrate the existence of generalizations of the Dirac identity[13] which emerge naturally by considering the operations of both groups, $U(n)$ and S_N, in the same space. In this way we are led to explicit expressions for the fully symmetrized Majorana operator and the analogous three-body exchange operator in terms of the invariants of $U(3)$.

group G of an arbitrary decomposable nonunitary group $M = G + Ga_0$ does not contain Type II irreducible representations; the direct product of a Type I and a Type II irreducible representation does not contain Type I irreducible representations.

Lemma 4: The direct product of two Type I or two Type II corepresentations of an arbitrary decomposable nonunitary group M does not contain Type II corepresentations; the direct product of a Type I and a Type II corepresentation does not contain Type I corepresentations.

Proof of Lemma 3: We take the direct product $\Delta = \Delta^i \times \Delta^j$, where Δ^i and Δ^j are each either Type I or Type II irreducible representations of the subgroup G of a nonunitary group M:

$$\Delta^i(a_0^{-1}ua_0)^* = \beta_i^{-1}\Delta^i(u)\beta_i, \quad \beta_i\beta_i^* = C_i\Delta^i(a_0^2),$$
$$\Delta^j(a_0^{-1}ua_0)^* = \beta_j^{-1}\Delta^j(u)\beta_j, \quad \beta_j\beta_j^* = C_j\Delta^j(a_0^2).$$

If the decomposition of the direct product contains the irreducible representation $\Delta^k(u)$ equivalent to $\Delta^k(a_0^{-1}ua_0)^*$, possibly more than once,

$$\Delta^k(a_0^{-1}ua_0)^* = \beta_k^{-1}\Delta^k(u)\beta_k, \quad \beta_k\beta_k^* = C_k\Delta^k(a_0^2),$$

then $C_k = C_iC_j$. The remainder of this proof is parallel to the proof of Lemma 2.

The proof of Lemma 4 follows immediately from Lemma 3 and Table I.

ACKNOWLEDGMENT

The author would like to thank Professor J. Zak for helpful discussions.

1 R. J. Elliott and R. Loudon, J. Phys. Chem. Solids **15**, 146 (1960).
2 A. P. Cracknell, Prog. Theoret. Phys. **38**, 1252 (1967).
3 E. P. Wigner, Amer. J. Math. **63**, 57 (1941).
4 M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1962), p. 139.
5 G. F. Koster, Phys. Rev. **109**, 227 (1958).
6 Reference 4, p. 98.
7 E. P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* (Academic , New York, 1959), Chap. 28.
8 G. F. Karavaev, Fiz. Tverd. Tela **6**, 3676 (1964) [Sov. Phys. Solid State **6**, 2943 (1965)].
9 C. J. Bradley and B. L. Davis, Rev. Mod. Phys. **40**, 359 (1968).

The Symmetric Group and the Gel'fand Basis of $U(3)$. Generalizations of the Dirac Identity

K. J. Lezuo
*Institut für Kernphysik, Universität Mainz, Mainz, Germany**
(Received 23 March 1972)

It is shown that the symmetrization of N particle states by means of the orthogonal units of the algebra of the symmetric group S_N yields the Gel'fand basis states of the irreducible representations of $U(3)$. The existence of generalizations of the Dirac identity is demonstrated, and a connection between the symmetrized two- and three-body exchange operators and the invariants of $U(3)$ is established.

INTRODUCTION

The study of the unitary group $U(3)$ and, more generally, of $U(n)$ is of great interest to present day physics. Most well known is the successful classification of the elementary particles according to the octet model as proposed by Gell-Mann and Ne'eman[1] in 1962. A physically different application of the theory of the unitary groups has been to the many particle system. In fact, a great deal of the development of the theory—associated with the names of Racah and Wigner[2]—has been done towards the goal of classifying the electronic states in the atom. More recently, the theory of the unitary groups has been used to obtain approximate solutions of the nuclear many-body problem.[3]

The study of the many-body system leads, in a rather natural manner, to consideration of the operations which permute the particles and, thus, to the introduction of the symmetric group S_N. The connection between the two groups $U(n)$ and S_N has been known since the work of Young and Frobenius around 1900. Later, recognizing the importance of the concepts for quantum mechanics, Weyl[4] continued research along these lines and laid the foundation for our present understanding of the subject. He formulated the concept of duality and gave it an expression in a number of theorems. These early investigations have been concerned with the irreducible representations and have used the characters as tools. It was only within the past decade that a systematic investigation of the basis states has been taken up, pursued mainly by Biedenharn[5,6] and also by Moshinsky[7,8] and their collaborators. Yet, the relevance of the symmetric group for the Gel'fand[9] states has been considered to a limited extent only. Moshinsky[10] showed that a certain class of Gel'fand states had a definite permutational symmetry, and Ciftan and Biedenharn[11] and Ciftan[12] used the concept of "hooks" (which originally has its proper meaning in the symmetric group) to construct the Gel'fand states of $U(4)$.

In the present paper we show that the duality between $U(n)$ and S_N can be extended to the individual basis states defined by the subgroup decomposition[6] $U(n) \supset U(n-1) \supset \cdots \supset U(1)$ on the one side and by an analogous chain on the other side. It will be shown that the Gel'fand states can be obtained by use of operations of S_N only, thus supplying a link to the understanding of the hook structure concept for the unitary groups. In addition to their transformation properties under the unitary groups, the Gel'fand states will be seen to transform like the basis states of the irreducible representations of S_N. The situation will be pictured by introducing a "combined Young–Weyl tableau."

As a first step we shall demonstrate the existence of generalizations of the Dirac identity[13] which emerge naturally by considering the operations of both groups, $U(n)$ and S_N, in the same space. In this way we are led to explicit expressions for the fully symmetrized Majorana operator and the analogous three-body exchange operator in terms of the invariants of $U(3)$.

The methods used draw from expositions of the properties of the algebra of S_N by Löwdin,[14] and some applications to branching diagram functions by Goddard.[15]

GENERALIZED DIRAC IDENTITIES FOR $U(3)$

The Dirac identity[13] constitutes a relation between operations of S_N and of $SU(2)$. It reads

$$P_{12} = \tfrac{1}{2}[1 + \sigma(1)\cdot\sigma(2)], \tag{1}$$

and expresses the permutation operator P_{12} in terms of the generators of $SU(2)$. A well-known application is to the theory of ferromagnetism where the exchange energy is represented by the scalar product of spin vectors leading to the Heisenberg Hamiltonian. The Dirac identity has been discussed in various papers, and we may, e.g., refer to Löwdin,[14] and to Biedenharn[16] for elementary discussion of the subject. Biedenharn, in fact, also indicates the possibility of generalization to three- and many-particle exchange operators.

A different interpretation can be given to the Dirac identity by considering the operator for the total spin, which for an N-particle system is

$$S^2 = \tfrac{3}{4}N + 2\sum_{i,j}(-\tfrac{1}{4} + \tfrac{1}{2}P_{ij}). \tag{2}$$

Since $\sum P_{ij}$ is the sum over all transpositions of S_N, it is useful to introduce the class operator $C^{[2]}$ as the "arithmetic mean" of all elements in the class, i.e.,

$$C^{[2]} = [2/N(N-1)]\sum_{i<j} P_{ij}, \tag{3}$$

and rewrite Eq. (2) as

$$S^2 = -\tfrac{1}{4}N(N-4) + \tfrac{1}{2}N(N-1)C^{[2]}. \tag{4}$$

Equation (4) constitutes an identity between the *fully symmetrized* exchange operator and the *invariant* of $SU(2)$. In the following we want to show that similar identities relate the class operators $C^{[2]}$ and $C^{[3]}$ to the invariants of $U(3)$.

Before we discuss the subject we define an explicit realization of the basis functions in terms of a Boson or a Fermion calculus. The possibility of this realization is well known from the work of Jordan[17] and has been used and further developed by Schwinger,[18] by Baird and Biedenharn[6] and by Moshinsky[7] and others. The generators of $U(3)$ can be expressed in terms of Boson or Fermion operators,

$$E_{ij} = \sum_{k=1}^{N} a_i^+(k)\,a_j(k), \qquad i,j = 1,2,3, \tag{5}$$

and, by use of the commutation relations

$$[a_i(k), a_j^+(l)]_\pm = \delta_{ij}\delta_{kl}, \tag{6}$$

can be shown to fulfill the defining relations for the Lie algebra

$$[E_{ij}, E_{kl}]_- = \delta_{jk}E_{il} - \delta_{il}E_{kj}. \tag{7}$$

An appropriate basis on which the generator as defined by Eq. (5) act is given by

$$|i_1 i_2 \cdots i_N\rangle = a_{i_1}^+(1)\,a_{i_2}^+(2)\cdots a_{i_N}^+(N)|0\rangle, \tag{8}$$

where $i_1, i_2, \ldots, i_N = 1, 2, 3$. In addition let us introduce the operations $U(\tau), \tau \in S_N$, which permute the particles. Explicitly[19]

$$U(\tau)|i_1 i_2 \cdots i_N\rangle = |i_{\tau(1)}, i_{\tau(2)}, \ldots, i_{\tau(N)}\rangle, \tag{9}$$

for all $\tau \in S_N$.

Returning now to the invariants of $U(3)$ we recall that there are two independent invariants, the second-order Casimir invariant

$$I_2^{(3)} = 2[H_1(H_1+1) + 3H_2(H_2+1) + E_{21}E_{12} + E_{31}E_{13} + E_{32}E_{23}], \tag{10}$$

where

$$H_1 = \tfrac{1}{2}(E_{11} - E_{22}), \qquad H_2 = \tfrac{1}{6}(E_{11} + E_{22} - 2E_{33}), \tag{11}$$

and the third-order Biedenharn[5] invariant

$$I_3^{(3)} = (2H_2+1)(H_1 - H_2)(H_1 + H_2 + 1) + (2H_1+1)E_{21}E_{12} + (H_1 - H_2)E_{31}E_{13} - (H_1 + H_2)E_{32}E_{23} + E_{12}E_{31}E_{23} + E_{21}E_{32}E_{13}, \tag{12}$$

defined by means of the symmetric coupling coefficients. By introducing the class operators

$$C^{[2]} = [2/N(N-1)]\sum_{i<j} U(ij), \tag{13}$$

$$C^{[3]} = [1/2\binom{N}{3}]\sum_{i<j<k}[U(ijk) + U(jik)], \tag{14}$$

it can be shown that the following operator identities hold:

$$I_2^{(3)} = -\tfrac{1}{3}N(N-9) + N(N-1)C^{[2]}, \tag{15}$$

$$I_3^{(3)} = \tfrac{1}{54}N(4N^2 - 27N + 63) - \tfrac{1}{6}N(N-1)(2N-9)C^{[2]} + \tfrac{1}{3}N(N-1)(N-2)C^{[3]}. \tag{16}$$

The proof of these relations can be given by explicit application to all basis states. Here, we only remark that it is sufficient to check the result for the states $|i_1 \leq i_2 \leq \cdots \leq i_N\rangle$. The rest of the basis can be obtained by the operations $U(\tau), \tau \in S_N$ which commute with the invariants on account of their being symmetric functions.

The relations (15) and (16) constitute generalizations of the Dirac identity equation (4). In physical application the emphasis may be put on reading the equations from right to left. Solving for $C^{[2]}$ and $C^{[3]}$ we obtain expressions for the symmetrized Majorana operator and for the three-body exchange operator in terms of the invariants of $U(3)$.[20]

THE GEL'FAND BASIS

The task of the representation theory is to reduce the N-particle space as defined by the basis states equation (8) into its irreducible components. It is well known that the reduction of the tensor space can be achieved by the operations of the symmetric group. Graphically this is expressed by using the Young frame to characterize the irreducible representations of $U(3)$. The basis states within the various irreducible representations are known as the Gel'fand

states and are defined by means of the subgroup decomposition[5] $U(3) \supset U(2) \supset U(1)$. The appropriate quantum numbers are defined by the weights and by the eigenvalues of the invariants of the group and its subgroups and may be pictured by means of the Gel'fand pattern

$$\begin{pmatrix} m_{13} & & m_{23} & & m_{33} \\ & m_{12} & & m_{22} & \\ & & m_{11} & & \end{pmatrix}, \tag{17}$$

where the m_{ij} are positive integers limited to the range $m_{i+1,j+1} \leq m_{ij} \leq m_{i,j+1}$. Alternatively, the m_{ij} may be visualized in the Weyl tableau, which is a Young frame lexically filled with the integers 1, 2, 3. For further development, it is important to note that the Weyl tableau as defined is a mnemonic device to label the states in a manner equivalent to the Gel'fand pattern and does not give a prescription for constructing the states explicitly as a boson or fermion operator polynomial. The constructive aspect, however, is an important one. For $U(3)$, Baird and Biedenharn[6] solved the problem by showing that the "semimaximal" state ($m_{11} = m_{12}$) is a single monomial in the boson creation operators and obtained the general state by use of the lowering operator E_{21}. More general methods have been devised by Moshinsky and Nagel,[8] and also by Ciftan and Biedenharn.[11]

In the present paper we want to show that the Gel'fand states can be obtained by appropriate symmetrization, thereby extending the concept of duality to the individual states of the basis. We make use of a set of idempotents which are known as the orthogonal units of the algebra of S_N. Discussion are found in many books[21,22] on the algebraic treatment of S_N. We briefly recall some of their properties. The orthogonal units are linear combinations of permutation operators, in which the matrix elements of the irreducible representations of S_N figure as coefficients, i.e.,

$$O^{\mu}_{rs} = f^{\mu} \sum_{\tau \in S_N} D^{\mu}_{rs}(\tau) U(\tau). \tag{18}$$

The notation, which is essentially that of Rutherford,[21] is as follows: μ is a partition of N and characterizes the irreducible representation of dimension f^{μ}. The individual matrix elements $D^{\mu}_{rs}(\tau)$ are labeled by means of lexical Young tableaux r and s. A basic property of these representation matrices (and hence of the orthogonal units) is that they decompose into *irreducible* submatrices upon restricting S_N to S_{N-1}. Numerical values for the matrices of the generating elements $(k-1, k)$ are given in the textbooks. The fundamental multiplication rule for the orthogonal units is the following:

$$O^{\mu}_{rs} O^{\nu}_{tv} = \delta^{\mu\nu} \delta_{st} O^{\mu}_{rv}. \tag{19}$$

We are now prepared to show that the states $O^{\mu}_{rs} | i_1 \cdots i_N \rangle$ are the Gel'fand basis states $|(m)\rangle$ apart from a normalization factor, i.e.,

$$|(m)\rangle = O^{\mu}_{rs} | i_1 i_2 \cdots i_N \rangle, \tag{20}$$

where (m) is a short-hand notation for the Gel'fand pattern (17). The proof of this assertion is straightforward and makes use of the operator identities (15) and (16) and the property of the orthogonal units to be eigenoperators of the class operators,[14] i.e.,

$$C^{[\lambda]} O^{\mu}_{rs} = (1/f^{\mu}) \chi^{\mu}_{[\lambda]} O^{\mu}_{rs}, \tag{21}$$

where $\chi^{\mu}_{[\lambda]}$ is the character of the class $[\lambda]$ in the irreducible representation μ of S_N. With the forms (15) and (16) of the $U(3)$ invariants, it is immediately clear that the states are eigenstates as should be. In order to identify the states as the correct Gel'fand states, the crucial test is to show that they are also eigenstates of the $U(2)$ invariant $I_2^{(2)}$. Since this does not seem to be as obvious we sketch the proof.

Proof: Let the state $O^{\mu}_{rs} | i_1 \cdots i_N \rangle$ be such that N_1 indices have the numerical value 1, N_2 the value 2, and N_3 of them the value 3. As shown in the Appendix we may choose $i_1 \leq i_2 \leq \cdots \leq i_N$ without loss of generality. Thus, the first set of $N' = N_1 + N_2$ indices has the numerical values 1 and 2. We should remark now that the class operator $C^{[2]}$ which occurs in the Dirac identity Eq. (4) is defined over the subgroup $S_{N'}$ of S_N, and not over S_N. We may write

$$C^{[2]} = (1/N'!) \sum_{\tau' \in S_{N'}} U(\tau') U(12) U(\tau'^{-1}), \tag{22}$$

where (12) stands typically for a transposition. With this form for $C^{[2]}$ and the expansion (18) for O^{μ}_{rs}, we can investigate the product $C^{[2]} O^{\mu}_{rs}$. Making use of the group property of $U(\tau)$, $U(\tau') U(12) U(\tau'^{-1}) U(\tau) = U[\tau'(12)\tau'^{-1}\tau]$, and changing variables $\sigma = \tau'(12)\tau'^{-1}\tau$, we are led to

$$C^{[2]} O^{\mu}_{rs} = (f^{\mu}/N'!) \sum_{\sigma, \tau'} \sum_{u} D^{\mu}_{ru}[\tau'(12)\tau'^{-1}] D^{\mu}_{us}(\sigma) U(\sigma). \tag{23}$$

For further evaluation we write this expression as

$$\sum_{\tau'} D^{\mu}_{ru}[\tau'(12)\tau'^{-1}] = \sum_{\tau'} \sum_{t,v} D^{\mu}_{rt}(\tau') D^{\mu}_{tv}(12) D^{\mu}_{vu}(\tau'^{-1}). \tag{24}$$

At this point it is essential to note that the representation matrices D^{μ} are so constructed that they decompose into irreducible submatrices—i.e., blocks on the diagonal, zeros elsewhere—upon restricting S_N to any of its subgroups in the chain $S_N \supset S_{N-1} \supset \cdots \supset S_1$. This property holds in particular for all $\tau' \in S_{N'}$. Thus, in the expression above the sums over t and v are effective only over the states of the irreducible representations of the subgroup $S_{N'}$ and can be worked out, using

$$\sum_{\tau'} D^{\mu'}_{rt'}(\tau') D^{\mu'}_{v'u}(\tau'^{-1}) = (N'!/f^{\mu'}) \delta_{t'v'} \delta_{ru}.$$

Here, μ' is the irreducible representation of $S_{N'}$ which occurs in the sum $\mu = \sum \mu'$ and has r among its basis vectors. The primes have been added to t and v to indicate their restricted ranges. Inserting the result into Eq. (23) we find

$$C^{[2]} O^{\mu}_{rs} = (1/f^{\mu'}) \chi^{\mu'}_{[2]} O^{\mu}_{rs}, \tag{25}$$

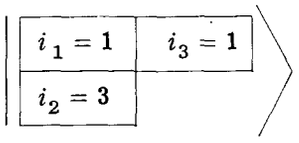
which proves O^{μ}_{rs} to be an eigenoperator of $C^{[2]}$ and, hence, the states $O^{\mu}_{rs} | i_1 \cdots i_N \rangle$ to be eigenstates of $I_2^{(2)}$.

Equation (20) states that the Gel'fand states can be obtained by means of the orthogonal units and, in fact, possess a further property: In addition to the m_{ij}

which identify the states within the unitary group, the Young tableau r defines their permutational symmetry. The situation may be pictured in a tableau which indicates the *symmetry properties* by the *name of the indices* ("Young tableau") and the *unitary properties* by their *numerical values* ("Weyl tableau"). For clarity we give an example

Young tableau Weyl tableau

Combined



We call the tableau defined above a *combined Young–Weyl tableau*. It defines the properties of the state under transformations of S_N as well as under transformations of $U(n)$ and stresses the concept of duality at the level of states.

We can now apply the operator identities (15) and (16) to the properly symmetrized states and obtain by use of Eq. (21) a relation between the eigenvalues of the invariants of $U(3)$ and the characters of S_N.

We conclude by remarking that we expect the results to generalize to the case of $U(n)$ also. For $U(4)$, in particular, the Gel'fand states have already been proven to have a definite permutational symmetry, since it is clear that the states $O^{\mu}_{rs}|i_1 \cdots i_N\rangle$, $i_1, \ldots, i_N = 1, 2, 3, 4$ belong to a $U(4)$ multiplet, the rest being deducible from the generalized Dirac identities of $U(3)$ and $U(2)$.

ACKNOWLEDGMENT

The author would like to thank Professor L. C. Biedenharn for critically reading the manuscript.

APPENDIX

For practical purposes it is useful to adopt several conventions. Given a certain state $|i_1 i_2 \cdots i_N\rangle$ the fully reduced basis is obtained by application of all possible O^{μ}_{rs}. An entirely equivalent reduction results, if instead we apply O^{μ}_{rs} to $U(\tau)|i_1 \cdots i_N\rangle = |i_{\tau(1)}, \ldots, i_{\tau(N)}\rangle$. In other words, we see that the reduction is unique only up to an equivalence. In order to make a specific choice we use

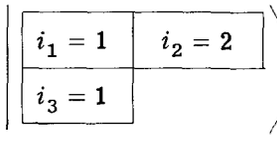
$$i_1 \le i_2 \le \cdots \le i_N. \tag{A1}$$

A further convention seems useful. First we note that the states $O^{\mu}_{rs}|i_1 \cdots i_N\rangle$ for fixed s transform according to irreducible representation of S_N, i.e.,

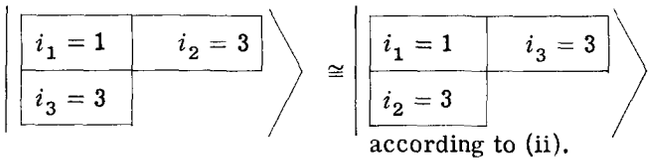
$$U(\sigma)O^{\mu}_{rs}|i_1 \cdots i_N\rangle = \sum_u D^{\mu}_{ur}(\sigma)O^{\mu}_{us}|i_1 \cdots i_N\rangle. \tag{A2}$$

Let the indices $i_1, i_2, \ldots, i_N$ take on all the values $1, 2, \ldots, N$. Then, it is clear that $O^{\mu}_{rs}|i_1, \ldots, i_N\rangle$ constitutes the reduction of the regular representation of S_N as s is varied over all possible standard Young tableaux (r of course also). In general, however, if the numerical values of $i_1, \ldots, i_N$ are restricted as in the $U(3)$ case, we cannot expect to find all possible symmetry classes present. More explicitly, it will be found that certain $O^{\mu}_{rs}|i_1 \cdots i_N\rangle$ may be zero; others may be proportional to one another. The rules are as follows and can be proven straightforwardly.[14] Let s stand for a standard Young tableau consisting of N boxes with the numbers $1, 2, \ldots, N$ inserted lexically. Now, if we insert the *numerical values* of $i_1, i_2, \ldots, i_N$ in place of $1, 2, \ldots, N$ a Weyl tableau is obtained. Moreover, (i) $O^{\mu}_{rs}|i_1 \cdots i_N\rangle = 0$ if the Weyl tableau has two identical integers in the same column, and (ii) $O^{\mu}_{rs}|i_1 \cdots i_N\rangle = \text{const}\, O^{\mu}_{rs}|i_1 \cdots i_N\rangle$ if the two Weyl tableaux resulting from s and s' are identical.

The situation may be visualized by using a tableau similar to the combined Young–Weyl tableau which shows both the name of the index and its numerical value in each box. We give two examples to demonstrate the rules above:

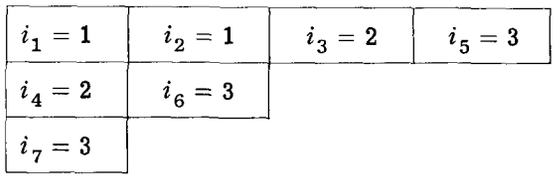


$= 0$ according to (i).



according to (ii).

In order to obtain a unique set of nonzero states, we adopt the following convention consistent with (A1): We draw a Weyl tableau according to the Gel'fand pattern (m) and insert the indices $i_1, i_2, \ldots, i_N$ to yield a Young–Weyl tableau. The first few indices will be associated with the 1's in the first row, the next set with the 1's in the second row, etc., until all the 1's are saturated. Then we start with the 2's in the first row and continue as before, and finally, we proceed with the 3's in the same manner. The following example should make this clear:



From the tableau so obtained, we deduce s and (m) such that

$$|(m)\rangle = O^{\mu}_{rs}|i_1 \cdots i_N\rangle \quad \text{for all } r.$$

---

* This work was carried out in part during the author's stay at the Sektion Physik, Universität München, Munich, Germany.

1 See M. Gell-Mann and Y. Ne'eman, *The Eightfold Way* (Benjamin, New York, 1965).

2 G. Racah, Phys. Rev. **76**, 1352 (1949); G. Racah, *Group Theory and Spectroscopy*, Ergebnisse der exakten Naturwissenschaften, Vol. 37 (Springer-Verlag, Berlin, 1965); E. P. Wigner, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren* (Academic, New York, 1959).

3 See M. Moshinsky, *Group Theory and the Many-Body Problem*, Documents on Modern Physics (Gordon and Breach, New York, 1968); P. Kramer, M. Moshinksy, in *Group Theory and Its Appli-*

cations, edited by E. M. Loebl (Academic, New York, 1968).
4 H. Weyl, *The Classical Groups* (Princeton U.P., Princeton, N.J., 1946); H. Weyl, *Gruppentheorie und Quantenmechanik* (Hirzel-Verlag, Leibzig, 1931).
5 L. C. Biedenharn, J. Math. Phys. **4**, 436 (1963); G. E. Baird and L. C. Biedenharn, J. Math. Phys. **5**, 1723, 1730 (1964); **6**,1847 (1965).
6 G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).
7 V. Bargmann and M. Moshinsky, Nucl. Phys. **23**, 177 (1961); M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).
8 J. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965).
9 I. M. Gel'fand and M. L. Zetlin, Dokl. Akad. Nauk SSSR **71**, 825 (1950).
10 M. Moshinsky, J. Math. Phys. **7**, 691 (1966).
11 M. Ciftan and L. C. Biedenharn, J. Math. Phys. **10**, 221 (1969).
12 M. Ciftan, J. Math. Phys. **10**, 1635 (1969).
13 P. A. M. Dirac, *Quantum Mechanics* (Oxford U.P., London, 1935).
14 P. O. Löwdin, Rev. Mod. Phys. **39**, 254 (1967); P. O. Löwdin and O. Goscinsky, Research Report No. 243, Quantum Chemistry Group (Uppsala, Sweden).
15 W. A. Goddard III, Phys. Rev. **157**, 81 (1967); **169**, 120 (1968); J. Chem. Phys. **48**, 1008, 5337 (1968).
16 L. C. Biedenharn, *Boulder Lectures in Theoretical Physics* Vol. V, edited by W. E. Brittin, B. W. Downs, and J. Downs (Interscience, New York, 1963).
17 P. Jordan, Z. Physik **94** (1925).
18 J. Schwinger (unpublished 1952), reprinted in *Selected Papers in Quantum Theory of Angular Momentum*, edited by L. C. Biedenharn and H. van Dam (Academic, New York, 1965).
19 S. S. Schweber, *Relativistic Quantum Field Theory* (Harper and Row, New York, 1962).
20 In a different context H. Han and L. C. Biedenharn [Phys. Rev. Letters **24**, 118 (1970)] give a formula for the diagonal matrix elements of the operators $\sum_{i<j} U(ij)$ and $\sum_{i<j<k} U(ijk)$ in $U(3)$ multiplets.
21 D. E. Rutherford, *Substitutional Analysis* (Edinburgh U.P., Edinburgh, 1948); see also Ref. 13.
22 G. de B. Robinson, *Representation Theory of the Symmetric Group* (U. of Toronto Press, Toronto, 1961); H. Boerner, *Darstellungen von Gruppen* (Springer-Verlag, Berlin, 1967); B. G. Wybourne, *Symmetry Principles and Atomic Spectroscopy* (Interscience, New York, 1970); M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1964).

Lie Theory and the Lauricella Functions F_D

Willard Miller, Jr.
School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455
(Received 2 March 1972)

It is shown that the Lauricella functions F_D in n variables transform as basis vectors corresponding to irreducible representations of the Lie algebra $sl(n+3,\mathbb{C})$. Group representation theory can then be applied to derive addition theorems, transformation formulas, and generating functions for the F_D. It is clear from this analysis that the use of $SL(m,\mathbb{C})$ symmetry in atomic and elementary particle physics will lead inevitably to the remarkable functions F_D.

INTRODUCTION

In a recent paper,[1] Ciftan has shown that the Appell function F_1 arises naturally from a study of the representation theory of the special linear groups. The author proved in Ref. 2 that this was due to the fact that $SL(5,\mathbb{C})$ was the dynamical symmetry group of F_1. Here we generalize this observation by demonstrating that $SL(n+3,\mathbb{C})$ is the dynamical symmetry group of the Lauricella functions F_D in n variables (Recall that F_1 is an F_D with $n=2$). We further show that exploitation of the $SL(n+3,\mathbb{C})$ symmetry yields elegant and simple derivations of addition theorems, transformation formulas, and generating functions for the F_D. It follows from this analysis that the implementation of $SL(m,\mathbb{C})$ symmetry in atomic and particle physics will necessarily lead to the functions F_D.

The methods employed in this paper are rather straightforward generalizations of those employed in Refs. 2 and 3.

1. THE DYNAMICAL SYMMETRY GROUP

The Lauricella function F_D is defined by the series

$$F_D(a;b_1,\ldots,b_n;c;x_1,\ldots,x_n) = \sum_{m_1\cdots m_j=0}^{\infty} \frac{(a,m_1+\cdots+m_n)(b_1,m_1)\cdots(b_n,m_n)}{(c,m_1+\cdots+m_n)m_1!\cdots m_n!} \times x_1^{m_1}\cdots x_n^{m_n}, \tag{1.1}$$

convergent for $|x_1|<1,\ldots,|x_n|<1$.[4,5]

Here,

$$(a,n) = a(a+1)\cdots(a+n-1) = (a)_n, \tag{1.2}$$

and it is assumed that $c \neq 0,-1,-2,\cdots$. We define the following partial differential operators acting on a space of functions of $2n+2$ complex variables, $s,u_1,\ldots,u_n,t,x_1,\ldots,x_n$:

$$E_\alpha = s\Big(\sum_{j=1}^n x_j\partial_{x_j} + s\partial_s\Big), \quad E_{\alpha\beta_k\gamma} = su_kt\partial_{x_k},$$

$$E_{\beta_k} = u_k(x_k\partial_{x_k} + u_k\partial_{u_k}), \quad E_{-\gamma} = t^{-1}\Big(\sum_{j=1}^n x_j\partial_{x_j} + t\partial_t - 1\Big),$$

$$E_{\alpha\gamma} = st\Big(\sum_{j=1}^n (1-x_j)\partial_{x_j} - s\partial_s\Big),$$

$$E_\gamma = t\Big(\sum_{j=1}^n (1-x_j)\partial_{x_j} + t\partial_t - s\partial_s - \sum_{j=1}^n u_j\partial_{u_j}\Big),$$

$$E_{-\alpha} = s^{-1}\Big(\sum_{j=1}^n x_j(1-x_j)\partial_{x_j} + t\partial_t - s\partial_s - \sum_{j=1}^n x_ju_j\partial_{u_j}\Big),$$

$$E_{-\beta_k} = u_k^{-1}\Big(x_k(1-x_k)\partial_{x_k} + x_k\sum_{j\neq k}(1-x_j)\partial_{x_j} + t\partial_t - x_ks\partial_s - \sum_{h=1}^n u_h\partial_{u_h}\Big),$$

$$E_{\beta_k\gamma} = u_kt\,((x_k-1)\partial_{x_k} + u_k\partial_{u_k}),$$

$$E_{-\alpha,-\gamma} = s^{-1}t^{-1}\Big(\sum_{j=1}^n x_j(1-x_j)\partial_{x_j} - \sum_{j=1}^n x_ju_j\partial_{u_j} + t\partial_t - 1\Big),$$

$$E_{-\alpha,-\beta_k,-\gamma} = s^{-1}u_k^{-1}t^{-1}\Big(\sum_{j=1}^n x_j(x_j-1)\partial_{x_j} - t\partial_t + x_ks\partial_s + \sum_{l\neq k} x_lu_l\partial_{u_l} - x_k + 1\Big),$$

$$E_{-\beta_k,-\gamma} = u_k^{-1}t^{-1}\Big(x_k(x_k-1)\partial_{x_k} + \sum_{j\neq k}(x_k-1)x_j\partial_{x_j} + x_ks\partial_s - t\partial_t + 1\Big),$$

$$E_{\beta_k,-\beta_p} = u_ku_p^{-1}((x_k-x_p)\partial_{x_k} + u_k\partial_{u_k}),$$

$$J_\alpha = s\partial_s - \tfrac{1}{2}t\partial_t, \quad J_{\beta_k} = u_k\partial_{u_k} - \tfrac{1}{2}t\partial_t + \tfrac{1}{2}\sum_{j\neq k} u_j\partial_{u_j},$$

cations, edited by E. M. Loebl (Academic, New York, 1968).
4 H. Weyl, *The Classical Groups* (Princeton U.P., Princeton, N.J., 1946); H. Weyl, *Gruppentheorie und Quantenmechanik* (Hirzel-Verlag, Leibzig, 1931).
5 L. C. Biedenharn, J. Math. Phys. **4**, 436 (1963); G. E. Baird and L. C. Biedenharn, J. Math. Phys. **5**, 1723, 1730 (1964); **6**,1847 (1965).
6 G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).
7 V. Bargmann and M. Moshinsky, Nucl. Phys. **23**, 177 (1961); M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).
8 J. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965).
9 I. M. Gel'fand and M. L. Zetlin, Dokl. Akad. Nauk SSSR **71**, 825 (1950).
10 M. Moshinsky, J. Math. Phys. **7**, 691 (1966).
11 M. Ciftan and L. C. Biedenharn, J. Math. Phys. **10**, 221 (1969).
12 M. Ciftan, J. Math. Phys. **10**, 1635 (1969).
13 P. A. M. Dirac, *Quantum Mechanics* (Oxford U.P., London, 1935).
14 P. O. Löwdin, Rev. Mod. Phys. **39**, 254 (1967); P. O. Löwdin and O. Goscinsky, Research Report No. 243, Quantum Chemistry Group (Uppsala, Sweden).
15 W. A. Goddard III, Phys. Rev. **157**, 81 (1967); **169**, 120 (1968); J. Chem. Phys. **48**, 1008, 5337 (1968).
16 L. C. Biedenharn, *Boulder Lectures in Theoretical Physics* Vol. V, edited by W. E. Brittin, B. W. Downs, and J. Downs (Interscience, New York, 1963).
17 P. Jordan, Z. Physik **94** (1925).
18 J. Schwinger (unpublished 1952), reprinted in *Selected Papers in Quantum Theory of Angular Momentum*, edited by L. C. Biedenharn and H. van Dam (Academic, New York, 1965).
19 S. S. Schweber, *Relativistic Quantum Field Theory* (Harper and Row, New York, 1962).
20 In a different context H. Han and L. C. Biedenharn [Phys. Rev. Letters **24**, 118 (1970)] give a formula for the diagonal matrix elements of the operators $\sum_{i<j} U(ij)$ and $\sum_{i<j<k} U(ijk)$ in $U(3)$ multiplets.
21 D. E. Rutherford, *Substitutional Analysis* (Edinburgh U.P., Edinburgh, 1948); see also Ref. 13.
22 G. de B. Robinson, *Representation Theory of the Symmetric Group* (U. of Toronto Press, Toronto, 1961); H. Boerner, *Darstellungen von Gruppen* (Springer-Verlag, Berlin, 1967); B. G. Wybourne, *Symmetry Principles and Atomic Spectroscopy* (Interscience, New York, 1970); M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1964).

Lie Theory and the Lauricella Functions F_D

Willard Miller, Jr.
School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455
(Received 2 March 1972)

It is shown that the Lauricella functions F_D in n variables transform as basis vectors corresponding to irreducible representations of the Lie algebra $sl(n+3,\mathbb{C})$. Group representation theory can then be applied to derive addition theorems, transformation formulas, and generating functions for the F_D. It is clear from this analysis that the use of $SL(m,\mathbb{C})$ symmetry in atomic and elementary particle physics will lead inevitably to the remarkable functions F_D.

INTRODUCTION

In a recent paper,[1] Ciftan has shown that the Appell function F_1 arises naturally from a study of the representation theory of the special linear groups. The author proved in Ref. 2 that this was due to the fact that $SL(5,\mathbb{C})$ was the dynamical symmetry group of F_1. Here we generalize this observation by demonstrating that $SL(n+3,\mathbb{C})$ is the dynamical symmetry group of the Lauricella functions F_D in n variables (Recall that F_1 is an F_D with $n=2$). We further show that exploitation of the $SL(n+3,\mathbb{C})$ symmetry yields elegant and simple derivations of addition theorems, transformation formulas, and generating functions for the F_D. It follows from this analysis that the implementation of $SL(m,\mathbb{C})$ symmetry in atomic and particle physics will necessarily lead to the functions F_D.

The methods employed in this paper are rather straightforward generalizations of those employed in Refs. 2 and 3.

1. THE DYNAMICAL SYMMETRY GROUP

The Lauricella function F_D is defined by the series

$$F_D(a;b_1,\ldots,b_n;c;x_1,\ldots,x_n) = \sum_{m_1\cdots m_j=0}^{\infty} \frac{(a,m_1+\cdots+m_n)(b_1,m_1)\cdots(b_n,m_n)}{(c,m_1+\cdots+m_n)m_1!\cdots m_n!} \times x_1^{m_1}\cdots x_n^{m_n}, \tag{1.1}$$

convergent for $|x_1|<1,\ldots,|x_n|<1$.[4,5]

Here,

$$(a,n) = a(a+1)\cdots(a+n-1) = (a)_n, \tag{1.2}$$

and it is assumed that $c \neq 0,-1,-2,\cdots$. We define the following partial differential operators acting on a space of functions of $2n+2$ complex variables, $s,u_1,\ldots,u_n,t,x_1,\ldots,x_n$:

$$E_\alpha = s\Big(\sum_{j=1}^n x_j\partial_{x_j} + s\partial_s\Big), \quad E_{\alpha\beta_k\gamma} = su_kt\partial_{x_k},$$

$$E_{\beta_k} = u_k(x_k\partial_{x_k} + u_k\partial_{u_k}), \quad E_{-\gamma} = t^{-1}\Big(\sum_{j=1}^n x_j\partial_{x_j} + t\partial_t - 1\Big),$$

$$E_{\alpha\gamma} = st\Big(\sum_{j=1}^n (1-x_j)\partial_{x_j} - s\partial_s\Big),$$

$$E_\gamma = t\Big(\sum_{j=1}^n (1-x_j)\partial_{x_j} + t\partial_t - s\partial_s - \sum_{j=1}^n u_j\partial_{u_j}\Big),$$

$$E_{-\alpha} = s^{-1}\Big(\sum_{j=1}^n x_j(1-x_j)\partial_{x_j} + t\partial_t - s\partial_s - \sum_{j=1}^n x_ju_j\partial_{u_j}\Big),$$

$$E_{-\beta_k} = u_k^{-1}\Big(x_k(1-x_k)\partial_{x_k} + x_k\sum_{j\neq k}(1-x_j)\partial_{x_j} + t\partial_t - x_ks\partial_s - \sum_{h=1}^n u_h\partial_{u_h}\Big),$$

$$E_{\beta_k\gamma} = u_kt\,((x_k-1)\partial_{x_k} + u_k\partial_{u_k}),$$

$$E_{-\alpha,-\gamma} = s^{-1}t^{-1}\Big(\sum_{j=1}^n x_j(1-x_j)\partial_{x_j} - \sum_{j=1}^n x_ju_j\partial_{u_j} + t\partial_t - 1\Big),$$

$$E_{-\alpha,-\beta_k,-\gamma} = s^{-1}u_k^{-1}t^{-1}\Big(\sum_{j=1}^n x_j(x_j-1)\partial_{x_j} - t\partial_t + x_ks\partial_s + \sum_{l\neq k} x_lu_l\partial_{u_l} - x_k + 1\Big),$$

$$E_{-\beta_k,-\gamma} = u_k^{-1}t^{-1}\Big(x_k(x_k-1)\partial_{x_k} + \sum_{j\neq k}(x_k-1)x_j\partial_{x_j} + x_ks\partial_s - t\partial_t + 1\Big),$$

$$E_{\beta_k,-\beta_p} = u_ku_p^{-1}((x_k-x_p)\partial_{x_k} + u_k\partial_{u_k}),$$

$$J_\alpha = s\partial_s - \tfrac{1}{2}t\partial_t, \quad J_{\beta_k} = u_k\partial_{u_k} - \tfrac{1}{2}t\partial_t + \tfrac{1}{2}\sum_{j\neq k} u_j\partial_{u_j},$$

$$J_\gamma = t\partial_t - \tfrac{1}{2}\Big(s\partial_s + \sum_{j=1}^{n} u_j\partial_{u_j} + 1\Big), \quad k,p = 1,2,\ldots,n. \tag{1.3}$$

We define basis functions $f_{\alpha,\beta_1,\ldots,\beta_n,\gamma}(s,u_1,\ldots,u_n, t,x_1,\ldots,x_n)$ in this space by

$$f_{\alpha,\beta_1,\ldots,\beta_n,\gamma}(s,u_1,\ldots,u_n,t,x_1,\ldots x_n) = f_{\alpha\beta_j\gamma}(s,u_j,t,x_j)$$
$$= [\Gamma(\gamma-\alpha)\Gamma(\alpha)/\Gamma(\gamma)]\, F_D(\alpha;\beta_1,\ldots,\beta_n;\gamma;x_1,\ldots,x_n) \times s^\alpha u_1^{\beta_1}\cdots u_n^{\beta_n} t^\gamma, \tag{1.4}$$

where $\gamma \neq 0,-1,-2,\cdots$ and $\Gamma(z)$ is the gamma function.[6] The action of the above operators on the basis functions is

$$E_\alpha f_{\alpha,\beta_j,\gamma} = (\gamma-\alpha-1) f_{\alpha+1,\beta_j,\gamma},$$
$$E_{\alpha\beta_k\gamma} f_{\alpha,\beta_j,\gamma} = \beta_k f_{\alpha+1,\hat\beta_k,\gamma+1},$$
$$E_{\beta_k} f_{\alpha,\beta_j,\gamma} = \beta_k f_{\alpha,\hat\beta_k,\gamma},$$
$$E_{-\gamma} f_{\alpha,\beta_j,\gamma} = (\gamma-\alpha-1) f_{\alpha,\beta_j,\gamma-1},$$
$$E_{\alpha\gamma} f_{\alpha,\beta_j,\gamma} = \Big(\sum_{j=1}^{n}\beta_j - \gamma\Big) f_{\alpha+1,\beta_j,\gamma+1},$$
$$E_\gamma f_{\alpha,\beta_j,\gamma} = \Big(\gamma - \sum_{j=1}^{n}\beta_j\Big) f_{\alpha,\beta_j,\gamma+1},$$
$$E_{-\alpha} f_{\alpha,\beta_j,\gamma} = (\alpha-1) f_{\alpha-1,\beta_j,\gamma},$$
$$E_{-\beta_k} f_{\alpha,\beta_j,\gamma} = \Big(\gamma - \sum_{j=1}^{n}\beta_j\Big) f_{\alpha,\tilde\beta_k,\gamma},$$
$$E_{\beta_k\gamma} f_{\alpha,\beta_j,\gamma} = \beta_k f_{\alpha,\hat\beta_k,\gamma+1},$$
$$E_{-\alpha,-\gamma} f_{\alpha,\beta_j,\gamma} = (\alpha-1) f_{\alpha-1,\beta_j,\gamma-1},$$
$$E_{-\alpha,-\beta_k,-\gamma} f_{\alpha,\beta_j,\gamma} = (1-\alpha) f_{\alpha-1,\tilde\beta_k,\gamma-1},$$
$$E_{-\beta_k,-\gamma} f_{\alpha,\beta_j,\gamma} = (\alpha-\gamma+1) f_{\alpha,\beta_k,\gamma-1},$$
$$E_{\beta_k,-\beta_p} f_{\alpha,\beta_j,\gamma} = \beta_k f_{\alpha,\beta_1,\ldots,\beta_k+1,\ldots,\beta_p-1,\ldots,\beta_n,\gamma},$$
$$J_\alpha f_{\alpha,\beta_j,\gamma} = (\alpha - \tfrac{1}{2}\gamma) f_{\alpha,\beta_j,\gamma},$$
$$J_{\beta_k} f_{\alpha,\beta_j,\gamma} = \Big(\beta_k - \tfrac{1}{2}\gamma + \tfrac{1}{2}\sum_{l\neq k}\beta_l\Big) f_{\alpha,\beta_j,\gamma},$$
$$J_\alpha f_{\alpha,\beta_j,\gamma} = \Big[\gamma - \tfrac{1}{2}\Big(\alpha + \sum_{l=1}^{n}\beta_l + 1\Big)\Big] f_{\alpha,\beta_j,\gamma}, \quad k,p = 1,2,\ldots,n. \tag{1.5}$$

(Here the E operators and the J operators are independent of the parameters α,β_j,γ. The subscripts merely indicate the action of these operators.) The symbols $\hat\beta_k$ and $\tilde\beta_k$ are defined by

$$\hat\beta_k = \beta_1,\ldots,\beta_{k-1},\beta_k+1,\beta_{k+1},\ldots,\beta_n,$$
$$\tilde\beta_k = \beta_1,\ldots,\beta_{k-1},\beta_k-1,\beta_{k+1},\ldots,\beta_n. \tag{1.6}$$

Relations (1.5) can be verified by routine computation. Furthermore, it is straightforward to show that the operators (1.3) form a basis for a simple Lie algebra of dimension $(n+3)^2-1$, i.e., a basis for $sl(n+3,\mathbb{C})$.

To determine the group action of $SL(n+3,\mathbb{C})$ induced by the operators (1.3), we note that each of the triplets

$$\{J^+,J^-,J^3\} \equiv \{E_\alpha, E_{-\alpha}, J_\alpha\},$$
$$\{E_{\beta_k}, E_{-\beta_k}, J_{\beta_k}\}, \quad \{E_\gamma, E_{-\gamma}, J_\gamma\},$$
$$\{E_{\alpha\beta_k\gamma}, E_{-\alpha,-\beta_k,-\gamma}, J_\alpha + J_{\beta_k} + J_\gamma\},$$
$$\{E_{\alpha\gamma}, E_{-\alpha,-\gamma}, J_\alpha + J_\gamma\}, \quad \{E_{\beta_k\gamma}, E_{-\beta_k,-\gamma}, J_{\beta_k} + J_\gamma\},$$
$$\{E_{\beta_l,-\beta_p}, E_{-\beta_l,\beta_p}, J_{\beta_l} - J_{\beta_p}\}, \quad k = 1,\ldots,n, 1 \le l < p \le n, \tag{1.7}$$

satisfies the commutation relations

$$[J^3,J^\pm] = \pm J^\pm, \quad [J^+,J^-] = 2J^3, \tag{1.8}$$

and forms a basis for a subalgebra of $sl(n+3,\mathbb{C})$ isomorphic to $sl(2,\mathbb{C})$. Furthermore, each triplet generates a local Lie subgroup of $SL(n+3,\mathbb{C})$ isomorphic to $SL(2,\mathbb{C})$ and the subgroups so obtained suffice to generate the full group action of $SL(n+3,\mathbb{C})$.

We pass from the Lie algebra action generated by $\{J^+,J^-,J^3\}$ to the group action via the relation

$$\mathbf{T}(A) = \exp[-(b/d)J^+]\exp(-cdJ^-)\exp(\tau J^3), \quad e^{\tau/2} = d^{-1}, \tag{1.9}$$

where

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{C}), \quad ad-bc = 1; \tag{1.10}$$

see Ref. 7. We find that the triplet $\{E_\alpha, E_{-\alpha}, J_\alpha\}$ generates the group action

$$\mathbf{T}_1(A) f(s,u_1,\ldots,u_n,t,x_1,\ldots,x_n)$$
$$= f\Big(\frac{as+c}{d+bs}, \frac{u_j(as+c)}{as+c(1-x_j)}, \frac{ts}{as+c}, \frac{x_j s}{(d+bs)(as-cx_j+c)}\Big) \tag{1.11}$$

and the triplet $\{E_{\beta_k} E_{-\beta_k}, J_{\beta_k}\}$

$$\mathbf{T}_{2,k}(A) f(s,u_j,u_k,t,x_j,x_k)$$
$$= f\Big(\frac{s(au_k+c)}{au_k+c(1-x_k)}, \frac{u_j}{u_k}(au_k+c), \frac{au_k+c}{d+bu_k}, \frac{u_k t}{au_k+c}, \frac{au_k x_j + c(x_j-x_k)}{au_k+c(1-x_k)}, \frac{x_k u_k}{(d+bu_k)(au_k-cx_k+c)}\Big). \tag{1.12}$$

In (1.11) the index j runs from 1 to n, but in (1.12) j runs from 1 to n excluding k. The triplet $\{E_\gamma, E_{-\gamma}, J_\gamma\}$ generates

$$\mathbf{T}_3(A) f(s,u_j,t,x_j)$$
$$= \Big(a+\frac{c}{t}\Big)^{-1} f\Big(s(d+bt), u_j(d+bt), \frac{at+c}{d+bt}, [dx_j - bt(1-x_j)]\Big(a+\frac{c}{t}\Big)\Big), \tag{1.13}$$

the triplet $\{E_{\alpha\beta_k\gamma}, E_{-\alpha,-\beta_k,-\gamma}, J_\alpha + J_{\beta_k} + J_\gamma\}$ generates

$$\mathbf{T}_{4,k}(A) f(s,u_j,u_k,t,x_j,x_k)$$
$$= \Big(a + \frac{c(1-x_k)}{u_k t_s}\Big)^{-1} f\Big(as - \frac{cx_k}{u_k t}, u_j\Big(\frac{asu_k t - cx_k}{asu_k t + c(x_j-x_k)}\Big),$$

$$au_k - \frac{cx_k}{st}, t\left(\frac{asu_kt + c(1-x_k)}{asu_kt - cx_k}\right), x_j\left(\frac{asu_kt + c(1-x_k)}{asu_kt + c(x_j - x_k)}\right),$$
$$(x_k d - bsu_k t)(a + c - cx_k)\Big), \tag{1.14}$$

the triplet $\{E_{\alpha\gamma}, E_{-\alpha,-\gamma}, J_\alpha + J_\gamma\}$ generates

$$\mathbf{T}_5(A)f(s,u_j,t,x_j) = \left(a - \frac{c}{st}\right)^{-1} f\left(\frac{s}{d-bst}, \frac{u_j st}{ast - cx_j}, at - \frac{c}{s}, \frac{(dx_j - bst)(ast - c)}{(ast - cx_j)(d - bst)}\right), \tag{1.15}$$

the triplet $\{E_{\beta_k\gamma}, E_{-\beta_k,-\gamma}, J_{\beta_k} + J_\gamma\}$ generates

$$\mathbf{T}_{6,k}(A)f(s,u_j,u_k,t,x_j,x_k) = \left(a + \frac{c}{ut}\right)^{-1} f\left(\frac{su_kt}{au_kt + cx_k}, u_j, \frac{u_k}{d + bu_kt}, at + \frac{c}{u_k}, \frac{x_j(au_kt + c)}{au_kt + cx_k}, \frac{(dx_k + bu_kt)(au_kt + c)}{(d + bu_kt)(au_kt + cx_k)}\right), \tag{1.16}$$

and the triplet $\{E_{\beta_k,-\beta_p}, E_{-\beta_k,\beta_p}, J_{\beta_k} - J_{\beta_p}\}$ generates

$$\mathbf{T}_{7,k,p}(A)f(s,u_j,u_p,t,x_j,x_k,x_p) = f\left(s, u_j, \frac{u_ku_p}{du_p + bu_k}, \frac{u_pu_k}{au_k + cu_p}, t, x_j, \frac{dx_ku_p + bx_pu_k}{du_p + bu_k}, \frac{ax_pu_k + cx_ku_p}{au_k + cu_p}\right), \quad 1 \le k < p \le n. \tag{1.17}$$

Let

$$C_k = E_\alpha E_{\beta_k} - E_{\alpha\beta_k\gamma}E_{-\gamma}, \quad 1 \le k \le n. \tag{1.18}$$

It is straightforward to check that the solution f of the simultaneous equations

$$J_\alpha f = (\alpha - \tfrac{1}{2}\gamma)f, \quad J_{\beta_k} f = \left(\beta_k - \tfrac{1}{2}\gamma + \tfrac{1}{2}\sum_{l\neq k}\beta_l\right)f,$$
$$J_\gamma f = \left[\gamma - \tfrac{1}{2}\left(\alpha + \sum_{l=1}^n \beta_l + 1\right)\right]f,$$
$$C_k f = 0, \ k = 1,\ldots,n, \tag{1.19}$$

analytic in a neighborhood of $x_1 = x_2 = \cdots = x_n = 0$ is

$$f = F_D(\alpha;\beta_1,\ldots,\beta_n;\gamma;x_1,\ldots,x_n)s^\alpha u_1^{\beta_1}\cdots u_n^{\beta_n}t^\alpha, \tag{1.20}$$

unique to within a multiplicative constant. In fact the first $n+2$ equations imply

$$f = F(x_1,\ldots,x_n)s^\alpha u_1^{\beta_1}\cdots u_n^{\beta_n}t^\gamma$$

and the last n imply

$$\left[\left(\sum_{j=1}^n x_j\partial_{x_j} + \alpha\right)(x_k\partial_{x_k} + \beta_k) - \partial_{x_k}\left(\sum_{j=1}^n x_j\partial_{x_j} + \gamma - 1\right)\right]F = 0, \quad k = 1,\ldots,n \tag{1.21}$$

which are the partial differential equations for F_D.[4] The operators C_k do not commute with all the elements of $sl(n+3,\mathbb{C})$, but each such element maps a solution f of $C_kf = 0$, $k = 1,\ldots,n$, into another solution. It follows that the operators $\mathbf{T}_j(A)$ also map solutions into solutions. Furthermore, if $f(s,u_j,t,x_j)$ is a solution of $C_kf = 0$, $k = 1,\ldots,n$, which has a Laurent expansion

$$f = \sum_{\alpha,\beta_j,\gamma} g_{\alpha\beta_j\gamma}(x_j)s^\alpha u_1^{\beta_1}\cdots u_n^{\beta_n}t^\gamma, \tag{1.22}$$

and if f is analytic at $x_1 = x_2 = \cdots = x_n = 0$, then it follows from the above remarks that

$$g_{\alpha\beta_j\gamma} = k(\alpha\beta_j\gamma)F_D(\alpha;\beta_1,\ldots,\beta_n;\gamma;x_1,\ldots,x_n), \tag{1.23}$$

where $k(\alpha\beta_j\gamma)$ is a constant.

Let $\alpha^0,\beta_j^0,\gamma^0$, $1 \le j \le n$, be fixed complex numbers, not integers, and let $\alpha = \alpha^0 + h$, $\beta_j = \beta_j^0 + n_j$, $\gamma = \gamma^0 + m$, where h, n_j, m run over all integers. The basis vectors $\{f_{\alpha\beta_j\gamma}\}$, (1.4), and operators (1.3) define an infinite-dimensional irreducible representation $P(\alpha^0,\beta_j^0,\gamma^0)$ of $sl(n+3,\mathbb{C})$. Using operators (1.10)–(1.17), we can extend this Lie algebra representation to a local group representation of $SL(n+3,\mathbb{C})$.

In order to compute the matrix elements of this group representation with respect to the basis functions $\{f_{\alpha\beta_j\gamma}\}$, it is useful to consider the following simple realization of $\rho(\alpha^0,\beta_j^0,\gamma^0)$ in terms of differential operators in $n+2$ complex variables: $s, u_1, \ldots, u_n, t$. The basis functions in this new model are

$$f_{\alpha\beta_j\gamma}(s,u_1,\ldots,u_n,t) = s^\alpha u_1^{\beta_1}\cdots u_n^{\beta_n}t^\gamma \tag{1.24}$$

and the Lie derivatives are

$$E_\alpha = s(t\partial_t - s\partial_s - 1), \quad E_{\alpha\beta_k\gamma} = su_k^2t\partial_{u_k},$$
$$E_{\beta_k} = u_k^2\partial_{u_k}, \quad E_{-\gamma} = t^{-1}(t\partial_t - s\partial_s - 1),$$
$$E_{\alpha\gamma} = st\left(\sum_{j=1}^n u_j\partial_{u_j} - t\partial_t\right), \quad E_\gamma = t\left(t\partial_t - \sum_{j=1}^n u_j\partial_{u_j}\right),$$
$$E_{-\alpha} = s^{-1}(s\partial_s - 1), \quad E_{-\beta_k} = u_k^{-1}\left(t\partial_t - \sum_{j=1}^n u_j\partial_{u_j}\right),$$
$$E_{\beta_k\gamma} = u_k^2t\partial_{u_k}, \quad E_{-\alpha,-\gamma} = s^{-1}t^{-1}(s\partial_s - 1),$$
$$E_{-\alpha,-\beta_k,-\gamma} = s^{-1}u_k^{-1}t^{-1}(1 - s\partial_s),$$
$$E_{-\beta_k,-\gamma} = u_k^{-1}t^{-1}(s\partial_s - t\partial_t + 1),$$
$$E_{\beta_k,-\beta_p} = u_k^2u_p^{-1}\partial_{u_k}, \quad J_\alpha = s\partial_s - \tfrac{1}{2}t\partial_t,$$
$$J_{\beta_k} = u_k\partial_{u_k} - \tfrac{1}{2}t\partial_t + \tfrac{1}{2}\sum_{l\neq k} u_l\partial_{u_k},$$
$$J_\gamma = t\delta_t - \tfrac{1}{2}\left(s\partial_s + \sum_{l=1}^n u_l\partial_u + 1\right). \tag{1.25}$$

As is simple to verify, these operators and basis functions satisfy relations (1.5), so that they determine a model of $\rho(\alpha^0,\beta_j^0,\gamma^0)$. We extend this model to encompass the group action by computing the operators $\mathbf{T}_j(A)$ analogous to (1.11)–(1.17):

$$\mathbf{T}_1(A)f(s,u_j,t) = (d - bs)^{-1}\left(a - \frac{c}{s}\right)^{-1} f\left(\frac{as - c}{d - bs}, u_j, t(d - bs)\right),$$

$$\mathbf{T}_{2,k}(A)f(s,u_j,u_k,t)=f\left(s,u_j\left(a+\frac{c}{u_k}\right),\frac{au_k+c}{d+bu_k},\frac{tu_k}{au_k+c}\right),$$

$$\mathbf{T}_3(A)f(s,u_j,t)=\left(a+\frac{c}{t}\right)^{-1}f\left(\frac{st}{at+c},u_j(d+bt),\frac{at+c}{d+bt}\right),$$

$$\mathbf{T}_{4,k}(A)f(s,u_j,u_k,t)=\left(a+\frac{c}{su_kt}\right)^{-1}f\left(as+\frac{c}{u_kt},\ u_j,\frac{u_k}{d+bu_kst},t\right),$$

$$\mathbf{T}_5(A)f(s,u_j,t)=\left(a-\frac{c}{st}\right)^{-1}f\left(as-\frac{c}{t},u_j(d-bst),\frac{t}{d-bst}\right),$$

$$\mathbf{T}_{6,k}(A)f(s,u_j,u_k,t)=\left(a+\frac{c}{u_kt}\right)^{-1}f\left(\frac{su_kt}{c+au_kt},\ u_j,\frac{u_k}{d+bu_kt},at+\frac{c}{u_k}\right),$$

$$\mathbf{T}_{7,k,p}(A)f(s,u_j,u_k,u_p,t)=\left(d-\frac{bu_k}{u_p}\right)^{-1}\left(a-c\frac{u_p}{u_k}\right)^{-1}\times f(s,u_j,au_k-cu_p,du_p-bu_k,t). \quad (1.26)$$

The matrix elements corresponding to a representation $\mathbf{T}(A)$ of $SL(2,\mathbb{C})$ induced by a triplet $\{J^+,J^-,J^3\}$ acting on a basis f_m according to the rule

$$J^\pm f_m=(-\omega\pm m)f_{m\pm1},\qquad J^3f_m=mf_m,\qquad \omega\in\mathbb{C}, \quad (1.27)$$

have been computed many times before.[7] The result is

$$T_{n'n}(A)=a^{\omega+m_0+n'}d^{\omega-m_0-n}c^{n-n'}\frac{\Gamma(\omega+m_0+n+1)}{\Gamma(\omega+m_0+n'+1)}\times\frac{{}_2F_1(-\omega-m_0-n',-\omega+m_0+n;n-n'+1;bc/ad)}{\Gamma(n-n'+1)}, \quad (1.28)$$

where

$$\mathbf{T}(A)f_{m_0+n}=\sum_{n'=-\infty}^{\infty}T_{n'n}(A)f_{m_0+n'},\qquad n=0,\pm1,\pm2,\ldots, \quad (1.29)$$

and A is in a sufficiently small neighborhood of the identity element. From this result it is easy to compute the matrix elements of each of the operators (1.11)–(1.17).

For more complicated group elements, however, the model (1.26) is very convenient. Consider the $(2n+3)$-dimensional complex Lie algebra G with basis $\{J^+,J^-,J^3,E_j^+,E_j^-,\ j=1,\ldots,n\}$ and commutation relations

$$\begin{aligned}&[J^3,J^\pm]=\pm J^\pm,[J^+,J^-]=2J^3,\\&[J^+,E_j^-]=-E_j^+,[J^-,E_j^+]=-E_j^-,\\&[J^+,E_j^+]=[J^-,E_j^-]=0,\\&[J^3,E_j^\pm]=\pm\tfrac12E_j^\pm,[E_j^+,E_k^+]=0,\\&[E_j^+,E_k^-]=[E_j^-,E_k^-]=0,\qquad j,k=1,\ldots,n. \end{aligned}\quad (1.30)$$

This is the Lie algebra of the group G of $(n+2)\times(n+2)$ matrices

$$\{A,\mathbf{g}^{(1)},\ldots,\mathbf{g}^{(n)}\}=\left(\begin{array}{c|c}A & \mathbf{g}^{(1)},\ldots,\mathbf{g}^{(n)}\\\hline 0 & E_n\end{array}\right),\qquad A\in SL(2,\mathbb{C}),\ g_l^{(k)}\in\mathbb{C}, \quad (1.31)$$

and multiplication law

$$\{A,\mathbf{g}^{(j)}\}\{A',\mathbf{g}'^{(j)}\}=\{AA',A\mathbf{g}'^{(j)}+\mathbf{g}^{(j)}\}. \quad (1.32)$$

Here $\mathbf{g}^{(j)}=\{g_1^{(j)},g_2^{(j)}\}$ is a column 2-vector and E_n is the $n\times n$ identity matrix. The Lie algebra $\mathcal{G}$ is related to G by the expression

$$\{A,\mathbf{g}^{(j)}\}=\exp(g_1^{(1)}E_1^++g_2^{(1)}E_1^-)\cdots\exp(g_1^{(n)}E_n^++g_2^{(n)}E_n^-)\times\exp[-(b/d)J^+]\exp(-cdJ^-)\exp(\tau J^3),\quad e^{\tau/2}=d^{-1}. \quad (1.33)$$

The Lie algebra $\mathcal{G}$ can be embedded as a subalgebra of $sl(n+3,\mathbb{C})$ in many distinct ways, but for purposes of illustration we consider only the example $\{J^+,J^-,J^3,E_j^+,E_j^-\}\equiv\{E_\alpha,E_{-\alpha},J_\alpha,E_{\alpha\beta_j\gamma},E_{\beta_j\gamma}\}$. Using this embedding, we compute the action of G in our $(n+2)$-variable model:

$$\mathbf{T}(A,\mathbf{g}^{(j)})f(s,u_j,t)=\exp(g_1^{(1)}E_{\alpha\beta,\gamma}+g_2^{(1)}E_{\beta,\gamma})\times\cdots\exp(g_1^{(n)}E_{\alpha\beta_n\gamma}+g_2^{(n)}E_{\beta_n\gamma})\mathbf{T}_1(A)f=(d-bs)^{-1}\left(a-\frac{c}{s}\right)^{-1}f\left(\frac{as-c}{d-bs},\frac{u_j}{1-g_1^{(j)}u_jst-g_2^{(j)}u_jt},t(d-bs)\right). \quad (1.34)$$

Applying $\mathbf{T}(A,\mathbf{g}^{(j)})$ to the basis vectors $f_{hn_jm}(s,u_j,t)=s^{\alpha^0+h}u_1^{\beta_1^0+n_1}\cdots u_n^{\beta_n^0+n_n}t^{\gamma^0+m}$, we obtain

$$\mathbf{T}\{A,\mathbf{g}^{(j)}\}f_{h,n_jm}=\sum_{h'n_k',m=-\infty}^{\infty}T(A,\mathbf{g}^{(j)})_{hn_jm}^{h'n_k'm'}f_{h'n_k'm'}, \quad (1.35)$$

or

$$T(A,\mathbf{g}^{(j)})_{hn_jm}^{h'n_k'm'}=0$$

unless $n_j'\geq n_j$, $j=1,\ldots,n$, and $m'-m=\sum_{j=1}^n(n_j'-n_j)$; (1.36)

$$\mathbf{T}(A,\mathbf{g}^{(j)})_{hn_jm}^{h+k,n_k+l_k,m+\Sigma l_j}=\binom{-\beta_1^0-n_1}{l_1}\cdots\binom{-\beta_n^0-n_n}{l_n}\times\frac{\Gamma(1-k-\alpha^0-h)}{\Gamma(1-\alpha^0-h)}[-g_2^{(1)}]^{l_1}\cdots[-g_2^{(n)}]^{l_n}a^{\alpha-1}d^{\gamma-\alpha-1}\times\frac{(a/c)^k}{\Gamma(1-k)}F_D\left(1-k-\alpha;-l_1,\ldots,-l_n,\alpha-\gamma+1;1-k;\frac{-g_1^{(1)}c}{g_2^{(1)}a},\ldots,\frac{-g_1^{(n)}c}{g_2^{(n)}a},\frac{bc}{ad}\right),\qquad l_j\geq0.$$

The group property (1.32) leads immediately to the addition theorem

$$T\{AB,A\mathbf{h}^{(j)}+\mathbf{g}^{(j)}\}_{hn_jm}^{h'n_k'm'}=\sum_{HN_SM}T\{A,\mathbf{g}^{(j)}\}_{HN_SM}^{h'n_k'm'}\times T\{B,\mathbf{h}^{(j)}\}_{hn_jm}^{HN_SM} \quad (1.37)$$

for the F_D.

Equation (1.35) with matrix elements (1.36) is also valid for the $(2n+2)$-variable model. In this case the basis functions are given by (1.4) and the operator $\mathbf{T}\{A,\mathbf{g}^{(j)}\}$ by

$$\mathbf{T}\{A,\mathbf{g}^{(j)}\}f(s,u_j,t,x_j)=f\Bigg(\frac{as+c}{d+bs},\ \frac{u_j(as+c)}{as(1-g_2^{(j)}u_jt)+c(1-x_j-g_1^{(j)}su_jt}},\ \frac{ts}{as+c},\ \frac{s(x_j+g_1^{(j)}su_jt-g_2^{(j)}u_jt)}{(d+bs)[as(1-g_2^{(j)}u_jt)+c(1-x_j-g_1^{(j)}su_jt)]}\Bigg). \tag{1.38}$$

A variety of addition theorems for the F_D can be obtained from (1.35) and (1.38) by specialization of the group parameters. Since this is routine, we give no examples.

2. TRANSFORMATION FORMULAS AND GENERATING FUNCTIONS

We next show that the transformation formulas for the F_D are consequences of the $SL(n+3,\mathbb{C})$ symmetry. Let

$$I=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\in SL(2,\mathbb{C}). \tag{2.1}$$

Expressions (1.4) and (1.11) imply

$$\mathbf{T}_1(I)f_{\alpha\beta_j\gamma}=(-1)^{\alpha+\gamma}[\Gamma(\gamma-\alpha)\Gamma(\alpha)/\Gamma(\gamma)]\times F_D(\alpha;\beta_j;\gamma,x_j/(x_j-1))(1-x_1)^{-\beta_1}\cdots\times(1-x_n)^{-\beta_n}s^{-\alpha+\gamma}u_1^{\beta_1}\cdots u_n^{\beta_n}t^{\gamma}. \tag{2.2}$$

However, $\mathbf{T}_1(I)f_{\alpha\beta_j\gamma}$ is a simultaneous eigenfunction of $J_\alpha,J_{\beta_1},\ldots,J_{\beta_n},J_\gamma$, analytic at $x_1=\cdots=x_n=0$. Thus,

$$\mathbf{T}_1(I)f_{\alpha\beta_j\gamma}=kF_D(\gamma-\alpha;\beta_j;\gamma;x_j)s^{-\alpha+\gamma}u_1^{\beta_1}\cdots u_n^{\beta_n}t^{\gamma}. \tag{2.3}$$

Setting $x_1=\cdots=x_n=0$ in (2.2) and (2.3), we can evaluate the constant k and obtain the transformation formula

$$(1-x_1)^{-\beta_1}\cdots(1-x_n)^{-\beta_n}F_D(\alpha;\beta_j;\gamma;x_j/(x_j-1))=F_D(\gamma-\alpha;\beta_j;\gamma;x_j).^4 \tag{2.4}$$

Similarly, $\mathbf{T}_{2,k}(I)f_{\alpha\beta_j\gamma}$ yields the formulas

$$(1-x_k)^{-\alpha}F_D(\alpha;\beta_j,\beta_k;\gamma;(x_k-x_j)/(x_k-1),x_k/(x_k-1))=F_D\Big(\alpha;\beta_j,\gamma-\sum_{l=1}^{n}\beta_l;\gamma;x_j,x_k\Big),\quad k=1,\ldots,n. \tag{2.5}$$

The remaining transformation formulas for the F_D can be obtained by composition from (2.4) and (2.5).

Computing $\mathbf{T}_3(I)f_{\alpha\beta_j\gamma}$, we find that

$$F_D(\alpha;\beta_j;\alpha+\sum_j\beta_j-\gamma+1;1-x_j) \tag{2.6}$$

is a solution of Eqs. (1.21), analytic at $x_1=\cdots=x_n=1$. Computing $\mathbf{T}_5(I)f_{\alpha\beta_j\gamma}$, we see that

$$x_1^{-\beta_1}\cdots x_n^{-\beta_n}F_D\Big(\sum_l\beta_l-\gamma+1;\beta_j;\sum_l\beta_l-\alpha+1;x_j^{-1}\Big) \tag{2.7}$$

is another solution of (1.21). Similarly, $\mathbf{T}_{6,k}(I)f_{\alpha\beta_j\gamma}$ yields the solution

$$x_k^{-\alpha}F_D(\alpha;\beta_j,\alpha-\gamma+1;\alpha-\beta_k+1;x_j/x_k,1/x_k). \tag{2.8}$$

For A close to the identity in $SL(2,\mathbb{C})$ the expressions $\mathbf{T}_{j,k}(A)f_{\alpha\beta_j\gamma}$ can be expanded by use of the matrix elements (1.28). However, for A far from the identity, say $A=I$, these expansions are no longer valid. For example,

$$(\exp cE_{\alpha\gamma})f(s,u_j,t,x_j)=f\Big(\frac{s}{1+cst},u_j,t,\frac{x_j+cst}{1+cst}\Big). \tag{2.9}$$

For $|c|$ small we find

$$(\exp cE_{\alpha\gamma})f_{\alpha\beta_j\gamma}=\sum_{h=0}^{\infty}\binom{\sum_l\beta_l-\gamma}{h}f_{\alpha+h,\beta_j,\gamma+h}c^h,$$

i.e.,

$$(1+c)^{-\alpha}F_D\Big(\alpha;\beta_j;\gamma;\frac{x_j+c}{1+c}\Big)=\sum_{h=0}^{\infty}\binom{\sum_l\beta_l-\gamma}{h}\times\frac{(\alpha)_h}{(\gamma)_h}F_D(\alpha+h;\beta_j;\gamma+h;x_j)c^h,\qquad |c|<1. \tag{2.10}$$

If $c=1$ and $|\tau|<1$, where $\tau=s^{-1}t^{-1}$, then $(\exp E_{\alpha\gamma})f_{\alpha\beta_j\gamma}$ is not analytic at $x_1=\cdots=x_n=\tau=0$. However, we can apply $\exp E_{\alpha\beta}$ to the solution (2.6) and use (1.22), (1.23) to obtain

$$(1+\tau)^{-\alpha}F_D(\alpha;\beta_j;\alpha+\sum\beta_l-\gamma+1;\tau(1-x_j)/(1+\tau))=\sum_{h=0}^{\infty}C_hF_D(-h;\beta_j;\gamma-\alpha-h;x_j)\tau^h. \tag{2.11}$$

To evaluate the constants C_h, we set $x_1=\cdots x_n=0$:

$$(1+\tau)^{-\alpha}F_D(\alpha,\beta_j;\alpha+\sum_l\beta_l-\gamma+1;\tau/(1+\tau))=(1+\tau)^{-\alpha}{}_2F_1(\alpha,\sum_l\beta_l;\alpha+\sum_l\beta_l-\gamma+1;\tau/(1+\tau))=\sum_{h=0}^{\infty}C_h\tau^h.$$

Thus,

$$C_h=\binom{-\alpha}{h}{}_2F_1(-h,\sum_l\beta_l;\alpha+\sum_l\beta_l-\gamma+1;1)=\binom{-\alpha}{h}\frac{(\alpha-\gamma+1)_h}{(\alpha+\sum\beta_l-\gamma+1)_h} \tag{2.12}$$

from Ref. 7, p. 211, and Vandermonde's theorem.

Expanding $\mathbf{T}_1(A)f_{\alpha\beta_j\gamma}$ as a power series in $\tau=s^{-1}$, we obtain

$$a^{\alpha-\gamma}b^{-\alpha}\Big(1+\frac{c\tau}{a}\Big)^{\alpha+\Sigma\beta_l-\gamma}\Big(1+\frac{d\tau}{b}\Big)^{-\alpha}\Big(1+\frac{c\tau(1-x_1)}{a}\Big)^{-\beta_1}\times\cdots\Big(1+\frac{c\tau(1-x_n)}{a}\Big)^{-\beta_n}\times F_D\Big(\alpha;\beta_j;\gamma;\frac{x_j\tau}{(b+dt)[a+c\tau(1-x_j)]}\Big)=\sum_{h=0}^{\infty}k_hF_D(-h;\beta_j;\gamma;x_j)\tau^h. \tag{2.13}$$

Setting $x_1=\cdots=x_n=0$ and using identity (5.124), Ref. 7, p. 206, we find

$$k_h=\Big(\frac{a}{b}\Big)^{\alpha}a^{-\gamma-h}c^h\binom{-\gamma}{h}{}_2F_1(-h,\alpha;\gamma;-1/bc),\qquad ad-bc=1. \tag{2.14}$$

If $a=d=b=1$ and $c=0$, the identity becomes

$$(1+\tau)^{-\alpha}F_D\left(\alpha;\beta_j;\gamma;\frac{x_j\tau}{1+\tau}\right)$$
$$=\sum_{h=0}^{\infty}\binom{-\alpha}{h}F_D(-h;\beta_j;\gamma;x_j)\tau^h,\quad |\tau|<1, \tag{2.15}$$

and, if $a=c=1, b=-w^{-1}$, it reduces to

$$(1+\tau)^{\alpha+\Sigma\beta_l-\gamma}[1+(1-w)\tau]^{-\alpha}\prod_{l=1}^{n}[1+(1-x_l)\tau]^{-\beta_l}$$
$$\times F_D\left(\alpha;\beta_j;\gamma;\frac{-x_j\tau w}{[1+(1-w)\tau][1+(1-x_j)\tau]}\right)$$
$$=\sum_{h=0}^{\infty}\binom{-\gamma}{h}{}_2F_1(-h,\alpha;\gamma;w)F_D(-h;\beta_j;\gamma;x_j)\tau^k,$$
$$|\tau|<\min(1,|1-x_j|^{-1},|1-w|^{-1}). \tag{2.16}$$

More generally, we can derive generating functions for the F_D through the characterization of a solution f of $C_jf=0$, $j=1,\ldots,n$, by the requirement that f is a simultaneous eigenfunction of $n+2$ independent operators constructed from $sl(n+3,\mathbb{C})$. Such a characterization of $f_{\alpha\beta_j\gamma}$ is given by (1.19).

As an example we compute the solution f of the simultaneous equations

$$E_\alpha f=f,\quad J_{\beta_k}f=\left(\beta_k+\tfrac{1}{2}\sum_{j\neq k}\beta_j-\tfrac{1}{2}\gamma\right)f,$$
$$(J_\gamma+\tfrac{1}{2}J_\alpha)f=\left(\tfrac{3}{4}\gamma-\tfrac{1}{2}\sum_l\beta_l-\tfrac{1}{2}\right)f,$$
$$C_kf=0,\ k=1,\ldots,n, \tag{2.17}$$

which is analytic at $x_1=\cdots=x_n=0$. The first $n+2$ equations have the general solution

$$f=h(x_1/s,\ldots,x_n/s)\exp(-s^{-1})u_1^{\beta_1}\cdots u_n^{\beta_n}t^\gamma,$$

where h is an arbitrary function. Substitution of this expression into $C_kf=0$, $1\le k\le n$, yields

$$h(x_1,\ldots,x_n)=\Phi(\beta_1,\ldots,\beta_n;\gamma;x_1,\ldots,x_n)$$
$$=\sum_{m_1\cdots m_n=0}^{\infty}\frac{(\beta_1)_{m_1}\cdots(\beta_n)_{m_n}}{(\gamma)_{m_1+\ldots+m_n}}\frac{x_1^{m_1}\cdots x_n^{m_n}}{m_1!\cdots m_n!}$$
$$=\lim_{\alpha\to\infty}F_D\left(\alpha;\beta_j;\gamma;\frac{x_j}{\alpha}\right), \tag{2.18}$$

unique up to a constant multiple. Expanding $\mathbf{T}_1(A)f$ as a power series in $\tau=s^{-1}$, we obtain

$$\exp\left[-\left(\frac{d\tau+b}{a+c\tau}\right)\right](a+c\tau)^{\Sigma\beta_l-\gamma}\prod_{l=1}^{n}[a+c\tau(1-x_l)]^{-\beta_l}$$
$$\times\Phi\left(\beta_j;\gamma;\frac{x_j\tau}{(a+c\tau)[a+c\tau(1-x_j)]}\right)$$
$$=\sum_{k=0}^{\infty}r_kF_D(-k;\beta_j;\gamma;x_j)\tau^k,\quad ad-bc=1. \tag{2.19}$$

Setting $x_1=\cdots=x_n=0$ and using the generating function for Laguerre polynomials [(5.101), Ref. 7, p. 190], we find

$$r_k=a^{-\gamma}e^{-b/a}\left(\frac{c}{a}\right)^k L_k^{(\gamma-1)}\left(\frac{1}{ac}\right), \tag{2.20}$$

where $L_n^{(\alpha)}(x)$ is a generalized Laguerre polynomial. If $b=c=0$, $a=d=1$, the identity simplifies to

$$\exp(-\tau)\Phi(\beta_j;\gamma;x_j\tau)=\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}F_D(-k;\beta_j;\gamma;x_j)\tau^k. \tag{2.21}$$

If $a=c=d^{-1}=w^{-1/2}$, $b=0$, we find

$$\exp\left(\frac{-w\tau}{1+\tau}\right)(1+\tau)^{\Sigma\beta_l-\gamma}[1+\tau(1-x_1)]^{-\beta_1}\cdots$$
$$\times[1+\tau(1-x_n)]^{-\beta_n}\Phi\left(\beta_j;\gamma;\frac{x_jw\tau}{(1+\tau)[1+\tau(1-x_j)]}\right)$$
$$=\sum_{k=0}^{\infty}L_k^{(\gamma-1)}(w)F_D(-k;\beta_j;\gamma;x_j)\tau^k,$$
$$|\tau|<\min(1,|x_j-1|^{-1}). \tag{2.22}$$

If $b=-c=1$, $a=d=0$, then $\mathbf{T}_1(A)f$ becomes

$$e^s(1-x_1)^{-\beta_1}\ldots(1-x_n)^{-\beta_n}s^\gamma\Phi\left(\beta_j;\gamma\,\frac{x_js}{1-x_j}\right)$$
$$\times u_1^{\beta_1}\ldots u_n^{\beta_n}t^\gamma.$$

Expanding this function in powers of s, we obtain

$$e^s(1-x_1)^{-\beta_1}\ldots(1-x_n)^{-\beta_n}\Phi\left(\beta_j;\gamma;\frac{x_js}{1-x_j}\right)$$
$$=\sum_{k=0}^{\infty}\frac{s^k}{k!}F_D(\gamma+k;\beta_j;\gamma;x_j). \tag{2.23}$$

Although the derivation of these generating functions is completely routine, an exhaustive classification of such generating functions awaits the classification of all algebraically irreducible representations of $sl(n+3,\mathbb{C})$.

The various confluent forms of the functions F_D have symmetry algebras corresponding to contractions of the algebra $gl(n+3,\mathbb{C})\cong sl(n+3,\mathbb{C})\oplus(\mathcal{E})$. For example, consider the confluent function

$$\psi(\alpha;\beta_1,\ldots,\beta_{n-1};\gamma;x_1,\ldots,x_n)=\sum_{m_1,\ldots,m_n=0}^{\infty}$$
$$\frac{(\gamma)_{m_1+\ldots+m_n}}{(\gamma)_{m_1+\ldots+m_n}}(\beta_1)_{m_1}\cdots(\beta_{n-1})_{m_{n-1}}\frac{x_1^{m_1}\cdots x_n^{m_n}}{m_1!\cdots m_n!}$$
$$=\lim_{\beta_n\to\infty}F_D\left(\alpha;\beta_1,\ldots,\beta_{n-1},\beta_n;\gamma;x_1,\ldots,x_{n-1},\frac{x_n}{\beta_n}\right). \tag{2.24}$$

To obtain the symmetry algebra, we introduce new operators

$$E'_{\alpha\beta_n\gamma}=\frac{1}{\beta_n}E_{\alpha\beta_n\gamma},\quad E'_{\alpha\gamma}=\frac{1}{\beta_n}E_{\alpha\gamma},\quad E'_\gamma=\frac{1}{\beta_n}E_\gamma$$

$$E'_{-\beta_k}=\frac{1}{\beta_n}E_{-\beta_k},\quad k\neq n,\quad E'_{\beta_n\gamma}=\frac{1}{\beta_n}E_{\beta_n\gamma},$$

$$E'_{\beta_n-\beta_k}=\frac{1}{\beta_n}E_{\beta_n,-\beta_k},$$

$J'_{\beta_n}=(1/\beta_n)J_{\beta_n}$ and $E'_\xi=E_\xi$ for all other elements of $sl(n+3,\mathbb{C})$.

Formally letting $\beta_n\to\infty$, we obtain a contracted Lie algebra not isomorphic to $sl(n+3,\mathbb{C})$. The operators which raise and lower u_n are now redundant. Dropping these operators, we are left with an $(n+2)^2$-dimensional non-semi-simple Lie algebra, the symmetry algebra of ψ. This algebra can be used to derive identities for the ψ functions in a manner analogous to that for F_D.

Vilenkin's method of integral transforms and the $(n+2)$-variable model (1.26) can be used to derive Mellin–Barnes integral identities for F_D and its confluent forms. The procedure is completely analogous to that given in Refs. 8, 2, and 3, but the details are complicated.

[1] M. Ciftan, J. Math. Phys. **10**, 1635 (1969).
[2] W. Miller, "Lie Theory and the Appell Functions F_1" (submitted to SIAM J. Math. Anal.).
[3] W. Miller, SIAM J. Math. Anal. 3, 31 (1972).
[4] P. Appell and J. Kampe De Feriet, *Functions hypergeometriques et hypersperiques* (Gauthiers-Villars, Paris, 1926).
[5] L. J. Slater, *Generalized Hypergeometric Functions* (Cambridge U. P., London, 1966).
[6] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. Tricomi, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 1.
[7] W. Miller, *Lie Theory and Special Functions* (Academic, New York, 1968), p. 21.
[8] N. Y. Vilenkin, *Special Functions and the Theory of Group Representations*, transl. (Amer. Math. Soc., Providence, R. I., 1968).

Averages of the Components of Random Unit Vectors

J. F. McDonald

Department of Mathematics, University of Windsor, Windsor, Ontario, Canada

(Received 15 December 1971; Revised Manuscript Received 18 April 1972)

It is shown that a parametrization of the orthogonal and unitary groups due to Hurwitz can be used to evaluate averages of components of random unit vectors for those two spaces. Explicit results are given for moments which are general enough to include most cases of interest in applications.

1. INTRODUCTION

Several years ago, Ullah developed a method for evaluating averages of components of random unit vectors.[1] The technique is applicable for an N-dimensional orthogonal, unitary, or symplectic space.[2] However, the method as given is restricted to the even moments of a single vector for the unitary and symplectic spaces and is restricted to moments involving at most two orthogonal unit vectors for the orthogonal space. Unfortunately, it does not seem possible to extend the method to averages which involve a larger number of vectors.

A possible alternative to Ullah's method is the explicit parametrization of the group of transformations involved. The advantage to this approach is that in principle there is no restriction on the number of vectors involved.

It would appear that the major obstacle is the parametrization itself. That is, one must parametrize the group in such a way that the calculation is tractable. In particular, one must be able to express any element of the rotation matrix explicitly in terms of the parameters, and one must be able to determine the corresponding volume element in the parameter space.

Fortunately, such a parametrization for the orthogonal and unitary groups is known. These parametrizations are due to Hurwitz.[3]

We shall show that these parametrizations are indeed satisfactory for the explicit evaluation of averages which involve any number of vectors.

2. THE GENERAL ROTATION MATRIX

The general rotation matrix for an N-dimension orthogonal or unitary space can be built up out of successive two-dimensional rotations as follows. Let the $N \times N$ matrices $\alpha_r^{(s)}$ be defined as

$$\begin{aligned}[\alpha_r^{(s)}]_{ij} &= \delta_{ij}, \quad i \neq r, r+1, \\ &= a_{N-1-r}^{(s)}\,\delta_{jr} + b_{N-1-r}^{(s)}\,\delta_{jr+1}, \quad i = r \\ &= c_{N-1-r}^{(s)}\,\delta_{jr} + d_{N-1-r}^{(s)}\,\delta_{jr+1}, \quad i = r+1, \end{aligned} \tag{1}$$

where $s = 1, 2, \ldots, N-1$ and $r = N-s, N-s-1, \ldots, N-1$.

The matrix $\alpha_r^{(s)}$ is a rotation in the corresponding two-dimensional subspace. For the orthogonal space $\alpha_r^{(s)}\tilde{\alpha}_r^{(s)} = 1$, and for the unitary space $\alpha_r^{(s)}\alpha_r^{(s)+} = 1$. These matrices will be parametrized below for the orthogonal and unitary spaces.

Next we define the matrices $E^{(s)}$ as

$$E^{(s)} = \overleftarrow{\prod_{r=1}^{N-1}}\ \alpha_{N-r}^{(s)}, \tag{2}$$

where $\overleftarrow{\prod}$ means that successive factors are to the left. It follows easily from mathematical induction that

$$E^{(s)} = \begin{bmatrix} I_{N-s-1} & 0 \\ 0 & T^{(s)} \end{bmatrix}, \tag{3}$$

where I_{N-s-1} is the $(N-s-1) \times (N-s-1)$ unit matrix, and $T^{(s)}$ is the $(s+1) \times (s+1)$ matrix with elements

$$T_{11}^{(s)} = a_{s-1}^{(s)}, \tag{4}$$

$$T_{1j}^{(s)} = a_{s-j}^{(s)} \prod_{r=s-j+1}^{s-1} b_r^{(s)}, \quad j \geq 2, \tag{5}$$

$$T_{ii}^{(s)} = a_{s-i}^{(s)}\, d_{s-i+1}^{(s)}, \quad i \geq 2, \tag{6}$$

$$T_{ij}^{(s)} = a_{s-j}^{(s)} d_{s-i+1}^{(s)} \prod_{r=s-j+1}^{s-i} b_r^{(s)}, \quad j > i \geq 2, \tag{7}$$

$$T_{i\,i-1}^{(s)} = c_{s-i+1}^{(s)}, \quad i \geq 2, \tag{8}$$

$$T_{ij}^{(s)} = 0, \quad i > j+1. \tag{9}$$

In the above equations $a_{-1} \equiv 1$.

Finally, the general rotation matrix A is given by

$$A = \overrightarrow{\prod_{s=1}^{N-1}}\ E^{(s)}, \tag{10}$$

where $\overrightarrow{\prod}$ means that successive factors are to the right. It follows easily from the definition of $T^{(s)}$ that

Vilenkin's method of integral transforms and the $(n+2)$-variable model (1.26) can be used to derive Mellin–Barnes integral identities for F_D and its confluent forms. The procedure is completely analogous to that given in Refs. 8, 2, and 3, but the details are complicated.

[1] M. Ciftan, J. Math. Phys. **10**, 1635 (1969).
[2] W. Miller, "Lie Theory and the Appell Functions F_1" (submitted to SIAM J. Math. Anal.).
[3] W. Miller, SIAM J. Math. Anal. 3, 31 (1972).
[4] P. Appell and J. Kampe De Feriet, *Functions hypergeometriques et hypersperiques* (Gauthiers-Villars, Paris, 1926).
[5] L. J. Slater, *Generalized Hypergeometric Functions* (Cambridge U. P., London, 1966).
[6] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. Tricomi, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 1.
[7] W. Miller, *Lie Theory and Special Functions* (Academic, New York, 1968), p. 21.
[8] N. Y. Vilenkin, *Special Functions and the Theory of Group Representations*, transl. (Amer. Math. Soc., Providence, R. I., 1968).

Averages of the Components of Random Unit Vectors

J. F. McDonald

Department of Mathematics, University of Windsor, Windsor, Ontario, Canada

(Received 15 December 1971; Revised Manuscript Received 18 April 1972)

It is shown that a parametrization of the orthogonal and unitary groups due to Hurwitz can be used to evaluate averages of components of random unit vectors for those two spaces. Explicit results are given for moments which are general enough to include most cases of interest in applications.

1. INTRODUCTION

Several years ago, Ullah developed a method for evaluating averages of components of random unit vectors.[1] The technique is applicable for an N-dimensional orthogonal, unitary, or symplectic space.[2] However, the method as given is restricted to the even moments of a single vector for the unitary and symplectic spaces and is restricted to moments involving at most two orthogonal unit vectors for the orthogonal space. Unfortunately, it does not seem possible to extend the method to averages which involve a larger number of vectors.

A possible alternative to Ullah's method is the explicit parametrization of the group of transformations involved. The advantage to this approach is that in principle there is no restriction on the number of vectors involved.

It would appear that the major obstacle is the parametrization itself. That is, one must parametrize the group in such a way that the calculation is tractable. In particular, one must be able to express any element of the rotation matrix explicitly in terms of the parameters, and one must be able to determine the corresponding volume element in the parameter space.

Fortunately, such a parametrization for the orthogonal and unitary groups is known. These parametrizations are due to Hurwitz.[3]

We shall show that these parametrizations are indeed satisfactory for the explicit evaluation of averages which involve any number of vectors.

2. THE GENERAL ROTATION MATRIX

The general rotation matrix for an N-dimension orthogonal or unitary space can be built up out of successive two-dimensional rotations as follows. Let the $N \times N$ matrices $\alpha_r^{(s)}$ be defined as

$$\begin{aligned}[\alpha_r^{(s)}]_{ij} &= \delta_{ij}, \quad i \neq r, r+1, \\ &= a_{N-1-r}^{(s)}\,\delta_{jr} + b_{N-1-r}^{(s)}\,\delta_{jr+1}, \quad i = r \\ &= c_{N-1-r}^{(s)}\,\delta_{jr} + d_{N-1-r}^{(s)}\,\delta_{jr+1}, \quad i = r+1, \end{aligned} \tag{1}$$

where $s = 1, 2, \ldots, N-1$ and $r = N-s, N-s-1, \ldots, N-1$.

The matrix $\alpha_r^{(s)}$ is a rotation in the corresponding two-dimensional subspace. For the orthogonal space $\alpha_r^{(s)}\tilde{\alpha}_r^{(s)} = 1$, and for the unitary space $\alpha_r^{(s)}\alpha_r^{(s)+} = 1$. These matrices will be parametrized below for the orthogonal and unitary spaces.

Next we define the matrices $E^{(s)}$ as

$$E^{(s)} = \overleftarrow{\prod_{r=1}^{N-1}}\ \alpha_{N-r}^{(s)}, \tag{2}$$

where $\overleftarrow{\prod}$ means that successive factors are to the left. It follows easily from mathematical induction that

$$E^{(s)} = \begin{bmatrix} I_{N-s-1} & 0 \\ 0 & T^{(s)} \end{bmatrix}, \tag{3}$$

where I_{N-s-1} is the $(N-s-1) \times (N-s-1)$ unit matrix, and $T^{(s)}$ is the $(s+1) \times (s+1)$ matrix with elements

$$T_{11}^{(s)} = a_{s-1}^{(s)}, \tag{4}$$

$$T_{1j}^{(s)} = a_{s-j}^{(s)} \prod_{r=s-j+1}^{s-1} b_r^{(s)}, \quad j \geq 2, \tag{5}$$

$$T_{ii}^{(s)} = a_{s-i}^{(s)}\, d_{s-i+1}^{(s)}, \quad i \geq 2, \tag{6}$$

$$T_{ij}^{(s)} = a_{s-j}^{(s)} d_{s-i+1}^{(s)} \prod_{r=s-j+1}^{s-i} b_r^{(s)}, \quad j > i \geq 2, \tag{7}$$

$$T_{i\,i-1}^{(s)} = c_{s-i+1}^{(s)}, \quad i \geq 2, \tag{8}$$

$$T_{ij}^{(s)} = 0, \quad i > j+1. \tag{9}$$

In the above equations $a_{-1} \equiv 1$.

Finally, the general rotation matrix A is given by

$$A = \overrightarrow{\prod_{s=1}^{N-1}}\ E^{(s)}, \tag{10}$$

where $\overrightarrow{\prod}$ means that successive factors are to the right. It follows easily from the definition of $T^{(s)}$ that

$$A_{1j} = T_{1j}^{(N-1)}, \tag{11}$$

$$A_{ij} = \sum_{k_1, k_2, \ldots, k_{i-1}} T_{1k_{i-1}}^{(N-i)} T_{k_{i-1}+1\,k_{i-2}}^{(N-i+1)} \cdots T_{k_2+1\,k_1}^{(N-2)} T_{k_1+1\,j}^{(N-1)}, \quad N > i \geq 2, \tag{12}$$

and

$$A_{Nj} = \sum_{k_1, \ldots, k_{N-2}} T_{2k_{N-2}}^{(1)} T_{k_{N-2}+1\,k_{N-3}}^{(2)} \cdots T_{k_2+1\,k_1}^{(N-2)} T_{k_1+1\,j}^{(N-1)}. \tag{13}$$

The A_{ij} can now be expressed in terms of the $a_r^{(s)}$, $b_r^{(s)}$, $c_r^{(s)}$, and $d_r^{(s)}$ by combining Eqs. (4)–(9) with Eqs. (11)–(13). The resulting expressions for those A_{ij} to be used below are

$$A_{11} = a_{N-2}^{(N-1)}, \tag{14}$$

$$A_{1j} = a_{N-1-j}^{(N-1)} \prod_{r=N-j}^{N-2} b_r^{(N-1)}, \quad j \geq 2, \tag{15}$$

$$A_{i1} = a_{N-i-1}^{(N-i)} \prod_{r=1}^{i-1} c_{N-r-1}^{(N-r)}, \quad N > i \geq 2, \tag{16}$$

$$A_{N1} = \prod_{r=1}^{N-1} c_{N-r-1}^{(N-r)}, \tag{17}$$

$$A_{22} = a_{N-3}^{(N-2)} a_{N-3}^{(N-1)} d_{N-2}^{(N-1)} - a_{N-4}^{(N-2)} b_{N-3}^{(N-2)} c_{N-3}^{(N-1)}, \tag{18}$$

$$\begin{aligned} A_{23} = {} & a_{N-3}^{(N-2)} a_{N-4}^{(N-1)} d_{N-2}^{(N-1)} b_{N-3}^{(N-1)} \\ & + a_{N-4}^{(N-2)} b_{N-3}^{(N-2)} a_{N-4}^{(N-1)} d_{N-3}^{(N-1)} \\ & - a_{N-5}^{(N-2)} b_{N-4}^{(N-2)} b_{N-3}^{(N-2)} c_{N-4}^{(N-1)}, \end{aligned} \tag{19}$$

and

$$\begin{aligned} A_{32} = {} & - a_{N-4}^{(N-3)} c_{N-3}^{(N-2)} a_{N-3}^{(N-1)} d_{N-2}^{(N-1)} \\ & - a_{N-4}^{(N-3)} a_{N-4}^{(N-2)} d_{N-3}^{(N-2)} c_{N-3}^{(N-1)} \\ & + a_{N-5}^{(N-3)} b_{N-4}^{(N-3)} c_{N-4}^{(N-2)} c_{N-3}^{(N-1)}. \end{aligned} \tag{20}$$

It should be noted that $A\bar{A} = 1$ for the orthogonal space and $AA^{+} = 1$ for the unitary space, because of the corresponding conditions on the $\alpha_r^{(s)}$.

3. THE ORTHOGONAL SPACE

We now introduce the explicit parametrization of the orthogonal group due to Hurwitz.[3] Let ϕ_r^s, $r = 0, 1, 2, \ldots, s-1$, $s = 1, 2, \ldots, N-1$, denote a set of $N(N-1)/2$ variables. The general rotation matrix is then given by the above expressions if we define

$$a_r^{(s)} = d_r^{(s)} = \cos\phi_r^s, \tag{21}$$

$$b_r^{(s)} = -c_r^{(s)} = \sin\phi_r^s, \tag{22}$$

where

$$0 \leq \phi_0^s < 2\pi, \quad 0 \leq \phi_r^s \leq \pi, \quad r > 0. \tag{23}$$

The explicit results for the nonzero $T_{ij}^{(s)}$ are

$$T_{11}^{(s)} = \cos\phi_{s-1}^s, \tag{24}$$

$$T_{1j}^{(s)} = \cos\phi_{s-j}^s \prod_{r=s-j+1}^{s-1} \sin\phi_r^s, \quad s+1 \geq j \geq 2, \tag{25}$$

$$T_{ii}^{(s)} = \cos\phi_{s-i}^s \cos\phi_{s-i+1}^s, \quad s+1 \geq i \geq 2, \tag{26}$$

$$T_{ii-1}^{(s)} = -\sin\phi_{s-i+1}^s, \tag{27}$$

$$T_{ij}^{(s)} = \cos\phi_{s-i+1}^s \cos\phi_{s-j}^s \prod_{r=s-j+1}^{s-1} \sin\phi_r^s, \quad s+1 \geq j > i \geq 2, \tag{28}$$

where $\phi_{-1}^{(s)} \equiv 0$. The resulting expressions for those A_{ij} to be used below are

$$A_{11} = \cos\phi_{N-2}^{N-1}, \tag{29}$$

$$A_{1j} = \cos\phi_{N-1-j}^{N-1} \prod_{r=N-j}^{N-2} \sin\phi_r^{N-1}, \quad N \geq j \geq 2, \tag{30}$$

$$A_{i1} = (-1)^{i-1} \cos\phi_{N-i-1}^{N-i} \prod_{r=1}^{i-1} \sin\phi_{N-r-1}^{N-r}, \quad N \geq i \geq 2, \tag{31}$$

$$\begin{aligned} A_{22} = {} & \cos\phi_{N-3}^{N-2} \cos\phi_{N-3}^{N-1} \cos\phi_{N-2}^{N-1} \\ & - \sin\phi_{N-3}^{N-2} \cos\phi_{N-4}^{N-2} \sin\phi_{N-3}^{N-1}, \end{aligned} \tag{32}$$

$$\begin{aligned} A_{23} = {} & \cos\phi_{N-3}^{N-2} \sin\phi_{N-3}^{N-1} \cos\phi_{N-4}^{N-1} \cos\phi_{N-2}^{N-1} \\ & + \cos\phi_{N-4}^{N-2} \sin\phi_{N-3}^{N-2} \cos\phi_{N-4}^{N-1} \cos\phi_{N-3}^{N-1} \\ & - \cos\phi_{N-5}^{N-2} \sin\phi_{N-4}^{N-2} \sin\phi_{N-3}^{N-2} \sin\phi_{N-4}^{N-1}, \end{aligned} \tag{33}$$

$$\begin{aligned} A_{32} = {} & - \cos\phi_{N-4}^{N-3} \sin\phi_{N-3}^{N-2} \cos\phi_{N-3}^{N-1} \cos\phi_{N-2}^{N-1} \\ & - \cos\phi_{N-4}^{N-3} \cos\phi_{N-4}^{N-2} \cos\phi_{N-3}^{N-2} \sin\phi_{N-3}^{N-1} \\ & + \sin\phi_{N-4}^{N-3} \cos\phi_{N-5}^{N-3} \sin\phi_{N-4}^{N-2} \sin\phi_{N-3}^{N-1}, \end{aligned} \tag{34}$$

where $\phi_{-1}^{0} \equiv 0$ in (31).

The volume element dV_N for this parametrization can be chosen as[3]

$$dV_N = \prod_{r,s} (\sin\phi_r^s)^r d\phi_r^s. \tag{35}$$

It is easily verified that

$$\eta_N \equiv \int dV_N = 2^{-N(N-1)/4}. \tag{36}$$

Further if $f(A)$ is any function of the A_{ij} we define its average as

$$\overline{f(A)} \equiv \eta_N^{-1} \int f(A)\, dV_N. \tag{37}$$

Here we wish to consider $f(A)$ of the form $\Pi_{i,j} A_{ij}^{n_{ij}}$, where the n_{ij} are nonnegative integers. For definiteness we consider only quantities of the form

$$Q \equiv A_{22}^n A_{23}^p A_{32}^q \prod_{i=1}^{N} A_{1i}^{l_i} A_{i1}^{m_i}, \tag{38}$$

where n, p, q, and all of the l_i and m_i are nonnegative integers. It should be noted that the power of A_{11} is $l_1 + m_1$.

From Eqs. (29)–(34) it easily follows that

$$\prod_{i=1}^{N} A_{1i}^{l_i} = \prod_{r=1}^{N-1} (\cos\phi_{r-1}^{N-1})^{l_{N-r}} (\sin\phi_{r-1}^{N-1})^{l'_{r-1}}, \tag{39}$$

$$\prod_{i=1}^{N} A_{i1}^{m_i} = (-1)^{\sigma m} \prod_{r=1}^{N-1} (\cos\phi_{r-1}^{r})^{m_{N-r}} (\sin\phi_{r-1}^{r})^{m'_{r-1}}, \tag{40}$$

$$\begin{aligned} A_{22}^n = {} & \sum_{n_1, n_2} (n; n_1, n_2)(-1)^{n_2} \\ & \times (\cos\phi_{N-3}^{N-2} \cos\phi_{N-3}^{N-1} \cos\phi_{N-2}^{N-1})^{n_1} \\ & \times (\sin\phi_{N-3}^{N-2} \cos\phi_{N-4}^{N-2} \sin\phi_{N-3}^{N-1})^{n_2}, \end{aligned} \tag{41}$$

$$\begin{aligned} A_{23}^p = {} & \sum_{p_1, p_2, p_3} (p; p_1, p_2, p_3)(-1)^{p_3} \\ & \times (\cos\phi_{N-3}^{N-2} \sin\phi_{N-3}^{N-1} \cos\phi_{N-4}^{N-1} \cos\phi_{N-2}^{N-1})^{p_1} \end{aligned}$$

TABLE I. Definition of the k_{rj}.

r \ j	$N-5$	$N-4$	$N-3$	$N-2$
$N-3$	q_3	$m_3+q_1+q_2$	—	—
$N-2$	p_3	$n_2+p_2+q_2$	$m_2+n_1+p_1+q_2$	—
$N-1$	—	$l_3+p_1+p_2$	$l_2+n_1+p_2+q_1$	$l_1+m_1+n_1+p_1+q_1$

TABLE II. Definition of the k^1_{rj}.

r \ j	$N-5$	$N-4$	$N-3$	$N-2$
$N-3$	0	$m'_{N-4}+q_3$	—	—
$N-2$	0	p_3+q_3	$m'_{N-3}+n_2+p_2+p_3+q_1$	—
$N-1$	—	$l'_{N-4}+p_3$	$l'_{N-3}+n_2+p_1+q_2+q_3$	$l'_{N-2}+m'_{N-2}$

$$\times(\cos\phi^{N-2}_{N-4}\sin\phi^{N-2}_{N-3}\cos\phi^{N-1}_{N-4}\cos\phi^{N-1}_{N-3})^{p_2}$$
$$\times(\cos\phi^{N-2}_{N-5}\sin\phi^{N-2}_{N-4}\sin\phi^{N-2}_{N-3}\sin\phi^{N-1}_{N-4})^{p_3}, \quad (42)$$

$$A^q_{32}=\sum_{q_1,q_2,q_3}(q;q_1,q_2,q_3)(-1)^{q_1+q_2}$$
$$\times[\cos\phi^{N-3}_{N-4}\sin\phi^{N-2}_{N-3}\cos\phi^{N-1}_{N-3}\cos\phi^{N-1}_{N-2}]^{q_1}$$
$$\times[\cos\phi^{N-3}_{N-4}\cos\phi^{N-2}_{N-4}\cos\phi^{N-2}_{N-3}\sin\phi^{N-1}_{N-3}]^{q_2}$$
$$\times[\sin\phi^{N-3}_{N-4}\cos\phi^{N-3}_{N-5}\sin\phi^{N-2}_{N-4}\sin\phi^{N-1}_{N-3}]^{q_3}, \quad (43)$$

where

$$n'_r\equiv\sum_{i=0}^{r}n_{N-i}, \quad (44)$$

$$\sigma_m\equiv\sum_{i=1}^{N}(i-1)m_i, \quad (45)$$

and

$$(n;n_1,n_2,\ldots,n_m)=\left(n!/\prod_{i=1}^{m}n_i!\right)\delta_{n,n_1+n_2+\cdots+n_m}. \quad (46)$$

Combining Eqs. (38)–(43) we obtain

$$Q=\sum_{\mathbf{n},\mathbf{p},\mathbf{q}}B_{\mathbf{n},\mathbf{p},\mathbf{q}}\prod_{r=1}^{N-1}Q^{(r)}, \quad (47)$$

where

$$B_{\mathbf{n},\mathbf{p},\mathbf{q}}\equiv(-1)^{\bar\sigma}(n;n_1,n_2)(p;p_1,p_2,p_3)(q;q_1,q_2,q_3), \quad (48)$$

$$Q^{(r)}=(\cos\phi^r_{r-1})^{m_{N-r}}[\sin\phi^r_{r-1}]^{m'_{r-1}}, \quad 1\le r\le N-4, \quad (49)$$

$$=\prod_{j=N-5}^{r-1}(\cos\phi^r_j)^{k_{rj}}(\sin\phi^r_j)^{k'_{rj}}, \quad r=N-3,N-2, \quad (50)$$

$$Q^{(N-1)}=\prod_{j=N-4}^{N-2}(\cos\phi^{N-1}_j)^{k_{N-1\,j}}(\sin\phi^{N-1}_j)^{k'_{N-1\,j}}$$
$$\times\prod_{r=1}^{N-4}(\cos\phi^{N-1}_{r-1})^{l_{N-r}}(\sin\phi^{N-1}_{r-1})^{l'_{r-1}}, \quad (51)$$

and

$$\bar\sigma\equiv\sigma_m+n_2+p_3+q_1+q_2. \quad (52)$$

The k_{rj} are given in Table I, and the k'_{rj} are given in Table II.

Since the ϕ^s_r are statistically independent, it follows that

$$\bar Q=\sum_{\mathbf{n},\mathbf{p},\mathbf{q}}B_{\mathbf{n},\mathbf{p},\mathbf{q}}\prod_{r=1}^{N-1}\bar Q^{(r)}, \quad (53)$$

where

$$\bar Q^{(1)}=J_{m_{N-1}m'_0}/J_{00}, \quad (54)$$

$$\bar Q^{(r)}=I_{m_{N-r}m'_{r-1}+r-1}/I_{0\,r-1}, \quad 2\le r\le N-4, \quad (55)$$

$$=\prod_{j=N-5}^{r-1}(I_{k_{rj}k'_{rj}+j}/I_{0j}), \quad r=N-3,N-2, \quad (56)$$

and

$$\bar Q^{(N-1)}=\prod_{j=N-4}^{N-2}(I_{k_{N-1\,j}k'_{N-1\,j}}+j/I_{0j})$$
$$\times(J_{l_{N-1}l'_0}/J_{00})\prod_{r=2}^{N-4}(I_{l_{N-r}l'_{r-1}+r-1}/I_{0\,r-1}). \quad (57)$$

In these equations

$$I_{pq}\equiv\int_0^{\pi}\cos^p\phi\,\sin^q\phi\,d\phi$$
$$=\beta_p\Gamma[\tfrac12(p+1)]\Gamma[\tfrac12(q+1)]/\Gamma[\tfrac12(p+q+2)], \quad (58)$$

and

$$J_{pq}=\int_0^{2\pi}\cos^p\phi\,\sin^q\phi\,d\phi=2\beta_qI_{pq}, \quad (59)$$

where

$$\beta_p\equiv[1+(-1)^p]/2. \quad (60)$$

Combining Eqs. (54)–(59) we obtain

$$\bar Q=C_N[\Gamma(\tfrac12(l'_{N-4}+N-3))\Gamma(\tfrac12(m'_{N-4}+N-3))]^{-1}$$
$$\times\prod_{r=4}^{N}\beta_{m_r}\beta_{l_r}\Gamma(\tfrac12(m_r+1))\Gamma(\tfrac12(l_r+1))$$
$$\times\sum_{\mathbf{n},\mathbf{p},\mathbf{q}}B_{\mathbf{n},\mathbf{p},\mathbf{q}}\prod_{r,j}\beta_{k_{rj}}\Gamma(\tfrac12(k_{rj}+1))\Gamma(\tfrac12(k'_{rj}+j+1))$$
$$\times\Gamma(\tfrac12(k_{rj}+k'_{rj}+j+2)), \quad (61a)$$

where

$$C_N\equiv\left(\prod_{i=0}^{3}\Gamma(\tfrac12(N-i))\right)/\Gamma^2(\tfrac12(N-4))\Gamma^{2N+2}(\tfrac12). \quad (62)$$

The product over r and j includes only those (r,j) for which k_{rj} and k'_{rj} were defined above. It should be noted that $\bar Q=0$ unless $l_r,m_r,r=4,5,\ldots,N$ are all even. If $n=p=q=0$, the result simplifies to

$$\bar Q_{n=p=q=0}=\beta_{l_1+m_1}\prod_{r=2}^{N}\beta_{l_r}\beta_{m_r}\Gamma[\tfrac12(l_r+1)]\Gamma[\tfrac12(m_r+1)]\frac{\Gamma[\tfrac12(l_1+m_1+1)]\Gamma[\tfrac12(l'_{N-2}+m'_{N-2}+N-1)]\Gamma(\tfrac12N)\Gamma[\tfrac12(N-1)]}{\Gamma[\tfrac12(l'_{N-2}+N-1)]\Gamma[\tfrac12(m'_{N-2}+N-1)]\Gamma[\tfrac12(l'_{N-1}+m'_{N-1}+N)]\Gamma^{2N-1}(\tfrac12)}. \quad (61b)$$

Clearly $\bar Q=0$ in this case unless l_1+m_1, and all the l_r and m_r, $r\ge2$ are even.

In Appendix A we have listed explicitly averages for $\Pi A^{n_{ij}}_{ij}$ with $\Sigma n_{ij}\le6$. In particular, those averages which are nonzero are given.

4. THE UNITARY SPACE

Again we use a parametrization due to Hurwitz.[3] Let ϕ_r^s, ψ_r^s, x^s, $s = 1, 2, \ldots, N-1$, $r = 0, 1, \ldots, s-1$, denote a set of $N^2 - 1$ variables. The general rotation matrix is then given by the above results if we define

$$a_r^{(s)} = d_r^{(s)*} = e^{i\psi_r^s} \cos\phi_r^s, \tag{63}$$

$$b_r^{(s)} = -c_r^{(s)*} = e^{ix^s\delta_{r0}} \sin\phi_r^s, \tag{64}$$

where

$$0 \leq x^s < 2\pi, \quad 0 \leq \psi_r^s < 2\pi, \quad 0 \leq \phi_r^s \leq \pi/2. \tag{65}$$

The explicit expressions for the nonzero $T_{ij}^{(s)}$ are

$$T_{11}^{(s)} = e^{i\psi_{s-1}^s} \cos\phi_{s-1}^s, \tag{66}$$

$$T_{1j}^{(s)} = e^{i\psi_{s-j}^s} \cos\phi_{s-j}^s \prod_{r=s-j+1}^{s-1} \sin\phi_r^s e^{ix^s\delta_{r0}}, \quad s+1 \geq j \geq 2, \tag{67}$$

$$T_{jj}^{(s)} = e^{i(\psi_{s-j}^s - \psi_{s-j+1}^s)} \cos\phi_{s-j}^s \cos\phi_{s-j+1}^s, \quad s+1 \geq j \geq 2, \tag{68}$$

$$T_{jk}^{(s)} = e^{i(\psi_{s-k}^s - \psi_{s-j+1}^s)} \cos\phi_{s-k}^s \cos\phi_{s-j+1}^s \times \prod_{r=s-k+1}^{s-j} e^{ix^s\delta_{r0}} \sin\phi_r^s, \quad s+1 \geq k > j \geq 2, \tag{69}$$

$$T_{jj-1}^{(s)} = -e^{-ix^s\delta_{s-j+1\,0}} \sin\phi_{s-j+1}^s, \quad s+1 \geq j \geq 2, \tag{70}$$

where $\phi_{-1}^s = \psi_{-1}^s \equiv 0$.

The volume element for this set of parameters can be chosen as

$$dV_N = \prod_{r,s} \cos\phi_r^s (\sin\phi_r^s)^{2r+1} \, d\phi_r^s \, d\psi_r^s \, dx^s, \tag{71}$$

where

$$\eta_N \equiv dV_N = 1/\sqrt{N!}\, 2^{N(N-1)/2}. \tag{72}$$

If now $f(A)$ is any function of A_{ij} and A_{ij}^*, we define its average as

$$\overline{f(A)} \equiv \eta_N^{-1} \int f(A) dV_N. \tag{73}$$

Here we wish to consider $f(A)$ of the form $\prod_{i,j} A_{ij}^{n_{ij}} A_{ij}^{*\bar{n}_{ij}}$, where the n_{ij} and $\bar{n}_{ij}$ are all nonnegative integers. For definiteness we consider only quantities of the form

$$Q \equiv A_{22}^{n} A^{*\bar{n}} \prod_{j=1}^{N} A_{1j}^{l_j} A_{1j}^{*\bar{l}_j} A_{j1}^{m_j} A_{j1}^{*\bar{m}_j}, \tag{74}$$

where $n, \bar{n}$, and $l_j, \bar{l}_j, m_j, \bar{m}_j$, $j = 1, \ldots, N$, are all nonnegative integers. Note that the power of A_{11} is $l_1 + m_1$ while the power of A_{11}^* is $\bar{l}_1 + \bar{m}_1$.

From Eqs. (11)–(13) and (66)–(70) it follows easily that

$$A_{11} = e^{i\psi_{N-2}^{N-1}} \cos\phi_{N-2}^{N-1}, \tag{75}$$

$$A_{1j} = e^{i\psi_{N-1-j}^{N-1}} \cos\phi_{N-1-j}^{N-1} \prod_{r=2}^{j} \sin\phi_{N-r}^{N-1}, \quad N-1 \geq j \geq 2, \tag{76}$$

$$A_{1N} = e^{ix^{N-1}} \prod_{r=2}^{N} \sin\phi_{N-r}^{N-1}, \tag{77}$$

$$A_{j1} = (-1)^{j-1} e^{i\psi_{N-j-1}^{N-j}} \cos\phi_{N-j-1}^{N-j} \prod_{r=2}^{j} \sin\phi_{N-r}^{N-r+1}, \quad N > j \geq 2, \tag{78}$$

$$A_{N1} = (-1)^{N-1} e^{-ix^1} \prod_{r=2}^{N} \sin\phi_{N-r}^{N-r+1}, \tag{79}$$

and

$$A_{22} = e^{i(\psi_{N-3}^{N-2} + \psi_{N-3}^{N-1} - \psi_{N-2}^{N-1})} \cos\phi_{N-3}^{N-2} \cos\phi_{N-3}^{N-1} \cos\phi_{N-2}^{N-1} - e^{i\psi_{N-4}^{N-2}} \cos\phi_{N-4}^{N-2} \sin\phi_{N-3}^{N-2} \sin\phi_{N-3}^{N-1}. \tag{80}$$

From these equations one easily obtains

$$\begin{aligned} Q = {} & (-1)^{\sigma} e^{i(\Delta l_N x^{N-1} - \Delta m_n x^1)} \\ & \times \prod_{r=0}^{N-2} (e^{i(\Delta l_{N-1-r}\psi_r^{N-1} + \Delta m_{M-1-r}\psi_r^{r+1})}) \\ & \times (\cos\phi_r^{N-1})^{L_{N-1-r}} (\sin\phi_r^{N-1})^{L'_r} (\cos\phi_r^{r+1})^{M_{N-1-r}} \\ & \times (\sin\phi_r^{r+1})^{M'_r} \sum_{k_1, k_2, \bar{k}_1, \bar{k}_2} B_{\mathbf{k}\bar{\mathbf{k}}} \\ & \times e^{i[\Delta k_1(\psi_{N-3}^{N-2} + \psi_{N-3}^{N-1} - \psi_{N-2}^{N-1}) + \Delta k_2 \psi_{N-4}^{N-2}]} \\ & \times (\cos\phi_{N-3}^{N-2} \cos\phi_{N-3}^{N-1} \cos\phi_{N-2}^{N-1})^{K_1} \\ & \times (\cos\phi_{N-4}^{N-2} \sin\phi_{N-3}^{N-2} \sin\phi_{N-3}^{N-1})^{K_2}, \end{aligned} \tag{81}$$

where

$$\sigma \equiv \sum_{j=1}^{N} (j-1) M_j, \tag{82}$$

$$\Delta l_r \equiv l_r - \bar{l}_r, \tag{83}$$

$$L_r \equiv l_r + \bar{l}_r, \tag{84}$$

$$L'_r = \sum_{j=0}^{r} L_{N-j}, \tag{85}$$

and

$$B_{k\bar{k}} \equiv (-1)^{K_2} (n; k_1, k_2)(\bar{n}; \bar{k}_1 \bar{k}_2). \tag{86}$$

Once again since all of the variables are statistically independent, the averaging process is straightforward. The result can be written as

$$\begin{aligned} \overline{Q} = {} & \delta_{\Delta l_1 \Delta n} \delta_{\Delta l_2 \Delta m_2} \delta_{\Delta l_1 - \Delta l_2} \prod_{r=3}^{N} \delta_{l_r \bar{l}_r} \delta_{m_r \bar{m}_r} \Gamma(l_r + 1)\Gamma(m_r + 1) \, \alpha_{\mathbf{LM}n\bar{n}} \sum B_{\mathbf{k}\bar{\mathbf{k}}} \delta_{k_2 \bar{k}_2} \\ & \times \frac{\Gamma[\frac{1}{2}(K_2+2)]\Gamma[\frac{1}{2}(L_1 + M_1 + K_1 + 2)]\Gamma[\frac{1}{2}(M_2 + K_1 + 2)]\Gamma[\frac{1}{2}(L_2 + K_1 + 2)]}{\Gamma[\frac{1}{2}(2N - 4 + K_2)]\Gamma[\frac{1}{2}(2N + K_1 + L'_{N-1} + M'_{N-1})]} \\ & \times \Gamma[\tfrac{1}{2}(2N - 4 + L'_{N-3} + K_2)]\Gamma[\tfrac{1}{2}(2N - 4 + M'_{N-3} + K_2)], \end{aligned} \tag{87}$$

$$\alpha_{\mathbf{LM}n\bar{n}} \equiv \frac{\Gamma(N)\Gamma(N-1)\Gamma(N-2)\Gamma(\frac{1}{2}(2N - 2 + L'_{N-2} + M'_{N-2}))}{\Gamma(N - 2 + l'_{N-3})\Gamma(N - 2 + m'_{N-3})\Gamma[\frac{1}{2}(2N - 2 + n + \bar{n} + L'_{N-2})]\Gamma[\frac{1}{2}(2N - 2 + n + \bar{n} + M'_{N-2})]} \tag{88}$$

If $n = m_1 = 0$, the result simplifies to

$$\overline{Q}_{n=m_1=0} = \delta_{l_1\bar{l}_1}\Gamma(l_1+1)\prod_{r=2}^{N}\delta_{l_r\bar{l}_r}\delta_{m_r\bar{m}_r}\Gamma(l_r+1)\,\Gamma(m_r+1)\,\frac{\Gamma(N)\Gamma(N-1)\Gamma[l'_{N-2}+m'_{N-2}+N-1]}{\Gamma[l'_{N-1}+m'_{N-1}+N]\Gamma[l'_{N-2}+N-1]\Gamma[m'_{N-2}+N-1]}. \quad (89)$$

Note that $\overline{Q}=0$ unless $l_r=\bar{l}_r$, $r=1,\ldots,N$, and $m_r=\bar{m}_r$, $r=2,3,\ldots,N$.

In Appendix B we have listed explicitly averages for $\Pi A_{ij}^{n_{ij}}A_{ij}^{\bar{n}_{ij}}$ with $\Sigma(n_{ij}+\bar{n}_{ij})\leq 4$. In particular, those averages which are nonzero are given.

5. DISCUSSION

It is clear from the calculations of the previous two sections that the evaluation of moments becomes more and more complex (for both the orthogonal and unitary spaces) as more and more A_{ij} $(i>1, j>1)$ are involved.[4] For this reason we have not attempted to evaluate the most general moments (i.e., all n_{ij} arbitrary). Instead, we restricted our discussion to moments which were special enough so that the calculations were tractable; but at the same time the moments were general enough so that the results given probably include most cases of interest in applications. Clearly, any particular case which is not included in the above results can be evaluated in a straightforward manner using the expressions given above for the parametrization of the A_{ij} and the volume elements.[5]

There are various possible applications of the results given in the previous two sections. Ullah used such averages to investigate the implications of the invariance hypothesis on matrix element correlations for an ensemble of random matrices.[1,6] The results given above enable one to consider the effects of the invariance hypothesis on higher correlations of Hamiltonian matrix elements.

Another obvious application of the above parametrizations is to the problem of the distribution of eigenvalues and eigenvector components of random matrices.[7]

For example, for an ensemble of Hamiltonian matrices which satisfies the invariance hypothesis, the distribution of widths for both the orthogonal and unitary spaces can be derived in a trivial manner. For such an ensemble the corresponding distribution of widths $p(X)$ is given by[8]

$$p(X)=\eta_N^{-1}\int\delta(X-N|A_{11}|^2)dV_N. \quad (90)$$

Via the parametrizations given above for the orthogonal and unitary spaces, one finds easily that

$$p_{ort}(X)=I_{0N-2}^{-1}\int_0^{\pi}\delta(X-N\cos^2\phi_{N-2}^{N-1})(\sin\phi_{N-2}^{N-1})^{N-2}\times d\phi_{N-2}^{N-1}=\frac{\Gamma(\frac{1}{2}N)}{\sqrt{\pi NX}\,\Gamma(\frac{1}{2}(N-1))}\left(1-\frac{X}{N}\right)^{(N-3)/2}, \quad (91)$$

and

$$p_{unit}(X)=J_1{}_{2N-3}^{-1}\int_0^{\pi/2}\delta(X-N\cos^2\phi_{N-2}^{N-1})\times\cos\phi_{N-2}^{N-1}(\sin\phi_{N-2}^{N-1})^{2N-3}d\phi_{N-2}^{N-1}=\frac{N-1}{N}\left(1-\frac{X}{N}\right)^{N-2}. \quad (92)$$

These results agree with those previously obtained.[9,10]

In addition there has been recent interest in the problem of how a small time-reversal odd term in the Hamiltonian of a complex system would influence the statistical properties of the energy levels and widths.[8,10-17] Such investigations have led to interest in ensembles which are not representationally invariant.

We feel that the above parametrizations (and/or moments) will turn out to be very valuable in future investigations of the distribution of eigenvalues and eigenvector components for such ensembles.

APPENDIX A: SOME PARTICULAR ORTHOGONAL AVERAGES

In this appendix we give the results for $\langle\pi A_{ij}^{n_{ij}}\rangle$ with $\Sigma n_{ij}\leq 6$. In particular,

$$\langle A_{11}^2\rangle=\frac{1}{N},\quad \langle A_{11}^4\rangle=\frac{3}{N(N+2)},$$

$$\langle A_{11}^6\rangle=\frac{16}{N(N+2)(N+4)},\quad \langle A_{11}^2A_{12}^2\rangle=\frac{1}{N(N+2)},$$

$$\langle A_{12}^2A_{21}^2\rangle=\frac{N+1}{N(N-1)(N+2)},$$

$$\langle A_{11}^4A_{12}^2\rangle=\frac{3}{N(N+2)(N+4)},$$

$$\langle A_{12}^4A_{21}^2\rangle=\frac{3(N+3)}{N(N-1)(N+2)(N+4)},$$

$$\langle A_{11}^2A_{12}^2A_{13}^2\rangle=\frac{1}{N(N+2)(N+4)},$$

$$\langle A_{11}^2A_{12}^2A_{21}^2\rangle=\frac{N+1}{N(N-1)(N+2)(N+4)},$$

$$\langle A_{12}^2A_{13}^2A_{21}^2\rangle=\frac{N+3}{N(N-1)(N+2)(N+4)},$$

$$\langle A_{31}^2A_{22}^2A_{13}^2\rangle=\frac{N^2+3N-2}{N(N-1)(N+2)(N-2)(N+4)},$$

$$\langle A_{11}A_{12}A_{21}A_{22}\rangle=-\frac{1}{N(N-1)(N+2)},$$

$$\langle A_{11}^2A_{12}A_{13}A_{22}A_{23}\rangle=-\frac{1}{N(N-1)(N+2)(N-2)(N-4)},$$

and

$$\langle A_{12}A_{13}A_{22}A_{23}A_{31}^2\rangle=-\frac{1}{N(N-1)(N-2)(N+4)}.$$

Every other *nonzero* average with $\Sigma n_{ij}\leq 6$ can be obtained from these by appropriate permutations of the labeling of rows and columns.

APPENDIX B: SOME PARTICULAR UNITARY AVERAGES

In this appendix we give the results for $\langle\pi A_{ij}^{n_{ij}}A_{ij}^{*\bar{n}_{ij}}\rangle$ with $\Sigma(n_{ij}+\bar{n}_{ij})\leq 4$. In particular,

$$\langle|A_{11}|^2\rangle = 1/N, \quad \langle|A_{11}|^4\rangle = 2/N(N+1),$$

$$\langle|A_{11}|^2|A_{12}|^2\rangle = 1/N(N+1),$$

$$\langle|A_{11}|^2|A_{22}|^2\rangle = 1/(N-1)(N+1),$$

and

$$\langle A_{11}A_{12}^*A_{21}^*A_{22}\rangle = -1/N(N-1)(N+1).$$

Every other *nonzero* average with $\Sigma(n_{ij}+\bar{n}_{ij}) \leq 4$ can be obtained from these by appropriate permutations of the labeling of rows and columns.

[1] N. Ullah, Nucl. Phys. **58**, 65 (1964).
[2] F. J. Dyson, J. Math. Phys. **3**, 140 (1962).
[3] A. Hurwitz, Math. Werke **2**, 546 (1933); reproduced in R. E. Bellman, *A Collection of Modern Mathematical Classics; Analysis* (Dover, New York, 1961).
[4] In principle the problem itself is not more difficult. However, the notation becomes extremely complex.
[5] It should be noted that the results are invariant under any permutation of the labeling of rows and columns. This can often be used to simplify the evaluation of a particular moment. For example, $\langle A_{22}^{q}A_{23}^{r}A_{32}^{p}A_{33}^{q}\rangle = \langle A_{11}^{q}A_{12}^{r}A_{21}^{p}A_{22}^{q}\rangle$, obviously the moment on the right involves a simpler calculation than the one on the left.
[6] N. Ullah and C. E. Porter, Phys. Letters **6**, 301 (1963).
[7] This problem arose in connection with the distribution of energy level spacings and level widths in nuclear spectra. See, for example, C. E. Porter, *Statistical Theories of Spectra: Fluctuations* (Academic, New York, 1965) (Note: Refs. 1, 2, 5, 6, and 9 are included in this collection).
[8] N. Rosenzweig, J. E. Monahan, and M. L. Mehta, Nucl. Phys. **A109**, 437 (1968).
[9] C. E. Porter and R. G. Thomas, Phys. Rev. **104**, 483 (1956).
[10] N. Ullah, J. Math. Phys. **4**, 1279 (1963).
[11] M. L. Mehta and N. Rosenzweig, Nucl. Phys. **A109**, 449 (1968).
[12] M. L. Mehta, Nuovo Cimento **65B**, 107 (1970).
[13] L. D. Favro and J. F. McDonald, Phys. Rev. Letters **19**, 1254 (1967).
[14] L. D. Favro and J. F. McDonald, J. Math. Phys. **9**, 1429 (1968).
[15] J. F. McDonald, J. Math. Phys. **10**, 1191 (1969).
[16] J. F. McDonald and L. D. Favro, J. Math. Phys. **11**, 3103 (1970).
[17] J. F. McDonald, J. Math. Phys. **12**, 542 (1971).

Wave Propagation in a Random Lattice. I*

Pao-Liu Chow and Joseph B. Keller
Department of Mathematics, University Heights and Courant Institute of Mathematical Sciences, New York University, New York, New York

(Received 12 January 1972; Revised Manuscript Received 20 January 1972)

The small amplitude periodic classical motion of a lattice of particles about their equilibrium positions in a lattice is considered. The effect of random masses and random spring constants upon the coherent or mean motion is treated by using an equation for the coherent motion derived previously by Keller and others. From this equation the dispersion equation for coherent wave motion is determined. It is solved for the case in which the spring constants are not random but the masses are random. Explicit results are obtained in the one-dimensional case for both uncorrelated and exponentially correlated mass defects. They show an alteration of frequency or of wavelength and of phase velocity, as well as an attenuation due to scattering by the defects. In addition new highly attenuated modes are found. These results are utilized in Part II in which various reflection and Green's function problems are treated.

INTRODUCTION

We consider the small amplitude time periodic classical motion of a collection of particles about their positions of equilibrium, which form a lattice. The masses of the particles and the coupling constants which determine the forces between them are assumed to be random quantities having a prescribed joint probability distribution. We call the ensemble of possible lattices, with this distribution, a random lattice, and consider it to be a model of a crystal with imperfections. However, the results also apply to other physical systems. Our purpose is to study wave propagation in such a lattice. To do so we assume that the masses and coupling constants differ by small random amounts from constant values, and we employ a perturbation theory to analyze the mean wave motion.

A suitable perturbation method for such problems has been devised by Keller,[1] Bourret,[2] and Tatarski and Gercenstein.[3] It yields an equation satisfied by the mean wave, which is correct through terms of second order in the random quantities. In Sec. 1 we use this method to derive the dispersion equation for the mean or coherent wave in a lattice with one particle per unit cell. In the Appendix we treat the general case. In Sec. 2 we simplify and solve this equation for certain one-, two-, and three-dimensional cubic lattices in which only the masses are random. In Sec. 3 we obtain more explicit results for the one-dimensional case. The results show that even in the pass bands of the perfect lattice the propagation constant or wavenumber of the mean wave is complex, so that the wave attenuates. The attenuation is a consequence of scattering by irregularities. The phase velocity of the mean wave is also affected by the random irregularities, and may be either greater or less than that in a perfect lattice, depending upon the frequency. The dispersion equation involves the correlation function of the masses, and we consider both correlated and uncorrelated mass variations.

In Paper II[4] we analyze the dispersion equation for two- and three-dimensional lattices. We also consider reflection of a plane wave from the plane interface between a perfect and an imperfect crystal and from a slab of imperfect crystal of finite thickness. In addition, we construct the Green's function for an infinite imperfect crystal and for a semi-infinite imperfect crystal joined to a semi-infinite perfect one.

Previously Koster,[5] Lifshitz,[6] and others have treated scattering of waves by localized irregularities in crystals. In addition various authors have considered the effect of irregularities on the vibration frequencies of crystals.[6] However, Rubin[7,8] seems to have been the first to have investigated wave propagation in an infinite one-dimensional crystal with randomly distributed mass defects. He has investigated the mean of the logarithm of the amplitude transmission

$$\langle|A_{11}|^2\rangle = 1/N, \quad \langle|A_{11}|^4\rangle = 2/N(N+1),$$

$$\langle|A_{11}|^2|A_{12}|^2\rangle = 1/N(N+1),$$

$$\langle|A_{11}|^2|A_{22}|^2\rangle = 1/(N-1)(N+1),$$

and

$$\langle A_{11}A_{12}^*A_{21}^*A_{22}\rangle = -1/N(N-1)(N+1).$$

Every other *nonzero* average with $\Sigma(n_{ij}+\bar{n}_{ij}) \leq 4$ can be obtained from these by appropriate permutations of the labeling of rows and columns.

[1] N. Ullah, Nucl. Phys. **58**, 65 (1964).
[2] F. J. Dyson, J. Math. Phys. **3**, 140 (1962).
[3] A. Hurwitz, Math. Werke **2**, 546 (1933); reproduced in R. E. Bellman, *A Collection of Modern Mathematical Classics; Analysis* (Dover, New York, 1961).
[4] In principle the problem itself is not more difficult. However, the notation becomes extremely complex.
[5] It should be noted that the results are invariant under any permutation of the labeling of rows and columns. This can often be used to simplify the evaluation of a particular moment. For example, $\langle A_{22}^{q}A_{23}^{r}A_{32}^{p}A_{33}^{q}\rangle = \langle A_{11}^{q}A_{12}^{r}A_{21}^{p}A_{22}^{q}\rangle$, obviously the moment on the right involves a simpler calculation than the one on the left.
[6] N. Ullah and C. E. Porter, Phys. Letters **6**, 301 (1963).
[7] This problem arose in connection with the distribution of energy level spacings and level widths in nuclear spectra. See, for example, C. E. Porter, *Statistical Theories of Spectra: Fluctuations* (Academic, New York, 1965) (Note: Refs. 1, 2, 5, 6, and 9 are included in this collection).
[8] N. Rosenzweig, J. E. Monahan, and M. L. Mehta, Nucl. Phys. **A109**, 437 (1968).
[9] C. E. Porter and R. G. Thomas, Phys. Rev. **104**, 483 (1956).
[10] N. Ullah, J. Math. Phys. **4**, 1279 (1963).
[11] M. L. Mehta and N. Rosenzweig, Nucl. Phys. **A109**, 449 (1968).
[12] M. L. Mehta, Nuovo Cimento **65B**, 107 (1970).
[13] L. D. Favro and J. F. McDonald, Phys. Rev. Letters **19**, 1254 (1967).
[14] L. D. Favro and J. F. McDonald, J. Math. Phys. **9**, 1429 (1968).
[15] J. F. McDonald, J. Math. Phys. **10**, 1191 (1969).
[16] J. F. McDonald and L. D. Favro, J. Math. Phys. **11**, 3103 (1970).
[17] J. F. McDonald, J. Math. Phys. **12**, 542 (1971).

Wave Propagation in a Random Lattice. I*

Pao-Liu Chow and Joseph B. Keller
Department of Mathematics, University Heights and Courant Institute of Mathematical Sciences, New York University, New York, New York

(Received 12 January 1972; Revised Manuscript Received 20 January 1972)

The small amplitude periodic classical motion of a lattice of particles about their equilibrium positions in a lattice is considered. The effect of random masses and random spring constants upon the coherent or mean motion is treated by using an equation for the coherent motion derived previously by Keller and others. From this equation the dispersion equation for coherent wave motion is determined. It is solved for the case in which the spring constants are not random but the masses are random. Explicit results are obtained in the one-dimensional case for both uncorrelated and exponentially correlated mass defects. They show an alteration of frequency or of wavelength and of phase velocity, as well as an attenuation due to scattering by the defects. In addition new highly attenuated modes are found. These results are utilized in Part II in which various reflection and Green's function problems are treated.

INTRODUCTION

We consider the small amplitude time periodic classical motion of a collection of particles about their positions of equilibrium, which form a lattice. The masses of the particles and the coupling constants which determine the forces between them are assumed to be random quantities having a prescribed joint probability distribution. We call the ensemble of possible lattices, with this distribution, a random lattice, and consider it to be a model of a crystal with imperfections. However, the results also apply to other physical systems. Our purpose is to study wave propagation in such a lattice. To do so we assume that the masses and coupling constants differ by small random amounts from constant values, and we employ a perturbation theory to analyze the mean wave motion.

A suitable perturbation method for such problems has been devised by Keller,[1] Bourret,[2] and Tatarski and Gercenstein.[3] It yields an equation satisfied by the mean wave, which is correct through terms of second order in the random quantities. In Sec. 1 we use this method to derive the dispersion equation for the mean or coherent wave in a lattice with one particle per unit cell. In the Appendix we treat the general case. In Sec. 2 we simplify and solve this equation for certain one-, two-, and three-dimensional cubic lattices in which only the masses are random. In Sec. 3 we obtain more explicit results for the one-dimensional case. The results show that even in the pass bands of the perfect lattice the propagation constant or wavenumber of the mean wave is complex, so that the wave attenuates. The attenuation is a consequence of scattering by irregularities. The phase velocity of the mean wave is also affected by the random irregularities, and may be either greater or less than that in a perfect lattice, depending upon the frequency. The dispersion equation involves the correlation function of the masses, and we consider both correlated and uncorrelated mass variations.

In Paper II[4] we analyze the dispersion equation for two- and three-dimensional lattices. We also consider reflection of a plane wave from the plane interface between a perfect and an imperfect crystal and from a slab of imperfect crystal of finite thickness. In addition, we construct the Green's function for an infinite imperfect crystal and for a semi-infinite imperfect crystal joined to a semi-infinite perfect one.

Previously Koster,[5] Lifshitz,[6] and others have treated scattering of waves by localized irregularities in crystals. In addition various authors have considered the effect of irregularities on the vibration frequencies of crystals.[6] However, Rubin[7,8] seems to have been the first to have investigated wave propagation in an infinite one-dimensional crystal with randomly distributed mass defects. He has investigated the mean of the logarithm of the amplitude transmission

coefficient for a section of the crystal containing N defects, for N large. A similar study was made by Matsuda and Ishii,[9] who also gave references to related work. One of the results of Rubin and Matsuda and Ishii is compared with one of ours in Sec. 4.

1. DERIVATION OF THE DISPERSION EQUATION

Let us consider an infinite crystal in n dimensions. Each particle of the crystal is labeled by an index q which is a set of n integers. The equilibrium position of particle q is the lattice site $\mathbf{x}(q)$. We suppose that the particles are undergoing small amplitude periodic motions with angular frequency ω about their equilibrium positions. Let the real part of $\mathbf{u}(q)e^{-i\omega t}$ be the displacement of particle q from its equilibrium position. We assume that $\mathbf{u}$ satisfies the linear equation of motion

$$-M(q)\omega^2\mathbf{u}(q) + \sum_{q'} \Phi(q,q')\mathbf{u}(q') = \mathbf{f}(q). \tag{1.1}$$

Here $M(q)$ is the mass of particle q, $\mathbf{f}(q)e^{-i\omega t}$ is the external force on particle q, and the matrix $\Phi(q,q')$ determines the force on particle q exerted by a displacement of particle q'.

We wish to study those solutions of (1.1) which represent either propagating or standing waves. In particular, we shall determine the effect of random defects or impurities in the crystal upon the waves. Thus we assume that the masses $M(q)$, the coupling matrices $\Phi(q,q')$, and the forces $\mathbf{f}(q)$ are random quantities with a given joint probability distribution. As a consequence a solution $\mathbf{u}(q)$ of (1.1) will also be random. The mean value of $\mathbf{u}(q)$, which we denote $\langle\mathbf{u}(q)\rangle$, is often referred to as the coherent wave and the difference $u - \langle u\rangle$ as the fluctuating or incoherent wave. We shall consider only the coherent wave and attempt to determine it.

For this purpose it is convenient to introduce the linear operator L defined by

$$L = \sum_{q'} [\Phi(q,q') - \omega^2 M(q)\delta_{qq'}]. \tag{1.2}$$

Then (1.1) can be written in the form

$$L\mathbf{u} = \mathbf{f}. \tag{1.3}$$

The solution of (1.3) can be written formally as $\mathbf{u} = L^{-1}\mathbf{f}$ and then $\langle\mathbf{u}\rangle = \langle L^{-1}\mathbf{f}\rangle$. We now assume that L and f are statistically independent so that $\langle\mathbf{u}\rangle = \langle L^{-1}\rangle\langle\mathbf{f}\rangle$. Next, following Keller we multiply by $\langle L^{-1}\rangle^{-1}$ to obtain

$$\langle L^{-1}\rangle^{-1}\langle\mathbf{u}\rangle = \langle\mathbf{f}\rangle. \tag{1.4}$$

This is an exact equation satisfied by $\langle\mathbf{u}\rangle$, although in this form it is not yet useful.

We now assume that the lattice is statistically homogeneous, which means that the operator $\langle L^{-1}\rangle$ is invariant under the translations which leave the lattice invariant. We also assume that $\langle\mathbf{f}\rangle = 0$. Then (1.4) possesses plane wave solutions. For a lattice with one particle per unit cell a plane wave has the form

$$\langle\mathbf{u}(q)\rangle = \mathbf{A}e^{i\mathbf{k}\cdot\mathbf{x}(q)}. \tag{1.5}$$

Here $\mathbf{x}(q)$ is the position of the qth lattice site. The case in which there are more than one particle in each unit cell is treated in the Appendix. By using (1.5) in (1.4) with $\langle\mathbf{f}\rangle = 0$ and multiplying by $e^{-i\mathbf{k}\cdot\mathbf{x}(q)}$, we obtain

$$e^{-i\mathbf{k}\cdot\mathbf{x}(q)}\langle L^{-1}\rangle^{-1}e^{i\mathbf{k}\cdot\mathbf{x}(q')}\mathbf{A} = 0. \tag{1.6}$$

In order that (1.6) have a solution $\mathbf{A}$ different from zero, the coefficient matrix must be singular,

$$\det(e^{-i\mathbf{k}\cdot\mathbf{x}(q)}\langle L^{-1}\rangle^{-1}e^{i\mathbf{k}\cdot\mathbf{x}(q)}) = 0. \tag{1.7}$$

Equation (1.7) is the exact dispersion equation relating the wave vector or propagation vector $\mathbf{k}$ of the mean wave to the frequency ω. The matrix in it is of nth order in n dimensions. It is independent of q as a consequence of the translational invariance of $\langle L^{-1}\rangle$; in fact, this independence could be taken as the definition of such invariance.

To make (1.7) useful we assume that L is the sum of a nonrandom operator L_0 and a small random operator ϵV, where ϵ is a small parameter. Thus we write

$$L = L_0 + \epsilon V. \tag{1.8}$$

We also assume that $\langle V\rangle = 0$, which can always be made so by choosing $L_0 = \langle L\rangle$. Then by using (1.8) we obtain, to second order in ϵ,

$$\langle L^{-1}\rangle^{-1} = L_0 - \epsilon^2\langle VL_0^{-1}V\rangle + O(\epsilon^3). \tag{1.9}$$

Upon using (1.9) in (1.7) and omitting $O(\epsilon^3)$ we obtain the dispersion equation in the following useful form:

$$\det[e^{-i\mathbf{k}\cdot\mathbf{x}(q)}(L_0 - \epsilon^2\langle VL_0^{-1}V\rangle)e^{i\mathbf{k}\cdot\mathbf{x}(q')}] = 0. \tag{1.10}$$

More terms can be obtained in (1.10) by keeping more terms in (1.9).

In order to apply the result (1.10) we assume that M and Φ are of the forms

$$M(q) = m[1 + \epsilon\mu(q)], \tag{1.11}$$

$$\Phi(q,q') = \Phi_0(q,q') + \epsilon\varphi(q,q'). \tag{1.12}$$

Here the constant m and the matrix $\Phi_0(q,q')$ are not random and pertain to the perfect crystal. The defects are described by the random functions $\mu(q)$ and $\varphi(q,q')$ which are assumed to have zero mean values:

$$\langle\mu(q)\rangle = 0, \quad \langle\varphi(q,q')\rangle = 0. \tag{1.13}$$

The small parameter ϵ is introduced to indicate that the random imperfections are small. We can now write L given by (1.2) in the form (1.8) with L_0 and V defined by

$$L_0 = \sum_{q'} [\Phi_0(q,q') - \omega^2 m\,\delta_{qq'}], \tag{1.14}$$

$$V = \sum_{q'} [\varphi(q,q') - \omega^2 m\,\mu(q)\delta_{qq'}]. \tag{1.15}$$

To use (1.10) we must calculate L_0^{-1}, which can be expressed by means of the Green's matrix $G(q,q',\omega)$ of the perfect crystal. This matrix is the outgoing solution of $L_0G = \delta_{qq'}I$, where I is the unit matrix. By using (1.14) we can write this equation as

$$-\omega^2 m G(q,q',\omega) + \sum_{q''} \Phi_0(q,q'') G(q'',q',\omega) = \delta_{qq'} I. \tag{1.16}$$

In terms of G we can write the dispersion equation (1.10) explicitly as follows:

$$\det\{-\omega^2 mI + \sum_{q'} \Phi_0(q,q') e^{i\mathbf{k}\cdot[\mathbf{x}(q')-\mathbf{x}(q)]} - \epsilon^2 \sum_{q'q''q'''} \langle[\varphi(q,q') - \omega^2 m\mu(q)\delta_{qq'}] \times G(q',q'',\omega)[\varphi(q'',q''') - \omega^2 m\mu(q'')\delta_{q''q'''}]\rangle \times e^{i\mathbf{k}\cdot[\mathbf{x}(q''')-\mathbf{x}(q)]}\} = 0. \tag{1.17}$$

This is the dispersion equation which we shall simplify and solve in the next section.

2. SIMPLIFICATION OF THE DISPERSION EQUATION

We shall now assume that the coupling matrices $\Phi(q,q')$ are not random so that $\varphi(q,q') = 0$. Then (1.17) becomes

$$\det\Big(-\omega^2 mI + \tilde{\Phi}(\mathbf{k}) - \epsilon^2\omega^4 m^2 \times \sum_{q'} G(q,q',\omega)\langle\mu(q)\mu(q')\rangle e^{i\mathbf{k}\cdot[\mathbf{x}(q')-\mathbf{x}(q)]}\Big) = 0. \tag{2.1}$$

Here we have introduced the matrix $\tilde{\Phi}(\mathbf{k})$, the discrete Fourier transform of $\Phi_0(q,q')$, defined by

$$\tilde{\Phi}(\mathbf{k}) = \sum_{q'} \Phi_0(q,q') e^{i\mathbf{k}\cdot[\mathbf{x}(q')-\mathbf{x}(q)]}. \tag{2.2}$$

Because Φ_0 is invariant under lattice translations, $\tilde{\Phi}(\mathbf{k})$ is independent of q.

It is convenient to introduce $R(q,q')$, the two point correlation function of μ, defined by

$$R(q,q') = \langle\mu(q)\mu(q')\rangle. \tag{2.3}$$

In view of the assumed statistical homogeneity of the crystal, R is invariant under lattice translations. Then the sum in (2.1) can be recognized as the discrete Fourier transform of GR. We shall denote it $\widetilde{GR}(\mathbf{k},\omega)$ since G depends upon ω and the transform variable is $\mathbf{k}$. Thus this matrix is defined by

$$\widetilde{GR}(\mathbf{k},\omega) = \sum_{q'} G(q,q',\omega)R(q,q')e^{i\mathbf{k}\cdot[\mathbf{x}(q')-\mathbf{x}(q)]}. \tag{2.4}$$

We note that $\widetilde{GR}$ is independent of q because both G and R are invariant under lattice translations. Now by using (2.4) we can write (2.1) in the compact form

$$\det\{\omega^2 mI - \tilde{\Phi}(\mathbf{k}) + \epsilon^2\omega^4 m^2 \widetilde{GR}(\mathbf{k},\omega)\} = 0. \tag{2.5}$$

To study propagating waves we must solve (2.5) for $\mathbf{k}$ as a function of ω, with ω real, while for standing waves we must solve for ω as a function of $\mathbf{k}$ with $\mathbf{k}$ real. In each case the solution will also depend upon ϵ^2, so we can write the solutions as $\mathbf{k}(\omega,\epsilon^2)$ and $\omega(\mathbf{k},\epsilon^2)$. When $\epsilon = 0$, (2.5) becomes the dispersion equation for the perfect crystal, which is

$$\det[\omega^2 mI - \tilde{\Phi}(\mathbf{k})] = 0. \tag{2.6}$$

Let us denote the solutions of (2.6) by $\mathbf{k}_0(\omega)$ and $\omega_0(\mathbf{k})$. Then because of the random defects these solutions will be modified, and we can write the modified solutions in the forms

$$\mathbf{k}(\omega,\epsilon^2) = \mathbf{k}_0(\omega)[1 + \epsilon^2 k_2(\omega) + O(\epsilon^4)], \tag{2.7}$$

$$\omega(\mathbf{k},\epsilon^2) = \omega_0(\mathbf{k}) + \epsilon^2\omega_2(\mathbf{k}) + O(\epsilon^4). \tag{2.8}$$

In (2.7) we have chosen the direction of $\mathbf{k}$ to be the same as that of $\mathbf{k}_0$, so that k_2 is a scalar.

In order to find k_2 we substitute (2.7) into (2.5), equate to zero the coefficient of ϵ^2, and solve the resulting equation. To find ω_2, we use (2.8) in (2.5) and proceed similarly. The results can be written in terms of derivatives of determinants or in terms of cofactors. Upon using them in (2.7) and (2.8) we can write the solutions in the following two forms:

$$\begin{aligned}\mathbf{k}(\omega,\epsilon^2) &= \mathbf{k}_0(\omega) - \epsilon^2\mathbf{k}_0(\omega)[\partial_{\epsilon^2}\det[\omega^2 mI - \tilde{\Phi}(\mathbf{k}_0)] \\ &\quad + \epsilon^2\omega^4 m^2 GR(\mathbf{k}_0,\omega)]_{\epsilon^2=0} \\ &\quad \times [\mathbf{k}_0\cdot\partial_{\mathbf{k}_0}\det[\omega^2 m^2 I - \tilde{\Phi}(\mathbf{k}_0)]]^{-1} + O(\epsilon^4) \\ &= \mathbf{k}_0(\omega) + \epsilon^2\mathbf{k}_0(\omega)\omega^4 m^2 \sum_{ij} [\widetilde{GR}(\mathbf{k}_0,\omega)]_{ij} \\ &\quad \times \operatorname{cof}[\omega^2 mI - \tilde{\Phi}(\mathbf{k}_0)]_{ij}\{\sum_{rn} [\mathbf{k}_0\cdot\partial_{\mathbf{k}_0}\tilde{\Phi}_{rn}(k_0) \\ &\quad \times \operatorname{cof}[\omega^2 mI - \tilde{\Phi}(\mathbf{k}_0)]_{rn}\}^{-1} + O(\epsilon^4),\end{aligned} \tag{2.9}$$

$$\begin{aligned}\omega(\mathbf{k},\epsilon^2) &= \omega_0(\mathbf{k}) - \epsilon^2[\partial_{\epsilon^2}\det[\omega_0^2 mI \\ &\quad - \tilde{\Phi}(\mathbf{k}) + \epsilon^2\omega_0^4 m^2\widetilde{GR}(\mathbf{k},\omega_0)]]_{\epsilon^2=0} \\ &\quad \times [\partial_{\omega_0}\det[\omega_0^2 mI - \tilde{\Phi}(\mathbf{k})]]^{-1} + O(\epsilon^4) \\ &= \omega_0(\mathbf{k}) - \tfrac{1}{2}\epsilon^2\omega_0^3 m\sum_{ij}[\widetilde{GR}(\mathbf{k},\omega_0)]_{ij}\operatorname{cof}[\omega_0^2 mI \\ &\quad - \tilde{\Phi}(k)]_{ij}\{\sum_n \operatorname{cof}[\omega_0^2 mI - \tilde{\Phi}(k)]_{nn}\}^{-1} + O(\epsilon^4).\end{aligned} \tag{2.10}$$

The ϵ^2 term in (2.9) is generally complex when ω is real. The real part of the correction represents a change in the wavelength of the wave and, therefore, also a change in its phase velocity. The imaginary part represents an attenuation of the coherent wave due to scattering of its energy into the incoherent wave, which gradually grows as the coherent wave decays. Similarly the real part of the ϵ^2 term in (2.10) represents a frequency shift due to scattering by the impurities and the imaginary part represents a corresponding damping of the coherent vibration.

The dispersion equation (2.5) and the results (2.9) and (2.10) can be simplified in the one-dimensional case. Then the matrix in (2.5) is a scalar so that (2.5) becomes

$$\omega^2 m - \tilde{\Phi}(k) + \epsilon^2\omega^4 m^2\widetilde{GR}(k,\omega) = 0. \tag{2.11}$$

The results (2.9) and (2.10) are now simply

$$k(\omega,\epsilon^2) = k_0(\omega) + \epsilon^2\omega^4 m^2\widetilde{GR}(k_0,\omega)\left(\frac{\partial\tilde{\Phi}(k_0)}{\partial k_0}\right)^{-1} + O(\epsilon^4), \tag{2.12}$$

$$\omega(k,\epsilon^2) = \pm\{[m^{-1}\tilde{\Phi}(k)]^{1/2} - (\epsilon^2/2m^{1/2})[\tilde{\Phi}(k)]^{3/2} \times\widetilde{GR}(k,\omega_0) + O(\epsilon^4)\}. \tag{2.13}$$

Here $\omega_0(k) = m^{-1/2}[\tilde{\Phi}(k)]^{1/2}$, and $k_0(\omega)$ is the solution of $\tilde{\Phi}(k_0) = \omega^2 m$. These results will be made more explicit in Sec. 4, where a crystal with nearest neigh-

bor interactions and two special forms of the correlation function R will be considered.

In addition to the above solutions of the dispersion equation, which are perturbations of the solutions for the perfect crystal, there are generally other solutions. They represent new waves or modes which occur only when imperfections are present in the crystal. We shall see an example of such solutions in Sec. 4.

3. UNCORRELATED DEFECTS

An important special case of an imperfect crystal is that in which the defects are uncorrelated. In that case the correlation function R is given by

$$R(q,q') = \langle\mu^2\rangle\delta_{qq'}. \tag{3.1}$$

Then if we note that $G(q,q,\omega)$ is independent of q, (2.4) yields

$$\widetilde{GR}(\mathbf{k},\omega) = \langle\mu^2\rangle G(0,0,\omega). \tag{3.2}$$

For a simple cubic crystal with suitable symmetry in n dimensions, the matrix $G(0,0,\omega)$ is a scalar multiple of the nth order unit matrix. This is also so for certain other crystals. Therefore we now assume that

$$G(0,0,\omega) = g(\omega)I,$$

where $g(\omega)$ is a scalar. Then the dispersion equation (2.5) becomes

$$\det\{[\omega^2 m + \epsilon^2\omega^4 m^2\langle\mu^2\rangle g(\omega)]I - \tilde{\Phi}(k)\} = 0. \tag{3.3}$$

This dispersion equation (3.3) is the same as Eq. (2.6) for a perfect crystal with m replaced by an effective mass $m_e(\omega)$ given by

$$m_e(\omega) = m + \epsilon^2\omega^2 m^2\langle\mu^2\rangle g(\omega). \tag{3.4}$$

Alternatively we may say that ω is replaced by $[\omega^2 + \epsilon^2\omega^4 m\langle\mu^2\rangle g(\omega)]^{1/2}$. Thus the solution of (3.3) for $\mathbf{k}(\omega,\epsilon^2)$ can be written as

$$\mathbf{k}(\omega,\epsilon^2) = \mathbf{k}_0\{[\omega^2 + \epsilon^2\omega^4 m\langle\mu^2\rangle g(\omega)]^{1/2}\}. \tag{3.5}$$

This result shows that no new modes of propagation occur in this case. Instead the solution for $\mathbf{k}$ in the presence of uncorrelated impurities is just that for a perfect crystal at a shifted and, generally, complex frequency. To expand (3.5) we set $\mathbf{k}_0(\omega) = \hat{\mathbf{k}}_0 k_0(\omega)$, where $\hat{\mathbf{k}}_0$ is a unit vector independent of ω and $k_0(\omega)$ is a scalar. Then (3.5) yields

$$\mathbf{k}(\omega,\epsilon^2) = \hat{\mathbf{k}}_0\left(k_0(\omega) + \frac{\epsilon^2}{2}\omega^3 m\langle\mu^2\rangle g(\omega)\frac{\partial k_0}{\partial\omega} + O(\epsilon^4)\right). \tag{3.6}$$

We note that the correction to $k_0(\omega)$ is proportional to the reciprocal of the group velocity $\partial\omega/\partial k$ for the perfect crystal.

In order to solve (3.3) for $\omega(\mathbf{k},\epsilon^2)$, we first use the solution $\omega_0(\mathbf{k})$ for the perfect crystal to write

$$\omega^2 + \epsilon^2\omega^4 m\langle\mu^2\rangle g(\omega) = \omega_0^2(\mathbf{k}). \tag{3.7}$$

Then by expanding $\omega(\mathbf{k},\epsilon^2)$ about $\omega_0(\mathbf{k})$ we obtain

$$\omega(\mathbf{k},\epsilon^2) = \omega_0(\mathbf{k}) - \epsilon^2 m\langle\mu^2\rangle\omega_0^3 g(\omega_0) + O(\epsilon^4). \tag{3.8}$$

In addition to the solution (3.8), (3.7) may have other solutions.

4. PROPAGATION CONSTANT FOR A ONE-DIMENSIONAL CRYSTAL

To illustrate the preceding results we shall apply them first to a one-dimensional crystal, i.e., to a linear chain with the qth particle at $x(q) = qd$, $q = 0, \pm 1, \cdots$. Let us assume that only nearest neighbors interact, with spring constant $\alpha \geq 0$. It follows that Φ is given by

$$\begin{aligned}\Phi(q,q') &= -\alpha \quad \text{if} \quad q' = q\pm 1,\\ &= +2\alpha \quad \text{if} \quad q' = q,\\ &= 0 \quad \text{otherwise.}\end{aligned} \tag{4.1}$$

Then (1.14) yields for L_0 the result

$$L_0 = -\alpha\Delta^2 - \omega^2 m^2. \tag{4.2}$$

Here Δ^2 is the second central difference operator defined by

$$\Delta^2 u(q) = u(q+1) - 2u(q) + u(q-1). \tag{4.3}$$

When $\varphi(q,q') = 0$, (1.15) yields $V = -\epsilon m_0\omega^2\mu(q)$.

The Fourier transform (2.2) becomes, when (4.1) is used in it,

$$\tilde{\Phi}(k) = -\alpha(e^{-ikd} - 2 + e^{ikd}) = 2\alpha(1 - \cos kd). \tag{4.4}$$

The Green's function of L_0 given by (4.2) is easily seen to be

$$G(|q-q'|,\omega) = \frac{-e^{ik_0|q-q'|d}}{2i\alpha\sin k_0 d}, \qquad \omega \neq 0, (4\alpha/m)^{1/2}. \tag{4.5}$$

Here $k_0(\omega)$ is the solution of the dispersion equation (2.6) for the perfect crystal, which becomes, when (4.4) is used,

$$\cos k_0 d = 1 - 2(\omega/\omega_c)^2, \qquad \omega_c = (4\alpha/m)^{1/2} \tag{4.6}$$

We see that the principal solution of (4.6) yields $k_0(\omega) = 0$ at $\omega = 0$, and that k_0 remains real and increases to the value π/d at the cut-off frequency $|\omega| = \omega_c = (4\alpha/m)^{1/2}$. For $|\omega| > \omega_c$, k_0 is imaginary. Thus in the perfect crystal with ω real, waves propagate for $0 < |\omega| < \omega_c$ and are evanescent or nonpropagating for $|\omega| > \omega_c$. The correlation function R defined by (2.3) is a function of $|q-q'|$ in the present case, $R = R(|q-q'|)$. Thus (2.4) becomes

$$\widetilde{GR}(k,\omega) = -(2i\alpha\sin k_0 d)^{-1}\sum_q R(|q|)\,e^{i(k_0|q| + kq)d}. \tag{4.7}$$

Now we use (4.7) and (4.4) in (2.11) to obtain the dispersion equation

$$\cos kd = 1 - 2(\omega/\omega_c)^2 - 8\epsilon^2\alpha(\omega/\omega_c)^4\widetilde{GR}(k,\omega). \tag{4.8}$$

By solving (4.8) for k or for ω, or by using (4.7) and (4.4) in (2.12) and (2.13), we obtain the results

$$k(\omega,\epsilon^2) = k_0 + \frac{i\epsilon^2}{d[(\omega_c/\omega)^2-1]}\sum_q R(|q|)e^{ik_0d(|q|+q)} + O(\epsilon^4), \quad (4.9)$$

$$\omega(k,\epsilon^2) = \pm\omega_c\left[\tfrac{1}{2}(1-\cos kd)^{1/2} - (i\epsilon^2/\sin kd) \times \tfrac{1}{2}(1-\cos kd)^{3/2}\sum_q R(|q|)e^{ikd(|q|+q)} + O(\epsilon^4)\right]. \quad (4.10)$$

Taking the real and imaginary parts of (4.9) and (4.10) yields

$$\mathrm{Re}k(\omega,\epsilon^2) = \mathrm{Re}k_0(\omega) - \frac{\epsilon^2}{d}\left(\frac{\omega_c^2}{\omega^2}-1\right)^{-1} \times \sum_{q=1}^{\infty} R(q)\sin 2qk_0d + O(\epsilon^4), \quad (4.11)$$

$$\mathrm{Im}k(\omega,\epsilon^2) = \mathrm{Im}R_0(\omega) + \frac{\epsilon^2}{d}\left(\frac{\omega_c^2}{\omega^2}-1\right)^{-1} \times\left(R(0) + \sum_{q=1}^{\infty} R(q)(1+\cos 2qk_0d)\right) + O(\epsilon^4), \quad (4.12)$$

$$\mathrm{Re}\omega(k,\epsilon^2) = \pm\omega_c\left[\left(\frac{1-\cos kd}{2}\right)^{1/2} + \frac{\epsilon^2}{\sin kd} \times\left(\frac{1-\cos kd}{2}\right)^{3/2}\sum_{q=1}^{\infty} R(q)\sin 2qkd + O(\epsilon^4)\right], \quad (4.13)$$

$$\mathrm{Im}\omega(k,\epsilon^2) = \pm\omega_c\left[-\frac{\epsilon^2}{\sin kd}\left(\frac{1-\cos kd}{2}\right)^{3/2} \times\left(R(0) + \sum_{q=1}^{\infty} R(q)(1+\cos 2qkd)\right) + O(\epsilon^4)\right]. \quad (4.14)$$

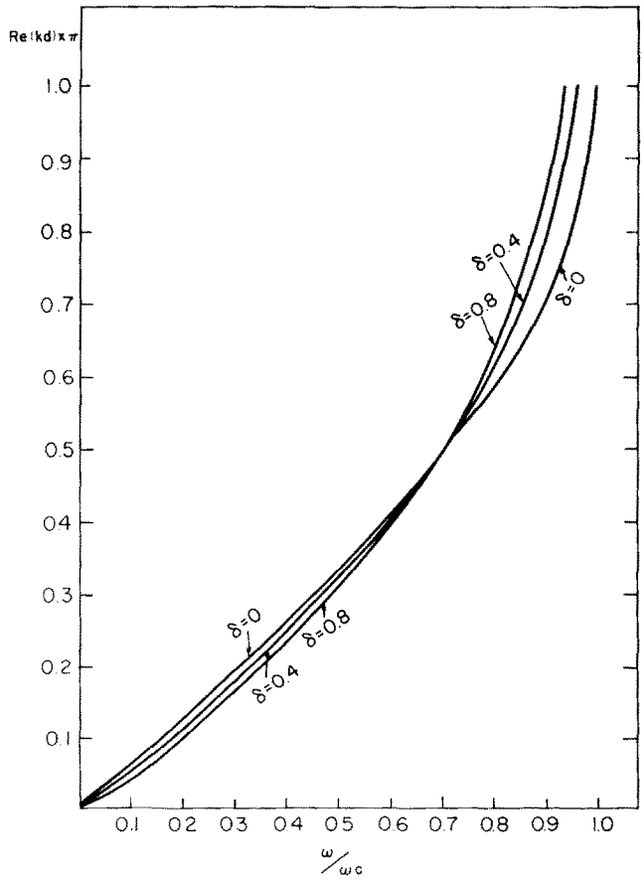


FIG. 1. Rekd is shown as a function of ω/ω_c for waves in a one-dimensional crystal with nearest neighbor interactions and uncorrelated mass defects for three values of $\delta = \epsilon^2\langle\mu^2\rangle$. The curves are based on the "exact" dispersion equation (4.15). The approximation (4.16) is independent of $\epsilon^2\langle\mu^2\rangle$ and yields only the perfect crystal curve for which $\epsilon = 0$.

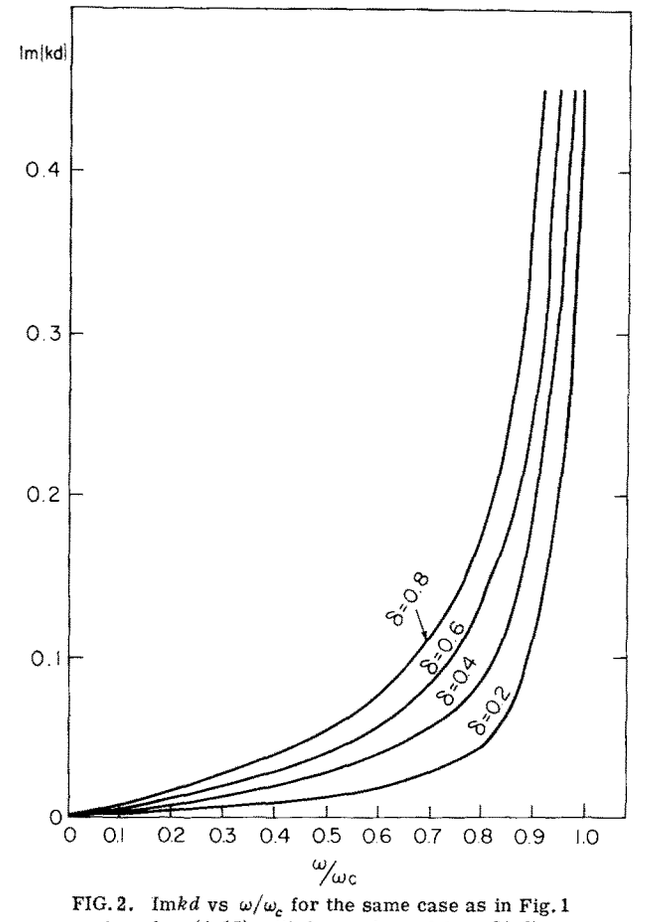


FIG. 2. Imkd vs ω/ω_c for the same case as in Fig. 1 based on (4.15), with four values of $\delta = \epsilon^2\langle\mu^2\rangle$.

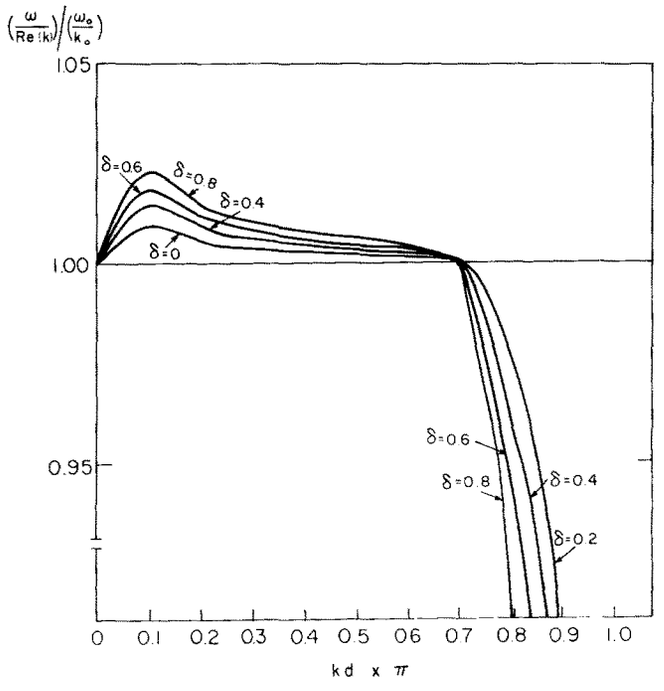


FIG. 3. The phase velocity $\omega/\mathrm{Re}k$, divided by the perfect crystal velocity $\omega/\mathrm{Re}k_0$, vs. ω/ω_c for the same cases as in Fig. 2 based on (4.15). Only velocities for frequencies below the cut-off frequency ω_c are shown.

The results (4.11) and (4.12) determine the phase velocity $\omega/\mathrm{Re}k$ and the attenuation coefficient $\mathrm{Im}k$ of the mean wave. For propagating waves the attenuation coefficient is positive because $\omega^2 < \omega_c^2$, and the expression in braces in (4.12) is also positive. This is so because it is the sum of $R(0)$ and the Fourier transform of the correlation function R evaluated at zero and at $2k_0d$, all of which are positive. Similarly (4.13) and (4.14) determine the frequency and decay rate of a standing wave.

As an example, let us consider uncorrelated mass defects. Then $R(q) = 0$ for $q \neq 0$, $R(0) = \langle\mu^2\rangle$ and the dispersion equation (4.8) becomes

$$\cos kd = 1 - 2\left(\frac{\omega}{\omega_c}\right)^2 - \frac{2i\epsilon^2\langle\mu^2\rangle(\omega/\omega_c)^3}{[1-(\omega/\omega_c)^2]^{1/2}}. \tag{4.15}$$

From (4.15), or from (4.11)–(4.14), or from (2.12) and (2.13), we obtain

$$\mathrm{Re}k(\omega,\epsilon^2) = \mathrm{Re}k_0(\omega) + O(\epsilon^4), \tag{4.16}$$

$$\mathrm{Im}k(\omega,\epsilon^2) = \mathrm{Im}k_0(\omega) + \frac{\epsilon^2\langle\mu^2\rangle}{d[(\omega_c/\omega)^2-1]} + O(\epsilon^4). \tag{4.17}$$

$$\mathrm{Re}\omega(k,\epsilon^2) = \pm\omega_c\,[\tfrac{1}{2}(1-\cos kd)^{1/2} + O(\epsilon^4)], \tag{4.18}$$

$$\mathrm{Im}\omega(k,\epsilon^2) = \pm\omega_c\,[-(\epsilon^2\langle\mu^2\rangle/\sin kd)\,\tfrac{1}{2}(1-\cos kd)^{3/2} + O(\epsilon^4)]. \tag{4.19}$$

Upon rationalizing (4.15), it becomes a cubic equation in $(\omega/\omega_c)^2$. Therefore in addition to the solutions $\pm\omega(k,\epsilon^2)$ given by (4.13) and (4.14), there are two other pairs of solutions. One of them, which corresponds to the vanishing of the denominator of the ϵ^2 term, is readily found to be

$$\omega(k,\epsilon^2) = \pm\left(\omega_c + \frac{2\epsilon^4\langle\mu^2\rangle^2\omega_c}{(\cos kd-1)^2} + O(\epsilon^6)\right). \tag{4.20}$$

The other is spurious, since it does not satisfy the unrationalized equation. We note that the solution (4.20) is real to the order shown.

Graphs of $\mathrm{Re}kd$ and $\mathrm{Im}kd$ as functions of ω/ω_c for several values of $\epsilon^2\langle\mu^2\rangle$ are shown in Figs. 1 and 2. Figure 3 shows the corresponding phase velocity $\omega/\mathrm{Re}k$. Figures 4 and 5 show $\mathrm{Re}\omega/\omega_c$ and $\mathrm{Im}\omega/\omega_c$ as functions of kd based on (4.15) and also on (4.18)–(4.20).

As a second example let us treat mass defects with the exponential correlation function

$$R(q) = \langle\mu^2\rangle e^{-|q|d/a}. \tag{4.21}$$

Here a is the correlation length of the random defects. Now (4.7) can be summed to yield

$$\widetilde{GR}(k,\omega) = \frac{-\langle\mu^2\rangle}{2i\alpha\,\sin k_0d}\,[(1-e^{-d/a+i(k_0+k)d})^{-1} + (e^{d/a+i(k_0+k)d}-1)^{-1}] = -\frac{\langle\mu^2\rangle\,\sin Kd}{2\alpha\,\sin k_0d\,(\cos kd-\cos Kd)}, \tag{4.22}$$

where K is defined by

$$K(\omega) = k_0(\omega) + (i/a). \tag{4.23}$$

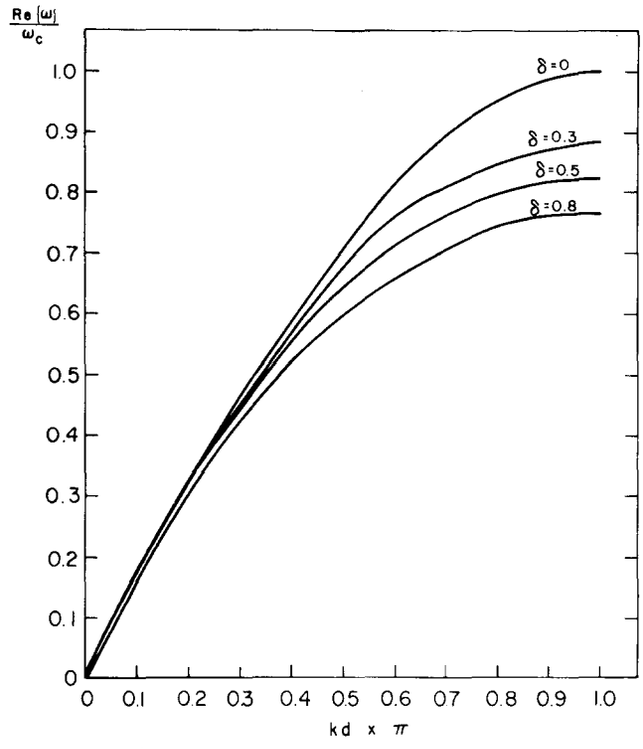


FIG. 4. $\mathrm{Re}\omega/\omega_c$ vs. kd for the same case as in Fig. 1 based on (4.15), with four values of $\delta = \epsilon^2\langle\mu^2\rangle$. The approximate result (4.18) yields only the perfect crystal curve for which $\epsilon = 0$.

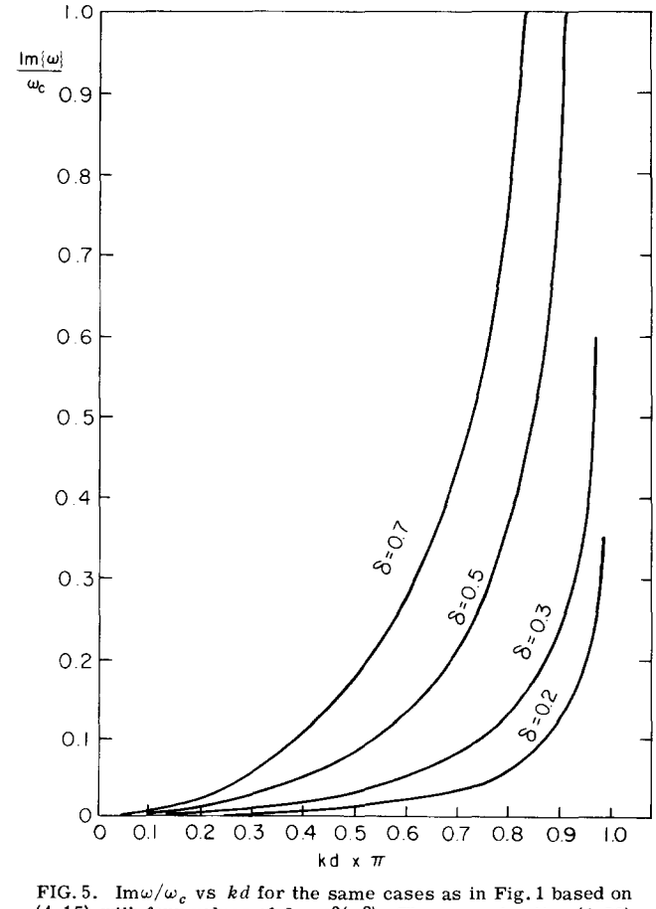


FIG. 5. $\mathrm{Im}\omega/\omega_c$ vs kd for the same cases as in Fig. 1 based on (4.15), with four values of $\delta = \epsilon^2\langle\mu^2\rangle$. The approximation (4.20) yields only the result $\mathrm{Im}\omega = 0$.

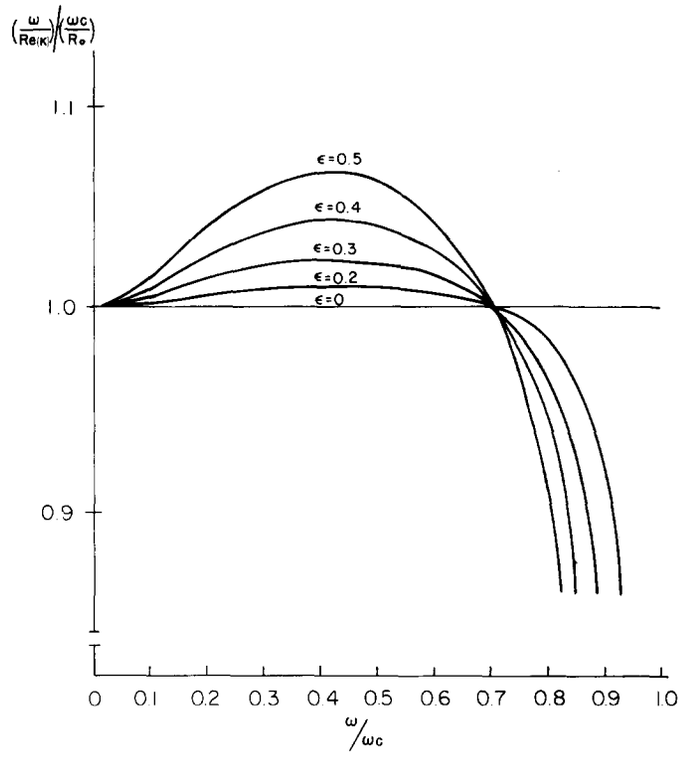


FIG. 6. The phase velocity $\omega/\mathrm{Re}k$ divided by the perfect crystal value $\omega/\mathrm{Re}k_0$ vs. ω/ω_c, for waves in a one-dimensional crystal with nearest neighbor interactions and exponentially correlated mass defects for various values of ϵ when $\langle\mu^2\rangle = 1$. The curves are based on the approximation (4.27) with $a = d$.

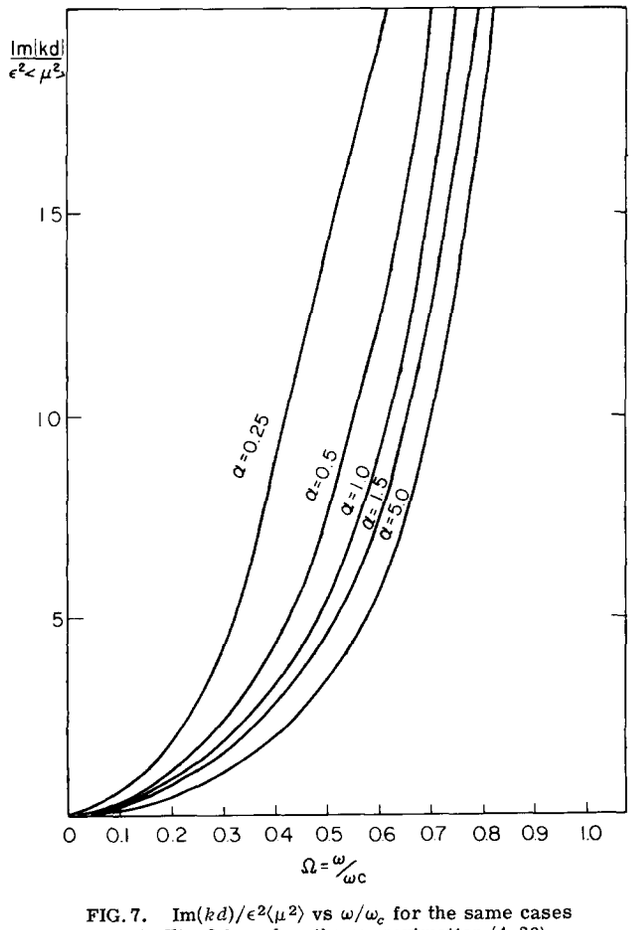


FIG. 7. $\mathrm{Im}(kd)/\epsilon^2\langle\mu^2\rangle$ vs ω/ω_c for the same cases as in Fig. 6, based on the approximation (4.28).

Then the dispersion equation (4.8) becomes

$$(\cos kd - \cos k_0 d)(\cos kd - \cos Kd) = 4\epsilon^2\langle\mu^2\rangle \times (\omega/\omega_c)^4 (\sin Kd/\sin k_0 d). \quad (4.24)$$

When $\epsilon = 0$, (4.24) has the two solutions $k = k_0(\omega)$ and $k = K(\omega)$. Then perturbation analysis yields for ϵ small the following two solutions:

$$k(\omega, \epsilon^2) = k_0(\omega) - \frac{4\epsilon^2\langle\mu^2\rangle (\omega/\omega_c)^4 \sin Kd}{d \sin^2 k_0 d (\cos k_0 d - \cos Kd)} + O(\epsilon^4), \quad (4.25)$$

$$k(\omega, \epsilon^2) = k_0(\omega) - \frac{i}{a} - \frac{4\epsilon^2\langle\mu^2\rangle (\omega/\omega_c)^4}{d \sin k_0 d (\cos Kd - \cos k_0 d)} + O(\epsilon^4). \quad (4.26)$$

Solution (4.25) represents the slightly attenuated modification of the mode of the perfect crystal, while (4.26) represents an additional strongly attenuated mode.

The real and imaginary parts of k given by (4.25) are

$$\mathrm{Re}k = k_0 - \frac{\epsilon^2\langle\mu^2\rangle}{d}\left(\frac{\omega_c^2}{\omega^2} - 1\right)^{-1} \times \frac{e^{-d/a} \sin 2k_0 d}{1 - 2e^{-d/a}\cos 2k_0 d + e^{-2d/a}} O(\epsilon^4), \quad (4.27)$$

$$\mathrm{Im}k = \frac{\epsilon^2\langle\mu^2\rangle}{d}\left(\frac{\omega_c^2}{\omega^2} - 1\right)^{-1}\left(\frac{e^{-d/a}}{1 - e^{-d/a}} + \frac{1 - e^{-d/a}\cos 2k_0 d}{1 - 2e^{-d/a}\cos 2k_0 d + e^{-2d/a}}\right) + O(\epsilon^4). \quad (4.28)$$

We see that $\mathrm{Im}k > 0$, and that $\mathrm{Re}k < k_0$ if $\sin 2k_0 d > 0$ and $\mathrm{Re}k > k_0$ if $\sin 2k_0 d < 0$. Thus the phase velocity $\omega/\mathrm{Re}k$ of the mean wave exceeds that of the unperturbed wave if $0 < k_0 < \pi/2d$ and is less than the unperturbed speed if $\pi/2d < k_0 < \pi/d$. Graphs of the phase velocity $\omega/\mathrm{Re}k$ and of $\mathrm{Im}k$ are shown in Figs. 6 and 7, based on (4.27) and (4.28). We note that $k_0 d = \pi/2$ at $\omega/\omega_c = 2^{-1/2} = .707$.

To solve (4.24) for $\omega(k, \epsilon^2)$, we use (4.4) and (4.22) in (2.13) to obtain

$$\omega(k, \epsilon^2) = \pm\omega_c\{\tfrac{1}{2}(1 - \cos kd)^{1/2} + (i\epsilon^2\langle\mu^2\rangle/\sin kd)^{3/2} \times [(1 - e^{-d/a + 2ikd})^{-1} + (e^{d/a} - 1)^{-1}] + O(\epsilon^4)\}. \quad (4.29)$$

In this case both $\mathrm{Re}\omega$ and $\mathrm{Im}\omega$ have corrections of order ϵ^2 to the solution for the perfect crystal. Additional solutions of (4.24) can also be obtained; but we shall not present them.

Our result (4.17) for uncorrelated defects can be compared with a result obtained by Rubin[8] and Matsuda and Ishii[9] for ω/ω_c small. When ω/ω_c is small, $\mathrm{Im}k_0(\omega) = 0$ and (4.17) yields $\mathrm{Im}k(\omega, \epsilon^2) \sim \epsilon^2\langle\mu^2\rangle d\omega_c^2$. Both Rubin and Matsuda and Ishii obtained one-half this result. The difference is evidently due to the different statistical properties of the solution determined by those authors and by us.

It is also of interest to compare our low frequency result with the result for a continuous elastic rod. In the one-dimensional case with nearest neighbor interaction, the equation of motion is

$$\alpha\Delta^2 u(q) + \omega^2 m[1 + \epsilon\mu(q)]u(q) = 0. \quad (4.30)$$

We now let the spacing d and the frequency ω tend to zero in a fixed ratio. Then we define $\hat{k}_0$, which is finite, by

$$\hat{k}_0 = (\omega/d)(m/\alpha)^{1/2}. \tag{4.31}$$

We also set $x = qd$ and then (4.30) becomes

$$\frac{d^2u(x)}{dx^2} + \hat{k}_0^2[1 + \epsilon\mu(x)]u(x) = 0. \tag{4.32}$$

This equation, with $\mu(x)$ a random function, has been analyzed by Kupiec et al.[10] When the correlation function of $\mu(x)$ is a delta function or an exponential, the results (22), (36), and (37) of Ref. 10 agree, respectively, with our results (4.15), (4.25), and (4.26) for ω small.

APPENDIX: NONSIMPLE LATTICES

Let us consider a lattice with p particles per cell. We shall label each cell by a vector $\mathbf{n} = (n_1, n_2, n_3)$ with integer components, and label each particle in a given cell by an index $r = 1, \ldots, p$. Thus $q = (\mathbf{n}, r)$ designates a particular particle in a particular cell. Let $\mathbf{a}_1, \mathbf{a}_2$ and $\mathbf{a}_3$ be the three primitive translation vectors of the lattice. Then $\mathbf{x}(\mathbf{n}) = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$ denotes a fixed point in cell $\mathbf{n}$ and $\mathbf{x}(\mathbf{n}, r) = \mathbf{x}(\mathbf{n}) + \boldsymbol{\xi}(r)$ denotes the equilibrium position of particle $\mathbf{n}, r$. The displacement $\mathbf{u}(q)e^{-i\omega t}$ satisfies (1.1), which can be written in the form (1.2) with L given by (1.3), and $\langle\mu\rangle$ satisfies (1.4).

A plane wave is a solution of the form

$$\langle\mathbf{u}(\mathbf{n}, r)\rangle = \mathbf{A}(r)e^{i\mathbf{k}\cdot\mathbf{x}(\mathbf{n})}, \quad r = 1, \ldots, p. \tag{A1}$$

By using (A1) in (1.4) with $\mathbf{f} = 0$, we obtain

$$e^{-i\mathbf{k}\cdot\mathbf{x}(\mathbf{n})}\langle L^{-1}\rangle^{-1}e^{i\mathbf{k}\cdot\mathbf{x}(\mathbf{n})}\mathbf{A}(r) = 0, \quad r = 1, \ldots, p. \tag{A2}$$

These are p vector equations for the p vectors $\mathbf{A}(r)$, and their determinant must vanish for the solution to be nontrivial. This determinant is of order $3p$ in three dimension, $2p$ in two dimensions, and p in one dimension. Its vanishing yields the dispersion equation for $\mathbf{k}$.

* This work was supported in part by NSF-GP-27209.

1 J. B. Keller, Proc. Symp. Appl. Math. **13**, 227 (1962).
2 R. C. Bourret, Nuovo Cimento **26**, 1 (1962).
3 V. I. Tatarskii and M. E. Gercenstein, Sov. Phys. JETP, Transl. no. 19, 946 (1963).
4 P. L. Chow and J. B. Keller, "Wave Propagation in a Random Lattice. II," to be published.
5 G. F. Koster, Phys. Rev. **95**, 1436 (1954).
6 I. M. Lifshitz, Nuovo Cimento Suppl. 10, **3**, 716 (1956).
7 R. J. Rubin, J. Math. Phys. **9**, 2252 (1968).
8 R. J. Rubin, J. Math. Phys. **11**, 1857 (1970).
9 H. Matsuda and K. Ishii, Progr. Theoret. Phys. (Supplement) **45**, 56 (1970).
10 I. Kupiec, L. B. Felsen, S. Rosenbaum, J. B. Keller, and P. Chow, Radio Sci. **4**, 1067 (1969).

Lattice Green's Functions for the Triangular and Honeycomb Lattices

Tsuyoshi Horiguchi
Department of Physics, Ohio University, Athens, Ohio 45701
(Received 10 April 1972)

The lattice Green's function for the triangular lattice at an arbitrary lattice site is expressed in terms of the complete elliptic integrals of the first and second kind. The lattice Green's function for the honeycomb lattice is shown to be expressed in terms of the one for the triangular lattice. The results obtained are shown by graphs.

1. INTRODUCTION

The triangular lattice is the two-dimensional closed packed lattice. The honeycomb lattice is composed of two triangular sublattices. The lattice sites on one of the sublattices, say A in Fig. 5, have bonds extending vertically downward, while the sites on the other sublattice B have bonds extending vertically upward. The lattice sites of one type cannot be found from those of the other type by a simple translation, which would be required if they were in equivalent position in the lattice. These situations resemble the relation of the fcc and diamond structure lattices. The density of state for the diamond structure lattice was recently calculated by Thorpe and Weaire[1] in connection with the problem of electronic properties of an amorphous solid of the Weaire model.[2] In that calculation, the one-band Hamiltonian of the Weaire model is used which is equivalent to the one of the electron in the solid on the tight-binding approximation. In order to see whether the structure of lattice reflects in its lattice Green's function or not, we investigate the behavior of the functions for the triangular and honeycomb lattices and discuss the similarity between the functions for the honeycomb and diamond structure lattices.

The theoretical studies of the lattice vibrations of graphite need the lattice Green's function for the lattice of the graphite structure. In that structure, the distance between layers is much larger than the distance between the nearest atoms within the same layer, and the forces between atoms within the same layer are usually much stronger than those between atoms in different layers.[3] If only the nearest neighbor interaction in the same layer is assumed, the lattice Green's function for the honeycomb lattice is needed for the studies on the lattice vibration of the graphite.[4]

In the present note, we investigate the lattice Green's functions for the triangular and honeycomb lattices. It is shown in Sec. 2 that the lattice Green's function for the triangular lattice is expressed in terms of the complete elliptic integrals of the first and second kind with a modulus of complex number, and in Sec. 3

We now let the spacing d and the frequency ω tend to zero in a fixed ratio. Then we define $\hat{k}_0$, which is finite, by

$$\hat{k}_0 = (\omega/d)(m/\alpha)^{1/2}. \tag{4.31}$$

We also set $x = qd$ and then (4.30) becomes

$$\frac{d^2u(x)}{dx^2} + \hat{k}_0^2[1 + \epsilon\mu(x)]u(x) = 0. \tag{4.32}$$

This equation, with $\mu(x)$ a random function, has been analyzed by Kupiec et al.[10] When the correlation function of $\mu(x)$ is a delta function or an exponential, the results (22), (36), and (37) of Ref. 10 agree, respectively, with our results (4.15), (4.25), and (4.26) for ω small.

APPENDIX: NONSIMPLE LATTICES

Let us consider a lattice with p particles per cell. We shall label each cell by a vector $\mathbf{n} = (n_1, n_2, n_3)$ with integer components, and label each particle in a given cell by an index $r = 1, \ldots, p$. Thus $q = (\mathbf{n}, r)$ designates a particular particle in a particular cell. Let $\mathbf{a}_1, \mathbf{a}_2$ and $\mathbf{a}_3$ be the three primitive translation vectors of the lattice. Then $\mathbf{x}(\mathbf{n}) = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$ denotes a fixed point in cell $\mathbf{n}$ and $\mathbf{x}(\mathbf{n}, r) = \mathbf{x}(\mathbf{n}) + \boldsymbol{\xi}(r)$ denotes the equilibrium position of particle $\mathbf{n}, r$. The displacement $\mathbf{u}(q)e^{-i\omega t}$ satisfies (1.1), which can be written in the form (1.2) with L given by (1.3), and $\langle\mu\rangle$ satisfies (1.4).

A plane wave is a solution of the form

$$\langle\mathbf{u}(\mathbf{n}, r)\rangle = \mathbf{A}(r)e^{i\mathbf{k}\cdot\mathbf{x}(\mathbf{n})}, \quad r = 1, \ldots, p. \tag{A1}$$

By using (A1) in (1.4) with $\mathbf{f} = 0$, we obtain

$$e^{-i\mathbf{k}\cdot\mathbf{x}(\mathbf{n})}\langle L^{-1}\rangle^{-1}e^{i\mathbf{k}\cdot\mathbf{x}(\mathbf{n})}\mathbf{A}(r) = 0, \quad r = 1, \ldots, p. \tag{A2}$$

These are p vector equations for the p vectors $\mathbf{A}(r)$, and their determinant must vanish for the solution to be nontrivial. This determinant is of order $3p$ in three dimension, $2p$ in two dimensions, and p in one dimension. Its vanishing yields the dispersion equation for $\mathbf{k}$.

* This work was supported in part by NSF-GP-27209.

1 J. B. Keller, Proc. Symp. Appl. Math. **13**, 227 (1962).
2 R. C. Bourret, Nuovo Cimento **26**, 1 (1962).
3 V. I. Tatarskii and M. E. Gercenstein, Sov. Phys. JETP, Transl. no. 19, 946 (1963).
4 P. L. Chow and J. B. Keller, "Wave Propagation in a Random Lattice. II," to be published.
5 G. F. Koster, Phys. Rev. **95**, 1436 (1954).
6 I. M. Lifshitz, Nuovo Cimento Suppl. 10, **3**, 716 (1956).
7 R. J. Rubin, J. Math. Phys. **9**, 2252 (1968).
8 R. J. Rubin, J. Math. Phys. **11**, 1857 (1970).
9 H. Matsuda and K. Ishii, Progr. Theoret. Phys. (Supplement) **45**, 56 (1970).
10 I. Kupiec, L. B. Felsen, S. Rosenbaum, J. B. Keller, and P. Chow, Radio Sci. **4**, 1067 (1969).

Lattice Green's Functions for the Triangular and Honeycomb Lattices

Tsuyoshi Horiguchi
Department of Physics, Ohio University, Athens, Ohio 45701
(Received 10 April 1972)

The lattice Green's function for the triangular lattice at an arbitrary lattice site is expressed in terms of the complete elliptic integrals of the first and second kind. The lattice Green's function for the honeycomb lattice is shown to be expressed in terms of the one for the triangular lattice. The results obtained are shown by graphs.

1. INTRODUCTION

The triangular lattice is the two-dimensional closed packed lattice. The honeycomb lattice is composed of two triangular sublattices. The lattice sites on one of the sublattices, say A in Fig. 5, have bonds extending vertically downward, while the sites on the other sublattice B have bonds extending vertically upward. The lattice sites of one type cannot be found from those of the other type by a simple translation, which would be required if they were in equivalent position in the lattice. These situations resemble the relation of the fcc and diamond structure lattices. The density of state for the diamond structure lattice was recently calculated by Thorpe and Weaire[1] in connection with the problem of electronic properties of an amorphous solid of the Weaire model.[2] In that calculation, the one-band Hamiltonian of the Weaire model is used which is equivalent to the one of the electron in the solid on the tight-binding approximation. In order to see whether the structure of lattice reflects in its lattice Green's function or not, we investigate the behavior of the functions for the triangular and honeycomb lattices and discuss the similarity between the functions for the honeycomb and diamond structure lattices.

The theoretical studies of the lattice vibrations of graphite need the lattice Green's function for the lattice of the graphite structure. In that structure, the distance between layers is much larger than the distance between the nearest atoms within the same layer, and the forces between atoms within the same layer are usually much stronger than those between atoms in different layers.[3] If only the nearest neighbor interaction in the same layer is assumed, the lattice Green's function for the honeycomb lattice is needed for the studies on the lattice vibration of the graphite.[4]

In the present note, we investigate the lattice Green's functions for the triangular and honeycomb lattices. It is shown in Sec. 2 that the lattice Green's function for the triangular lattice is expressed in terms of the complete elliptic integrals of the first and second kind with a modulus of complex number, and in Sec. 3

that the lattice Green's function for the honeycomb lattice is calculated with the aid of the one for the triangular lattice. The analytic properties become clear for these functions. The discussion is given in Sec. 4 for the singularities of the function and for the behavior of the function for the honeycomb and diamond structure lattices.

2. TRIANGULAR LATTICE

The lattice Green's function for the triangular lattice with the nearest neighbor interaction is the solution of the following difference equation which involves the δ-function type inhomogeneous term

$$\begin{aligned} 2tG(la, mb) &- G(la + 2a, ma) - G(la - 2a, mb) \\ &- G(la + a, mb + b) - G(la + a, mb - b) \\ &- G(la - a, mb + b) - G(la - a, mb - b) \\ &= 2\delta_{l,0}\delta_{m,0}, \end{aligned} \tag{2.1a}$$

where $l + m$ is an even integer, and a and b are equal to $\frac{1}{2}$ and $\frac{1}{2}\sqrt{3}$ times the length of the edge of the triangles: $a = \frac{1}{2}$ and $b = \frac{1}{2}\sqrt{3}$ if the nearest neighbor distance is chosen to be equal to 1. The boundary value of the function is required to be equal to zero as $l^2 + m^2$ tends to infinity. The solution of this equation under this boundary condition is given as follows:

$$G(la, mb) = \frac{ab}{(2\pi)^2} \int_{-\pi/a}^{\pi/a} dx \int_{-\pi/b}^{\pi/b} dy \times \frac{e^{i(lax+mby)}}{t - \cos 2ax - 2\cos ax \cos by}. \tag{2.2a}$$

The function is unchanged under the rotation by an angle $\frac{1}{3}(n\pi)$, $n = 1, 2, 3, 4, 5$, of the coordinate axes around the origin,

$$G(la, mb) = G(la\cos\tfrac{1}{3}(n\pi) - mb\sin\tfrac{1}{3}(n\pi), \; la\sin\tfrac{1}{3}(n\pi) + mb\cos\tfrac{1}{3}(n\pi)), \tag{2.3}$$

and under the inversion on the la axis and mb axis,

$$G(la, mb) = G(la, -mb) = G(-la, mb); \tag{2.4}$$

cf. Fig. 1. Using Eq. (2.3) for $n = 5$ and Eq. (2.4) for the inversion on the mb axis, one obtains the following equation by which the function $G(la, mb)$ for $mb > 3la$ is expressed in terms of the one for $mb < 3la$:

$$G(la, mb) = G(\tfrac{1}{2}la + \tfrac{1}{2}\sqrt{3}\, mb, \tfrac{1}{2}\sqrt{3}\, la - \tfrac{1}{2}mb). \tag{2.5a}$$

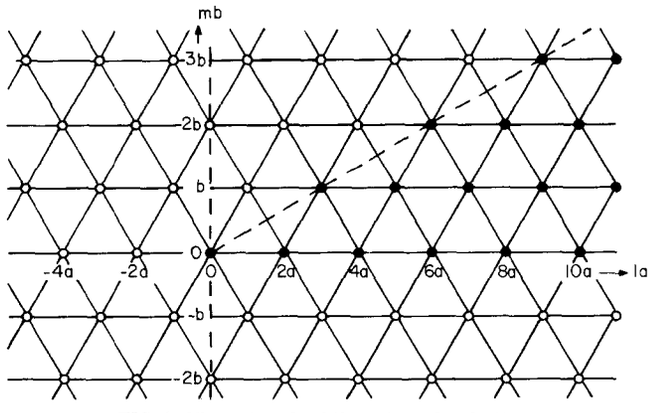


FIG. 1. The network of the triangular lattice.

Thus it is sufficient to calculate only the values at lattice sites shown by a black point in Fig. 1.

Equations (2.1a) and (2.2a) are simplified to

$$\begin{aligned} 2tG(l, m) &- G(l + 2, m) - G(l - 2, m) \\ &- G(l + 1, m + 1) - G(l + 1, m - 1) \\ &- G(l - 1, m + 1) - G(l - 1, m - 1) \\ &= 2\delta_{l,0}\delta_{m,0}, \end{aligned} \tag{2.1b}$$

$$G(l, m) = \frac{1}{\pi^2} \int_0^{\pi} dx \int_0^{\pi} dy \frac{\cos lx \cos my}{t - \cos 2x - 2\cos x \cos y}. \tag{2.2b}$$

(2.2b) may be obtained either by solving (2.1b) or by changing the variables x and y in (2.2a) by x/a and y/b, respectively. To show the variable t explicitly, we use the notation $G(t; l, m)$ in place of $G(l, m)$.

We shall see later that the value of $G(l, m)$ at an arbitrary lattice site is calculated by using some recurrence formulas if the values of $G(0, 0), G(2, 0)$, and $G(4, 0)$ are known. At first, we investigate these three functions.

If t is real and $t > 3$, the values of $G(l, 0)$ for $l = 0, 2, 4$ are obtained by the standard formulas[5] as follows:

$$G(0, 0) = (1/2\pi) g K(k), \tag{2.6}$$

$$G(2, 0) = (1/2\pi) g \{2[1 - (1/\alpha^2)]\, \Pi\, (\alpha^2, k) + [(2/\alpha^2) - 1] K(k)\}, \tag{2.7}$$

$$\begin{aligned} G(4, 0) = (1/2\pi) g \{ &[(\alpha^2 - 2)^2/\alpha^4] K(k) \\ &+ [4(\alpha^2 - 1)/(k^2 - \alpha^2)\alpha^2] E(k) \\ &+ [4(\alpha^2 - 1)/(k^2 - \alpha^2)\alpha^4](k^2 - 2\alpha^2 + \alpha^4) \\ &\times \Pi\, (\alpha^2, k)\}, \end{aligned} \tag{2.8}$$

where

$$g = 8/[(2t + 3)^{1/2} - 1]^{3/2}[(2t + 3)^{1/2} + 3]^{1/2}, \tag{2.9}$$

$$\alpha^2 = 4/[1 - (2t + 3)^{1/2}]^2, \tag{2.10}$$

$$k = 4(2t + 3)^{1/4}/[(2t + 3)^{1/2} - 1]^{3/2}[(2t + 3)^{1/2} + 3]^{1/2}, \tag{2.11}$$

and $K(k), E(k)$, and $\Pi\, (\alpha^2, k)$ are the complete elliptic integrals of the first, second, and third kind, respectively:

$$K(k) = \int_0^{\pi/2} \frac{1}{(1 - k^2 \sin^2\theta)^{1/2}} d\theta, \tag{2.12}$$

$$E(k) = \int_0^{\pi/2} (1 - k^2 \sin^2\theta)^{1/2} d\theta, \tag{2.13}$$

$$\Pi(\alpha^2, k) = \int_0^{\pi/2} \frac{1}{(1 - \alpha^2 \sin^2\theta)(1 - k^2 \sin^2\theta)^{1/2}} d\theta \tag{2.14}$$

On the other hand, substituting (2.6) into Eq. (2.1b) for $l = 0$ and $m = 0$ and using the symmetry relations (2.3) and (2.4), we obtain another expression for $G(2, 0)$;

$$G(2, 0) = (t/6\pi) g K(k) - \tfrac{1}{3}. \tag{2.15}$$

By comparing two expressions (2.7) and (2.15) for $G(2, 0)$, we obtain the following relation which expres-

ses the complete elliptic integral of the third kind (2.14) in terms of the one of the first kind (2.12) for the special modulus k and parameter α^2 given by (2.11) and (2.10):

$$\Pi(\alpha^2,k) = [1/6(\alpha^2-1)]\{[(3+t)\alpha^2-6]K(k) - (2\pi\alpha^2/g)\}. \quad (2.16)$$

By using this relation in (2.8), $G(4,0)$ is given by

$$G(4,0) = \frac{1}{2\pi} g\left(\frac{3(\alpha^2-2)(k^2+\alpha^2-2)+2(k^2-2\alpha^2+\alpha^4)t}{3(k^2-\alpha^2)\alpha^2}K(k) + \frac{4(\alpha^2-1)}{(k^2-\alpha^2)\alpha^2}E(k)\right) - \frac{2(k^2-2\alpha^2+\alpha^4)}{3(k^2-\alpha^2)\alpha^2}. \quad (2.17)$$

Equations (2.6), (2.15), and (2.17) express $G(0,0)$, $G(2,0)$, and $G(4,0)$ in terms of the complete elliptic integrals of the first and second kind.

We have derived these expressions for the case in which t is real and $t > 3$. Because of the analyticity of the lattice Green's function, these expressions are valid for any complex value t. When $t = s \pm i\epsilon$, where s is real and ϵ is a positive infinitesimal number, the factor $(s \pm i\epsilon - a)^{1/n}$ for $s > a$ changes to $e^{\pm i\pi/n}(a - s)^{1/n}$ for $s < a$ in Eqs. (2.9)–(2.11). Figures 2(a) and 2(b) show the curves for modulus k for $t = s \pm i\epsilon$, $-\infty < s < \infty$, and $\epsilon > 0$. The upper half of the complex t plane is mapped into the region which is the second and third quadrants outside of the loop and the fourth quadrant for $t = s + i\epsilon$ in Fig. 2(a), and the lower half into the region which is the second and third quadrants outside of the loop and the first quadrant for $t = s - i\epsilon$ in Fig. 2(b). The function $K(k)$ and $E(k)$ have branch points at $k = \pm 1$. The expressions (2.12) and (2.13) are analytic on the Riemann surface excluding the branch cuts connecting + 1 and + ∞, and − 1 and − ∞, respectively, on the real axis. We call this part of the Riemann surface as sheet I. The Riemann surface which is the upper quarter plane reached through the cut connecting − 1 and − ∞ from the sheet I is denoted as sheet II in Fig. 2(a), and the one which is the lower quarter plane reached through the cut connecting − 1 and − ∞ is denoted as sheet III in Fig. 2(b). In Figs. 2(a) and 2(b), sheets I and II and sheets I and III, respectively, are written in the same figure by drawing the branch cut connecting − 1 and + 1. For k on the sheets II and III, the analytic continuations of the function $K(k)$ and $E(k)$ on the sheet I to the sheets II and III are considered, and the obtained expressions are used in place of $K(k)$ and $E(k)$. They are given as follows[6,7]:

$$\begin{aligned} K^{II}(k) &= K(k) + 2iK'(k), \\ E^{II}(k) &= E(k) + 2i[K'(k) - E'(k)], \\ K^{III}(k) &= K(k) - 2iK'(k), \\ E^{III}(k) &= E(k) - 2i[K'(k) - E'(k)], \end{aligned} \quad (2.18)$$

where $K'(k)$ and $E'(k)$, respectively, are the complete elliptic integral of the first and second kind with the complementary modulus of k. One sees from (2.18) that the analytically continued function has the branch points at the origin and at the infinity. In Fig. 2(a), these branch points are connected along the positive portion of the imaginary axis and in Fig. 2(b) along the negative portion of it. Hence sheet II in Fig. 2(a) is restricted to the left upper quarter of the complex k plane and sheet III in Fig. 2(b) to the left lower quarter.

We have the expressions of the functions $G(0,0)$, $G(2,0)$, and $G(4,0)$ for any complex variable t:

$$G(0,0) = (1/2\pi)g\tilde{K}(k), \quad (2.19)$$

$$G(2,0) = (t/6\pi)g\tilde{K}(k) - \tfrac{1}{3}, \quad (2.20)$$

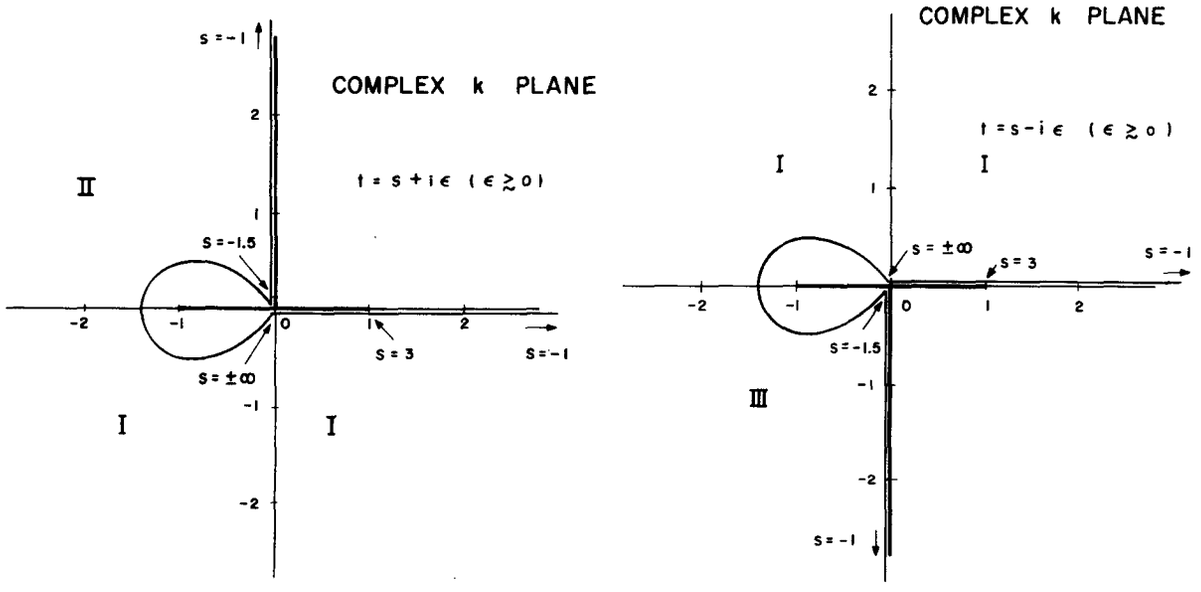


FIG. 2. (a) The curve of the modulus k in the complex k plane when $t = s + i\epsilon$, s is real and ϵ is a positive infinitesimal number. The bold solid lines between − 1 and + 1, and 0 and + $i\infty$ denote the branch cuts of the complete elliptic integrals of the first and second kind. The lower half-plane represents the sheet I and the left upper quarter of the plane the sheet II. (b) The curve of the modulus k in the complex k-plane when $t = s - i\epsilon$. The bold solid lines between − 1 and + 1, and 0 and − $i\infty$ denote the branch cuts of the complete elliptic integrals of the first and second kind. The upper half-plane represents the sheet I and the left lower quarter of the plane the sheet III.

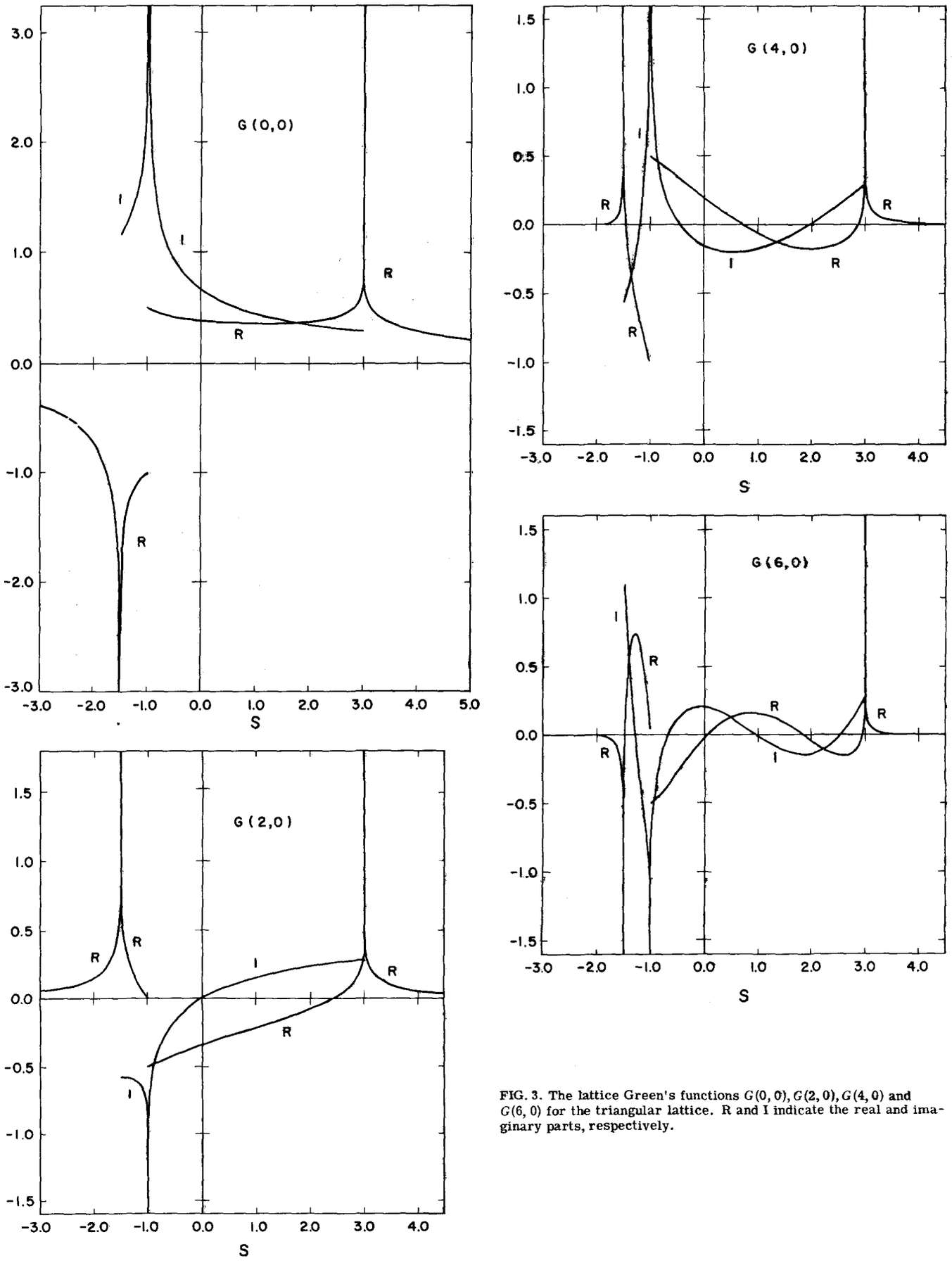


FIG. 3. The lattice Green's functions $G(0, 0)$, $G(2, 0)$, $G(4, 0)$ and $G(6, 0)$ for the triangular lattice. R and I indicate the real and imaginary parts, respectively.

$$G(4,0)=\frac{1}{2\pi}\times g\left(\frac{3(\alpha^2-2)(k^2+\alpha^2-2)+2(k^2-2\alpha^2+\alpha^4)t}{3(k^2-\alpha^2)\alpha^2}\tilde{K}(k)+\frac{4(\alpha^2-1)}{(k^2-\alpha^2)\alpha^2}\tilde{E}(k)\right)-\frac{2(k^2-2\alpha^2+\alpha^4)}{3(k^2-\alpha^2)\alpha^2}, \quad (2.21)$$

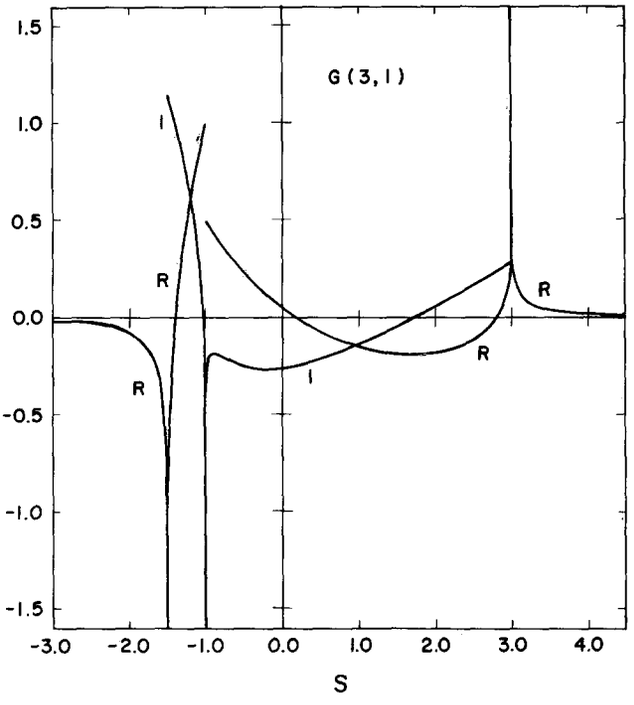
where $\tilde{K}(k)$ and $\tilde{E}(k)$ are given as follows:

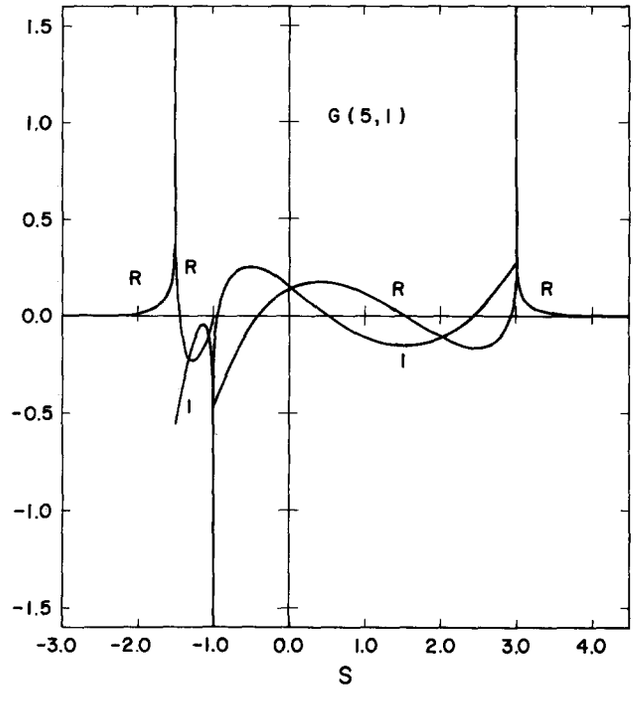
$$\tilde{K}(k)=\begin{cases} K(k) & \text{for } \mathrm{Im}t>0 \text{ and } \mathrm{Im}k<0 \\ & \text{or } \mathrm{Im}t<0 \text{ and } \mathrm{Im}k>0 \\ K^{\mathrm{II}}(k) & \text{for } \mathrm{Im}t>0 \text{ and } \mathrm{Im}k>0 \\ K^{\mathrm{III}}(k) & \text{for } \mathrm{Im}t<0 \text{ and } \mathrm{Im}k<0 \end{cases}, \quad (2.22)$$

$$\tilde{E}(k)=\begin{cases} E(k) & \text{for } \mathrm{Im}t>0 \text{ and } \mathrm{Im}k<0 \\ & \text{or } \mathrm{Im}t<0 \text{ and } \mathrm{Im}k>0 \\ E^{\mathrm{II}}(k) & \text{for } \mathrm{Im}t>0 \text{ and } \mathrm{Im}k>0 \\ E^{\mathrm{III}}(k) & \text{for } \mathrm{Im}t<0 \text{ and } \mathrm{Im}k<0 \end{cases}. \quad (2.23)$$

It is known that the complete elliptic integrals of the first and second kind with the modulus of complex number are easily calculated by using the arithemetic geometric means.[8] Thus, one is able to know the Green's functions $G(l,0)$ for $l=0,2,4$ numerically as well as analytically.

We shall derive some recurrence formulas in order to obtain the values of $G(l,m)$ at an arbitrary lattice site. In the first place, we consider the function $G(l,0)$. In Refs. 9 and 10, a method of deriving a re-





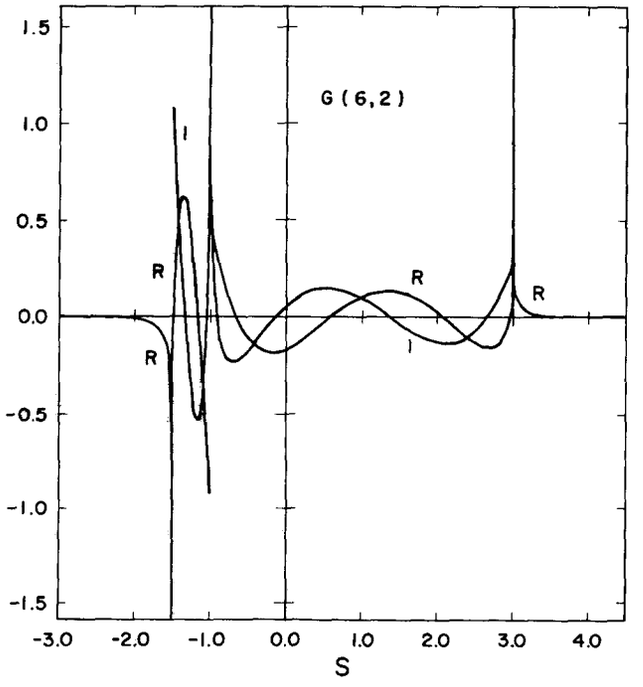


FIG. 4. The lattice Green's functions $G(3,1)$, $G(5,1)$, and $G(6,2)$ for triangular lattice. R and I indicate the real and imaginary parts, respectively.

currence formula connecting the lattice Green's function along an axis is discussed for the square lattice with the interaction up to the second neighbors. By the same method, one obtains the following recurrence formula for $G(l,0)$:

$$G(l+4,0) = [1/(l+2)][4(l+1)(t+1)G(l+2,0) - 2l(2t^2-3)G(l,0) + 4(l-1)(t+1) \times G(l-2,0) - (l-2)G(l-4,0)], \quad (2.24)$$

where l is even. From this equation, one sees that the values of $G(l,0)$ for $l \geq 6$ are calculated from the knowledge of $G(0,0)$, $G(2,0)$, and $G(4,0)$. Thus one is able to calculate the lattice Green's function at an arbitrary lattice site on the la axis in Fig. 1. The graphs for the $G(l,0)$ $(l = 0, 2, 4, 6)$ are given in Fig. 3 for $t = s - i\epsilon$, $-\infty < s < \infty$ and $\epsilon \gtrsim 0$.

In the next place we consider the function $G(l,1)$. From Eq. (2.1b), one obtains the following equation by taking symmetry properties (2.4) into account:

$$G(l+1,1) = tG(l,0) - \tfrac{1}{2}[G(l+2,0) + G(l-2,0)] - G(l-1,1), \quad (2.25)$$

where l is even and greater than or equal to 2, and

$$G(1,1) = G(2,0). \quad (2.26)$$

For $G(l,m)$ where $l \geq 4$ and $m \geq 2$, the following recurrence formula is obtained by applying Eq. (2.1b) for lattice site $(l-1, m-1)$:

$$G(l,m) = 2tG(l-1,m-1) - G(l-2,m) - G(l-3,m-1) - G(l+1,m-1) - G(l,m-2) - G(l-2,m-2). \quad (2.27)$$

When Eq. (2.27) is used, we note that the function $G(l,m)$ for $m > 3l$ is expressed by the one for $m < 3l$ by using the relation (2.5a):

$$G(l,m) = G(\tfrac{1}{2}(l+3m), \tfrac{1}{2}(l-m)). \quad (2.5b)$$

The graph for $G(3,1)$, $G(5,1)$, and $G(6,2)$ are given in Fig. 4 for $t = s - i\epsilon$, $-\infty < s < \infty$ and $\epsilon \gtrsim 0$.

Now the analytic properties and values of the lattice Green's function for the triangular lattice at an arbitrary lattice site can be discussed from those of $G(0,0)$, $G(2,0)$, and $G(4,0)$ by using the recurrence formulas (2.24), (2.25), and (2.27).

3. HONEYCOMB LATTICE

We consider the monatomic lattice Green's function for the honeycomb lattice with the nearest neighbor interaction. The honeycomb network is shown in Fig. 5. Here we distinguish the lattice site A or B according as it is an upper site or lower site of a vertical bond. Each of the sublattice constitutes a triangular lattice.

At first, the case in which the initial lattice site belongs to the sublattice A is considered. In that case, the lattice Green's function is the solution of the following set of two difference equations:

$$2tG^{AA}(la,mb) - G^{BA}(la-a,mb+b) - G^{BA}(la+a,mb+b) - G^{BA}(la,mb-2b) = 2\delta_{l,0}\delta_{m,0}, \quad (3.1)$$

$$2tG^{BA}(la,mb) - G^{AA}(la+a,mb-b) - G^{AA}(la-a,mb-b) - G^{AA}(la,mb+2b) = 0, \quad (3.2)$$

where

$$la = (\mathbf{r}_f - \mathbf{r}_i)_x, \quad (3.3)$$

$$mb = (\mathbf{r}_f - \mathbf{r}_i)_y, \quad (3.4)$$

and $\mathbf{r}_i$ and $\mathbf{r}_f$ are the initial and final lattice site, respectively. $a = \sqrt{3}$ and $b = 1$, if lattice constant is chosen to be equal to 2. The solution of these equations under the boundary condition that $G(la,mb)$ is equal to zero as $l^2 + m^2$ tends to infinity, is given by

$$G^{AA}(la,mb) = 4t\frac{ab}{4\pi^2}\int_{-\pi/a}^{\pi/a} dx \int_{-\pi/b}^{\pi/b} dy \times \frac{e^{i(lax+mby)}}{4t^2 - 3 - 2\cos 2ax - 4\cos ax \cos 3by}, \quad (3.5)$$

$$G^{BA}(la,mb) = \frac{ab}{4\pi^2}\int_{-\pi/a}^{\pi/a} dx \int_{-\pi/b}^{\pi/b} dy \times \frac{e^{i(lax+mby)}[e^{i2bx} + e^{i(-ax-by)} + e^{i(ax-by)}]}{4t^2 - 3 - 2\cos 2ax - 4\cos ax \cos 3by}. \quad (3.6)$$

To show the variable t explicitly, we use the notation, for example, $G^{AA}(t; la, mb)$ instead of (3.5).

Next, we consider the case in which the initial site belongs to the sublattice B. In a similar way to the foregoing paragraph, the following relations are obtained:

$$G^{BB}(la,mb) = G^{AA}(la,mb), \quad (3.7)$$

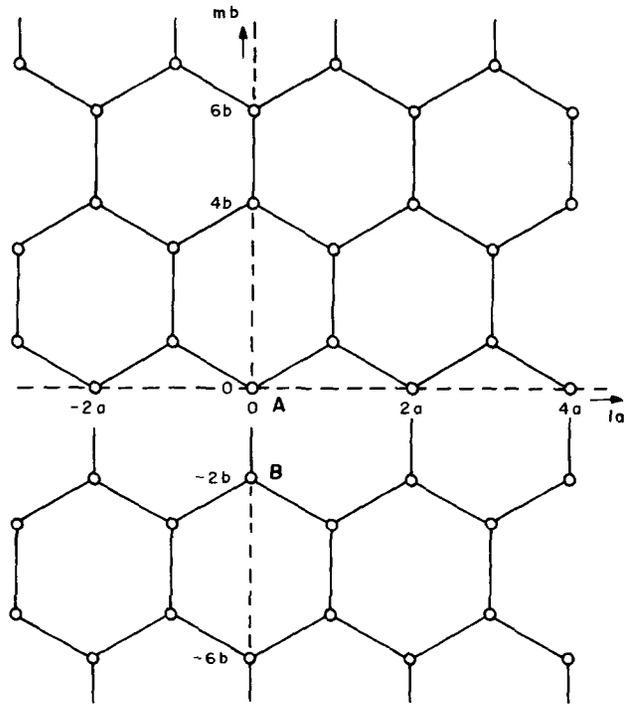


FIG. 5. The network of the honeycomb lattice. Distinction is given on the lattice site as A or B according as it is an upper site or lower site of a vertical bond.

$$G^{AB}(la, mb) = G^{BA}(-la, -mb). \tag{3.8}$$

Then, it is needed to consider the detailed knowledge of $G^{AA}(la, mb)$ and $G^{BA}(la, mb)$ in the following. We shall drop the superscript on the function when we do not need to indicate them.

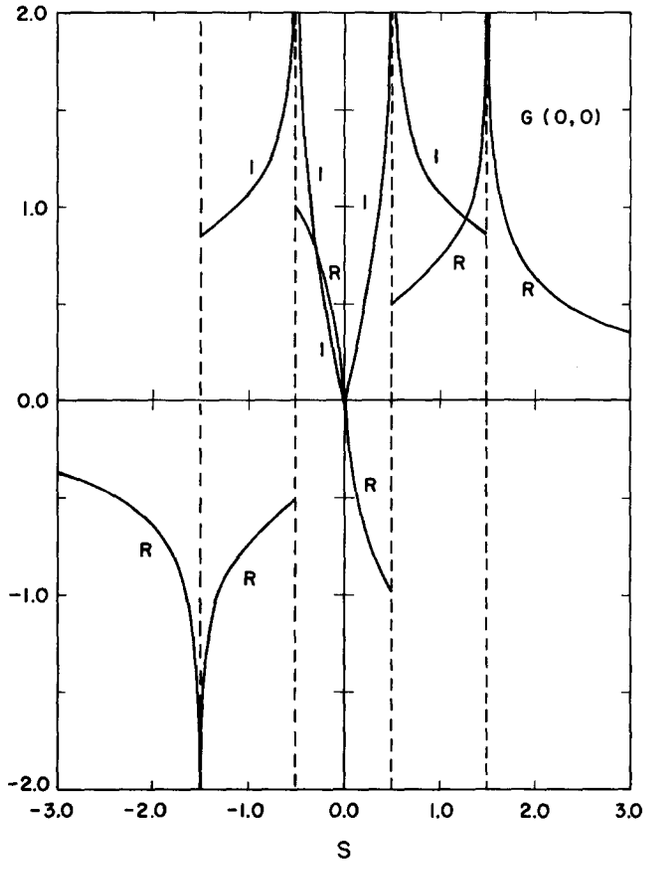
The functions $G^{AA}(la, mb)$ and $G^{BA}(la, mb)$ are unchanged under the rotation $\frac{2}{3}n\pi$, $n = 1, 2$, of the coordinate axes around the initial lattice site:

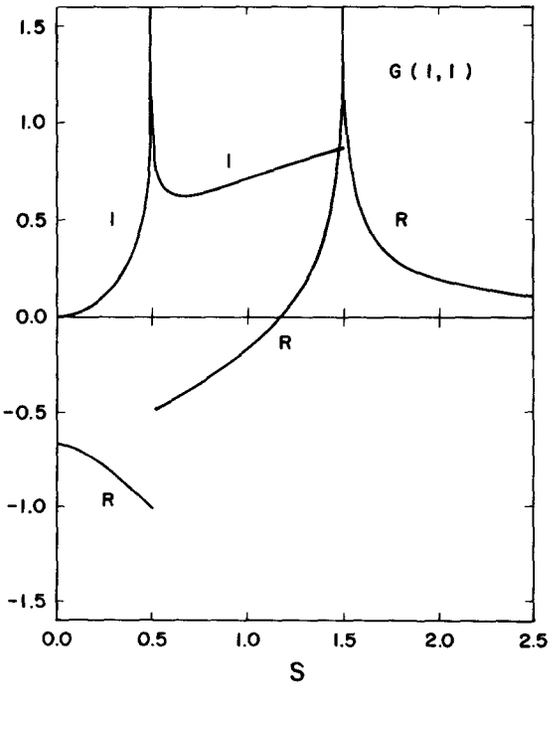
$$G(la, mb) = G(la\cos(\tfrac{2}{3}n\pi) - mb\sin(\tfrac{2}{3}n\pi), \quad la\sin(\tfrac{2}{3}n\pi) + mb\cos(\tfrac{2}{3}n\pi)), \tag{3.9}$$

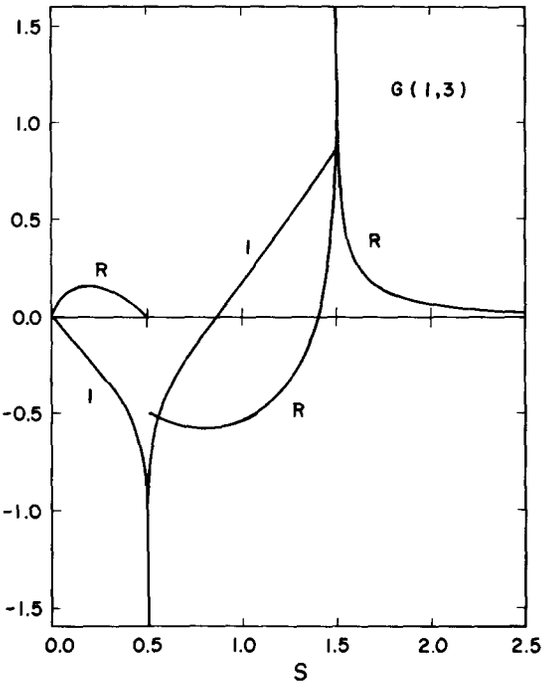
and under the inversion on the vertical axis

$$G(la, mb) = G(-la, mb). \tag{3.10}$$

As we put $t = s - i\epsilon$, where s is real and ϵ is a positive infinitesimal number, we can easily confirm that







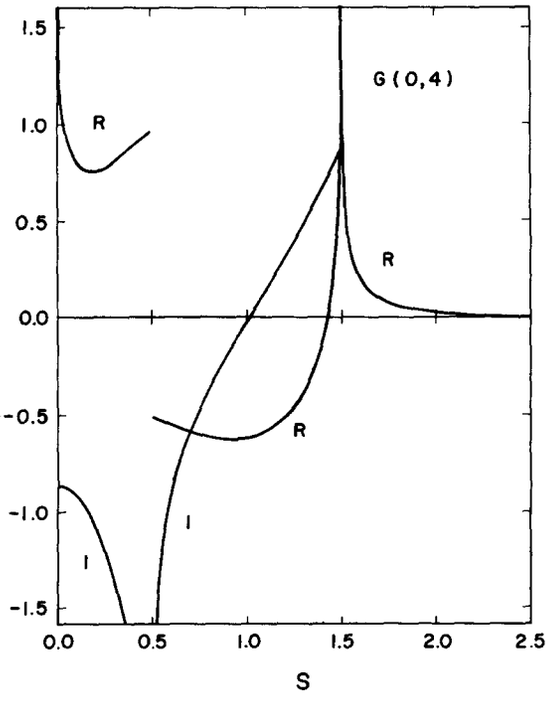


FIG. 6. The lattice Green's functions $G(0, 0)$, $G(1, 1)$, $G(1, 3)$, and $G(0, 4)$ for the honeycomb lattice. R and I indicate the real and imaginary parts, respectively. It is noticed that the imaginary parts of $G(0, 0)$ and $G(1, 3)$ and the real parts of $G(1, 1)$ and $G(0, 4)$ are even functions of s, and the real parts of $G(0, 0)$ and $G(1, 3)$ and the imaginary parts of $G(1, 1)$ and $G(0, 4)$ are odd functions of s.

the real and imaginary parts of $G^{AA}(t; la, mb)$ are an odd and an even function, respectively, and the real and imaginary parts of $G^{BA}(t; la, mb)$ are an even and an odd function, respectively;

$$\mathrm{Re}G(s - i\epsilon; la, mb) = -\Delta_{lm}\mathrm{Re}G(-s - i\epsilon; la, mb), \tag{3.11}$$

$$\mathrm{Im}G(s - i\epsilon; la, mb) = \Delta_{lm}\mathrm{Im}G(-s - i\epsilon; la, mb), \tag{3.12}$$

where Δ_{lm} is plus or minus unity according as the sites l and m belong to the same sublattice or to different sublattices. This fact corresponds to the symmetry properties discussed generally in Ref. 11. for the "alternating lattice."

After the parameters a and b are deleted by the variable transformation, these functions $G^{AA}(la, mb)$ and $G^{BA}(la, mb)$ are expressed in terms of the lattice Green's function for the triangular lattice

$$G^{AA}(l, m) = 2tG_t(\tfrac{1}{2}(4t^2 - 3); l, \tfrac{1}{3}m), \tag{3.13}$$

$$\begin{aligned} G^{BA}(l, m) = {} & G_t(\tfrac{1}{2}(4t^2 - 3); l, \tfrac{1}{3}(m + 2)) \\ & + G_t(\tfrac{1}{2}(4t^2 - 3); l - 1, \tfrac{1}{3}(m - 1)) \\ & + G_t(\tfrac{1}{2}(4t^2 - 3); l + 1, \tfrac{1}{3}(m - 1)), \end{aligned} \tag{3.14}$$

where $G_t(t; l, m)$ is defined by Eq. (2. 2b). Thus, the analytic properties and the values of the lattice Green's function for the honeycomb lattice are obtained from the knowledge of the one for the triangular lattice. The graphs for the honeycomb lattice are given in Fig. 6 for $G(0, 0), G(1, 1), G(1, 3)$, and $G(0, 4)$.

The imaginary part of $G(0, 0)$ is equal to the frequency distribution function of the lattice vibration without a factor π. The curve for the frequency distribution was first given by Hobson and Nierenberg.[4]

4. DISCUSSIONS

The exact expressions for the lattice Green's function for the lattice sites $(0, 0), (2, 0)$, and $(4, 0)$ are derived for the triangular lattice. The function at an arbitrary lattice site is shown to be calculated by using the recurrence formulas (2. 24), (2. 25), and (2. 27) from the knowledge of these three functions. The lattice Green's function for the honeycomb lattice is presented in terms of that for the triangular lattice. These expressions are also useful for the case when the parameter t is a complex number.

For $t = s - i\epsilon$ where s is real and ϵ is a positive infinitesimal number, the graphs of the functions for the several lattice sites are shown. The singularities are found at the singular points which are determined by the critical points. These critical points are easily seen to be nondegenerate, and, hence, the behaviors of the function are expected to be logarithmic divergent from the general discussion of Ref. 12 at the singular points. However, some cancellations of the singularity of the lattice Green's function occur at the middle of the band for the honeycomb lattice. In fact, the value of the function at the lattice sites which belong to the same sublattice is equal to zero at the middle of the band, cf. (3. 13) and (3. 7). In this case, the derivative of the imaginary part of the function with respect to s has a finite jump but the derivative of the real part of the function with respect to s shows a logarithmic divergence. For the function at the lattice sites which are on different sublattices, no cancellation occurs and the real part of the function diverges logarithmically, cf. (3. 14). Only exception is function $G(1, 1)$. This function is easily expressed as

$$G(1, 1) = \tfrac{4}{3}t^2G_t(\tfrac{1}{2}(4t^2 - 3); 0, 0) - \tfrac{2}{3}.$$

It is interesting to compare the graphs for $G(0, 0)$ and $G(1, 3)$ in Fig. 6 with the graphs for $G(0, 0)$ and $G(2, 0)$ in Fig. 3. The sites $(0, 0)$ and $(1, 3)$ of the honeycomb lattice correspond to the sites $(0, 0)$ and $(2, 0)$, respectively, of the triangular lattice. The region which is defined by $t = s - i\epsilon, s > 0$, and $\epsilon \gtrsim 0$, for the honeycomb lattice is mapped into the region which is defined by $t = s - i\epsilon, s > -1.5$, and $\epsilon \gtrsim 0$ for the triangular lattice. The region which is defined by $t = s - i\epsilon$, $s < 0$, and $\epsilon \gtrsim 0$ for the honeycomb lattice is mapped into the region which is defined by $t = s + i\epsilon, s > -1.5$, and $\epsilon \gtrsim 0$ for the triangular lattice. The values $s = -1.5$ and -1.0 and 3.0 which are the singular points of the function for the triangular lattice (cf. Fig. 3) change to the values $s = 0, \pm 0.5$, and ± 1.5 which are the singular points of the function for the honeycomb lattice (cf. Fig. 6). The top of the band, $s = -1.5$, for the triangular lattice corresponds to the middle of the band, $s = 0$, for the honeycomb lattice and, as mentioned above, at that point the singular behavior of the function for the honeycomb lattice cancels out. The portion corresponding to the one above the top of the band for the triangular lattice disappears for the honeycomb lattice. The function for the triangular lattice does not satisfy any symmetry properties with respect to s; however, the function for the honeycomb lattice satisfies the symmetry properties (3. 11) and (3. 12). We notice a similar correspondence between the graphs for $G(0, 0)$ in Fig. 6 and $G(0,0)$ in Fig. 3 and also between the graphs for $G(1,3)$ in Fig. 6 and $G(2, 0)$ in Fig. 3.

Thorpe and Weaire have recently shown that the imaginary part of the lattice Green's function for the diamond structure lattice at the origin is expressed in terms of the one for the fcc lattice. The result was shown by a graph. The diamond lattice is composed of two fcc sublattices and fcc lattice is one of the three-dimensional closed packed lattices. As discussed in Ref. 13, if the lattice sites on one of the sublattices have the bonds extending vertically upward, the sites on the other sublattice have bonds extending vertically downward, cf. Fig. 2–7 in Ref. 13. One cannot get a site on one of sublattices from that on the other sublattice only by a simple translation. These relations for the diamond structure and fcc lattices are quite similar to the ones for the honeycomb and triangular lattices. The top of the band for the fcc lattice corresponds to the middle of the band for the diamond structure lattice and the singularity of the function at the origin for the diamond lattice is cancelled out at this point of the band. Other singular points are due to the nondegenerate critical points then singular behaviors are square root as discussed generally in Ref. 12. The singular behaviors of the function for the honeycomb lattice are logarithmic divergent or finite jump at the singular point due to the nondegenerate critical points. Taking these facts into account, we see that the similarity of the lattice reflects to the lattice Green's function if we compare

Fig. 6 with Fig. 2 in Ref. 1. On the other hand, the linear, square, sc, and bcc lattices are composed of the linear, square, fcc, and sc sublattices, respectively. The relations for these lattices and their sublattices are not similar to the ones for the diamond and fcc lattices or the honeycomb and triangular lattices. Especially, one can get a site on one of the sublattices from that on the other sublattice only by a simple translation, and the behaviors of the lattice Green's functions for these lattices are different from the ones of the functions for the honeycomb and diamond structure lattices.

ACKNOWLEDGMENTS

The author would like to express his sincere thanks to Professor T. Morita for valuable suggestions and discussions and a critical reading of the manuscript. The numerical calculations were carried out by IBM 360 Model 44 in the Computer Center of Ohio University.

1 M. F. Thorpe and D. Weaire, Phys. Rev. **4**, 3518 (1971).
2 D. Weaire, Phys. Rev. Letters **21**, 1541 (1971).
3 G. F. Newell, J. Chem. Phys. **23**, 2431 (1955).
4 J. P. Hobson and W. A. Nierenberg, Phys. Rev. **89**, 662 (1953).
5 P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals for Engineers and Scientist* (Springer-Verlag, Berlin, 1971).
6 T. Morita and T. Horiguchi, J. Math. Phys. **12**, 981, 986 (1971).
7 T. Horiguchi, J. Phys. Soc. Japan, **30**, 1261 (1971).
8 T. Morita and T. Horiguchi, to be published.
9 T. Morita, J. Math. Phys. **12**, 1744 (1971).
10 T. Morita and T. Horiguchi, unpublished work. See T. Horiguchi, thesis (Tohoku University, 1971, unpublished, in Japanese).
11 T. Morita and T. Horiguchi, J. Math. Phys. **13**, 1243 (1972).
12 T. Morita and T. Horiguchi, J. Phys. **A5**, 67 (1972).
13 J. C. Slater, *Quantum Theory of Molecules and Solids* (McGraw-Hill, New York, 1965), Vol. 2.

An Approximation Technique for Solving Multiparticle and Bound-State Scattering Processes

Tsu-huei Liu
Department of Physics, Portland State University, Portland, Oregon
(Received 13 March 1972)

A modified version of the Lee model by Bronzan is investigated by the LSZ formalism approach. The existence of $(V\theta)$ bound state in the Bronzan's model is discussed. An iterative technique is applied for solving the tau functions in the $V-2\theta$ sector. Each term in the series preserves (a) $(V\theta)$ bound state properties, (b) analytic structure, and (c) symmetry properties of the tau function. The S-matrix elements for both nonbound-state and $(V\theta)$ bound state scattering processes in the $V-2\theta$ sector are obtained. Comparison of the result obtained by our iterative expansion method with that obtained by Bronzan's variational principle method is made.

I. INTRODUCTION

In addition to the V, N, and θ particles, Bronzan[1] introduced another fermion called U particle with the coupling $U \leftrightarrow V + \theta$ to the Lee model.[2] We will refer to this modified Lee model as B model here after. Bronzan investigated the $U + \theta$ elastic scattering process[3] using a variational principle technique.

In the previous articles,[4,5] we studied the $V-2\theta$ and general higher sectors of the Lee model using the LSZ formalism. An iterative expansion method was used for calculating the multiparticle and bound state scattering processes. Each term in the series was shown to preserve (a) the properties of the bound state, (b) the analytic structures, and (c) the symmetry properties of the tau function. Because our iterative expansion preserves the above properties, it is a useful technique for approximating the complex Green's function.

It will be shown in this article that, in analogy to the Lee model, a $(V\theta)$ bound state can exist in the B model for suitably chosen parameters. Therefore, the B model also offers the opportunity for the study of bound state scattering processes.

The purpose of this article is thus to apply our iterative expansion technique to solve the $V-2\theta$ sector of the B model. From the iterative solution of the tau functions, we can then calculate all the S-matrix elements for both nonbound-state and $(V\theta)$ bound-state scattering processes in the $V-2\theta$ sector. A comparison is made of our result for the $U + \theta$ elastic scattering amplitude and that obtained by using the variational principle method in Ref. 3.

The outline of the article is as follows. The description of the B model and LSZ formalism for the reduction formula of the S-matrix are given in Sec. II. In Sec. III, the existence of a $(V\theta)$ bound state in the B model is discussed. The iterative expansion method is illustrated in Sec. IV for solving the tau functions in the $V-2\theta$ sector. Our solution preserves the analytic structure and symmetry properties of the tau function and most important, the properties of the $(V\theta)$ bound state. Solution to the tau functions furnish all the information for the $V-2\theta$ sector. In Sec. V, S-matrix elements for the nonbound state scattering processes

$$U + \theta \rightarrow U + \theta,$$
$$U + \theta \leftrightarrow V + \theta + \theta,$$
$$U + \theta \leftrightarrow N + \theta + \theta + \theta,$$
$$V + \theta + \theta \rightarrow V + \theta + \theta,$$
$$V + \theta + \theta \leftrightarrow N + \theta + \theta + \theta,$$
$$N + \theta + \theta + \theta \rightarrow N + \theta + \theta + \theta$$

are calculated. Comparison between our result of $U + \theta$ elastic scattering amplitude and that obtained by variationed principle in the dispersion theory approach is made. The calculation of the S matrix for all the bound state processes in the $V-2\theta$ sector,

$$(V\theta) + \theta \leftrightarrow (V\theta) + \theta,$$
$$(V\theta) + \theta \leftrightarrow V + \theta + \theta,$$
$$(V\theta) + \theta \leftrightarrow U + \theta,$$
$$(V\theta) + \theta \leftrightarrow N + \theta + \theta + \theta,$$

is discussed. The conclusion follows in Sec. VI.

Fig. 6 with Fig. 2 in Ref. 1. On the other hand, the linear, square, sc, and bcc lattices are composed of the linear, square, fcc, and sc sublattices, respectively. The relations for these lattices and their sublattices are not similar to the ones for the diamond and fcc lattices or the honeycomb and triangular lattices. Especially, one can get a site on one of the sublattices from that on the other sublattice only by a simple translation, and the behaviors of the lattice Green's functions for these lattices are different from the ones of the functions for the honeycomb and diamond structure lattices.

ACKNOWLEDGMENTS

The author would like to express his sincere thanks to Professor T. Morita for valuable suggestions and discussions and a critical reading of the manuscript. The numerical calculations were carried out by IBM 360 Model 44 in the Computer Center of Ohio University.

1 M. F. Thorpe and D. Weaire, Phys. Rev. **4**, 3518 (1971).
2 D. Weaire, Phys. Rev. Letters **21**, 1541 (1971).
3 G. F. Newell, J. Chem. Phys. **23**, 2431 (1955).
4 J. P. Hobson and W. A. Nierenberg, Phys. Rev. **89**, 662 (1953).
5 P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals for Engineers and Scientist* (Springer-Verlag, Berlin, 1971).
6 T. Morita and T. Horiguchi, J. Math. Phys. **12**, 981, 986 (1971).
7 T. Horiguchi, J. Phys. Soc. Japan, **30**, 1261 (1971).
8 T. Morita and T. Horiguchi, to be published.
9 T. Morita, J. Math. Phys. **12**, 1744 (1971).
10 T. Morita and T. Horiguchi, unpublished work. See T. Horiguchi, thesis (Tohoku University, 1971, unpublished, in Japanese).
11 T. Morita and T. Horiguchi, J. Math. Phys. **13**, 1243 (1972).
12 T. Morita and T. Horiguchi, J. Phys. **A5**, 67 (1972).
13 J. C. Slater, *Quantum Theory of Molecules and Solids* (McGraw-Hill, New York, 1965), Vol. 2.

An Approximation Technique for Solving Multiparticle and Bound-State Scattering Processes

Tsu-huei Liu
Department of Physics, Portland State University, Portland, Oregon
(Received 13 March 1972)

A modified version of the Lee model by Bronzan is investigated by the LSZ formalism approach. The existence of $(V\theta)$ bound state in the Bronzan's model is discussed. An iterative technique is applied for solving the tau functions in the $V-2\theta$ sector. Each term in the series preserves (a) $(V\theta)$ bound state properties, (b) analytic structure, and (c) symmetry properties of the tau function. The S-matrix elements for both nonbound-state and $(V\theta)$ bound state scattering processes in the $V-2\theta$ sector are obtained. Comparison of the result obtained by our iterative expansion method with that obtained by Bronzan's variational principle method is made.

I. INTRODUCTION

In addition to the V, N, and θ particles, Bronzan[1] introduced another fermion called U particle with the coupling $U \leftrightarrow V + \theta$ to the Lee model.[2] We will refer to this modified Lee model as B model here after. Bronzan investigated the $U + \theta$ elastic scattering process[3] using a variational principle technique.

In the previous articles,[4,5] we studied the $V-2\theta$ and general higher sectors of the Lee model using the LSZ formalism. An iterative expansion method was used for calculating the multiparticle and bound state scattering processes. Each term in the series was shown to preserve (a) the properties of the bound state, (b) the analytic structures, and (c) the symmetry properties of the tau function. Because our iterative expansion preserves the above properties, it is a useful technique for approximating the complex Green's function.

It will be shown in this article that, in analogy to the Lee model, a $(V\theta)$ bound state can exist in the B model for suitably chosen parameters. Therefore, the B model also offers the opportunity for the study of bound state scattering processes.

The purpose of this article is thus to apply our iterative expansion technique to solve the $V-2\theta$ sector of the B model. From the iterative solution of the tau functions, we can then calculate all the S-matrix elements for both nonbound-state and $(V\theta)$ bound-state scattering processes in the $V-2\theta$ sector. A comparison is made of our result for the $U + \theta$ elastic scattering amplitude and that obtained by using the variational principle method in Ref. 3.

The outline of the article is as follows. The description of the B model and LSZ formalism for the reduction formula of the S-matrix are given in Sec. II. In Sec. III, the existence of a $(V\theta)$ bound state in the B model is discussed. The iterative expansion method is illustrated in Sec. IV for solving the tau functions in the $V-2\theta$ sector. Our solution preserves the analytic structure and symmetry properties of the tau function and most important, the properties of the $(V\theta)$ bound state. Solution to the tau functions furnish all the information for the $V-2\theta$ sector. In Sec. V, S-matrix elements for the nonbound state scattering processes

$$U + \theta \rightarrow U + \theta,$$
$$U + \theta \leftrightarrow V + \theta + \theta,$$
$$U + \theta \leftrightarrow N + \theta + \theta + \theta,$$
$$V + \theta + \theta \rightarrow V + \theta + \theta,$$
$$V + \theta + \theta \leftrightarrow N + \theta + \theta + \theta,$$
$$N + \theta + \theta + \theta \rightarrow N + \theta + \theta + \theta$$

are calculated. Comparison between our result of $U + \theta$ elastic scattering amplitude and that obtained by variationed principle in the dispersion theory approach is made. The calculation of the S matrix for all the bound state processes in the $V-2\theta$ sector,

$$(V\theta) + \theta \leftrightarrow (V\theta) + \theta,$$
$$(V\theta) + \theta \leftrightarrow V + \theta + \theta,$$
$$(V\theta) + \theta \leftrightarrow U + \theta,$$
$$(V\theta) + \theta \leftrightarrow N + \theta + \theta + \theta,$$

is discussed. The conclusion follows in Sec. VI.

II. MODEL AND REDUCTION FORMULA

The renormalized Hamiltonian for the B model in momentum space is[1]

$$H = mZ_U\psi_U^+\psi_U + mZ_V\psi_V^+\psi_V + m\psi_N^+\psi_N + \sum_k w a_k^+ a_k + g[\psi_V^+\psi_N A + \psi_N^+ A^+\psi_V] + \lambda Z_V[\psi_U^+\psi_V A + \psi_V^+ A^+\psi_U] + \delta m_U Z_U\psi_U^+\psi_U + \delta m_V\psi_V^+\psi_V, \quad (1)$$

where

$$A = \sum_k [u(w)/(2w)^{1/2}]a_k, w = (\mu^2 + k^2)^{1/2},$$
$$[a_k, a_k^+] = \delta_{kk'}, \{\psi_U, \psi_U^+\} = 1/Z_U$$
$$\{\psi_V, \psi_V^+\} = 1/Z_V, \{\psi_N, \psi_N^+\} = 1, \quad (2)$$

and all other commutation relations vanish. For simplicity, we let the renormalized masses for U, V, and N particles have the same value m.

From Eqs. (1) and (2), the field equations can be derived as

$$\frac{(2w)^{1/2}}{u(w)}\left(i\frac{d}{dt} - w\right)a_k(t) = g\psi_N^+(t)\psi_V(t) + \lambda Z_V\psi_V^+(t)\psi_U(t)$$

$$Z_V\left(i\frac{d}{dt} - m - \delta m_V\right)\psi_V(t) = g\psi_N(t)A(t),$$

$$\left(i\frac{d}{dt} - m\right)\psi_N(t) = g\psi_V(t)A^+(t),$$

$$Z_U\left(i\frac{d}{dt} - m - \delta m_U\right)\psi_U(t) = \lambda Z_V\psi_V(t)A(t). \quad (3)$$

There are two number operators in this B model which commute with H:

$$Q_1 = Z_U\psi_U^+\psi_U + Z_V\psi_V^+\psi_V + \psi_N^+\psi_N,$$
$$Q_2 = Z_U\psi_U^+\psi_U + \sum_k a_k^+ a_k - \psi_N^+\psi_N, \quad (4)$$

Thus the model breaks up into sectors designated by the eigenvalues of Q_1 and Q_2, namely q_1 and q_2.

The reduction formula for the B model can be derived similarly as in Ref. 6. Assume the following conditions for the Heisenberg fields:

$$\lim_{t\to\mp\infty} \langle\!\langle \beta; n' | e^{-imt}\psi_\alpha^+(t) | \alpha; n\rangle\!\rangle = \langle\!\langle \beta; n' | \psi^+_{\alpha_{\substack{\text{in}\\\text{out}}}} | \alpha; n\rangle\!\rangle, \quad (5)$$

$$\lim_{t\to\pm\infty} \sum_{k'} f(w', w)e^{-iw't}\langle\!\langle \beta; n' | a_{k'}^+(t) | \alpha; n\rangle\!\rangle = \langle\!\langle \beta; n' | a^+_{k_{\substack{\text{in}\\\text{out}}}} | \alpha; n\rangle\!\rangle, \quad (6)$$

where $f(w', w)$ is a good function of w', centered about the point $w' = w$, satisfying the condition

$$\sum_{k''} f^*(w'', w)f(w'', w') = \begin{cases} 0 & k \neq k' \\ 1 & k = k' \end{cases} \text{ as } \cdot \quad (7)$$

$\langle\!\langle$ | represents the physical state, α, β represents U or V or N particle, n and n' are the number of θ particles. The S-matrix element can then be shown to be

$$S_{\alpha\beta} \equiv \langle\!\langle \beta; n'_{\text{out}} | \alpha; n_{\text{in}}\rangle\!\rangle = \delta_{\alpha\beta}\delta_{nn'} + \frac{2\pi i}{(n!n'!)^{1/2}}\delta\left(\sum_{\nu=1}^{n} w_\nu - \sum_{\nu=1}^{n'} w'_\nu\right) \times \left(\sum_{\nu=1}^{n} w_\nu - \sum_{\nu=1}^{n'} w'_\nu\right)^2 \times [\tau_{\alpha\beta}(W)]|_{W=m+\sum_{\nu=1}^{n} w_\nu}, \quad (8)$$

where the tau function $\tau_{\alpha\beta}(W)$ is the Fourier transform of $\tau_{\alpha\beta}(t)$,

$$\tau_{\alpha\beta}(W) \equiv \frac{1}{i}\int_{-\infty}^{\infty} dt e^{iWt}\tau_{\alpha\beta}(t), \quad (9)$$

and

$$\tau_{\alpha\beta}(t) \equiv \left\langle 0 | T\left(\psi_\beta(t)\prod_{\nu=1}^{n'} a'_{k_\nu}(t)\psi_\alpha^+(0)\prod_{\nu=1}^{n} a^+_{k_\nu}(0)\right) | 0\right\rangle. \quad (10)$$

III. ($V\theta$) BOUND STATE IN THE B MODEL

In the following presentation, we will make use of the results of the V sector ($q_1 = 1, q_2 = 0$) and $V-\theta$ sector ($q_1 = 1, q_2 = 1$) solved by the dispersion theory,[1,3] and some of our results by the LSZ formalism which are listed in Appendix A.

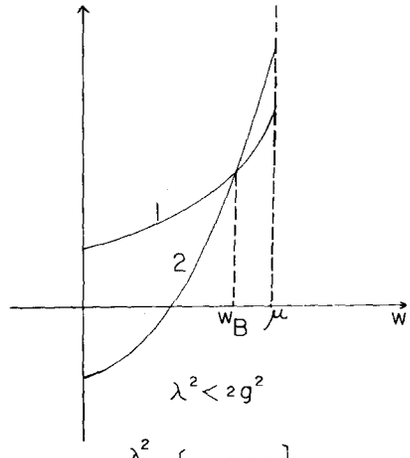
To study the existence of a ($V\theta$) bound state in the B model, we write down the $V + \theta$ elastic scattering amplitude which has been shown to be

$$S^{V\theta}_{kk'} = \delta_{kk'} + 2\pi i\frac{u^2(w)g^2}{2wh(w)} \times \frac{(2g^2 - \lambda^2)[1 + h(w)A(w)] - \lambda^2[1 - \beta(w)]}{\{(2g^2 - \lambda^2)[1 - h(w)A(w)] + \lambda^2[1 - \beta(w)]\}}, \quad (11)$$

where

$$h(w) \equiv w[1 - \beta(w)],$$
$$[1 - \beta(w) \equiv 1 + \frac{g^2 w}{4\pi^2}\int_\mu^\infty \frac{dw' u^2(w')(w'^2 - \mu^2)^{1/2}}{w'^2(w' - w - i\epsilon)} \quad (12)$$

and



Curve 1: $\frac{\lambda^2}{2g^2}\left[1 - \beta(w)\right]$

Curve 2: $\left(\frac{\lambda^2}{2g^2} - 1\right)\left[1 - h(w)A(w)\right]$

FIG. 1. Diagrams of $\lambda^2[1 - \beta(w)]/2g^2$ and $[(\lambda^2 - 2g^2)/2g^2][1 - h(w)A(w)]$ for $\lambda^2 < 2g^2$.

$$A(w) \equiv -I_w(w),$$

$$I_W(z) \equiv \frac{1}{\pi}\int_\mu^\infty dw' \,\mathrm{Im}\left(\frac{1}{h(w')}\right)\times\frac{1}{(w'-z)[1-\beta(W-w')]}. \tag{13}$$

Or, in terms of LSZ formalism, $S^{V\theta}_{k\,k'}$ can be derived as

$$S^{V\theta}_{k\,k'} = \delta_{k\,k'} + 2\pi i\,\frac{u^2(w)}{2w}\,\delta(w-w')(w-w')^2 g^2 \times\left[\frac{X^-(W,w,w')}{h(W-w)h(W-w')}\right]\Bigg|_{W=w}. \tag{14}$$

With the solution of $X^-(W,w,w')$ given in Eq. (A9), we can reobtain Eq. (11). From Eq. (12), we find

$$\left.\begin{array}{llll} 1-\beta(w) & >0, & h(w) & \geq 0 \\ 1-\beta(0) & =1, & h(0) & =0 \\ \dfrac{d}{dw}[1-\beta(w)] & >0, & \dfrac{d}{dw}h(w) & >0 \\ \dfrac{d^2}{dw^2}[1-\beta(w)] & >0, & \dfrac{d^2}{dw^2}h(w) & >0 \end{array}\right\}, \quad 0<w<\mu. \tag{15}$$

From Eq. (13), we find

$$\left.\begin{array}{ll} A(w) & >0 \\ \dfrac{d}{dw}A(w) & >0 \\ \dfrac{d^2}{dw^2}A(w) & >0 \end{array}\right\}, \quad 0<w<\mu. \tag{16}$$

Hence, $1-\beta(w)$ and $h(w)$ are real and monotonically increasing functions of w for $0<w<\mu$.

Moreover, it has been shown[7] that

$$h(\mu)A(\mu) > 1 \quad \text{(cf. Fig. 1)} \tag{17}$$

for sufficiently large coupling constant g.

Through this analysis we see that if $\lambda^2 < 2g^2$, $S^{V\theta}_{k\,k'}$ in Eq. (11) will have a pole at w_B where $0<w_B<\mu$, provided that the following condition is satisfied:

$$(\lambda^2/2g^2)[1-\beta(\mu)] < [(\lambda^2/2g^2)-1][1-h(\mu)A(\mu)]. \tag{18}$$

The occurence of this pole w_B in $S^{V\theta}_{k\,k'}$ corresponds to the existence of a $(V\theta)$ bound state in the B model with mass M_B, where

$$M_B \equiv m + w_B, \quad 0<w_B<\mu. \tag{19}$$

In order to study the higher order scattering processes and the ones involving the $(V\theta)$ bound state, we investigate the V-2θ sector in the next section.

IV. V-2θ SECTOR

The V-2θ is characterized by $q_1 = 1$ and $q_2 = 2$. The appropriate tau functions are the following[8]:

$$\tau_6(t,w,w') = \frac{(4ww')^{1/2}}{u(w)u(w')} \times\langle 0|\,T(\psi_U(t)a_k(t)\psi^+_U(0)a^+_{k'}(0)\,|0\rangle, \tag{20a}$$

$$\tau_7(t,w,w',w'') = \frac{(8ww'w'')^{1/2}}{u(w)u(w')u(w'')} \times\langle 0|\,T(\psi_V(t)a_k(t)a_{k'}(t)\psi^+_U(0)a^+_{k''}(0))\,|0\rangle, \tag{20b}$$

$$\tau_8(t,w,w',w'') = \frac{(8ww'w'')^{1/2}}{u(w)u(w')u(w'')} \times\langle 0|\,T(\psi_U(t)a_k(t)\psi^+_V(0)a^+_{k'}(0)a^+_{k''}(0))\,|0\rangle, \tag{20c}$$

$$\tau_9(t,w,w',w'',w''') = \frac{(16ww'w''w''')^{1/2}}{u(w)u(w')u(w'')u(w''')} \times\langle 0|\,T(\psi_N(t)a_k(t)a_{k'}(t)a_{k''}(t)\psi^+_U(0)a^+_{k'''}(0))\,|0\rangle, \tag{20d}$$

$$\tau_{10}(t,w,w',w'',w''') = \frac{(16ww'w''w''')^{1/2}}{u(w)u(w')u(w'')u(w''')} \times\langle 0|\,T(\psi_U(t)a_k(t)\psi^+_N(0)a^+_{k'}(0)a^+_{k''}(0)a^+_{k'''}(0))\,|0\rangle \tag{20e}$$

and

$$\tau^9(t,w_1,w_2,w_3,w_4) = \prod_{i=1}^{4}\frac{(2w_i)^{1/2}}{u(w_i)} \times\langle 0|\,T(\psi_V(t)a_{k_1}(t)a_{k_2}(t)\psi^+_V(0)a^+_{k_3}(0)a^+_{k_4}(0))\,|0\rangle, \tag{21a}$$

$$\tau^{10}(t,w_1,w_2,w_3,w_4,w_5) = \prod_{i=1}^{5}\frac{(2w_i)^{1/2}}{u(w_i)} \times\langle 0|\,T(\psi_N(t)a_{k_1}(t)a_{k_2}(t)a_{k_3}(t)\psi^+_V(0)a^+_{k_4}(0)a^+_{k_5}(0))\,|0\rangle, \tag{21b}$$

$$\tau^{11}(t,w_1,w_2,w_3,w_4,w_5) = \prod_{i=1}^{5}\frac{(2w_i)^{1/2}}{u(w_i)} \times\langle 0|\,T(\psi_V(t)a_{k_1}(t)a_{k_2}(t)\psi^+_N(0)a^+_{k_3}(0)a^+_{k_4}(0)a^+_{k_5}(0))\,|0\rangle, \tag{21c}$$

$$\tau^{12}(t,w_1,w_2,w_3,w_4,w_5,w_6) = \prod_{i=1}^{6}\frac{(2w_i)^{1/2}}{u(w_i)}\langle 0|\,T(\psi_N(t)a_{k_1}(t)a_{k_2} \times(t)a_{k_3}(t)\psi^+_N(0)a^+_{k_4}(0)a^+_{k_5}(0)a^+_{k_6}(0))\,|0\rangle. \tag{21d}$$

With the aid of field equations in Eq. (3) and commutation relations in Eq. (2), we can derive the following equations:

$$(W-m-\delta m_U-w)\tau_6(W,w,w') = \frac{2w\delta_{kk'}}{u^2(w)Z_U} + \frac{\lambda Z_V}{Z_U}\sum_{k''}\frac{u^2(w'')}{2w''}\,\tau_7(W,w'',w,w'), \tag{22a}$$

$$(W-m-\delta m_V-w-w')\tau_7(W,w,w',w'') = \frac{g}{Z_V}\sum_{k'''}\frac{u^2(w''')}{2w'''}\,\tau_9(W,w'''.w,w',w'') + \lambda[\tau_6(W,w',w'')+\tau_6(W,w,w'')], \tag{22b}$$

$$(W-m-\delta m_V-w'-w'')\tau_8(W,w,w',w'') = \frac{g}{Z_V}\sum_{k'''}\frac{u^2(w''')}{2w'''}\,\tau_{10}(W,w,w',w'',w''') + \lambda[\tau_6(W,w,w'')+\tau_6(W,w,w')], \tag{22c}$$

$$(W-m-w-w'-w'')\tau_9(W,w,w',w'',w''') = g[\tau_7(W,w',w'',w''')+\tau_7(W,w,w'',w''') + \tau_7(W,w,w',w''')], \tag{22d}$$

$$(W-m-w'-w''-w''')\tau_{10}(W,w,w',w'',w'''),$$
$$= g[\tau_8(W,w,w',w'') + \tau_8(W,w,w',w''') + \tau_8(W,w,w'',w''')] \quad (22e)$$

and

$$(W-m-\delta m_V - w_1 - w_2)\tau^9(W,w_1,w_2,w_3,w_4)$$
$$= \frac{4w_1w_2}{u^2(w_1)u^2(w_2)Z_V}[\delta_{k_1k_3}\delta_{k_2k_4} + \delta_{k_1k_4}\delta_{k_2k_3}]$$
$$+ \lambda[\tau_8(W,w_2,w_3,w_4) + \tau_8(W,w_1,w_3,w_4)]$$
$$+ \frac{g}{Z_V}\sum_k \frac{u^2(w)}{2w}\tau^{10}(W,w,w_1,w_2,w_3,w_4), \quad (23a)$$

$$(W-m-w_1-w_2-w_3)\tau^{10}(W,w_1,w_2,w_3,w_4,w_5)$$
$$= g[\tau^9(W,w_2,w_3,w_4,w_5) + \tau^9(W,w_1,w_3,w_4,w_5) + \tau^9(W,w_1,w_2,w_4,w_5)], \quad (23b)$$

$$(W-m-w_3-w_4-w_5)\tau^{11}(W,w_1,w_2,w_3,w_4,w_5)$$
$$= g[\tau^9(W,w_1,w_2,w_3,w_4) + \tau^9(W,w_1,w_2,w_3,w_5) + \tau^9(W,w_1,w_2,w_4,w_5)], \quad (23c)$$

$$(W-m-w_1-w_2-w_3)\tau^{12}(W,w_1,w_2,w_3,w_4,w_5,w_6)$$
$$= \frac{8w_1w_2w_3}{u^2(w_1)u^2(w_2)u^2(w_3)}\sum_{Pfi}\delta^3_{fi}$$
$$+ g[\tau^{11}(W,w_1,w_2,w_4,w_5,w_6) + \tau^{11}(W,w_1,w_3,w_4,w_5,w_6) + \tau^{11}(W,w_2,w_3,w_4,w_5,w_6)], \quad (23d)$$

where

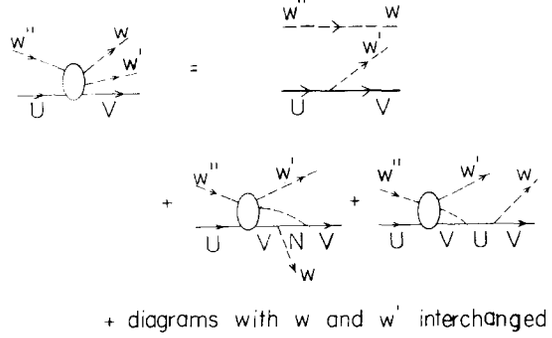
$$\sum_{Pfi}\delta^3_{fi} \equiv \delta_{k_1k_4}\delta_{k_2k_5}\delta_{k_3k_6} + \delta_{k_1k_4}\delta_{k_2k_6}\delta_{k_3k_5} + \delta_{k_1k_5}\delta_{k_2k_4}\delta_{k_3k_6} + \delta_{k_1k_5}\delta_{k_2k_6}\delta_{k_3k_4} + \delta_{k_1k_6}\delta_{k_2k_4}\delta_{k_3k_5} + \delta_{k_1k_6}\delta_{k_2k_5}\delta_{k_3k_4}. \quad (24)$$

Due to the coupling between the tau functions we find that the entire sector is solved if we can solve the tau functions $\tau_7(-)$ and $\tau^9(-)$.

The integral equation for $\tau_7(-)$ can be obtained by substituting Eqs. (22a) and (22d) into Eq. (22b):

$$h(W-w-w')\tau_7(W+m,w,w',w'')$$
$$= \frac{\lambda Z_V 2w''(\delta_{k\,k''} + \delta_{k'k''})}{Z_U u^2(w'')(W-\delta m_U - w'')}$$
$$+ \frac{1}{\pi}\int_\mu^\infty dw'''\,\mathrm{Im}h(w''')\tau_7(W+m,w''',w',w'')$$
$$\times\left(\frac{1}{W-w-w'-w'''} + \frac{\lambda^2 Z_V^2}{Z_U g^2(W-\delta m_U - w')}\right)$$
$$+ \frac{1}{\pi}\int_\mu^\infty dw'''\,\mathrm{Im}h(w''')\tau_7(W+m,w''',w,w'')$$
$$\times\left(\frac{1}{W-w-w'-w'''} + \frac{\lambda^2 Z_V^2}{Z_U g^2(W-\delta m_U - w)}\right) \quad (25)$$

(cf. Fig. 2).



+ diagrams with w and w' interchanged

FIG. 2. Diagram of the integral equation for $\tau_7(W,w,w',w'')$.

By defining

$$Y^-(W,w,w',w'') \equiv \lim_{z\to w-i\epsilon} Y(W,z,w',w'')$$
$$\equiv h(W-w-w')\tau_7(W+m,w,w',w'') \quad (26)$$

and continuing w into complex z plane, Eq. (25) becomes

$$Y(W,z,w',w'') = \frac{\lambda Z_V 2w''(\delta_{k\,k''} + \delta_{k'k''})}{Z_U u^2(w'')(W-\delta m_U - w'')}$$
$$+ \frac{1}{\pi}\int_\mu^\infty \frac{dw'''\,\mathrm{Im}h(w''')}{h(W-w'-w''')}Y^-(W,w''',w',w'')$$
$$\times\left(\frac{1}{W-z-w'-w'''} + \frac{\lambda^2 Z_V^2}{Z_U g^2(W-\delta m_U - w')}\right)$$
$$+ \frac{1}{\pi}\int_\mu^\infty \frac{dw'''\,\mathrm{Im}\,h(w''')}{h(W-z-w''')}Y^-(W,w''',z,w'')$$
$$\times\left(\frac{1}{W-z-w'-w'''} + \frac{\lambda^2 Z_V^2}{Z_U g^2(W-\delta m_U - z)}\right). \quad (27)$$

Notice that $Y^-(W,w,w',w'')$ in Eq. (27) is symmetric under the interchange of w and w'. We define

$$Y^-(W,w,w',w'') \equiv F^-(W,w,w',w'') + F^-(W,w',w,w''), \quad (28)$$

where

$$F(W,z,w',w'') = \frac{\lambda Z_V 2w''\delta_{k'k''}}{Z_U u^2(w'')(W-\delta m_U - w'')}$$
$$+ \frac{1}{\pi}\int_\mu^\infty \frac{dw'''\,\mathrm{Im}h(w''')}{h(W-w'-w''')}$$
$$\times\left(\frac{1}{W-z-w'-w'''} + \frac{\lambda^2 Z_V^2}{g^2 Z_U(W-\delta m_U - w')}\right)$$
$$\times Y^-(W,w''',w',w''). \quad (29)$$

Substituting Eq. (28) into Eq. (29), we get an integral equation for $F(W,z,w',w'')$,

$$F(W,z,w',w'') = \frac{\lambda Z_V 2w''\delta_{k'k''}}{Z_U u^2(w'')(W-\delta m_U - w'')}$$
$$+ \frac{1}{\pi}\int_\mu^\infty \frac{dw'''\,\mathrm{Im}h(w''')}{h(W-w'-w''')}$$
$$\times\left[\frac{1}{W-z-w'-w'''} + \frac{\lambda^2 Z_V^2}{Z_U g^2(W-\delta m_U - w')}\right]$$
$$\times[F^-(W,w''',w',w'') + F^-(W,w',w''',w'')]. \quad (30)$$

We can read off the analytic structure of $F(W,z,w'w'')$ as a function of z for fixed W,w', and w''. It has a branch cut along the real axis for $-\infty \le z \le W-w'-w$.

The singular integral equation (30) is very similar to Eq. (22) obtained in Ref. 4. Due to the complicated

structure of the tau functions involved in this sector, also of those in the higher sectors, a closed solution is not possible. We use our iterative expansion technique to solve this integral equation.

Starting by introducing a parameter α in front of the last term of Eq. (30), we get

$$F(W,z,w',w'') = \frac{\lambda Z_V 2w''\delta_{k'k''}}{Z_U u^2(w'')(W-\delta m_U - w'')} + \frac{1}{\pi}\int_\mu^\infty \frac{dw'''\ \mathrm{Im}h(w''')}{h(W-w'-w''')} F^-(W,w''',w',w'') \times\left(\frac{1}{W-z-w'-w'''} + \frac{\lambda^2 Z_V^2}{Z_U g^2(W-w'-\delta m_U)}\right) + \frac{\alpha}{\pi}\int_\mu^\infty \frac{dw'''\ \mathrm{Im}h(w''')}{h(W-w'-w''')} F^-(W,w',w''',w'') \times\left(\frac{1}{W-z-w'-w'''} + \frac{\lambda^2 Z_V^2}{Z_U g^2(W-\delta m_U-w')}\right). \quad (31)$$

Next, expressing $F(-)$ as a power series in α, i.e.,

$$F(W,z,w',w'') = \sum_{n=0}^{\infty} \alpha^n F_n(W,z,w',w''), \quad (32)$$

substituting Eq. (32) into Eq. (31), and equating terms with the same order in α, we get

$$F_0(W,z,w',w'') = \frac{\lambda Z_V 2w''\delta_{k'k''}}{Z_U u^2(w'')(W-\delta m_U - w'')} + \frac{1}{\pi}\int_\mu^\infty \frac{dw'''\ \mathrm{Im}h(w''')}{h(W-w'-w''')} \times\left(\frac{1}{W-z-w'-w'''} + \frac{\lambda^2 Z_V^2}{Z_U g^2(W-\delta m_U-w')}\right) \times F_0^-(W,w''',w',w''), \quad (33)$$

$$F_1(W,z,w',w'') = \frac{1}{\pi}\int_\mu^\infty \frac{dw'''\ \mathrm{Im}h(w''')}{h(W-w'-w''')} \times\left(\frac{1}{W-z-w'-w'''} + \frac{\lambda^2 Z_V^2}{Z_U g^2(W-\delta m_U-w')}\right) \times [F_1^-(W,w''',w',w'') + F_0^-(W,w',w''',w'')]. \quad (34)$$

$$F_n(W,z,w',w'') = \frac{1}{\pi}\int_\mu^\infty \frac{dw'''\ \mathrm{Im}h(w''')}{h(W-w'-w''')} \times\left(\frac{1}{W-z-w'-w'''} + \frac{\lambda^2 Z_V^2}{Z_U g^2(W-\delta m_U-w')}\right) \times [F_n^-(W,w''',w',w'') + F_{n-1}^-(W,w',w''',w'')]. \quad (35)$$

At the end of calculation, we set $\alpha = 1$, then Eq. (31) reduced back to the original form in Eq. (30).

From Eqs. (33)–(35), it is obvious that the analytic structure and symmetry properties will be preserved in the solutions.

To solve for $F_0(-)$, notice that Eq. (33) is an equation of the variable z for fixed W, w', and w''. Comparing it with the equation for the vertex function $\tau_2(W,z)$ in Eq. (A3), we find

$$F_0^-(W,w,w',w'') = \frac{2w'\delta_{k'k''}}{u^2(w')} h(W-w-w')\tau_2(W-w'+m,w). \quad (36)$$

In Appendix B, Eq. (34) is solved. The result is

$$F_1(W,z,w',w'') = \frac{1}{\pi}\int_\mu^\infty \frac{dw_1\ \mathrm{Im}h(w_1)}{h(W-w_1-w')} \times X(W-w',z,w_1)F_0^-(W,w',w_1,w''). \quad (B4)$$

Substituting Eq. (36) into Eq. (B4), we get

$$F_1^-(W,w,w',w'') = g^2 X^-(W-w',w,w'')\tau_2(W-w''+m,w'). \quad (37)$$

In analogy with equation for $F_1(-)$ in (34), we can write down the solution for the general term in the series F_n in (35) as

$$F_n(W,z,w',w'') = \frac{1}{\pi}\int_\mu^\infty \frac{dw_1\ \mathrm{Im}h(w_1)}{h(W-w_1-w')} \times X(W-w',z,w_1)F_{n-1}^-(W,w',w_1,w''). \quad (38)$$

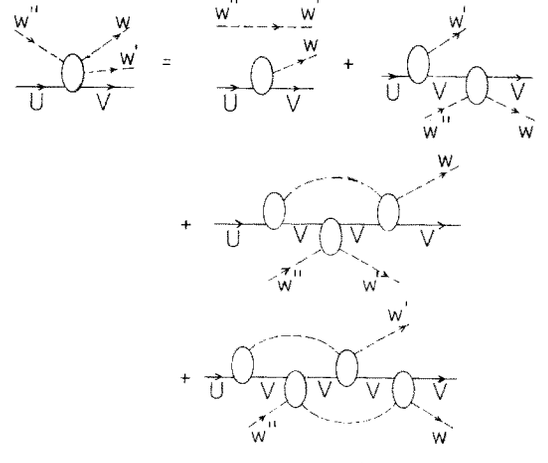
Substituting Eq. (37) into Eq. (38), we obtain the solution for $F_2(-)$. Analogously, we can get F_3 and all higher order terms in the series. The result of F_2 and F_3 are

$$F_2^-(W,w,w',w'') = \frac{g^2}{\pi}\int_\mu^\infty \frac{dw_1\ \mathrm{Im}h(w_1)}{h(W-w_1-w')} \times X^-(W-w',w,w_1)X^-(W-w_1,w',w'') \times \tau_2(W-w''+m,w_1). \quad (39)$$

$$F_3^-(W,w,w',w'') = \frac{g^2}{\pi^2}\int_\mu^\infty\int_\mu^\infty \frac{dw_1 dw_2\ \mathrm{Im}h(w_1)\ \mathrm{Im}h(w_2)}{h(W-w'-w_2)h(W-w_1-w_2)} \times X^-(W-w',w,w_2)X^-(W-w_2,w',w_1) \times X^-(W-w_1,w_2,w'')\tau_2(W-w''+m,w_1). \quad (40)$$

With the aid of Eqs. (26), (28), (36), (37), (39), and (40) the first four order terms of $\tau_7(-)$ are (cf. Fig. 3)

$$\tau_7(W+m,w,w',w'') = \frac{2w'}{u^2(w')}\delta_{k'k''}\tau_2(W-w'+m,w) + \frac{g^2 X^-(W-w',w,w'')\tau_2(W-w''+m,w')}{h(W-w-w')}$$



+ diagrams with w and w' interchanged

FIG. 3. Diagrams of the first four terms in $\tau_7(W+m,w,w',w'')$.

$$+\frac{g^2}{\pi}\int_\mu^\infty \frac{dw_1\ \mathrm{Im}h(w_1)X^-(W-w',w,w_1)}{h(W-w'-w_1)h(W-w-w')}$$
$$\times X^-(W-w_1,w',w'')\tau_2(W-w''+m,w_1)$$
$$+\frac{g^2}{\pi^2}\int_\mu^\infty\int_\mu^\infty \frac{dw_1dw_2\ \mathrm{Im}h(w_1)\ \mathrm{Im}h(w_2)}{h(W-w'-w_2)}$$
$$\times \frac{X^-(W-w',w,w_2)X^-(W-w_2,w',w_1)}{h(W-w_1-w_2)h(W-w-w')}$$
$$\times X^-(W-w_1,w_2,w'')\tau_2(W-w''+m,w_1) \qquad (41)$$

+ terms with w and w' interchanged.

From Eq. (41), we see that each term in the iterative solution contains the function $X^-(W-w,w'w'')$ which has a pole at the $(V\theta)$ bound state for suitable chosen parameters [cf. Eqs. (11) and (14)]. This confirms that our iterative expansion technique preserves the properties of the bound state.

Substituting the solution of $\tau_7(-)$ in (41) into Eq. (22), we get the functions $\tau_6(-)$, $\tau_8(-)$, $\tau_9(-)$, and $\tau_{10}(-)$.

To complete the solution for this sector, we will now solve the integral equation for $\tau^9(-)$. Substituting Eq. (23b) into Eq. (23a), we get

$$h(W-w_1-w_2)\tau^9(W+m,w_1,w_2,w_3,w_4)$$
$$=\frac{4w_1w_2}{u^2(w_1)u^2(w_2)}(\delta_{k_1k_3}\delta_{k_2k_4}+\delta_{k_1k_4}\delta_{k_2k_3})$$
$$+\lambda Z_V[\tau_8(W+m,w_2,w_3,w_4)+\tau_8(W+m,w_1,w_3,w_4)]$$
$$+\frac{1}{\pi}\int_\mu^\infty \frac{dw\ \mathrm{Im}h(w)}{(W-w-w_1-w_2)}[\tau^9(W+m,w,w_2,w_3,w_4)$$
$$+\tau^9(W+m,w,w_1,w_3,w_4)]. \qquad (42)$$

Following the same reason, we again solve Eq. (42) by the iterative expansion method.

The algebra is outlined in Appendix C. The result is (cf. Fig. 4)

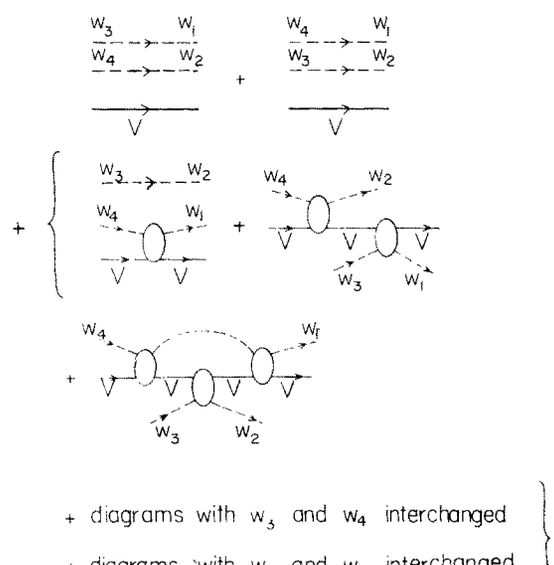


FIG. 4. Diagrams of the first four terms in $\tau^9(W+m,w_1,w_2,w_3,w_4)$.

$$\tau^9(W+m,w_1,w_2,w_3,w_4)$$
$$=\frac{4w_1w_2(\delta_{k_1k_3}\delta_{k_2k_4}+\delta_{k_1k_4}\delta_{k_2k_3})}{u^2(w_1)u^2(w_2)h(W-w_1-w_2)}$$
$$+\frac{1}{h(W-w_1-w_2)h(W-w_3-w_4)}$$
$$\times\Big\{\frac{2w_2}{u^2(w_2)}\delta_{k_2k_3}g^2X^-(W-w_2,w_1,w_4)$$
$$+\frac{g^4X^-(W-w_2,w_1,w_3)X^-(W-w_3,w_2,w_4)}{h(W-w_2-w_3)}$$
$$+\frac{g^4}{\pi}\int_\mu^\infty \frac{dw\ \mathrm{Im}h(w)X^-(W-w_2,w_1,w)X^-(W-w,w_2,w_3)}{h(W-w_2-w_3)h(W-w-w_2)}$$
$$\times X^-(W-w_3,w,w_4)$$
$$+\text{terms with } w_3 \text{ and } w_4 \text{ interchanged}$$
$$+\text{terms with } w_1 \text{ and } w_2 \text{ interchanged}\Big\}. \qquad (43)$$

By using Eqs. (23b)–(23d) the solution to the tau functions $\tau^{10}(-)$, $\tau^{11}(-)$, and $\tau^{12}(-)$ can be obtained.

With solutions of all the tau functions in the $V-2\theta$ sector which are new results, we are ready to investigate the S matrix for all relevant processes.

V. S-MATRIX ELEMENTS

In the $V-2\theta$ sector, the nonbound-state scattering processes are

(1) $U+\theta_{\kappa'}\to U+\theta_\kappa$,
(2) $U+\theta_{\kappa''}\leftrightarrow V+\theta_\kappa+\theta_{\kappa'}$,
(3) $U+\theta_{\kappa'''}\leftrightarrow N+\theta_\kappa+\theta_{\kappa'}+\theta_{\kappa''}$,
(4) $V+\theta_{\kappa_3}+\theta_{\kappa_4}\to V+\theta_{\kappa_1}+\theta_{\kappa_2}$,
(5) $V+\theta_{\kappa_4}+\theta_{\kappa_5}\leftrightarrow N+\theta_{\kappa_1}+\theta_{\kappa_2}+\theta_{\kappa_3}$.
(6) $N+\theta_{\kappa_4}+\theta_{\kappa_5}+\theta_{\kappa_6}\to N+\theta_{\kappa_1}+\theta_{\kappa_2}+\theta_{\kappa_3}$.

Using the reduction formula in (8) and definition of the tau functions in (20)–(21), the corresponding S-matrix elements can be derived to be

$$(1)\ S^{U\theta\to U\theta}_{k'k}=\delta_{kk'}+2\pi i\delta(w-w')[u^2(w)/2w](w-w')^2$$
$$\times\tau_6(W+m,w,w')|_{W=w}, \qquad (44)$$

$$(2)\ S^{U\theta\leftrightarrow V\theta\theta}_{k''k'k}=2\pi i\delta(w''-w-w')\frac{u(w)u(w')u(w'')}{(8ww'w'')^{1/2}}$$
$$\times\frac{(w''-w-w')^2}{\sqrt{2}}\tau_7(W+m,w,w',w'')|_{W=w''}, \qquad (45)$$

$$(3)\ S^{U\theta\to N\theta\theta\theta}_{k'''k''k'k}=2\pi i\delta(w'''-w-w'-w'')$$
$$\times\frac{u(w)u(w')u(w'')u(w''')}{(16ww'w''w''')^{1/2}}\frac{(w'''-w-w'-w'')^2}{\sqrt{6}}$$
$$\times\tau_9(W+m,w,w',w'',w''')|_{W=w'''} \qquad (46)$$

$$(4)\ S^{V\theta\theta\to V\theta\theta}_{k_4k_3k_2k_1}=\tfrac{1}{2}(\delta_{k_1k_3}\delta_{k_2k_4}+\delta_{k_1k_4}\delta_{k_2k_3})+2\pi i\delta$$
$$\times(w_1+w_2-w_3-w_4)\tfrac{1}{2}(w_1+w_2-w_3-w_4)^2$$
$$\times\prod_{i=1}^{4}\frac{u(w_i)}{(2w_i)^{1/2}}\tau^9(W+m,w_1,w_2,w_3,w_4)|_{W=w_1+w_2}. \qquad (47)$$

$$(5)\ S^{V\theta\theta\leftrightarrow N\theta\theta\theta}_{k_5k_4k_3k_2k_1} = \pi i\delta(w_4+w_5-w_1-w_2-w_3) \times \frac{(w_4+w_5-w_1-w_2-w_3)^2}{\sqrt{3}} \prod_{i=1}^{5} \frac{u(w_i)}{(2w_i)^{1/2}}\, \tau^{10} \times (W+m, w_1, w_2, w_3, w_4, w_5)\,|_{W=w_4+w_5}, \quad (48)$$

$$(6)\ S^{N\theta\theta\theta\to N\theta\theta\theta}_{k_6k_5k_4k_3k_2k_1} = \frac{1}{6}\sum_{Pfi}\delta^3_{fi} + 2\pi i\delta(w_4+w_5+w_6-w_1-w_2-w_3) \times \frac{1}{6}(w_4+w_5+w_6-w_1-w_2-w_3)^2 \prod_{i=1}^{6} \frac{u(w_i)}{(2w_i)^{1/2}} \times \tau^{12}(W+m, w_1, w_2, w_3, w_4, w_5, w_6)\,|_{W=w_1+w_2+w_3}. \quad (49)$$

The specific solution of all these S-matrix elements can be obtained straightforwardly by using Eqs. (22), (23) and the results of $\tau_7(-)$ and $\tau^9(-)$ in Eqs. (44) and (43). For example, the S matrix for $U+\theta$ elastic scattering is (cf. Fig. 5)

$$S^{U\theta\to U\theta}_{k'k} = \delta_{kk'} + 2\pi i\delta(w-w')\frac{u^2(w)}{2w}\left\{\frac{\lambda^2[1-\beta(w)]}{w} - \frac{\lambda^2}{\pi}\int_\mu^\infty \frac{dw_1\ \mathrm{Im}[1-\beta(w_1)]}{w_1} X^-(w-w_1, w, w) + \cdots\right\} \quad (50)$$

We will now compare our new result for $U+\theta$ elastic scattering amplitude with that obtained by Bronzan.[3] The function $T^3_{22}(w)$ calculated by Bronzan's dispersion approach corresponds to the on-mass shell value of τ_6 in our LSZ formalism. His result of $T^3_{22}(w)$ involves four functionals F_i which are functions of an unknown function ϕ. Instead of getting the final result, Bronzan conjectured a trial solution for ϕ. We pursued his conjecture and found that his method of calculation is too complicated and very impracticable. Nevertheless, we can show that as $g\to 0$, to the lowest order in λ, $T^3_{22}(w)$ reduces to

$$T^3_{22}(w) = \lambda^2/w \quad \text{(dispersion approach with the variational principle method)}, \quad (51)$$

and our result for $S^{U\theta}_{kk'}\to U\theta$ in Eq. (50) reduces to

$$S^{U\theta\to U\theta}_{kk'} = \delta_{kk'} + 2\pi i\delta(w-w')u^2(w)\lambda^2/2w^2$$

(LSZ formalism approach with iterative expansion method). (52)

Since the relation between S matrix and T is

$$S^{U\theta\to U\theta}_{kk'} = \delta_{kk'} + 2\pi i\delta(w-w')[u^1(w)/2w]T^3_{22}(w). \quad (53)$$

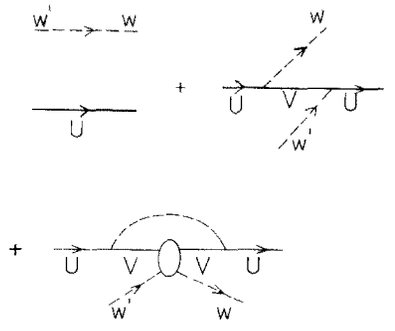


FIG. 5. Diagrams of the first three lowest order terms of the S matrix for $U+\theta_{k'}\to U+\theta_k$.

We see that to the lowest order in the coupling constant result of $U+\theta$ elastic scattering amplitude agrees with that obtained by Bronzan.

We now consider the bound state scattering processes in the V–2θ sector. They are

$$(V\theta)+\theta\to(V\theta)+\theta,$$
$$(V\theta)+\theta\leftrightarrow V+\theta+\theta,$$
$$(V\theta)+\theta\leftrightarrow U+\theta,$$
$$(V\theta)+\theta\leftrightarrow N+\theta+\theta+\theta.$$

The $(V\theta)$ bound state field operator $B_0(t)$ can be constructed as[4,9]

$$B_0(t) = C_1\psi_U(t) + C_2\psi_V(t)\int d^3k a_k(t) + C_3\psi_N(t)\iint d^3k_1 d^3k_2 a_{k_1}(t)a_{k_2}(t), \quad (54)$$

where C_1, C_2, and C_3 are c-numbers.

The renormalized field operator $B(t)$ is defined as

$$B(t) = B_0(t)/\sqrt{Z_B}, \quad (55)$$

where

$$(Z_B)^{1/2} = \langle 0|B_0(0)|B\rangle\!\rangle \quad (56)$$

and $|B\rangle\!\rangle$ is the eigenstate of the $(V\theta)$ bound state.

Also, the asymptotic condition for the field $B(t)$ is assumed to be

$$\lim_{t\to\pm\infty}\langle\!\langle\alpha|e^{-iMt}B^+(t)|\beta\rangle\!\rangle = \langle\!\langle\alpha|B^+_{\substack{\text{in}\\\text{out}}}|\beta\rangle\!\rangle. \quad (57)$$

By using the asymptotic condition and the reduction formula, we can obtain all the S-matrix elements for the $(V\theta)$ bound state scattering processes listed above. With the aid of Eqs. (20)–(23), we see that the matrix elements can all be expressed in terms of $\tau_7(-)$ or $\tau^9(-)$. Due to the lengthy algebra, all the details will be left out here. We would refer the interested readers to our articles[4,5] on the $(V\theta)$ bound state scattering in the Lee model.

VI. CONCLUSION

From the study of the $V+\theta$ elastic scattering amplitude in the B model we concluded that a $(V\theta)$ bound state can exist in the B model for suitable chosen parameters. In the V–2θ sector, we investigated not only the multiparticle reactions but also various cases of scattering a θ particle off the $(V\theta)$ bound state.

The iterative solution preserved the analytic structure, symmetry properties and the properties of the $(V\theta)$ bound state.

By using the LSZ reduction formula and the iterative solutions of the tau functions, S-matrix elements for various relevant processes including both nonbound-state and bound-state scattering were obtained. Comparison of our new result for $U+\theta$ elastic scattering amplitude with that obtained by Bronzan's variational principle method was made. To the lowest order in the coupling constant, the two results agree with each other.

ACKNOWLEDGMENT

The author wishes to express her sincere thanks to Professor Robert L. Zimmerman, Department of

Physics and Institute of Theoretical Science, University of Oregon, for many valuable discussions and for his criticism of this manuscript.

APPENDIX A

The $U-V-\theta$ vertex function $\tau_2(W,w)$ is the Fourier transform of $\tau_2(t,w)$ defined as

$$\tau_2(W,w) \equiv \frac{1}{i}\int_{-\infty}^{\infty} dt\ e^{iWt}\tau_2(t,w), \tag{A1}$$

where

$$\tau_2(t,w) \equiv [(2w)^{1/2}/u(w)]\langle 0|T(\psi_V(t)a_k(t)\psi_U^+(0))|0\rangle. \tag{A2}$$

It can be shown that $\tau_2(W,w)$ satisfies the following equation:

$$h(W-z)\tau_2(W+m,z) = \frac{\lambda Z_V}{Z_U(W-\delta m_U)} + \frac{1}{\pi}\int_\mu^\infty dw'\ \mathrm{Im}h(w') \times \tau_2(W+m,w')\left(\frac{1}{W-w'-z} + \frac{\lambda^2 Z_V^2}{Z_U g^2(W-\delta m_U)}\right), \tag{A3}$$

where $h(w)$ is defined in Eq. (12).

Let $\tau^5(t,w,w')$ be the tau function corresponding to the process $V+\theta_\kappa \to V+\theta_{\kappa'}$:

$$\tau^5(t,w,w') \equiv [(4ww')^{1/2}/u(w)u(w')] \times \langle 0|T(\psi_V^+(t)a_k(t)\psi_V^+(0)a_{k'}^+(0))|0\rangle. \tag{A4}$$

Its Fourier transform is

$$\tau^5(W,w,w') = (1/i)\int_{-\infty}^{\infty} dt\ e^{iWt}\tau^5(t,w,w') \tag{A5}$$

By defining

$$X^-(W,w,w') \equiv \frac{2w}{u^2(w)}\frac{\delta_{kk'}}{g^2}\,h(W-w') + \frac{h(W-w)h(W-w')}{g^2}\,\tau^5(W+m,w,w'), \tag{A6}$$

the following equation can be derived:

$$X(W,z,w') = \frac{1}{W-z-w'} + \frac{\lambda Z_V}{g^2}\,\tau_2(W+m,w')h(W-w') + \frac{1}{\pi}\int_\mu^\infty \frac{dw''\ \mathrm{Im}h(w'')X^-(W,w'',w')}{h(W-w'')(W-z-w'')}. \tag{A7}$$

By using the method introduced by Maxon,[7] both Eqs. (A3) and (A7) can be solved to be

$$\tau_2(W,w) = \frac{2g^2\lambda\{1/(W-m-w) + [h(W-m)/w][I_{W-m}(W-m-w) - I_{W-m}(W-m)]\}}{(W-m)\{\lambda^2[1-\beta(W-m)] + (2g^2-\lambda^2)[1+h(W-m)I_{W-m}W-m)]\}} \tag{A8}$$

and

$$\begin{aligned} X^-(W,w,w') &= \lim_{z\to w-it} X(W,z,w') \\ &= \left[\frac{\lambda}{g^2}\,\tau_2(W+m,w')h(W-w') + [1-\beta(W-w')]\left(I_{W-w'}(w') - \frac{W-w'}{w'}I_{W-w'}(W-w')\,\frac{1}{h(w')[1-\beta(W-w')]}\right)\right] \\ &\quad\times \frac{\{[1-\beta(W-w)] + [h(W-w)h(W)/w][I_{W-w'}(W-w) - I_{W-w'}(W)]\}}{1+h(W)I_{W-w'}(W)} + \frac{h(W-w)}{h(w')(W-w-w'+i\epsilon)} \\ &\quad - \frac{h(W-w)I_{W-w'}(w')[1-\beta(W-w')]}{W-w-w'+i\epsilon} + \frac{h(W-w')h(W-w)I_{W-w'}(W-w')}{w'(w'-w+i\epsilon)} \\ &\quad + \frac{[1-\beta(W-w')](W-2w')[1-\beta(W-w)](w-W)^2 I_{W-w'}(W-w)}{(w'-w+i\epsilon)(w-W+w'-i\epsilon)w'}, \end{aligned} \tag{A9}$$

where $I_W(z)$ is defined in Eq. (13).

APPENDIX B

To solve Eq. (34) for $F_1(-)$, we define

$$\begin{aligned} R_1(W,z,w',w'') &\equiv \frac{F_1(W,z,w',w'') - C_1(W,w',w'')}{h(W-z-w')} \\ &\quad - \frac{1}{\pi}\int_\mu^\infty \frac{dw'''\ \mathrm{Im}h(w''')}{w'''[1-\beta(W-w'-z)} \\ &\quad\times \frac{F_0^-(W,w',w''',w'')}{h(W-w'-w''')(W-z-w'-w''')}, \end{aligned} \tag{B1}$$

where

$$C_1(W,w',w'') \equiv F_1(W,W-w+i\epsilon,w',w''). \tag{B2}$$

An integral equation for $R_1(W,z,w',w'')$ can be obtained by using Eqs. (34) and (B1):

$$\begin{aligned} &[1-\beta(W-z-w')R_1(W,z,w',w'') \\ &\quad = \frac{C_1(W,w',w'')}{\pi}\int_\mu^\infty \frac{dw'''\ \mathrm{Im}[1-\beta(w''')}{h(W-w'-w''')(W-z-w'-w''')} \\ &\quad - \frac{1}{\pi}\int_\mu^\infty \frac{dw'''\ \mathrm{Im}(1-\beta(w'''))}{[1-\beta(W-w'-w''')](z+w'+w'''-W)} \\ &\quad\times \int_\mu^\infty \frac{dw_1\ \mathrm{Im}[1-\beta(w_1)]F_0^-(W,w',w_1,w'')}{h(W-w'-w_1)(W-w'''-w'-w_1+i\epsilon)} \\ &\quad - \frac{1}{\pi}\int_\mu^\infty \frac{dw'''\ \mathrm{Im}[1-\beta(w''')]R_1^-(W,w''',w',w'')}{(z+w'+w'''-W)}. \end{aligned} \tag{B3}$$

Following Maxon's method in Ref. 7, Eq. (B3) can be solved and $C_1(W,w',w'')$ can be determined. Substituting both solutions of R_1 and C_1 into Eq. (B1), we finally have

$$F_1(W,z,w',w'') = \frac{1}{\pi}\int_\mu^\infty \frac{dw_1\ \mathrm{Im}h(w_1)}{h(W-w_1-w')} F_0^-(W,w',w_1,w'') \times X(W-w',z,w_1), \quad \text{(B4)}$$

where $X^-(W,w,w')$ is given in Eq. (A9).

APPENDIX C

We now solve the function $\tau^9(-)$ by the same iterative technique used to solve $\tau_7(-)$ in the text.

By defining

$$V^-(W,w_1,w_2,w_3,w_4) \equiv \lim_{z_1\to w_1^-i\epsilon} V(W,z_1,w_2,w_3,w_4)$$
$$= -\frac{4w_1w_2h(W-w_3-w_4)}{g^2u^2(w_1)u^2(w_2)}(\delta_{k_1k_3}\delta_{k_2k_4}+\delta_{k_1k_4}\delta_{k_2k_3}) + \frac{h(W-w_1-w_2)h(W-w_3-w_4)}{g^2} \times \tau^9(W+m,w_1,w_2,w_3,w_4), \quad \text{(C1)}$$

Eq. (42) becomes

$$V(W,z_1,w_2,w_3,w_4) = \frac{2w_2(\delta_{k_2k_3}+\delta_{k_2k_4})}{u^2(w_2)(W-z_1-w_3-w_4)} + \frac{2w_1(\delta_{k_1k_3}+\delta_{k_1k_4})}{u^2(w_1)(W-w_2-w_3-w_4)} + \frac{1}{\pi}\int_\mu^\infty \frac{dw\ \mathrm{Im}h(w)}{(W-w-z_1-w_2)}\left[\frac{V^-(W,w,w_2,w_3,w_4)}{h(W-w-w_2)} + \frac{V^-(W,w,z_1,w_3,w_4)}{h(W-w-z_1)}\right] + \frac{\lambda Z_V h(W-w_3-w_4)}{g^2} \times [\tau_7(W+m,w_3,w_4,w_5)+\tau_7(W+m,w_3,w_4,z_1)]. \quad \text{(C2)}$$

Due to its symmetric property in w_1 and w_2, we rewrite $V^-(W,w_1,w_2,w_3,w_4)$ as

$$V^-(W,w_1,w_2,w_3,w_4) = G^-(W,w_1,w_2,w_3,w_4) + G^-(W,w_2,w_1,w_3,w_4), \quad \text{(C3)}$$

where

$$G(W,z_1,w_2,w_3,w_4) = \frac{2w_2(\delta_{k_2k_3}+\delta_{k_2k_4})}{u^2(w_2)(W-z_1-w_3-w_4)} + \frac{\lambda Z_V}{g^2}h(W-w_3-w_4)\tau_7(W+m,w_3,w_4,w_2) + \frac{1}{\pi}\int_\mu^\infty \frac{dw\ \mathrm{Im}h(w)V^-(W,w,w_2,w_3,w_4)}{(W-w-z_1-w_2)(W-w-w_2)}. \quad \text{(C4)}$$

Substituting Eq. (C3) into (C4), we get an integral equation for $G(-)$,

$$G(W,z_1,w_2,w_3,w_4) = \frac{2w_2(\delta_{k_2k_3}+\delta_{k_2k_4})}{u^2(w_2)(W-z_1-w_3-w_4)} + \frac{\lambda Z_V h(W-w_3-w_4)}{g^2}\tau_7(W+m,w_3,w_4,w_2) + \frac{1}{\pi}\int_\mu^\infty \frac{dw\ \mathrm{Im}h(w)G^-(W,w,w_2,w_3,w_4)}{(W-w-z_1-w_2)h(W-w-w_2)} + \frac{1}{\pi}\int_\mu^\infty \frac{dw\ \mathrm{Im}h(w)G^-(W,w_2,w,w_3,w_4)}{(W-w-z_1-w_2)h(W-w-w_2)}. \quad \text{(C5)}$$

We now solve Eq. (C5) by its iterative expansion:

(1) Introduce a parameter α to the last term in Eq. (C5).

(2) Write a series expansion for G,

$$G^-(W,w_1,w_2,w_3,w_4) = \sum_n \alpha^n G_n^-(W,w_1,w_2,w_3,w_4). \quad \text{(C6)}$$

(3) From Eqs. (26), (28), and (32), we can write

$$\tau_7(-) = \sum_n [\tau_7(-)]_n \alpha^n. \quad \text{(C7)}$$

(4) Equating the terms of same order in α in Eq. (C5), we get

$$G_0(W,z_1,w_2,w_3,w_4) = \frac{2w_2(\delta_{k_2k_3}+\delta_{k_2k_4})}{u^2(w_2)(W-z_1-w_3-w_4)} + \frac{\lambda Z_V}{g^2}h(W-w_3-w_4) \times [\tau_7(W+m,w_3,w_4,w_2)]_0 + \frac{1}{\pi}\int_\mu^\infty \frac{dw\ \mathrm{Im}h(w)G_0^-(W,w,w_2,w_3,w_4)}{(W-w-z_1-w_2)h(W-w-w_2)}, \quad \text{(C8)}$$

$$G_1(W,z_1,w_2,w_3,w_4) = \frac{\lambda Z_V}{g^2}h(W-w_3-w_4)[\tau_7(W+m,w_3,w_4,w_2)]_1 + \frac{1}{\pi}\int_\mu^\infty dw\ \mathrm{Im}h(w) \times \frac{[G_1^-(W,w,w_2,w_3,w_4)+G_0^-(W,w_2,w,w_3,w_4)]}{(W-w-z_1-w_2)h(W-w-w_2)}, \quad \text{(C9)}$$

and

$$G_n^-(W,z_1,w_2,w_3,w_4) = \frac{\lambda Z_V}{g^2}h(W-w_3-w_4)[\tau_7(W+m,w_3,w_4,w_2)]_n + \frac{1}{\pi}\int_\mu^\infty dw\ \mathrm{Im}h(w) \times \frac{[G_n^-(W,w,w_2,w_3,w_4)+G_{n-1}^-(W,w_2,w,w_3,w_4)]}{h(W-w-w_2)(W-w-z_1-w_2)}, \quad \text{(C10)}$$

where [cf. Eq. (41)]

$$[\tau_7(W+m,w_3,w_4,w_2)]_0 = \frac{2w_2}{u^2(w_2)}\delta_{k_2k_4}\tau_2(W-w_2+m,w_3) + [2w_2/u^2(w_2)]\ \delta_{k_2k_3}\tau_2(W-w_2+m,w_4) \quad \text{(C8')}$$

and

$$[\tau_7(W+m,w_3,w_4,w_2)]_1 = \frac{g^2}{h(W-w_3-w_4)}\tau_2(W-w_2+m,w_4)X^-(W-w_4,w_3,w_2) + \frac{g^2}{h(W-w_3-w_4)}\tau_2(W-w_2+m,w_3)X^-(W-w_3,w_4,w_2). \quad \text{(C9')}$$

Notice that Eq. (C8) is an integral equation of variable z_1 for fixed W, w_2, w_3, and w_4. As we compare this equation with Eq. (A7), we find that

$$G_0^-(W,w_1,w_2,w_3,w_4) = \frac{2w_2\delta_{k_2k_3}}{u^2(w_2)}X^-(W-w_2,w_1,w_4) + \frac{2w_2\delta_{k_2k_4}}{u^2(w_2)}X^-(W-w_2,w_1,w_3). \quad \text{(C11)}$$

Using the same method for solving Eq. (34), $G_1(-)$ in Eq. (C9) can be solved to be

$$G_1^-(W,w_1,w_2,w_3,w_4) = \frac{1}{\pi}\int_\mu^\infty \frac{dw\,\mathrm{Im}h(w)}{h(W-w-w_2)}X^-(W-w_2,w_1,w) \times G_0^-(W,w_2,w,w_3,w_4). \quad \text{(C12)}$$

The general terms in the series expansion is then

$$G_n^-(W,w_1,w_2,w_3,w_4) = \frac{1}{\pi}\int_\mu^\infty \frac{dw\,\mathrm{Im}\,h(w)}{h(W-w-w_2)}X^-(W-w_2,w_1,w) \times G_{n-1}^-(W,w_2,w,w_3,w_4). \quad \text{(C13)}$$

The solution of $\tau^9(-)$ can then be obtained by using Eqs. (C1) and (C3) together with the results in Eqs. (C11)–(C13). The first few terms in the expansion solution for $\tau^9(-)$ are written in Eq. (43).

1 J. B. Bronzan, Phys. Rev. **139**, B751 (1965).
2 T. D. Lee, Phys. Rev. **95**, 1329 (1954); S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Harper & Row, New York, 1962), Sec. 12.
3 J. B. Bronzan, Phys. Rev. **172**, 1429 (1968).
4 T. H. Liu and R. L. Zimmerman, Suppl. Nuovo Cimento **6**, 1297 (1968).
5 T. H. Liu and R. L. Zimmerman, J. Math Phys. **11**, 1941 (1970).
6 M. S. Maxon and R. B. Curtis, Phys. Rev. **137**, B996 (1965).
7 M. S. Maxon, Phys. Rev. **149**, 1273 (1966).
8 In this article we use subscript on the tau functions which involve U particle and superscript on the ones without U particle.
9 R. L. Zimmerman, Nuovo Cimento **56A**, 391 (1968).

Two-Time Spin-Pair Correlation Function of the Heisenberg Magnet at Infinite Temperature. III. bcc Lattice and Arbitrary Spin

Tohru Morita*
Department of Physics, Ohio University, Athens, Ohio 45701

(Received 24 March 1972)

The expansion coefficients in powers of time of the two-time spin-pair correlation function are obtained up to terms of order t^8 for the bcc Heisenberg magnet of spin $\frac{1}{2}$ at infinite temperature and up to terms of order t^6 for the linear, square, sc, and bcc Heisenberg magnets of an arbitrary spin at infinite temperature. The result is applicable to the isotropic as well as the anisotropic Heisenberg magnet where the exchange integrals in the z direction and in the orthogonal plane are assumed to be different. Analysis of the results are given in separate articles.

1. INTRODUCTION

In the first paper of this series,[1] the numerical values of the short-time expansion coefficients of the two-time spin-pair correlation function of the Heisenberg magnet of spin $\frac{1}{2}$ at infinite temperature are provided up to $O(t^{10})$ for the linear chain and up to $O(t^8)$ for the square and sc lattices. The analysis of the results obtained by the expansion[2,3] have been found to reproduce the exact solution for the one-dimensional Heisenberg magnet of finite length of spin $\frac{1}{2}$,[4] and also the results of the computer simulation calculation for the Heisenberg magnet of classical spin,[5] and the neutron diffraction data for Rb Mn F_3,[5] with a suitable choice of time scale for different spins.

The only existing Heisenberg magnet of spin $\frac{1}{2}$ will be solid He^3, which crystalizes in the bcc structure. If one considers magnets with larger spin, one has Mn F_2 in which the magnetic ions constitute the bcc structure, and extensive experimental data are available. For the investigation of these systems, the short-time expansion coefficients are provided up to terms of $O(t^8)$ for the bcc Heisenberg magnet of spin $\frac{1}{2}$ and up to $O(t^6)$ for the linear, square, sc, and bcc magnets of an arbitrary spin in this paper.

The calculation of the short-time expansion coefficients $\sigma_{2l}^{(2n)}(R_{if})$ in the preceding paper was performed in two steps. In the first step, the expansion coefficients $\gamma_{2l}^{(2n)}$ (diagram IF) for small diagrams are computed. In the second step, the number of different ways $n(R_{if}$; diagram IF) by which the small diagrams with the initial and final sites I and F occur in the lattices with the difference of the final site and initial site R_{if} are obtained. The sum of the products of those two sets of numbers gives the coefficients $\sigma_{2l}^{(2n)}(R_{if})$ for the lattice under consideration. In order to obtain the coefficients of the term of $O(t^8)$ $\sigma_{2l}^{(8)}(R_{if})$ for the bcc lattice, we need the coefficients $\gamma_{2l}^{(8)}$ (diagram IF) for the diagram 10 of Fig. 1 in addition to those for diagram 1–9 which are given in Ref. 1. In Sec. 2 of the present paper, these coefficients for spin $\frac{1}{2}$ are provided, and then the computation of the values $n(R_{if}$; diagram IF) is described. Combining them with the values of $\gamma_{2l}^{(2n)}$ (diagram IF) given in the previous paper, we obtain the values of $\sigma_{2l}^{(2n)}(R_{if})$ for the bcc Heisenberg magnet of spin $\frac{1}{2}$. In Sec. 3, $\gamma_{2l}^{(2n)}$ (diagram IF) for $2n \leq 6$ are calculated for an arbitrary spin. Those values are combined with the values of $n(R_{if}$; diagram IF) given in the preceding paper[1] and Sec. 2 to give the coefficients $\sigma_{2l}^{(2n)}(R_{if})$ up to $O(t^6)$ for the linear, square, sc, and bcc lattices. The numerical values are provided. Analysis of the results are not given in the present paper.

Notice that Eq. (C8) is an integral equation of variable z_1 for fixed W, w_2, w_3, and w_4. As we compare this equation with Eq. (A7), we find that

$$G_0^-(W,w_1,w_2,w_3,w_4) = \frac{2w_2\delta_{k_2k_3}}{u^2(w_2)}X^-(W-w_2,w_1,w_4) + \frac{2w_2\delta_{k_2k_4}}{u^2(w_2)}X^-(W-w_2,w_1,w_3). \quad \text{(C11)}$$

Using the same method for solving Eq. (34), $G_1(-)$ in Eq. (C9) can be solved to be

$$G_1^-(W,w_1,w_2,w_3,w_4) = \frac{1}{\pi}\int_\mu^\infty \frac{dw\,\mathrm{Im}h(w)}{h(W-w-w_2)}X^-(W-w_2,w_1,w) \times G_0^-(W,w_2,w,w_3,w_4). \quad \text{(C12)}$$

The general terms in the series expansion is then

$$G_n^-(W,w_1,w_2,w_3,w_4) = \frac{1}{\pi}\int_\mu^\infty \frac{dw\,\mathrm{Im}\,h(w)}{h(W-w-w_2)}X^-(W-w_2,w_1,w) \times G_{n-1}^-(W,w_2,w,w_3,w_4). \quad \text{(C13)}$$

The solution of $\tau^9(-)$ can then be obtained by using Eqs. (C1) and (C3) together with the results in Eqs. (C11)–(C13). The first few terms in the expansion solution for $\tau^9(-)$ are written in Eq. (43).

1 J. B. Bronzan, Phys. Rev. **139**, B751 (1965).
2 T. D. Lee, Phys. Rev. **95**, 1329 (1954); S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Harper & Row, New York, 1962), Sec. 12.
3 J. B. Bronzan, Phys. Rev. **172**, 1429 (1968).
4 T. H. Liu and R. L. Zimmerman, Suppl. Nuovo Cimento **6**, 1297 (1968).
5 T. H. Liu and R. L. Zimmerman, J. Math Phys. **11**, 1941 (1970).
6 M. S. Maxon and R. B. Curtis, Phys. Rev. **137**, B996 (1965).
7 M. S. Maxon, Phys. Rev. **149**, 1273 (1966).
8 In this article we use subscript on the tau functions which involve U particle and superscript on the ones without U particle.
9 R. L. Zimmerman, Nuovo Cimento **56A**, 391 (1968).

Two-Time Spin-Pair Correlation Function of the Heisenberg Magnet at Infinite Temperature. III. bcc Lattice and Arbitrary Spin

Tohru Morita*
Department of Physics, Ohio University, Athens, Ohio 45701

(Received 24 March 1972)

The expansion coefficients in powers of time of the two-time spin-pair correlation function are obtained up to terms of order t^8 for the bcc Heisenberg magnet of spin $\frac{1}{2}$ at infinite temperature and up to terms of order t^6 for the linear, square, sc, and bcc Heisenberg magnets of an arbitrary spin at infinite temperature. The result is applicable to the isotropic as well as the anisotropic Heisenberg magnet where the exchange integrals in the z direction and in the orthogonal plane are assumed to be different. Analysis of the results are given in separate articles.

1. INTRODUCTION

In the first paper of this series,[1] the numerical values of the short-time expansion coefficients of the two-time spin-pair correlation function of the Heisenberg magnet of spin $\frac{1}{2}$ at infinite temperature are provided up to $O(t^{10})$ for the linear chain and up to $O(t^8)$ for the square and sc lattices. The analysis of the results obtained by the expansion[2,3] have been found to reproduce the exact solution for the one-dimensional Heisenberg magnet of finite length of spin $\frac{1}{2}$,[4] and also the results of the computer simulation calculation for the Heisenberg magnet of classical spin,[5] and the neutron diffraction data for Rb Mn F_3,[5] with a suitable choice of time scale for different spins.

The only existing Heisenberg magnet of spin $\frac{1}{2}$ will be solid He^3, which crystalizes in the bcc structure. If one considers magnets with larger spin, one has Mn F_2 in which the magnetic ions constitute the bcc structure, and extensive experimental data are available. For the investigation of these systems, the short-time expansion coefficients are provided up to terms of $O(t^8)$ for the bcc Heisenberg magnet of spin $\frac{1}{2}$ and up to $O(t^6)$ for the linear, square, sc, and bcc magnets of an arbitrary spin in this paper.

The calculation of the short-time expansion coefficients $\sigma_{2l}^{(2n)}(R_{if})$ in the preceding paper was performed in two steps. In the first step, the expansion coefficients $\gamma_{2l}^{(2n)}$ (diagram IF) for small diagrams are computed. In the second step, the number of different ways $n(R_{if}$; diagram IF) by which the small diagrams with the initial and final sites I and F occur in the lattices with the difference of the final site and initial site R_{if} are obtained. The sum of the products of those two sets of numbers gives the coefficients $\sigma_{2l}^{(2n)}(R_{if})$ for the lattice under consideration. In order to obtain the coefficients of the term of $O(t^8)$ $\sigma_{2l}^{(8)}(R_{if})$ for the bcc lattice, we need the coefficients $\gamma_{2l}^{(8)}$ (diagram IF) for the diagram 10 of Fig. 1 in addition to those for diagram 1–9 which are given in Ref. 1. In Sec. 2 of the present paper, these coefficients for spin $\frac{1}{2}$ are provided, and then the computation of the values $n(R_{if}$; diagram IF) is described. Combining them with the values of $\gamma_{2l}^{(2n)}$ (diagram IF) given in the previous paper, we obtain the values of $\sigma_{2l}^{(2n)}(R_{if})$ for the bcc Heisenberg magnet of spin $\frac{1}{2}$. In Sec. 3, $\gamma_{2l}^{(2n)}$ (diagram IF) for $2n \leq 6$ are calculated for an arbitrary spin. Those values are combined with the values of $n(R_{if}$; diagram IF) given in the preceding paper[1] and Sec. 2 to give the coefficients $\sigma_{2l}^{(2n)}(R_{if})$ up to $O(t^6)$ for the linear, square, sc, and bcc lattices. The numerical values are provided. Analysis of the results are not given in the present paper.

2. COMPUTATION

The two-time spin-pair correlation function $\sigma(R_{if}, t)$ is defined by

$$\sigma(R_{if}, t) = \langle s_i^z(t) s_f^z(0)\rangle - \langle s_i^z\rangle\langle s_f^z\rangle. \quad (2.1)$$

The short-time expansion coefficients of $\sigma(R_{if}, t)$ is denoted as $\sigma^{(2n)}(R_{if})$:

$$\sigma(R_{if}, t) = \sigma^{(0)}(R_{if}) + \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n)!}\sigma^{(2n)}(R_{if}) t^{2n}. \quad (2.2)$$

When the exchange interaction is of nearest neighbors and the values coupling the $z-z$ components and the orthogonal direction are $J_{\parallel}$ and $J_{\perp}$, respectively, $\sigma^{(2n)}(R_{if})$ is expressed as follows:

$$\sigma^{(2n)}(R_{if}) = \sum_{l=0}^{n} J_{\parallel}^{2l} J_{\perp}^{2n-2l} \sigma_{2l}^{(2n)}(R_{if}). \quad (2.3)$$

For the isotropic Heisenberg magnet where $J_{\parallel} = J_{\perp} \equiv J$, $\sigma^{(2n)}(R_{if})$ are given by

$$\sigma^{(2n)}(R_{if}) = J^{2n}\sigma_t^{(2n)}(R_{if}), \quad (2.4)$$

$$\sigma_t^{(2n)}(R_{if}) = \sum_l \sigma_{2l}^{(2n)}(R_{if}). \quad (2.5)$$

The coefficients $\sigma_{2l}^{(2n)}(R_{if})$ are calculated from the corresponding expansion coefficients $\gamma_{2l}^{(2n)}$ (diagram IF) for the small diagrams by the following formula:

$$\sigma_{2l}^{(2n)}(R_{if}) = \sum_{\text{diagram}} \sum_{\text{IF}} n(R_{if}; \text{diagram IF}) \times \gamma_{2l}^{(2n)}(\text{diagram IF}). \quad (2.6)$$

The diagrams to be considered in the calculation of the expansion coefficients for the bcc lattice up to $O(t^8)$ are given in Fig. 1. The values $\gamma_{2l}^{(2n)}$(diagram IF) for diagrams 1–9 have been given up to $2n = 8$ in Table II of Ref. 1. The values $\gamma_{2l}^{(2n)}$(diagram 10, IF) are given in Table I.

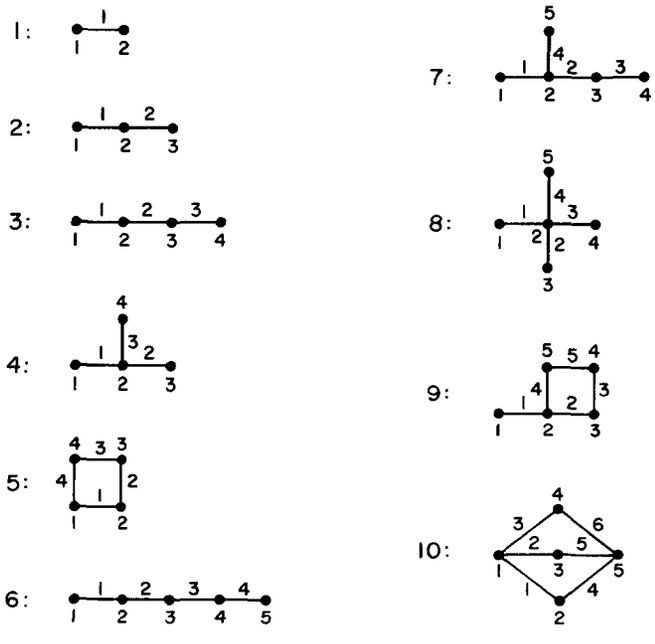


FIG. 1

TABLE I. The coefficients $\gamma_t^{(2n)}$ (diagram IF) and $\gamma_{2l}^{(2n)}$ (diagram IF) for diagram 10 of Fig. 1. This table supplements Table II of Ref. 1 in the case of the bcc lattice.

diagram	IF	$\gamma_t^{(8)}$	$\gamma_0^{(8)}$	$\gamma_2^{(8)}$	$\gamma_4^{(8)}$	$\gamma_6^{(8)}$
10	11	−78.0	246.0	−522.0	198.0	0.0
10	12	−142.0	−264.0	398.0	−276.0	0.0
10	15	504.0	546.0	−672.0	630.0	0.0
10	22	424.0	248.0	−40.0	216.0	0.0
10	23	−70.0	140.0	−378.0	168.0	0.0

The quantity $n(R_{if}$; diagram IF) is the total number of different ways by which the small diagram with the initial and final sites IF occurs in the lattice with the difference of the final and initial sites R_{if}. It is calculated by a computer as follows. The sites and bonds are labeled as in Fig. 1. The topologies of the diagrams are memorized in the computer, as follows. For instance, for diagram 4, the following numbers are memorized in addition to the total numbers of bonds and sites in the diagram:

1	1	2	0
2	2	3	0
3	3	4	0
4	1	4	3

The first row means that the first bond starts at site 1 and ends at 2; the second row that the second bond starts at 2 and ends at 3; etc. Except for the first bond, each bond must start with a site which is connected with one of the preceding bonds. The nonzero number at the last column means that the corresponding bond ends at a site which has already appeared in the preceding bonds. If that number is $n(n \geq 1)$, that bond ends at the same site as the nth bond which has already appeared. The position of the site 1 is set at the origin of the lattice. Directions of the first through nth bond are chosen in z^n ways, where z is the coordination number of the lattice. After each choice of the direction of the bond, it is checked whether the ending site is not on any of the sites which have already appeared or whether it is on the same site as the nth bond, according as the final number on the respective row is 0 or $n(n \geq 1)$. For each set of the choice of the directions which gives a figure in the required topological form, we calculate the difference R_{if} of the coordinates on the lattice for each pair IF of sites of the diagram and add one to the register for $n(R_{if}$; diagram IF). At the end of z^n choices,[6] we get $n(R_{if}$; diagram IF). The values are needed for the topologically different cases when the labeling on the sites and bonds are erased and hence the resulting numbers are divided by the symmetry number of the diagram in which the initial and final sites are distinguished. This program was used in checking the values in Table III of Ref. 1. The result for the bcc lattice is given in Table II.

By using the values in Tables I and II and the coefficients $\gamma_{2l}^{(2n)}$ (diagram IF) given in Table II of Ref. 1, we calculate the sums of the products given in (2.6) and obtain the values of $\sigma_{2l}^{(2n)}(R_{if})$ and $\sigma_t^{(2n)}(R_{if})$ for the bcc Heisenberg magnet of spin $\frac{1}{2}$. They are listed in Table III.

TABLE II. The number of different ways $n(R_{if}$; diagram IF) by which diagrams with the initial and final sites IF occur in the bcc lattice with the difference of the final and initial sites $R_{if} = (l, m, n)$.

$R_{if} = (0, 0, 0)$

diagram	IF	(0, 0, 0)
1	11	8
2	11	56
2	22	28
3	11	392
3	22	392
4	11	168
4	22	56
5	11	48
6	11	2648
6	22	2648
6	33	1324
7	11	2352
7	22	1176
7	33	1176
7	44	1176
8	11	280
8	22	70
9	11	288
9	22	288
9	33	576
9	44	288
10	11	24
10	22	36

$R_{if} = (1, 1, 1)$

diagram	IF	(1, 1, 1)
1	12	1
2	12	7
2	21	7
3	12	49
3	21	49
3	14	12
3	23	49
4	12	21
4	21	21
5	12	12
6	12	331
6	21	331
6	14	72
6	41	72
6	23	331
6	32	331
7	12	294
7	21	294
7	14	72
7	41	72
7	23	147
7	32	147
7	34	147
7	43	147
8	12	35
8	21	35
9	12	36
9	21	36
9	14	9
9	41	9
9	23	72
9	32	72
9	34	72
9	43	72
10	12	9
10	21	9

$R_{if} = (2, 0, 0), (2, 2, 0), (2, 2, 2)$

diagram	IF	(2, 0, 0)	(2, 2, 0)	(2, 2, 2)
2	13	4	2	1
3	13	28	14	7
3	31	28	14	7
4	13	24	12	6
5	13	6	1	
6	13	184	96	49
6	31	184	96	49
6	15	56	52	42
6	24	184	96	49
7	13	168	84	42
7	31	168	84	42
7	15	168	84	42
7	24	84	42	21
7	42	84	42	21
8	13	60	30	15
9	13	36	22	12
9	31	36	22	12
9	24	36	6	
9	42	36	6	
9	35	72	12	
10	15	4		
10	23	8	2	

$R_{if} = (3, 1, 1), (3, 3, 1), (3, 3, 3)$

diagram	IF	(3, 1, 1)	(3, 3, 1)	(3, 3, 3)
3	14	9	3	1
6	14	63	21	7
6	41	63	21	7
7	14	54	18	6
7	41	54	18	6
9	14	8	1	
9	41	8	1	

$R_{if} = (4, 0, 0), (4, 2, 0), (4, 2, 2)$

diagram	IF	(4, 0, 0)	(4, 2, 0)	(4, 2, 2)
6	15	36	24	16

$R_{if} = (4, 4, 0), (4, 4, 2), (4, 4, 4)$

diagram	IF	(4, 4, 0)	(4, 4, 2)	(4, 4, 4)
6	15	6	4	1

3. COMPUTATION FOR ARBITRARY SPIN

The short-time expansion coefficients $\sigma^{(2n)}(R_{if})$ is expressed as follows:

$$\begin{aligned}
\sigma^{(0)}(R_{if}) &= \langle s_i^z s_f^z \rangle , \\
\sigma^{(2)}(R_{if}) &= -\langle [H, s_i^z][H, s_f^z] \rangle , \\
\sigma^{(4)}(R_{if}) &= \langle [H, [H, s_i^z]][H, [H, s_f^z]] \rangle , \\
\sigma^{(6)}(R_{if}) &= -\langle [H, [H, [H, s_i^z]]][H, [H, [H, s_f^z]]] \rangle ,
\end{aligned} \tag{3.1}$$

where H is the Hamiltonian of the system. In computing $\sigma^{(2n)}(R_{if})$, we first calculate coefficients $\gamma^{(2n)}$ (diagram IF) for the small diagrams as shown in Fig. 1. $\gamma^{(2n)}$(diagram IF) is the sum of all those contributions from

$$\langle [H, [H, \ldots, [H, s_I^z] \cdots]] s_F^z \rangle \tag{3.2}$$

involving $2n$ commutators, which involves at least one term of H associated with each bond in the diagram, where H is the Hamiltonian for the diagram. It is argued in Ref. 1 that we have a contribution only when $2(m-1) \leq 2n$ for the case of spin $\frac{1}{2}$, where m is the total number of sites in the diagram. It is shown in the Appendix that it is true for an arbitrary spin. As a consequence, we have only to consider diagrams with the total number of sites $m \leq 2, 3$, and 4 for $2n = 2, 4$, and 6, respectively. For instance in the computation of $\sigma^{(4)}(R_{if})$, we first compute the coefficients $\gamma^{(4)}$ (diagram IF) for each of diagrams 1 and 2 shown in Fig. 1. In the calculation, we use the form of (3.1) instead of (3.2). All the nonzero terms of the commutation $[H, [H, s_i^z]]$ for all the sites i in the diagram are computed and stored in the memory of the computer. All the pairs of such terms with $i =$ I and $i =$ F, respectively, are considered. If every bond in the diagram is used in either of the commutations, the trace of the product of the two terms is calculated for spins from $\frac{1}{2}$ to $\frac{11}{2}$ and summed to $\gamma_{2l}^{(2n)}$ (diagram IF); for each spin, they are expanded in powers of $S(S+1)/3$ as follows:

TABLE III. Expansion coefficients $\sigma_t^{(2n)}(R_{if})$ and $\sigma_{2l}^{(2n)}(R_{if})$ for the bcc Heisenberg magnet of spin $\frac{1}{2}$ at infinite temperature.

R_{if}	$2n$	$\sigma_t^{(2n)}$	$\sigma_0^{(2n)}$	$\sigma_2^{(2n)}$	$\sigma_4^{(2n)}$	$\sigma_6^{(2n)}$
(0, 0, 0)	0	0.25	0.25	0.0	0.0	0.0
	2	4.0	4.0	0.0	0.0	0.0
	4	212.0	156.0	56.0	0.0	0.0
	6	20 140.0	10 144.0	7 756.0	2 240.0	0.0
	8	2 960 756.0	1 047 076.0	1 102 688.0	668 976.0	142 016.0
(1, 1, 1)	2	−0.5	−0.5	0.0	0.0	0.0
	4	−37.0	−30.0	−7.0	0.0	0.0
	6	−3 870.0	−2 343.0	−1 247.0	−280.0	0.0
	8	−568 240.0	−261 116.0	−188 824.0	−100 548.0	−17 752.0
(2, 0, 0)	4	6.0	6.0	0.0	0.0	0.0
	6	840.0	720.0	120.0	0.0	0.0
	8	124 096.0	94 528.0	20 496.0	9 072.0	0.0
(2, 2, 0)	4	3.0	3.0	0.0	0.0	0.0
	6	450.0	360.0	90.0	0.0	0.0
	8	72 016.0	46 256.0	20 188.0	5 572.0	0.0
(2, 2, 2)	4	1.5	1.5	0.0	0.0	0.0
	6	232.5	180.0	52.5	0.0	0.0
	8	38 962.0	23 254.0	12 432.0	3 276.0	0.0
(3, 1, 1)	6	−45.0	−45.0	0.0	0.0	0.0
	8	−10 794.0	−8 820.0	−1 526.0	−448.0	0.0
(3, 3, 1)	6	−15.0	−15.0	0.0	0.0	0.0
	8	−3 738.0	−2 940.0	−742.0	−56.0	0.0
(3, 3, 3)	6	−5.0	−5.0	0.0	0.0	0.0
	8	−1 274.0	−980.0	−294.0	0.0	0.0
(4, 0, 0)	8	630.0	630.0	0.0	0.0	0.0
(4, 2, 0)	8	420.0	420.0	0.0	0.0	0.0
(4, 2, 2)	8	280.0	280.0	0.0	0.0	0.0
(4, 4, 0)	8	105.0	105.0	0.0	0.0	0.0
(4, 4, 2)	8	70.0	70.0	0.0	0.0	0.0
(4, 4, 4)	8	17.5	17.5	0.0	0.0	0.0

$$\gamma_{2l}^{(2n)}\text{(diagram IF)} = \sum_p \gamma_{2l,p}^{(2n)}\text{(diagram IF)}[S(S+1)/3]^p. \quad (3.3)$$

The obtained values for $\gamma_{2l,p}^{(2n)}$ (diagram IF) are given in Table IV.

For diagram 1, the complete set of the eigenfunctions and eigenvalues are well known for the isotropic case $J_{\|} = J_{\perp}$. By using the complete set, the coefficients $\gamma_t^{(2n)}$ (diagram 1 IF) can be expressed as

$$\gamma_t^{(2n)}\text{(diagram 1, 11)} = -\gamma_t^{(2n)}\text{(diagram 1, 12)}$$
$$= \frac{1}{(2S+1)^2} \sum_{j=0}^{2S} \sum_{j'=0}^{2S} \sum_{\substack{m=-j \\ (-j'\le m\le j')}}^{j} \sum_{m_1=-S}^{S} \sum_{m_1'=-S}^{S} C(SSj\,|m_1', m-m_1')m_1'$$
$$\times C(SSj'\,|m_1', m-m_1')C(SSj'\,|m_1, m-m_1)m_1\, C(SSj\,|m_1, m-m_1)(j'(j'+1)-j(j+1))^{2n}, \quad (3.4)$$
$$(-S\le m-m_1\le S,\ -S\le m-m_1'\le S)$$

where $C(j_1 j_2 j\,|m_1, m-m_1)$ are the Clebsch–Gordon coefficients.[7] This equation was used to check the coefficients $\gamma_{t,p}^{(2n)}$ (diagram 1, IF) in Table IV.

For an arbitrary spin, the coefficients $\sigma_{2l}^{(2n)}(R_{if})$ are expanded as follows:

$$\sigma_{2l}^{(2n)}(R_{if}) = \sum_p \sigma_{2l,p}^{(2n)}(R_{if})[S(S+1)/3]^p. \quad (3.5)$$

$\sigma_{2l,p}^{(2n)}(R_{if})$ are calculated by the formula (2.6) from $\gamma_{2l,p}^{(2n)}$ (diagram IF) with the aid of the numbers $n(R_{if}$; diagram IF) given in Ref. 1 and the preceding section. The obtained coefficients for the linear, square, sc, and bcc lattices are listed in Tables V–VIII.

For the isotropic Heisenberg magnet, $\sigma_t^{(2n)}(R_{if})$ appearing in (2.4) is expanded as

$$\sigma_t^{(2n)}(R_{if}) = \sum_p \sigma_{t,p}^{(2n)}(R_{if})[S(S+1)/3]^p. \quad (3.6)$$

Recalling the relation (2.5), we have

$$\sigma_{t,p}^{(2n)}(R_{if}) = \sum_l \sigma_{2l,p}^{(2n)}(R_{if}). \quad (3.7)$$

The values of $\sigma_{t,p}^{(2n)}(R_{if})$ are also listed in Tables V–VIII.

4. SUMMARY

The two-time spin-pair correlation function for the Heisenberg magnet at infinite temperature is expanded in powers of time as Eq. (2.2). The coefficients $\sigma^{(2n)}(R_{if})$ are calculated from $\sigma_t^{(2n)}(R_{if})$ and $\sigma_{2l}^{(2n)}(R_{if})$ by (2.3) and (2.4) for spin $\frac{1}{2}$. The numerical values of $\sigma_t^{(2n)}(R_{if})$ and $\sigma_{2l}^{(2n)}(R_{if})$ for the bcc Heisenberg magnet of spin $\frac{1}{2}$ are computed for $0 \le 2n \le 8$ and listed in Table III. For the case of an arbitrary spin, the coefficients $\sigma^{(2n)}(R_{if})$ are obtained by (2.3) and

TABLE IV. The coefficients $\gamma_{t,p}^{(2n)}$ (diagram IF) and $\gamma_{2l,p}^{(2n)}$ (diagram IF) of the terms $t^{2n}[S(S+1)/3]^p$ for the diagrams shown in Fig. 1, where IF denote the initial and final sites.

$2n = 2$

diagram	IF	p	$\gamma_{t,p}^{(2)}$	$\gamma_{0,p}^{(2)}$
1	11	2	8.0	8.0
1	12	2	−8.0	−8.0

$2n = 4$

diagram	IF	p	$\gamma_{t,p}^{(4)}$	$\gamma_{0,p}^{(4)}$	$\gamma_{2,p}^{(4)}$
1	11	2	−16.0	−6.4	−9.6
1	11	3	192.0	153.6	38.4
1	12	2	16.0	6.4	9.6
1	12	3	−192.0	−153.6	−38.4
2	11	3	64.0	32.0	32.0
2	12	3	−160.0	−128.0	−32.0
2	13	3	96.0	96.0	0.0
2	22	3	320.0	256.0	64.0

$2n = 6$

diagram	IF	p	$\gamma_{t,p}^{(6)}$	$\gamma_{0,p}^{(6)}$	$\gamma_{2,p}^{(6)}$	$\gamma_{4,p}^{(6)}$
1	11	2	89.6	68.04 + 5/7	−19.32 − 3/7	39.88 + 5/7
1	11	3	−1 228.8	−710.52 − 3/7	−236.84 − 1/7	−280.44 − 3/7
1	11	4	5 529.6	3 791.44 + 2/7	1 263.48 + 3/7	473.68 + 2/7
1	12	2	−89.6	−68.04 − 5/7	19.32 + 3/7	−39.88 − 5/7
1	12	3	1 228.8	710.52 + 3/7	236.84 + 1/7	280.44 + 3/7
1	12	4	−5 529.6	−3 791.44 − 2/7	−1 263.48 − 3/7	−473.68 − 2/7
2	11	3	−960.0	−384.0	−320.0	−256.0
2	11	4	7 680.0	3 840.0	2 688.0	1 152.0
2	12	3	1 280.0	−768.0	256.0	256.0
2	12	4	−13 440.0	−9 216.0	−3 072.0	−1 152.0
2	13	3	−320.0	−384.0	64.0	0.0
2	13	4	5 760.0	5 376.0	384.0	0.0
2	22	3	−2 560.0	−1 536.0	−512.0	−512.0
2	22	4	26 880.0	18 432.0	6 144.0	2 304.0
3	11	4	512.0	128.0	384.0	0.0
3	12	4	−1 792.0	−768.0	−1 024.0	0.0
3	13	4	2 560.0	1 920.0	640.0	0.0
3	14	4	−1 280.0	−1 280.0	0.0	0.0
3	22	4	6 144.0	2 176.0	3 200.0	768.0
3	23	4	−6 912.0	−3 328.0	−2 816.0	−768.0
4	11	4	2 560.0	1 024.0	768.0	768.0
4	12	4	−8 960.0	−6 144.0	−2 048.0	−768.0
4	13	4	3 200.0	2 560.0	640.0	0.0
4	22	4	26 880.0	18 432.0	6 144.0	2 304.0
5	11	4	−1 792.0	0.0	−1 792.0	0.0
5	12	4	2 816.0	0.0	2 816.0	0.0
5	13	4	−3 840.0	0.0	−3 840.0	0.0

TABLE V. Expansion coefficients $\sigma_{t,p}^{(2n)}(R_{if})$ and $\sigma_{2l,p}^{(2n)}(R_{if})$ for the linear chain.

R_{if}	$2n$	p	$\sigma_{t,p}^{(2n)}$	$\sigma_{0,p}^{(2n)}$	$\sigma_{2,p}^{(2n)}$	$\sigma_{4,p}^{(2n)}$
(0)	0	1	1.0	1.0	0.0	0.0
	2	2	16.0	16.0	0.0	0.0
	4	2	−32.0	−12.8	−19.2	0.0
	4	3	832.0	627.2	204.8	0.0
	6	2	179.2	137.08 + 3/7	−38.64 − 6/7	80.76 + 3/7
	6	3	−6 937.6	−3 725.04 − 6/7	−1 625.68 − 2/7	−1 584.88 − 6/7
	6	4	66 611.2	38 302.88 + 4/7	21 214.96 + 6/7	7 091.36 + 4/7
(1)	2	2	−8.0	−8.0	0.0	0.0
	4	2	16.0	6.4	9.6	0.0
	4	3	−512.0	−409.6	−102.4	0.0
	6	2	−89.6	−68.04 − 5/7	19.32 + 3/7	−39.88 − 5/7
	6	3	3 788.8	2 246.52 + 3/7	748.84 + 1/7	792.44 + 3/7
	6	4	−42 905.6	−27 087.44 − 2/7	−12 271.48 − 3/7	−3 545.68 − 2/7
(2)	4	3	96.0	96.0	0.0	0.0
	6	3	−320.0	−384.0	64.0	0.0
	6	4	10 880.0	9 216.0	1 664.0	0.0
(3)	6	4	−1 280.0	−1 280.0	0.0	0.0

TABLE VI. Expansion coefficients $\sigma_{t,p}^{(2n)}(R_{if})$ and $\sigma_{2l,p}^{(2n)}(R_{if})$ for the square lattice.

R_{if}	$2n$	p	$\sigma_{t,p}^{(2n)}$	$\sigma_{0,p}^{(2n)}$	$\sigma_{2,p}^{(2n)}$	$\sigma_{4,p}^{(2n)}$
(0, 0)	0	1	1.0	1.0	0.0	0.0
	2	2	32.0	32.0	0.0	0.0
	4	2	−64.0	−25.6	−38.4	0.0
	4	3	3 456.0	2 534.4	921.6	0.0
	6	2	358.4	274.16 + 6/7	−78.28 − 5/7	161.52 + 6/7
	6	3	−31 795.2	−16 667.08 − 5/7	−7 859.36 − 4/7	−7 266.76 − 5/7
	6	4	646 246.4	340 798.76 + 1/7	229 822.92 + 5/7	75 623.72 + 1/7
(1, 0)	2	2	−8.0	−8.0	0.0	0.0
	4	2	16.0	6.4	9.6	0.0
	4	3	−1 152.0	−921.6	−230.4	0.0
	6	2	−89.6	−68.04 − 5/7	19.32 + 3/7	−39.88 − 5/7
	6	3	8 908.8	5 318.52 + 3/7	1 772.84 + 1/7	1 816.44 + 3/7
	6	4	−231 321.6	−142 287.44 − 2/7	−70 127.48 − 3/7	−18 905.68 − 2/7
(1, 1)	4	3	192.0	192.0	0.0	0.0
	6	3	−640.0	−768.0	128.0	0.0
	6	4	51 200.0	44 032.0	7 168.0	0.0
(2, 0)	4	3	96.0	96.0	0.0	0.0
	6	3	−320.0	−384.0	64.0	0.0
	6	4	27 520.0	22 016.0	5 504.0	0.0
(2, 1)	6	4	−3 840.0	−3 840.0	0.0	0.0
(3, 0)	6	4	−1 280.0	−1 280.0	0.0	0.0

TABLE VII. Expansion coefficients $\sigma_{t,p}^{(2n)}(R_{if})$ and $\sigma_{2l,p}^{(2n)}(R_{if})$ for the sc lattice.

R_{if}	$2n$	p	$\sigma_{t,p}^{(2n)}$	$\sigma_{0,p}^{(2n)}$	$\sigma_{2,p}^{(2n)}$	$\sigma_{4,p}^{(2n)}$
(0, 0, 0)	0	1	1.0	1.0	0.0	0.0
	2	2	48.0	48.0	0.0	0.0
	4	2	−96.0	−38.4	−57.6	0.0
	4	3	7 872.0	5 721.6	2 150.4	0.0
	6	2	537.6	412.24 + 2/7	−117.92 − 4/7	243.28 + 2/7
	6	3	−74 572.8	−38 825.12 − 4/7	−18 701.04 − 6/7	−17 044.64 − 4/7
	6	4	2334 873.6	1190 109.64 + 5/7	865 438.88 + 4/7	279 323.08 + 5/7
(1, 0, 0)	2	2	−8.0	−8.0	0.0	0.0
	4	2	16.0	6.4	9.6	0.0
	4	3	−1 792.0	−1 433.6	−358.4	0.0
	6	2	−89.6	−68.04 − 5/7	19.32 + 3/7	−39.88 − 5/7
	6	3	14 028.8	8 390.52 + 3/7	2 796.84 + 1/7	2 840.44 + 3/7
	6	4	−575 385.6	−345 551.44 − 2/7	−183 279.48 − 3/7	−46 553.68 − 2/7
(1, 1, 0)	4	3	192.0	192.0	0.0	0.0
	6	3	−640.0	−768.0	128.0	0.0
	6	4	84 480.0	69 632.0	14 848.0	0.0
(1, 1, 1)	6	4	−7 680.0	−7 680.0	0.0	0.0
(2, 0, 0)	4	3	96.0	96.0	0.0	0.0
	6	3	−320.0	−384.0	64.0	0.0
	6	4	44 160.0	34 816.0	9 344.0	0.0
(2, 1, 0)	6	4	−3 840.0	−3 840.0	0.0	0.0
(3, 0, 0)	6	4	−1 280.0	−1 280.0	0.0	0.0

(2.4) and (3.5) and (3.6) from $\sigma_{t,p}^{(2n)}(R_{if})$ and $\sigma_{2l,p}^{(2n)}(R_{if})$. The numerical values of these coefficients are calculated for $0 \leq 2n \leq 6$ for the linear, square, sc, and bcc lattices and listed in Tables V–VIII.

ACKNOWLEDGMENT

The author is grateful to Professor E. R. Hunt for valuable discussions on magnetic properties of solid He^3, which stimulated the present computation. The computation was performed with the aid of computer IBM 360/44 of Computer Center, Ohio University.

APPENDIX: DIAGRAMS CONTRIBUTING TO ORDER t^{2n}

The value $\gamma^{(2n)}$ (diagram IF) for a diagram shown in Fig. 1 is the sum of all those contributions from

$$\langle [H, [H, \ldots, [H, s_I^z] \cdots]] s_F^z \rangle \tag{A1}$$

involving $2n$ commutators, which involve at least one term of H associated with each bond in the diagram. Here H is the Hamiltonian for the diagram. In this appendix, we prove that $\gamma^{(2n)}$ (diagram IF) can be nonzero only when

$$2(m-1) \leq 2n, \tag{A2}$$

where m is the total number of sites in the diagram.

Before taking commutators, operator s_I^z is associated with site I and operator unity with all the other sites. By taking commutation with H, the operator associated to each site becomes a product of a number of spin operators; e.g., a product of three operators like $s_1^+ s_1^+ s_1^z$ for site 1. Initially we have only one spin

TABLE VIII. Expansion coefficients $\sigma_{l,p}^{(2n)}(R_{if})$ and $\sigma_{2l,p}^{(2n)}(R_{if})$ for the bcc lattice.

R_{if}	$2n$	p	$\sigma_{l,p}^{(2n)}$	$\sigma_{0,p}^{(2n)}$	$\sigma_{2,p}^{(2n)}$	$\sigma_{4,p}^{(2n)}$
(0, 0, 0)	0	1	1.0	1.0	0.0	0.0
	2	2	64.0	64.0	0.0	0.0
	4	2	−128.0	−51.2	−76.8	0.0
	4	3	14080.0	10188.8	3891.2	0.0
	6	2	716.8	549.32 + 5/7	−157.56 − 3/7	324.04 + 5/7
	6	3	−135270.4	−70199.16 − 3/7	−34151.72 − 1/7	−30918.52 − 3/7
	6	4	5685452.8	2868861.52 + 2/7	2124670.84 + 3/7	691919.44 + 2/7
(1, 1, 1)	2	2	−8.0	−8.0	0.0	0.0
	4	2	16.0	6.4	9.6	0.0
	4	3	−2432.0	−1945.6	−486.4	0.0
	6	2	−89.6	−68.04 − 5/7	19.32 + 3/7	−39.88 − 5/7
	6	3	19148.8	11462.52 + 3/7	3820.84 + 1/7	3864.44 + 3/7
	6	4	−1065881.6	−644559.44 − 2/7	−334831.48 − 3/7	−86489.68 − 2/7
(2, 0, 0)	4	3	384.0	384.0	0.0	0.0
	6	3	−1280.0	−1536.0	256.0	0.0
	6	4	220160.0	190464.0	29696.0	0.0
(2, 2, 0)	4	3	192.0	192.0	0.0	0.0
	6	3	−640.0	−768.0	128.0	0.0
	6	4	117760.0	95232.0	22528.0	0.0
(2, 2, 2)	4	3	96.0	96.0	0.0	0.0
	6	3	−320.0	−384.0	64.0	0.0
	6	4	60800.0	47616.0	13184.0	0.0
(3, 1, 1)	6	4	−11520.0	−11520.0	0.0	0.0
(3, 3, 1)	6	4	−3840.0	−3840.0	0.0	0.0
(3, 3, 3)	6	4	−1280.0	−1280.0	0.0	0.0

operator s_1^z. We notice that (i) the operator for each site other than I becomes s^z, s^+, or s^- when it is involved in a commutator for the first time, and (ii) at each commutation the total number of the spin operators for all the sites in the diagram is increased by one. If there are m sites in the diagram under consideration and a term of the Hamiltonian to each of the bonds in the diagram is required to occur at least once, a nonvanishing contribution is expected only when a product of at least two spin operators is associated with each site excluding F and at least one spin operator with site F. In order to get such a product of at least $2(m-1)+1$ spin operators for the whole diagram, we need at least $2(m-1)$ commutations starting from the initial one spin operator s_I^z. Hence we have a contribution only when the total number of commutations $2n$ is equal to or greater than $2(m-1)$, namely (A2).

---

* Senior Foreign Scientist of National Science Foundation and Ohio University (1971-1972). Present address: Department of Applied Science, Faculty of Engineering, Tohoku University, Sendai, Japan.

1 T. Morita, J. Math. Phys. **12**, 2062 (1971).

2 T. Morita, J. Math. Phys. **13**, 714 (1972); T. Horiguchi and T. Morita, in preparation.

3 R. A. Tahir-Kheli and D. G. McFadden, Phys. Rev. **182**, 604 (1969); B **1**, 3649 (1970). In these papers, Tahir-Kheli and McFadden made a number of analyses with the aid of the expansion including terms up to $O(t^4)$ for an arbitrary spin.

4 F. Carboni and P. M. Richards, Phys. Rev. **177**, 889 (1969).

5 C. G. Windsor, in *Neutron Inelastic Scattering*, Vol. II (IAEA, Vienna, 1968), p. 83.

6 In actual computation, we can save the computer time by fixing the direction of the first bond to one of z directions and using the symmetry property of the lattice to get the final values of $n(R_{if}$; diagram IF).

7 E.g., M. E. Rose, *Elementary Theory of Angular Momentum* (Wiley, New York, 1957).

Energy of Gravitational Shock Waves

David C. Robinson
Department of Mathematics, King's College, London, England

and

Jeffrey Winicour
General Physics Research Laboratory, Aerospace Research Laboratories, Wright-Patterson AFB, Ohio 45433
(Received 27 January 1972)

A model of a gravitational shock wave is developed using the null hypersurface approach to the initial value problem. Interior to the shock front is a flat Minkowski region. This is joined to a shear-free but curved exterior. The geometric properties of the shock front determine the physical properties of the wave. The gravitational energy measured at null infinity is a functional of the 2-geometry of the shock front. The positive energy conjecture reduces to a simple statement concerning the geometry of 2-surfaces. Explicit numerical calculations carried out for various 2-surfaces have led to positive energy. However, no inequality is known to show that the energy is positive for all 2-surfaces.

1. INTRODUCTION

In this paper, we investigate the energy content of a simple model of a gravitational shock wave. This work has been motivated by recent attempts to answer the question, "Can the energy of an asymptotically flat space-time be negative?"[1-5] In the weak field case, this question has been decided in favor of positive energy.[2,4] However, these efforts have failed to produce any conclusive results in the case of strong gravitational fields.

Here we treat the energy as formulated in the null infinity description of asymptotically flat space-times.[6-9] Our program is from the point of view of the characteristic initial value problem analogous to the treatments of Brill[10] and of Arnowitt, Deser, and Misner[11] based upon Cauchy data on a space-like hypersurface. These latter space-like models have established that the Hamiltonian energy measured at spatial infinity is positive for the special cases in which the Cauchy data can be classified in terms of the canonical formalism as either "axially symmetric pure coordinate" or "pure momenta."

We have not been able to draw any such definite conclusions from our model. The energy content of the gravitational shock wave depends upon the geometrical properties of a two-dimensional surface characterizing the shock front in a way which does not make the sign of the energy evident. We have numerically calculated the energy for several explicit choices of shock front geometry. All these calculations have led to positive energy. Hence the question of the possibility of negative energy remains unresolved. However, the shock wave model does furnish valuable physical insight into the properties of gravitational energy.

Our signature for the space-time metric is minus two. Greek indices range from zero to three, and the four-dimensional covariant derivative is denoted by ∇.

2. MODEL OF A GRAVITATIONAL SHOCK WAVE

Asymptotically flat solutions of Einstein's vacuum field equations can be specified by characteristic initial value data on a null hypersurface N of topology $S^2 \times E^1$ which extends from an interior cross-over region B to a sphere at future null infinity.[12] In the interior region of N the complex shear, which can be given freely, forms the complete characteristic initial value data. To complete the data, additional information must be given about the geometry of the cross-over region B. We shall construct a model of a gravitational shock wave by specifying the data for N and B in an especially simple way which guarantees that the data generate a nonsingular asymptotically flat space-time.

In order to do this it is convenient to introduce a null polar coordinate system with retarded time $x^0 = u$, affine parameter along the null rays $x^1 = r$, and ray labels $x^A (A = 2, 3)$. The line element takes the form

$$ds^2 = g_{00}du^2 + 2g_{01}dudr + 2g_{0A}dudx^A + g_{AB}\,dx^A dx^B, \quad (2.1)$$

and the contravariant form of the metric satisfies the algebraic conditions

$$g^{00} = g^{0A} = 0, \quad g^{01} = 1, \quad \text{and} \quad g^{AB}g_{BC} = \delta^A_C.$$

In the formulation of the characteristic initial value problem for the vacuum Einstein field equations, the null hypersurface constraint equations take the form

$$R_{11} = p_{,11} + \tfrac{1}{4}(p_{,1})^2 = 0, \quad (2.2)$$

$$R_{1A} = \tfrac{1}{2}g_{AB}e^{-p}(e^p g^{1B}{}_{,1})_{,1} + \tfrac{1}{4}p_{,1;A} = 0, \quad (2.3)$$

and

$$g^{AB}R_{AB} = {}^{(2)}R + e^{-p/2}[e^{p/2}(\tfrac{1}{2}g^{11}p_{,1} + p_{,0})]_{,1} + g^{1B}{}_{,1;B} + \tfrac{1}{2}g_{AB}g^{1A}{}_{,1}g^{1B}{}_{,1} + (g^{1B}p_{,1})_{;B} = 0. \quad (2.4)$$

Here $p \equiv \ln(-g)$, ${}^{(2)}R$ is the Ricci scalar corresponding to the 2-surface metric g_{AB}, the semi-colon indicates two-dimensional covariant differentiation with respect to g_{AB}, and a comma indicates partial differentiation.

Once the shear is specified on the initial surface N given by $u = u_0$, these equations can be integrated to give a solution of Einstein's equations in terms of the integration constants that arise. These integration constants are determined by additional data that must be given at the interior cross-over region B. It is important and nontrivial to prescribe data at B so as not to produce any singularities in the 4-geometry of space-time. We carry this out in the following way.

We let the data on the inner portion of N correspond to flat space data; that is the inner portion of N corresponds to a shearing null hypersurface embedded in Minkowski space. An example is the null hypersurface interior to an ellipsoid. Next we choose some two-dimensional cross section Σ_0 of this inner portion of N such that Σ_0 lies exterior to the cross-over region B. For further simplicity we take Σ_0 to

be embedded in a surface of constant Minkowski time. We fix the affine parameter r so that Σ_0 is given by $r = 1$ and complete our data by demanding that the shear vanish for $r > 1$. As long as the divergence ρ of the null rays is positive at Σ_0, the null hypersurface N will extend to future null infinity without developing caustics and, since the shear is zero exterior to Σ_0, the solution will be asymptotically flat. Unless the flat inner portion of N is a *null cone* its shear will not vanish and there will be a discontinuity in both the shear and curvature across Σ_0, the shock front of an imploding spherical gravitational wave.[13] However, these discontinuities are of the allowable type which may be removed by a smearing of the region Σ_0 until the usual continuity conditions of the Lichnerowicz type are satisfied.[14] The energy of the shock front varies continuously under such a smoothing operation so that these discontinuities may be considered to be nonpathological. Setting the data this way has also ensured that there are no singularities associated with the cross-over region B, since spacetime is flat in the neighborhood of B.

In order to calculate the energy of the shock wave we first integrate the constraint equations in the shear free region $1 < r \leq \infty$. Since the shear vanishes in this region, we have

$$g_{AB,1} = \tfrac{1}{2} g_{AB} p_{,1}. \tag{2.5}$$

The boundary conditions at Σ_0 imply that at $r = 1$,

$$g_{AB} = h_{AB}, \tag{2.6}$$

where h_{AB} is the metric of the 2-surface Σ_0, and that also at $r = 1$,

$$g^{1B} = g^{1B}{}_{,1} = 0 \tag{2.7}$$

and

$$\tfrac{1}{2} g^{11} p_{,1} + p_{,0} = -\tfrac{1}{2} p_{,1}. \tag{2.8}$$

Equations (2.7) and (2.8) follow from the fact that Σ_0 is embedded in a flat 3-surface of constant time in Minkowski space.

From Eqs. (2.2), (2.5), and (2.6) it follows that

$$g_{AB} = [1 + (r - 1)B]^2 h_{AB}, \tag{2.9}$$

where $B(x^A)$ is an arbitrary function whose geometrical significance can be seen by calculating the divergence of the null rays on N. If $l_\mu = u_{,\mu}$, then the divergence ρ is given by

$$\rho = \tfrac{1}{2} \nabla_\mu l^\mu. \tag{2.10}$$

By virtue of Eqs. (2.9) and (2.10) we have

$$\rho = B[1 + (r - 1)B]^{-1}, \tag{2.11}$$

so that at $r = 1$, $\rho = B$. Hence to ensure that N extends to future null infinity without developing caustics exterior to Σ_0, we require that B be a smooth *positive* function so that ρ is continuous and *positive* at Σ_0.

By using Eqs. (2.7) and (2.9), the constraint Eq. (2.3) can be integrated to give

$$g^{1B} = h^{AB} B_{,A} B^{-4}[(r - 1 + B^{-1})^{-2} - \tfrac{2}{3} B^{-1}(r - 1 + B^{-1})^{-3} - \tfrac{1}{3} B^2]. \tag{2.12}$$

We can now integrate the remaining constraint Eq. (2.4); but as we shall see in the next section this is unnecessary as we already have sufficient information about the metric on N to calculate the energy at the retarded time u_0.

3. THE ENERGY OF THE WAVE

We shall use the asymptotic symmetry linkages to calculate the energy of the shock wave at the retarded time u_0.[7-9] These quantities are functionals of the spherical cross-sections Σ of N and may be conveniently written[7]

$$L_\xi(\Sigma) = -\oint dS l^\nu n_\mu \nabla_\nu \xi^\mu. \tag{3.1}$$

Here

$$dS = (4\pi)^{-1}(-g)^{1/2} dx^2 dx^3,$$

and n^μ is an arbitrary vector field on Σ which satisfies $l^\mu n_\mu = -1$. The Bondi–Metzner–Sachs symmetry descriptors ξ^μ are determined asymptotically by applying Killing's equations at future null infinity and are then propagated along N by the projection of Killing's equations

$$l_\nu \nabla^{(\nu} \xi^{\mu)} = \tfrac{1}{2} l^\mu \nabla_\rho \xi^\rho. \tag{3.2}$$

The total energy of a system is usually calculated by evaluating the linkage which corresponds to an asymptotic time translation and then taking the limit to future null infinity $\mathcal{I}^+$

$$E = \lim_{\Sigma \to \Sigma^+} L_\xi(\Sigma), \tag{3.3}$$

where $\Sigma^+ = N \cap \mathcal{I}^+$. When the shear is zero, however, the values of the linkages are independent of the particular slice Σ on which they are evaluated.[15] We shall use this fact to calculate the energy by evaluating at Σ_0 the linkage which corresponds to an asymptotic time translation. Hence for this model

$$E = L_\xi(\Sigma_0). \tag{3.4}$$

In the null polar coordinate system given by Eq. (2.1) the propagation equations (3.2) take the form

$$\xi^0{}_{,1} = 0, \tag{3.5}$$

$$p_{,1} \xi^1 = 2 g^{1A} \xi^0{}_{,A} - 2\xi^A{}_{;A} - p_{,0} \xi^0, \tag{3.6}$$

and

$$\xi^A{}_{,1} = -g^{AB} \xi^0{}_{,B}. \tag{3.7}$$

By using Eqs. (2.2), (3.5), and (3.6) it is a straightforward matter to show that at $r = 1$:

$$-n_\mu l^\nu \nabla_\nu \xi^\mu = -(p_{,1})^{-1}[p_{,0} + \tfrac{1}{2} g^{11} p_{,1}]_{,1} \xi^0 - 2(p_{,1})^{-1} \xi^A{}_{;A,1} - \tfrac{1}{2} \xi^A{}_{;A} - \tfrac{1}{4}[p_{,0} + \tfrac{1}{2} g^{11} p_{,1}] \xi^0. \tag{3.8}$$

We can rewrite this expression by using Eq. (2.4), Eq. (3.7), and the boundary conditions given by Eqs. (2.7) and (2.8). We obtain at $r = 1$

$$-n_\mu l^\nu \nabla_\nu \xi^\mu = (p_{,1})^{-1}\, {}^{(2)}R \xi^0 - \tfrac{1}{2} \xi^A{}_{;A} - \tfrac{1}{8} p_{,1} \xi^0 + 2(p_{,1})^{-1} \xi^{0;A}{}_A - (p_{,1})^{-1} p_{,1;A} \xi^A \tag{3.9}$$

The asymptotic form of the time translation is given by

$$\xi^\mu \sim B V \nabla^\mu r, \tag{3.10}$$

where V is the conformal factor relating the metric of Σ_0 to the metric of the unit sphere q_{AB},

$$h_{AB} = - V^2 q_{AB}. \tag{3.11}$$

(The minus sign arises because h_{AB} is negative definite and q_{AB} is positive definite.) Note that although this form for ξ^μ could be found by solving Killing's equations asymptotically, it can also be obtained directly from known results[8] by recalling the behavior of asymptotic Killing vectors under conformal transformations and by noting from Eq. (2.9) that asymptotically

$$g_{AB} \sim - r^2 B^2 V^2 q_{AB}.$$

By using Eqs. (2.9), (2.12), and (3.10), we can integrate Eqs. (3.5) and (3.7) to find ξ^0 and ξ^A at $r = 1$. We are then in a position to evaluate the energy from Eq. (3.4). It follows that

$$E = \oint_{\Sigma_0} [\tfrac{1}{4}{}^{(2)}R - \tfrac{1}{2}B^2 - \tfrac{1}{6}(\ln B)_{;A}(\ln B)^{;A} + \tfrac{1}{2}(\ln B)^{;A}{}_{A}] V dS. \tag{3.12}$$

Since Σ_0 is embedded in a hypersurface of constant Minkowski time and B is equal to the divergence of N Σ_0, we can use a result of Kantowski[16] to identify B with the mean curvature of Σ_0 (up to a constant determined by convention). In this way the energy of our model becomes a functional of the 2-geometry of a smooth (say C^5) surface with spherical topology embedded in flat Euclidean 3-space.

Therefore in the remainder of this paper we shall regard the problem of calculating the energy of the gravitational shock wave as a problem in the theory of 2-surfaces embedded in flat 3-space. For this reason it is convenient to introduce a positive-definite metric η_{AB} to describe the intrinsic geometry of the 2-surface Σ_0

$$\eta_{AB} = - h_{AB},$$

so that using Eq. (3.11), the line element may be written as

$$ds^2 = \eta_{AB} dx^A dx^B = V^2 q_{AB} dx^A dx^B. \tag{3.13}$$

We then rewrite the energy expression Eq. (3.12) in terms of the Gaussian curvature

$$K = \tfrac{1}{2}{}^{(2)}R$$

and the mean curvature[17]

$$H = 2B > 0$$

of a surface embedded in flat 3-space to obtain

$$E = \tfrac{1}{2}\oint [K - \tfrac{1}{4}H^2 - (\ln H)^{;A}{}_{A} + \tfrac{1}{3}(\ln H)_{;A}(\ln H)^{;A}] V dS, \tag{3.14}$$

where the colon represents covariant differentiation with respect to the metric η_{AB}. The conformal relationship to the unit sphere given in Eq. (3.13) implies that

$$dS = V^2 d\Omega,$$

where $d\Omega$ is the element of solid angle on the unit sphere, and[18]

$$V^2 K = 1 - V V^{;A}{}_{A} + V^{;A} V_{;A}. \tag{3.15}$$

The sign properties of the energy of this model clearly cannot be determined at a glance. If the energy of the gravitational field could indeed become negative, then the present model would seem to offer a reasonable possibility for constructing a negative energy space-time. The energy would have been zero if we had not set the shear equal to zero at Σ_0, but rather allowed N to extend to future null infinity as a shearing null hypersurface embedded in Minkowski space. However, elimination of the shear exterior to Σ_0 introduces a pulse of Weyl curvature at Σ_0 which make the energy nonzero. Considerations of curved space geometrical optics[13] indicate that the presence of shear causes positive focussing and, in that sense, simulates the effects of positive mass. This heuristically suggests that the elimination of shear from the Minkowski null surface might lead to negative energy. Indeed, the first two terms of the integrand in Eq. (3.14) may be written as

$$K - \tfrac{1}{4}H^2 = - \tfrac{1}{4}(k_1 - k_2)^2 = - A^2, \tag{3.16}$$

where k_1 and k_2 are the two principle curvature[17] of Σ_0, so that A is a measure of the astigmatism or shear[16] of the null rays of N interior to Σ_0. We see, however, that although the contribution of A to the energy is negative definite, the contribution from the term

$$\tfrac{1}{3}(\ln H)_{;A}(\ln H)^{;A} \tag{3.17}$$

due to inhomogeneities in the mean curvature is positive definite. On the other hand, the term

$$- (\ln H)^{;A}{}_{A}$$

has no definite sign. The weighting of these terms by the conformal factor V adds to the difficulty in determining the sign of the integral. The remainder of this paper will be devoted to this problem.

4. THE WEAK FIELD ENERGY

It is known that the energy of a weak gravitational field evaluated on a good null cone is positive definite.[4] In fact, this energy, considered as a functional of the geometry, has a strict local minimum at flat space. Although the energy of the present model is not evaluated on a good cone, it is reasonable to suppose that a good cone does exist in the domain of dependence of N in the flat space limit, so that in this limit the energy evaluated at this later retarded time must be positive. It would then follow from the Bondi–Sachs[19,20] mass loss theorem that the weak field energy evaluated at N was positive. In this section, we shall show that the weak-field energy of our model does indeed have these properties.

In our model, flat space-time corresponds to the choice of a metric sphere for Σ_0. In this case N is a null cone and the energy is zero. We calculate the weak field energy by evaluating the functional $E(\Sigma_0)$ for a two-dimensional surface with spherical topology in Euclidean 3-space which differs from a metric sphere by a small amount. Because of the scaling

behavior of the energy,[5] it is sufficient to consider perturbations of the *unit* metric sphere.

The use of spherical polar coordinates in E^3, $x^1 = R$, $x^2 = \theta$, $x^3 = \varphi$, enables the representation of such a surface by the simple equation

$$R = 1 + \epsilon\Psi(\theta, \varphi), \tag{4.1}$$

where ϵ is a small parameter, and Ψ generates the first order deformation of the unit sphere. Since the line element of E^3 takes the form

$$ds^2 = dR^2 + R^2(d\theta^2 + \sin^2\theta d\varphi^2), \tag{4.2}$$

the line element of the 2-surface is given by

$$ds^2 = (1 + \epsilon\Psi)^2(d\theta^2 + \sin^2\theta d\varphi^2) + O(\epsilon^2). \tag{4.3}$$

Hence the conformal factor is given by

$$V = 1 + \epsilon\Psi + O(\epsilon^2). \tag{4.4}$$

In order to calculate H and K we introduce in E^3 an orthonormal triad $[n^i, t_{(A)}{}^i]$, where n^i is the unit normal to Σ_0 and $t_{(A)}{}^i$ are unit vectors tangential to the surface. In spherical polar coordinates we have

$$n_i = \delta_i^1 - \epsilon\Psi_{,A}\delta_i{}^A + O(\epsilon^2), \tag{4.5a}$$

$$t_{(B)}{}^i = R^{-1}q_{(B)}{}^A(\delta_A{}^i + \epsilon\delta_1{}^i\Psi_{,A}) + O(\epsilon^2), \tag{4.5b}$$

where $q_{(B)}{}^A$ form an orthonormal dyad on the unit sphere, the conventional choice in polar coordinates being $q_{(2)}{}^A = (1, 0)$ and $q_{(3)}{}^A = (0, 1/\sin\theta)$. Calculation of the second fundamental form then gives

$$H_{(AB)} \equiv t_{(A)}{}^i t_{(B)}{}^j D_j n_i = \delta_{(AB)}(1 - \epsilon\Psi) - \epsilon\Psi_{;CD}q_{(A)}{}^C q_{(B)}{}^D + O(\epsilon^2), \tag{4.6}$$

where D represents covariant differentiation in E^3, and the semicolon represents covariant differentiation on the unit sphere. We may now calculate the curvature invariants from[17]

$$H = \mathrm{Tr}[H_{(AB)}] \tag{4.7a}$$

and

$$K = \det[H_{(AB)}]. \tag{4.7b}$$

The calculation of the energy integral is facilitated by the use of a complex dyad

$$\iota^{(A)} = (2)^{-1/2}[\delta_2{}^{(A)} + i\delta_3{}^{(A)}] \tag{4.8}$$

and the use of the differential operator $\eth$.[15] The action of the operator $\eth$ is defined by relations of the type

$$\eth(\eta_{ABC}q^A q^B \bar{q}^C) = \sqrt{2}\,\eta_{ABC;D}\,q^A q^B \bar{q}^C q^D, \tag{4.9}$$

where $q^A = -q_{(B)}{}^A \iota^{(B)}$. We have

$$\begin{aligned} &\delta^{(AB)} = \iota^{(A)}\bar{\iota}^{(B)} + \bar{\iota}^{(A)}\iota^{(B)} \\ &\epsilon^{(AB)} = i(\iota^{(A)}\bar{\iota}^{(B)} - \bar{\iota}^{(A)}\iota^{(B)}), \\ &\delta_{(AB)}\iota^{(A)}\iota^{(B)} = 0, \end{aligned} \tag{4.10}$$

so that

$$H = \delta^{(AB)}H_{(AB)} = 2 - 2\epsilon\Psi - \epsilon\eth\eth^*\Psi \tag{4.11}$$

and

$$K - \tfrac{1}{4}H^2 = \tfrac{1}{2}\epsilon^{(AB)}\epsilon^{(CD)}H_{(AC)}H_{(BD)} - \tfrac{1}{4}H^2 = -\tfrac{1}{4}\epsilon^2\eth^2\Psi\,\eth^{*2}\Psi. \tag{4.12}$$

By using the commutation relation for spin-weight s quantities,

$$(\eth^*\eth - \eth\eth^*)\eta = 2s\eta, \tag{4.13}$$

and by using the properties of $\eth$ for integration over the unit sphere, it is straightforward to show that

$$E(\Sigma_0) = \tfrac{1}{2}\epsilon^2\oint[\tfrac{5}{4}\eth^2\Psi\,\eth^{*2}\Psi + \tfrac{1}{3}\eth(\Psi + \tfrac{1}{2}\eth\eth^*\Psi)\eth^*(\Psi + \tfrac{1}{2}\eth\eth^*\Psi)]d\Omega + O(\epsilon^3). \tag{4.14}$$

Hence the lowest-order term is quadratic in ϵ and positive definite.

5. AXIALLY SYMMETRIC CALCULATIONS

The energy integral in Eq. (3.14) is in general quite difficult to calculate numerically because of the problem of determining the conformal factor V. However, in the axially symmetric case an explicit expression for the conformal factor may be readily obtained. In this case, the surface Σ_0 is generated by revolution of a curve in E^3.

To set up an axially symmetric formalism for numerical computation, we introduce cylindrical coordinates (ρ, z, φ) so that the line element of E^3 is

$$ds^2 = d\rho^2 + dz^2 + \rho^2 d\varphi^2.$$

We represent the generating curve by $\rho = f(z)$, where z is the axis of revolution. The line element in Σ_0 is then given by

$$ds^2 = (1 + f'^2)dz^2 + f^2 d\varphi^2, \tag{5.1}$$

where the prime denotes partial differentiation with respect to z. In this coordinate system, it is straightforward to obtain

$$K = -f''/f(1 + f'^2), \tag{5.2a}$$

$$H = (1 + f'^2 - ff'')/f(1 + f'^2)^{3/2}, \tag{5.2b}$$

and

$$V = f\cosh\int_0^z [k + f^{-1}(1 + f'^2)^{1/2}]dz. \tag{5.2c}$$

The constant k reflects the freedom of the group of conformal transformations of the sphere onto itself. This group is in correspondence with the 6-parameter Lorentz group.[21] Because of axial symmetry, only Lorentz transformations in the z direction are involved. All our calculations are for surfaces which are symmetric under reflections of the z axis about $z = 0$. For this reason we take $k = 0$ so that the energy corresponds to the rest energy of the system.

We normalize our generating curve to range between endpoints at $z = -1$ and $z = +1$, with reflection symmetry about $z = 0$. We write $A = f(z)$, $B = f'(z)$, $C = f''(z)$, $D = f'''(z)$, and $E = f''''(z)$. By using Eqs. (5.1) and (5.2), the energy integral in Eq. (3.14) reduces to

$$E = \int_0^1 dz N\cosh\int_0^z dz'(1 + B^2)^{1/2}/A, \tag{5.3}$$

where

$$N = -\frac{1}{8(1+B^2)^{1/2}}\left(1+\frac{AC}{1+B^2}\right)^2 - \frac{1}{2}B\left(\frac{3A^2BC^2}{(1+B^2)^{5/2}} - \frac{A(AD+BC)}{(1+B^2)^{3/2}} - \frac{B}{(1+B^2)^{1/2}}\right) + \frac{2A^2}{3(1+B^2)^{1/2}(1+B^2-AC)^2} \times \left(\frac{3ABC^2}{1+B^2} - AD - BC - \frac{B(1+B^2)}{A}\right)^2 - \frac{A^2(1+B^2)^{1/2}}{2(1+B^2-AC)}\left(-\frac{15AB^2C^3}{(1+B^2)^3} + \frac{3(AC^3+B^2C^2+3ABCD)}{(1+B^2)^2} + \frac{2B^2C-AC^2-ABD-A^2E}{A(1+B^2)} + \frac{2B^2-AC}{A^2}\right). \quad (5.4)$$

Equation (5.3) expresses the energy in a form suitable for numerical calculation. The smooth curve $f(z)$ determines the functions A, B, C, D, E, and N through straightforward algebraic and differential relations. The double integral in Eq. (5.3) can then be numerically integrated. Some care is necessary in treating the endpoint $z = 1$. Because $f(z)$ is a smooth closed curve, its slope at $z = 1$ must be infinite. This singularity is of course a purely formal one. It does not lead to any difficulties in the numerical calculation as long as the terms are properly grouped.

A. Ellipse

Our first calculations were for a one-parameter family of ellipsoids. We represent the generating ellipse by the function

$$f = a(1-z^2)^{1/2},$$

where the parameter a ranges between 0 and ∞. The case $a = a_0 = 1$ generates the sphere. The prolate case is given by $a < 1$ with $a \to 0$ giving the limit of a thin rod. The oblate case is given by $a > 1$ with $a \to \infty$ giving the limit of a thin disc.

In the case of the ellipsoid, we can analytically integrate Eq. (5.2c) to obtain an explicit expression for the conformal factor. Putting

$$P = 1 + z^2(a^2-1),$$

We find for $a < 1$,

$$V = \tfrac{1}{2}a(az + P^{1/2})\exp\{a^{-1}(1-a^2)^{1/2} \times \sin^{-1}[z(1-a^2)^{1/2}]\} + \tfrac{1}{2}a(1-z^2)(az+P^{1/2})^{-1} \times \exp\{-a^{-1}(1-a^2)^{1/2}\sin^{-1}[z(1-a^2)^{1/2}]\}$$

and for $a > 1$,

$$V = \tfrac{1}{2}a(az+P^{1/2})[z(a^2-1)^{1/2}+P^{1/2}]^{-(a^2-1)^{1/2}/a} + \tfrac{1}{2}a(1-z^2)(az+P^{1/2})^{-1} \times [z(a^2-1)^{1/2}+P^{1/2}]^{(a^2-1)^{1/2}/a}.$$

The qualitative features of numerical integration of the energy integral are shown in Fig. 1. At $a = a_0$ corresponding to the sphere, we have

$$E = 0 \quad \text{and} \quad \frac{dE}{da} = 0,$$

in agreement with the weak field results of Sec. 4. The energy then increases monotonically as $|a - a_0|$ increases and becomes infinite at the two critical values $a_c = 0$ or $a_c = \infty$ at which curvature singularities arise.

B. Ovals of Cassini

The one-parameter family of surfaces are generated by the ovals of Cassini[22] described by the function

$$f = \{-a - z^2 + [(1-a)^2 + 4az^2]^{1/2}\}^{1/2} \quad (5.5)$$

with $0 \le a < 1$. The shape of these curves for various values of the parameter a are illustrated in Fig. 2 in which the solid arrow indicates the axis of revolution. The case $a = a_0 = 0$ generates the sphere [Fig. 2(a)]. For $a = a_c = 1$ the generating curve in the lemniscate of Bernoulli [Fig. 2(d)]. In this limit the corresponding surface has a curvature singularity at the equator. Although these surfaces are qualitatively very different from the ellipsoids, the numerical calculations for the energy give exactly the same results shown in Fig. 1. The energy rises monotonically from its minimum at the sphere and approaches infinity as $a \to a_c$.

Calculations were also carried out for the ovals of Cassini for the case in which the axis of revolution (z axis) corresponds to the broken arrow in Fig. 2. In this case, the generating curves can still be described by Eq. (5.5), but now with $-1 < a \le 0$. The case $a = a_0 = 0$ again generates the sphere; but the case $a = a_c = -1$ corresponds to the curve in Fig. 2(b). In this limit, the mean curvature vanishes at the pole. This provides a limit to the physical reasonableness of this case since for values $a \le -1$ the

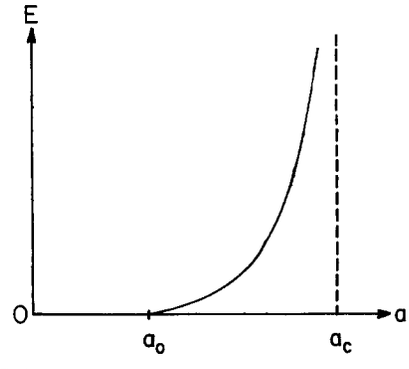


FIG. 1. Numerical results for the energy integral are plotted vs curve parameter.

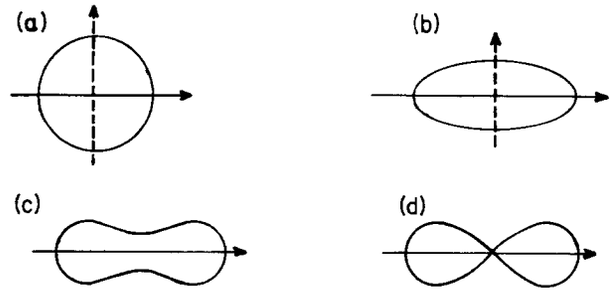


FIG. 2. A sequence of four ovals of Cassini are illustrated. The arrows indicate possible choices of axes of revolution.

corresponding space-time would not necessarily be asymptotically flat. The numerical results for the energy are given as before by Fig. 1. The infinite energy in the limit $a = a_c$ now arises from the flat spots where $H = 0$ at the poles and not from curvature singularities.

C. Miscellaneous

Calculations were also made for two additional families of surfaces. These surfaces do not shed any further light on the subject except to show that the preceding examples were not statistical anomalies.

In the first case, the generating curve is described by the function

$$f = (1 - z^2)^{1/2} - a(1 - z^2)$$

with $a_0 = 0 \leq a < 1 = a_c$. The corresponding surfaces are only of differentiability class C^2 at the poles (except when $a = 0$). Due to axial symmetry, the curvature is continuous, but not differentiable at the poles. This leads to a δ-function contribution to the energy integral from the poles. The integral, however, still converges except in the limit $a = a_c$. The numerical results are again described by Fig. 1 except for $a = a_0$:

$$E = 0 \quad \text{but} \quad \frac{dE}{da} \neq 0.$$

The behavior at the poles invalidates the usual weak field limit.

In the second case, the function $\rho(z)$ describing the generating curve is given in parametric form by

$$\rho = (1 - \mu^2)^{1/2}(1 - a\mu^2)(1 + a)^{-1},$$
$$z = \mu[1 + (1 - \mu^2)a(1 + a)^{-1}],$$

where $-1 \leq \mu \leq 1$. For $a = a_0 = 0$, the curve is a circle. For $a = a_c = 1$, a curvature singularity arises at the poles. We were able to obtain the explicit expression for the conformal factor

$$V = \frac{(1 - a\mu^2)}{2(1 + a)} \times \left[(1 + \mu)\left(\frac{1 + \mu\sqrt{a}}{1 - \mu\sqrt{a}}\right)^{\sqrt{a}} + (1 - \mu)\left(\frac{1 - \mu\sqrt{a}}{1 + \mu\sqrt{a}}\right)^{\sqrt{a}}\right]$$

The energy integral was calculated using a different computer program than in the previous examples. However, the results again turned out to be described by Fig. 1.

6. DISCUSSION

The calculations of the preceding section constitute the first explicit determination of the energy of vacuum gravitational fields from the null hypersurface point of view. Their common features exhibited in Fig. 1 are consistent with the conjecture that the total energy of a nonpathological gravitational system must be positive. The positive-definite term (3.17) in the energy integral arising from curvature in homogeneities dominates in the weak field limit over the negative-definite astigmatism term (3.16). Apparently, this headstart is sufficient to keep the energy positive throughout the strong field domain. However, we have not exhausted a wide enough spectrum of allowable surfaces to warrant any firm conclusions. In particular, our calculations are limited to surfaces having axial symmetry. The chief difficulty in extending these calculations to the case of no symmetry is the problem of determining the conformal factor V. This is also the chief obstacle to formulating analytical arguments to determine the behavior of the energy.

To emphasize the importance of the role of the conformal factor, let us consider the functional $E^*(\Sigma)$ obtained by setting $V = 1$ in the energy integral Eq. (3.14). Let the surface Σ be a cylinder to whose ends arbitrary end caps are smoothly joined to produce spherical topology.

The contribution to the functional E^* is strictly negative from the cylindrical region of Σ where the mean curvature is constant. No matter how large a positive value the end caps might contribute, the total value of E^* may always be made negative by lengthening the cylindrical region while keeping the same end caps.

For the energy functional $E(\Sigma)$ the situation is quite different. As the cylinder is lengthened, the conformal factor, which is a global quantity, weights the end caps more heavily so that what results in the limit of infinite length is not obvious. We have examined this limit for end caps generated by the Cassini ovals given in Eq. (5.5) with $a = \frac{1}{3}$. For this choice of parameter a, the shape corresponds to Fig. 2b with the solid arrow indicating the axis of revolution. The equatorial radius $f(0) = \frac{1}{3}$ matches the radius of the coaxial cylinder to which the equators of the ovaloids are joined at each end. The surface is of differentiability class C^3 at the join; but the curvature invariants are of class C^1. Accordingly, the join does not make a contribution to the energy integral. For a cylinder of length $2l$ the conformal factor Eq. (5.2c) is given by

$$V_l = \tfrac{1}{3}\cosh 3z \quad \text{for } |z| \leq l,$$
$$V_l = W \cosh 3l + (W^2 - F^2)^{1/2} \sinh 3l \quad \text{for } 1 \leq z \leq l + 1,$$

where

$$W(z) = V_0(|z| - l)$$

and

$$F(z) = f(|z| - l).$$

Here $V_0(z)$ and $f(z)$ are, respectively, the conformal factor and the profile function for the Cassini oval by itself ($l = 0$). The energy functional Eq. (3.14) takes the form

$$E_l = -\tfrac{1}{24}\sinh 3l + E_0 \cosh 3l + R \sinh 3l.$$

The first term is the negative contribution from the cylinder. In the second term, E_0 is the positive energy of the Cassini oval without the cylinder; the factor $\cosh 3l$ accounts for the displacement along the axis. The third term is more difficult to interpret, and its sign is not obvious. However, it is sufficient to compare the first two terms to understand why the previous construction which led to negative values for $E^*(\Sigma)$ does not work in such a simple way for $E(\Sigma)$. The computer calculations for the Cassini oval with parameter $a = \frac{1}{3}$ give $E_0 \approx 0.1165$. Hence the contribution from the end caps (neglecting the third

term) overrides the contribution from the cylinder even in the limit $l \to \infty$.

The most attractive feature of our shock wave model is that it reduces the study of gravitational energy to a study of two-dimensional surfaces. The conjecture that gravitational energy must be positive reduces to a simple geometrical statement.

Reduced Energy Conjecture. For all smooth surfaces Σ in Euclidean 3-space with spherical topology and positive mean curvature, the functional $E(\Sigma)$ defined in Eq. (3.14) is positive. Furthermore, $E(\Sigma) = 0$ implies that Σ is a metric sphere.

Although many inequalities concerning the global differential geometry of 2-surfaces are known,[23] they bear no similarity to the above conjecture. Resolution of this conjecture would be a major step toward understanding the properties of strong gravitational fields.

ACKNOWLEDGMENTS

We are deeply grateful to J. Caslin and H. Fettis of the Applied Mathematics Laboratory, ARL, for carrying out the computer integrations reported here.

[1] D. Brill and S. Deser, Phys. Rev. Letters **20**, 75 (1968).
[2] D. Brill and S. Deser, Ann. Phys. (N.Y.) **50**, 548 (1968).
[3] D. Brill, S. Deser, and L. Faddeev, Phys. Letters **26A**, 538 (1968).
[4] D. C. Robinson and J. Winicour, Phys. Rev. **3**, 840 (1971).
[5] D. C. Robinson and J. Winicour, J. Math. Phys. **12**, 995 (1971).
[6] R. Penrose, Phys. Rev. Letters **10**, 66 (1963).
[7] L. Tamburino and J. Winicour, Phys. Rev. **150**, 1039 (1966).
[8] J. Winicour, J. Math. Phys. **9**, 861 (1968).
[9] J. Winicour, in *Relativity*, edited by M. Carmeli, S. Fickler, and L. Witten (Plenum, New York, 1970).
[10] D. Brill, Ann. Phys. (N.Y.) **7**, 466 (1959).
[11] R. Arnowitt, S. Deser, and C. W. Misner, Ann. Phys. (N.Y.) **11**, 116 (1960).
[12] R. Penrose, "Null Hypersurface Initial Data for Classical Fields of Arbitrary Spin and for General Relativity," ARL Technical Report **63-56** (1963).
[13] R. Penrose, in *Perspectives in Geometry and Relativity*, edited by B. Hoffman (Indiana U. P., Bloomington, 1966).
[14] A. Lichnerowicz, *Théories Relativiste de la Gravitation et de l'Electromagnétisme* (Masson, Paris, 1955).
[15] L. Derry, R. Isaacson, and J. Winicour, Phys. Rev. **150**, 1039 (1966).
[16] R. Kantowski, J. Math. Phys. **9**, 336 (1968).
[17] We use the conventions of N. J. Hicks, *Notes on Differential Geometry* (Van Nostrand, New York, 1965).
[18] L. P. Eisenhart, *Riemannian Geometry* (Princeton U.P., Princeton, N.J., 1926), p. 90.
[19] H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner, Proc. Roy. Soc. (London) **A269**, 21 (1962).
[20] R. K. Sachs, Proc. Roy. Soc. (London) **A270**, 103 (1962).
[21] R. K. Sachs, Phys. Rev. **128**, 2851 (1962).
[22] R. C. Yates, *A Handbook on Curves and Their Properties* (Edwards, Ann Arbor, 1952), p. 8.
[23] See, for example, E. F. Beckenbach, in *Inequalities*, edited by O. Shisha (Academic, New York, 1967), p. 17.