The construction of solution of nonlinear relativistic wave equation in $\lambda:\Phi_4^4$: theory

R. Rączka

Institute of Nuclear Research, Warsaw, Hoża 69, Poland

The nonlinear equation $(\Box + m^2)\Phi(x) = \lambda:\Phi^3:$ for the quantum scalar field with special form of ordering of the interaction term, is considered. The unique soultion $\Phi(x)$ of this equation is constructed. It satisfies the relativistic covariance, asymptotic, irredicibility and the relativistic primitive causality conditions in the sense of sesquilinear form. The spectral condition is also satisfied.

1. INTRODUCTION

The construction of an interacting quantum field which satisfies a system of Wightman, Haag-Schroer or Haag-Kastler axioms is a central problem in quantum field theory. It seems that the best starting point would be a relativistic dynamical equation of motion for the field, determined in a specific carrier vector space.[1,2] However, because of difficulties associated with the definition of the product of fields one solves this problem in roundabout manner; one first constructs the interacting field using, e.g., Hamiltonian approach and then one defines the equation of motion which this field satisfies (see Glimm and Jaffe,[3] Dell'Antonio,[4] or Federbush[5] for specific examples). This approach which works in cases of quantum field theory models in two-dimensional space-time is not fully satisfactory. In fact the role of the dynamical equation of motion is reduced in this approach to the identity, which in principle has no physical consequences. On the other hand, the development of theoretical physics points out that the dynamical equation, together with initial or boundary conditions contains whole information on a physical system. Therefore, it seems that it would be more reasonable to start with certain relativistic dynamical equations for field satisfying a specific initial condition and then construct the field by solving the given field equation.

In the present paper we try to realize this program for self-interacting scalar field considering the special form of ordering of the product of field operators. In Sec. 2 we establish the form of field equations, initial conditions and the carrier Hilbert space H. Next in Sec. 3 we construct explicitly the unique solution $\Phi(x)$ of the field equation and we discuss the meaning of $\Phi(x)$ as a sesquilinear form in H. We prove in next sections that the field $\Phi(x)$ satisfies asymptotic, relativistic covariance, spectral, irreducibility and relativistic primitive causality conditions and possesses the unique vacuum. We construct in Sec. 8 the transformation which connects Φ_{out} and Φ_{in} field and we discuss the problem of existence of a nontrivial unitary relativistic scattering operator in the present model. Finally in Sec. 9 we discuss some generalizations of the present method for a class of nonpolynomial analytic interactions and for n-dimensional space-time, $n = 2, 3, 4, \cdots$.

2. THE FIELD EQUATION IN $\lambda:\Phi_4^4$: THEORY

Let $\Phi(x)$ be a scalar quantum field in four-dimensional Minkowski space-time satisfying the following field equation:

$$(\Box + m^2)\,\Phi(x) = \lambda\,{:}\Phi^3(x){:}, \qquad \lambda < 0, \tag{2.1}$$

with *a priori* nonspecified meaning of the nonlinear term. The meaning of this term with a particular rule of ordering will be given later on.

The problem which we solve in this paper is to find a solution $\Phi(x)$ of Eq. (2.1) which in the sense specified below satisfies the asymptotic condition $\Phi(x) \underset{t\to-\infty}{\longrightarrow} \Phi_{\text{in}}(x)$, where $\Phi_{\text{in}}(x)$ is a free relativistic local field. For that purpose it is convenient to use the equivalent Yang-Feldman form of Eq. (2.1) given by

$$\Phi(x) = \Phi_{\text{in}}(x) + \lambda \int \Delta_R(x-y)\,{:}\Phi^3(y){:}\,d^4y \tag{2.2}$$

We shall look for a solution $\Phi(x)$ of Eq. (2.2) defined as a sesquilinear form in the Fock space H_{in} associated with the field Φ_{in}. Let $\{a(p), a^*(p)\}$ be the set of annihilation and creation operators of the Φ_{in} field satisfying the canonical commutation relations. Let $\{e_l(p)\}_{l=1}^{\infty}$ be a complete orthonormal set of Hermite polynomials defined on the mass hyperboloid $p^2 = m^2$ (see, e.g., Ref. 6, p. 141). The smeared out annihilation and creation operators a_l and a_k^*, $l, k = 1, 2, \cdots$, satisfy the canonical commutation relations $[a_l, a_k^*] = \delta_{lk}$ and form an infinite-dimensional nilpotent Lie algebra which is irreducible in the space H_{in}. In Ref. 7 we have constructed the interacting field $\Phi(x)$ using the integral representation for $\Phi(x)$ in terms of the irreducible set $a_l, a_k^*, l, k = 1, 2, \cdots, .$ Here we present an alternative method.

The distributional nature of the interacting field operator $\Phi(x)$ is *a priori* not known. We know, however, that the solution $\Phi(x)$ of the nonlinear equations like (2.2) may usually be given as some functional of the Φ_{in} field in the form of the limit of an iterative series. For instance the nth order iteration of Eq. (2.2) has the form

$$\Phi^{(n)}(x) = \{\Phi_{\text{in}} + \lambda N_R\{\Phi_{\text{in}} + \lambda N_R\{\cdots\{\Phi_{\text{in}} + \lambda N_R\{\Phi_{\text{in}} + \lambda N_R(\Phi_{\text{in}})\}\}\cdots\}\}\}_{(n)}, \tag{2.3}$$

where for functional $\Psi(x)$ of the creation and annihilation operators in H_{in} the operation $N_R(\Psi)(x)$ is given by the formula

$$N_R(\Psi)(x) = \int \Delta_R(x-y)\,{:}\Psi^3(y){:}\,d^4y. \tag{2.4}$$

To simplify the notation, we shall use for the limit of the iterative series (2.3) the notation

$$\lim_{n\to\infty} \Phi^{(n)}(x) \equiv [(I - \lambda N_R)^{-1}(\Phi_{\text{in}})](x). \tag{2.5}$$

We see that $\Phi^{(n)}(x)$ is a functional of the creation and annihilation operators. It was shown by Kristensen, Mejlbo, and Poulsen[8] that the operators $a(p)$ and $a^*(p)$ have a natural dense domain of definition $\sigma \subset H_{\text{in}}$ which is some analog to the Schwartz S-space. They also showed that $a(p)$: $\sigma \to \sigma$ whereas $a^*(p)$: $\sigma' \to \sigma'$, where σ' is the dual of σ. Hence only normally ordered functions of creation and annihilation operators may be well-defined in H_{in}. In particular if we define the quotation marks in (2.1) as the normal ordering with respect to

the creation and annihilation operators of the Φ_{in} field, we see that $\Phi^{(n)}(x)$ is the well-defined sesquilinear form on the domain $D(\Phi)$ being the linear envelope of the coherent state vectors in H_{in}. This suggests that the ordering "$\Phi^3(x)$" =: $\Phi^3(x)$: may lead to a well-defined dynamical equation. It is, however, not known for a time being how this particular ordering rule influences the underlying dynamics described by Eq. (2.1).

3. THE CONSTRUCTION OF THE INTERACTING FIELD $\Phi(x)$

We first construct a convenient dense subset $D(\Phi)$ contained in H_{in}. Let

$$|z> = \exp(-\tfrac{1}{2}\|z\|^2)\exp(za^*)|0>, \tag{3.1}$$

where $za^* = \sum_1^\infty z_k a_k^*$ and $z = \{z_k\}_{k=1}^\infty$ is an element of the Schwartz S-space of quickly decreasing sequences. The vectors $|z\rangle, z \in S$, are called the coherent state vectors. It follows from Eq. (3.1) that

$$\langle z'|z\rangle = \exp\left\{-\sum_{k=1}^{\infty}[\tfrac{1}{2}|z'_k - z_k|^2 - i\,\mathrm{Im}(\bar{z}'_k z_k)]\right\}. \tag{3.2}$$

It is well known that if $\mathcal{Z}$ is a dense subset in the Hilbert space l^2 of complex sequences $\{z_k\}_1^\infty$, then the set $D(\Phi)\{|z\rangle, z \in \mathcal{Z}\}$ is the complete set in H_{in}.[9] Hence the set of all finite linear combinations of coherent states $|z\rangle, z \in S$, is dense in H_{in}.

The following proposition summarizes the main results of the present paper.

Proposition 1: The dynamical equation (2.2) has the unique solution $\Phi(x)$ given by the formula

$$\Phi(x) = [(I - \lambda N_R)^{-1}(\Phi_{\text{in}})](x), \tag{3.3}$$

where for a functional $\Psi(x)$ of the creation and annihilation operators in H_{in} the operator $N_R(\Psi)$ is given by the formula

$$N_R(\psi)(x) = \int \Delta_R(x-y)\colon\psi^3(y)\colon d^4y. \tag{3.4}$$

The field $\Phi(x)$ is a sesquilinear form defined on the dense set $D(\Phi)$ constructed above. The series (3.3) is weakly convergent on $D(\Phi)$ with respect to the F-norm given by the formula [for $u, v \in D(\Phi)$ we denote $\langle u|\Phi(x)|v\rangle \equiv \hat{\Phi}(x)$]

$$\|\hat{\Phi}\|_F^2 = \sup_t \|\hat{\Phi}(t)\|_E^2 + \sup_{t,\mathbf{x}}|\hat{\Phi}(t,\mathbf{x})|^2 + \int \sup_{\mathbf{x}}|\hat{\Phi}(t,\mathbf{x})|^2 dt, \tag{3.5}$$

where $\|\cdot\|_E$ is the energy norm

$$\|\Phi(t)\|_E^2 = \int_{R^3}[\dot{\hat{\Phi}}^2(t,\mathbf{x}) + |\nabla\hat{\Phi}(t,\mathbf{x}|^2 + m^2\Phi^2(t,\mathbf{x})]d^3\mathbf{x}. \tag{3.6}$$

Proof: The nth order iterative approximation of Eq. (2.2) has the form (2.3) where the nonlinear operator $N_R(\cdot)$ is given by the formula (3.4). The field $\Phi_{\text{in}}(x)$ has the following form in terms of a_l and a_k^* operators:

$$\Phi_{\text{in}}(x) = \varphi_k(x)a_k^* + \overline{\varphi_k(x)}a_k, \tag{3.7}$$

where

$$\varphi_k(x) = N\int \exp(ipx)e_k(p)d\nu(p), \qquad d\nu(p) = d^3\mathbf{p}/p_0, \quad N = [2(2\pi)^3]^{-1/2} \tag{3.8}$$

if $|z'\rangle, |z\rangle \in D(\Phi)$, then

$$\frac{\langle z'|\Phi_{\text{in}}(x)|z\rangle}{\langle z'|z\rangle} = \varphi_k(x)\bar{z}'_k + \bar{\varphi}_k(x)z_k \equiv \hat{\Phi}_{\text{in}}(x,\bar{z}',z) \tag{3.9}$$

Moreover, for $m = 2, 3, 4, \cdots$, we have

$$\langle z'|\colon\Phi_{\text{in}}^m(x)\colon|z\rangle/\langle z'|z\rangle = \hat{\Phi}_{\text{in}}^m(x,\bar{z}',z).$$

Since $\Phi^{(n)}(x)$ contains only normally ordered powers of Φ_{in}, we obtain

$$\begin{aligned}\hat{\Phi}^{(n)}(x,\bar{z}',z) &\equiv \langle z'|\Phi^{(n)}(x)|z\rangle\langle z'|z\rangle^{-1}\\ &= \{\hat{\Phi}_{\text{in}} + \lambda\hat{N}_R\{\hat{\Phi}_{\text{in}} + \lambda\hat{N}_R\{\cdots\{\hat{\Phi}_{\text{in}} + \lambda\hat{N}_R\{\Phi_{\text{in}}\\ &\quad + \lambda\hat{N}_R(\hat{\Phi}_{\text{in}})\}\}\cdots\}\},\end{aligned} \tag{3.10}$$

where $\hat{N}_R$ is a nonlinear map which for elements $\hat{\psi}$ with a finite F-norm is given by the formula

$$\hat{N}_R(\hat{\psi})(x) = \int \Delta_R(x-y)\hat{\psi}^3(y)d^4y.$$

The function $\hat{\Phi}^{(n)}(x,\bar{z}',z)$ is the nth order iterative approximation of the classical equation

$$\hat{\Phi}(x,\bar{z}',z) = \hat{\Phi}_{\text{in}}(x,\bar{z}',z) + \lambda\int\Delta_R(x-y)\hat{\Phi}^3(y,\bar{z},z)d^4y. \tag{3.11}$$

The nonlinear equations (3.11) were extensively studied by Segal,[10] Brodsky,[11] Moravetz and Strauss.[12] Let $\hat{\Phi}_0(x,\bar{z}',z)$ be a solution of the free Klein-Gordon equation and let $\hat{\Phi}_0(t=0,\mathbf{x},\bar{z}',z) \equiv \psi_1(\mathbf{x})$ and $\hat{\Phi}_0(t=0,\mathbf{x},\bar{z}',z) \equiv \psi_2(\mathbf{x})$ Define F_1 as the space of the free solutions such that ψ_1 has third derivatives in $L_1(R^3)$ and second derivatives in $L_2(R^3)$ while $\psi_2(\mathbf{x})$ has second derivatives in $L_1(R^3)$ and first derivatives in $L_2(R^3)$. Let F denote a completion of F_1 in the F-norm. Then the Moravetz-Strauss theorem[12 iv] assures that if $\hat{\Phi}_{\text{in}}(x,\bar{z}',z) \in F$, then there exists a unique solution $\hat{\Phi}(x,\bar{z}',z)$ of Eq. (3.11), which has finite F-norm and is given as the limit in F-norm of the iterative series. In addition there exists a free solution $\hat{\Phi}_{\text{out}}(x,\bar{z}',z)$ such that

$$\hat{\Phi}_{\text{in}}(t,\mathbf{x},\bar{z}',z) \underset{t\to-\infty}{\longleftarrow} \hat{\Phi}(t,\mathbf{x},\bar{z}',z) \underset{t\to\infty}{\longrightarrow} \hat{\Phi}_{\text{out}}(t,\mathbf{x},\bar{z}',z), \tag{3.12}$$

in the energy norm.

Now, because $z = \{z_k\}_1^\infty$ and $z' = \{z'_k\}_1^\infty$ are elements of the nuclear S-space of sequences the functions $e_k(p)\bar{z}'_k p_0^{-1}$ and $\bar{e}_k(p)z_k p_0^{-1}$ are elements of the Schwartz S-space of functions in the momentum space; hence $\hat{\Phi}_{\text{in}}(0,\mathbf{x},\bar{z}',z)$ and the transform $\dot{\hat{\Phi}}_{\text{in}}(0,\mathbf{x},\bar{z}',z)$ corresponding to the time derivative $\dot{\Phi}_{\text{in}}(x)$ satisfies the regularity conditions imposed on elements of F-space. Consequently the classical solution $\hat{\Phi}_{\text{in}}(x,\bar{z}',z)$ of the free Klein-Gordon equation given by Eq. (3.9) belongs to F-space. Therefore, there exists a unique solution of Eq. (3.11) in the form:

$$\hat{\Phi}(x,\bar{z}',z) = [(I - \lambda\hat{N}_R)^{-1}(\hat{\Phi}_{\text{in}})](x,\bar{z}',z). \tag{3.13}$$

The iterative approximation (2.3) of $\Phi(x)$ by virtue of Eq. (3.13) is weakly convergent in F-norm on the set of coherent state vectors. If $u = u_i|\overset{i}{z'}\rangle$ and $v = v_k|\overset{k}{z}\rangle$ are arbitrary elements in $D(\Phi)$, then we have

$$\begin{aligned}\langle u|\Phi^{(n)}(x)|v\rangle &= \sum_{i,k}\bar{u}_i v_k\langle\overset{i}{z'}|\overset{k}{z}\rangle\hat{\Phi}^{(n)}(x,\bar{z}',z)\\ &\xrightarrow[n\to\infty]{} \sum_{i,k}\bar{u}_i v_k\langle\overset{i}{z'}|\overset{k}{z}\rangle\hat{\Phi}(x,\overset{\bar{i}}{z'},\overset{k}{z}) \qquad \text{in the F-norm.}\end{aligned}$$

Hence the solution $\Phi(x)$ given by Eq. (3.3) is the sesquilinear form on $D(\Phi)$. QED

From the formula (3.3) it follows that

$$\langle 0|\Phi(x)|0\rangle = 0 \quad \text{and} \quad \langle p|\Phi(x)|0\rangle = [2(2\pi)^3]^{-1/2}\exp(ipx).$$

4. ASYMPTOTIC CONDITIONS

The present approach enables a new precise formulation of asymptotic conditions. Indeed we have

Proposition 2: There exists a quantum field $\Phi_{out}(x)$ which satisfies the free Klein-Gordon equation such that

$$\Phi_{in}(t,\mathbf{x}) \underset{t\to-\infty}{\longleftarrow} \Phi(t,\mathbf{x}) \underset{t\to\infty}{\rightarrow} \Phi_{out}(t,\mathbf{x}) \tag{4.1}$$

in the sense of the sesquilinear forms on $D(\Phi)$. More precisely, for $u, v \in D(\Phi)$ the limits

$$\langle u|\Phi_{in}(t,\mathbf{x})|v\rangle \underset{t\to-\infty}{\longleftarrow} \langle u|\Phi(t,\mathbf{x})|v\rangle \underset{t\to\infty}{\rightarrow} \langle u|\Phi_{out}(t,\mathbf{x})|v\rangle, \tag{4.2}$$

hold in the energy norm (3.6).

Proof: For $|z'\rangle, |z\rangle$ in $D(\Phi)$ the function (3.9) is in F-space. Hence $\hat{\Phi}(t,\mathbf{x},\bar{z}',z)$ given by (3.13) satisfies (3.12). Consequently, for $u = u_i|\overset{i}{z}\rangle$ and $v = v_k|\overset{k}{z}\rangle$ in $D(\Phi)$ we have

$$\lim_{t\to-\infty} \|\langle u|(\Phi(t,\mathbf{x}) - \Phi_{in}(t,\mathbf{x}))|v\rangle\|_E \leq \sum_{i,k} |\bar{u}_i v_k \langle \overset{i}{z}|\overset{k}{z}\rangle| \lim_{t\to-\infty} \|(\hat{\Phi}(t,\mathbf{x},\overset{\bar{i}}{\bar{z}},\overset{k}{z}) - \hat{\Phi}_{in}(t,\mathbf{x},\overset{\bar{i}}{z},\overset{k}{z}))\|_E = 0.$$

Similarly we show using (3.12) that Φ_{out} field given by

$$\Phi_{out}(x) = \{[I - \lambda N_A]\}[(I - \lambda N_R)^{-1}(\Phi_{in})]\}(x), \tag{4.4}$$

where

$$N_A(\Psi)(x) = \int \Delta_A(x-y) {:}\Psi^3(y){:}\, d^4y \tag{4.5}$$

satisfies the condition (4.2) QED

5. RELATIVISTIC COVARIANCE, SPECTRAL CONDITION AND THE UNIQUENESS OF VACUUM

Proposition 3: Let $(a,\Lambda) \to U^{in}_{(a,\Lambda)}$ be the unitary representation of the Poincaré group in the Fock space H_{in}. Then

$$U^{in}_{(a,\Lambda)}\Phi(x)(U^{in}_{(a,\Lambda)})^{-1} = \Phi(\Lambda x + a), \tag{5.1}$$

$$U^{in}_{(a,\Lambda)}\Phi_{out}(x)(U^{in}_{(a,\Lambda)})^{-1} = \Phi_{out}(\Lambda x + a). \tag{5.2}$$

These formulae are understood as the equality of sesquilinear forms on $D(\Phi)$.

Proof: If $|z\rangle, z \in S$, is a coherent state, then $|z_{(a,\Lambda)}\rangle \equiv (U^{in}_{(a,\Lambda)})^{-1}|z\rangle$ is also a coherent state with $z_{(a,\Lambda)} \in S$. Hence Eq. (3.10) implies

$$\hat{\Phi}^{(n)}(x,\bar{z}'_{(a,\Lambda)}, z_{(a,\Lambda)}) = \hat{\Phi}^{(n)}(\Lambda x + a, \bar{z}', z).$$

For $n \to \infty$ both sides of this equality converge in F-norm. This implies Eq. (5.1). Similarly, using (4.4), we derive (5.2). QED

The formula (5.1) implies that the total 4-momentum P_μ, $\mu = 0, 1, 2, 3$, coincides with P^{in}_μ. Because $H = H_{in}$ and $U_{(a,\Lambda)} = U^{in}_{(a,\Lambda)}$ there is exactly one state in H which is invariant under all Poincaré transformation: this is the Fock vacuum $|0\rangle$. The operator $M = (P_\mu P^\mu)^{1/2}$ has discrete eigenvalues 0 and m and the continuum of mass values above $2m$. Hence the spectral condition in the Wightman or the Haag-Ruelle form is automatically satisfied in this model.

6. IRREDUCIBILITY

Let $\Psi(x)$ be a sesquilinear form, defined on the dense set $D(\Psi)$ in a Hilbert space H. we say that $\Psi(x)$ is irreducible if for every self-adjoint bounded operator B with the range $R(B) \subset D(\Psi)$ the equality

$$\langle u|\Psi(x)B|v\rangle = \langle u|B\Psi(x)|v\rangle \quad \text{for all } u, v \in D(\Psi) \tag{6.1}$$

implies $B = bI$, $b \in R$. In our case we have

Proposition 4: The field $\Phi(x)$ is irreducible.

Proof: Let $\Phi(x)$ satisfies (6.1). Then for $u, v \in D(\Phi)$ we have

$$\langle u|(\Box + m^2)\Phi(x)B|v\rangle = \langle u|B(\Box + m^2)\Phi(x)|v\rangle.$$

Therefore the sesquilinear form: $\Phi^3(x)$: commutes with B. Hence by virtue of Eq. (2.2) we conclude that B commutes with $\Phi_{in}(x)$ on $D(\Phi)$. Hence $B = bI$. QED

7. RELATIVISTIC PRIMITIVE CAUSALITY

One distinguishes in quantum field theory at least three concepts of causality[13]:

(i) Einstein causality (i.e., local commutativity),

(ii) primitive causality,

(iii) relativistic primitive causality.

The relativistic primitive causality for sesquilinear form $\Psi(x)$ demands that if a region O_2 in R^4 is causally dependent on a region O_1 (i.e., every ray in the backward cone originating from O_2 passes through O_1), then all matrix elements $\langle u|\Psi(x)|v\rangle$, $(t,\mathbf{x}) \in O_2$, $u, v \in \boldsymbol{D}(\Psi)$, are determined by the matrix elements of $\Psi(x)$ for $(t,\mathbf{x}) \in O_1$.

In our case we have

Proposition 5: The field $\Phi(x)$ given by Eq. (3.3) satisfies the relativistic primitive causality condition.

Proof: For $|z'\rangle, |z\rangle$ in $D(\Phi)$ the function $\hat{\Phi}(t,\mathbf{x},\bar{z}',z)$ satisfies the local hyperbolic equation (3.11). Hence the assertion of Proposition 5 follows from the causality of classical solutions. QED

8. SCATTERING OPERATOR

The scattering operator is defined in the Heisenberg picture as the transformation which connects Φ_{out} and Φ_{in} fields, i.e., $\Phi_{out} = \tilde{S}(\Phi_{in})$. In conventional approach it is assumed that Φ_{out} field is canonical, irreducible and possesses the same set of smearing test functions as Φ_{in} field; in such a case transformation $\tilde{S}$ is unitary implemented in H, i.e., $\Phi_{out} = S^{-1}\Phi_{in}S$. In our case we have

Proposition 6: The transformation $\tilde{S}$ is given by the formula

$$\Phi_{out}(x) = \tilde{S}(\Phi_{in})(x) = \{[I + \lambda N(I - \lambda N_R)^{-1}](\Phi_{in})\}(x), \tag{8.1}$$

where for functional $\Psi(x)$ of creation and annihilation operators the operation $N(\Psi)$ is defined by the formula

$$N(\Psi)(x) = \int \Delta(x-y) {:}\Psi^3(y){:}\, d^4y. \tag{8.2}$$

The transformation $\tilde{S}$ has the following properties:

(i) $\tilde{S} \neq I$; (8.3)

(ii) $\tilde{S}[U^{\text{in}}_{(a,\Lambda)} \Phi_{\text{in}} (U^{\text{in}}_{(a,\Lambda)})^{-1}] = U^{\text{in}}_{(a,\Lambda)} \tilde{S}(\Phi_{\text{in}})(U^{\text{in}}_{(a,\Lambda)})^{-1}$; (8.4)

(iii) $\tilde{S}$ possesses the inverse given by the formula

$$\Phi_{\text{in}}(x) = \tilde{S}^{-1}(\Phi_{\text{out}})(x) = \{[I - \lambda N(I - \lambda N_A)^{-1}](\Phi_{\text{out}})\}(x). \quad (8.5)$$

The convergence of series (8.1) and (8.5) is the weak convergence on $D(\Phi)$ with respect to E-norm (3.6).

Proof: Taking the matrix elements of (8.1) between coherent states vectors $|z'\rangle$, $|z\rangle$ in $D(\Phi)$, we reduce (8.1) to the corresponding equality for classical wavefunctions $\hat{\Phi}(x, \bar{z}', z)$ and $\hat{\Phi}_{\text{in}}(x, \bar{z}', z)$, which holds by virtue of Moravetz-Strauss theorem.[12v] Similarly one proves Eq. (8.1) and all remaining assertions for arbitrary vectors u, v in $D(\Phi)$. QED

We are not able to prove or disprove for the time being that the transformation $\tilde{S}$ is unitary implemented in the carrier space H. If $\tilde{S}$ would be unitary implemented, then the scattering operator $S \neq I$.

9. DISCUSSION

A. The results of the present paper may be extended to a class of nonpolynomial interactions λ: $F(\Phi)$:, where $F(\cdot)$ satisfies the conditions:

(i) $F(\cdot)$ is an odd analytic function,

(ii) $F'(0) = 0$,

(iii) $|F(z)z^{-5}| \to 0$ as $|z| \to \infty$.

The extension of the present results may be proven by using in the proof of Proposition 1 the corresponding results for the classical nonlinear relativistic wave equations with analytic nonlinear term.[12iv,12v]

B. The present formulation may be easily extended for the case of n-dimensional space-time, $n = 2, 3, 4, \cdots$. Again it is sufficient to use the corresponding results for the classical equations.[10iii]

C. To complete the present approach one should clarify the problem of local commutativity and unitary implementation of the scattering operator (8.1). These problems for a time being are open.

ACKNOWLEDGMENTS

I should like to thank Professors A. O. Barut, I. Białynicki-Birula, M. Flato, E. Kapuścik, R. Streater, and S. Woronowicz for interesting discussions and valuable suggestions.

[1]P. A. M. Dirac, Physics Today 23, No. 4, 29 (1970).
[2]I. Segal, cf., e.g., *Functional Analysis and Related Fields*, edited by F. Browder (Springer-Verlag, Berlin, 1970), pp. 188-210.
[3]J. Glimm, and A. Jaffe, "Quantum Field Theory Models," in *Proceedings of International Summer School on Quantum Field Theory Les Houses 1971*, edited by C. DeWitt and R. Stora (Gordon and Breach, New York, 1972) and references contained therein.
[4]G. F. Dell'Antonio, *Proceedings of XII International Winter School on Mathematical Physics. Schladming, 1973*, edited by P. Urban, Acta Physica Austriaca, Suppl. 11, (1973).
[5]P. G. Federbush, Progr. Theoret. Phys. 26, 148 (1961).
[6]N. N. Bogolubov, A. Logunov and I. Todorov, *Principle of Axiomatic Approach in Quantum Theory* (Moscow, 1969).
[7]R. Rączka, (i) "Global Solutions of Nonlinear Equations for the Scalar Quantum Field for Polynomial and Nonpolynomial Interactions", preprint IC/72/106, Trieste. (ii) K. Jezuita, and R. Rączka, "Integral Representations for Unbounded Operators and Bilinear Forms," preprint 1973.
[8]P. Kristensen, L. Mejlbo, and E. T. Poulsen, (I). Comm. Math. Phys. 1, 175 (1965); (II) Math. Scand. 14, 129 (1964); (III) Comm. Math. Phys. 6, 29 (1967).
[9]J. R. Klauder and G. Sudarshan, Fundamentals of Quantum Optics (Benjamin, New York, 1968), Chap. VII, Sec. 4.
[10]I. E. Segal, (i) *Quantum Fields and Analysis in the Solution Manifolds of Differential Equations, Analysis in Function Space*, edited by T. W. Martin and I. Segal (M.I.T. Press, Cambridge, 1965), pp. 129-153; (ii) *The Global Cauchy Problem for a Relativistic Scalar Field with Power Interaction, Proceedings of International Congress of Mathematicians* (Mir, Moscow, 1966) pp. 681-690 (iii) Ann. Sci. Ecole Norm. Sup. (4) 1, 459 (1968).
[11]A. R. Brodsky, Pacific J. Math. 19, 1 (1966).
[12](i) C. S. Moravetz, Comm. Pure Appl. Math. 16, 353 (1963); (ii) Proc. Roy. Soc. A 306, 291 (1968); (iii) W. A. Strauss, J. Funct. Anal. 2, 409, 457 (1968); (iv) C. S. Moravetz and W. A. Strauss, Comm. Pure Appl. Math. 25, 1 (1972); (v) C. S. Moravetz and W. A. Strauss, "Some Properties of Nonlinear Relativistic Scattering Operator," preprint.
[13]R. Haag and B. Schroer, J. Math. Phys. 3, 248 (1962).

Application of nonstandard analysis to quantum mechanics

M. O. Farrukh

Physics Department, Imperial College of Science and Technology, London, S.W. 7, England
(Received 8 June 1973)

Quantum mechanics is formulated using a nonstandard Hilbert space. The concept of an eigen vector of a linear operator, which applies to standard as well as nonstandard Hilbert spaces, is replaced by the more general concept of an ultra eigen vector, which applies to nonstandard Hilbert spaces alone. Ultra eigen vectors corresponding to all spectral points of internal self-adjoint operators are proved to exist. This result enables us to set up a formalism which is equally valid for the discrete, as well as the continuous spectrum. Finally, Dirac's formalism is reproduced, in a rigorous form within the nonstandard Hilbert space structure.

1. INTRODUCTION

Conventional quantum mechanics is usually formulated in a separable complex Hilbert space.[1] States are represented by unit rays in the space, and observables are represented by densely defined self-adjoint operators. The relation between the measuring process and the spectral properties of self-adjoint operators can be established as follows. Let $\mathcal{H}$ be the separable complex Hilbert space. If $f, g \in \mathcal{H}$, then their scalar product is denoted by the complex number $\langle f,g\rangle \in C$, and taken to be antilinear in the first argument, but linear in the second one. The norm $\|f\|$ is defined as $\|f\| = \sqrt{\langle f,f\rangle}$.

If A is a densely defined self-adjoint operator, then there exists a spectral measure E, which assigns to every Borel subset Δ of the real line R, a projection operator E_Δ, and such that, if $E_\lambda \equiv E_{(-\infty,\lambda)}$, then by the spectral theory[2] we have

$$A = \int \lambda \, dE_\lambda . \tag{1.1}$$

Now, if the system under consideration is in a state $f(\|f\| = 1)$, then the probability that the measurement of the observable A yields a result in the Borel set $\Delta \subset R$ is:[1]

$$\text{prob}(A,f,\Delta) = \langle f, E_\Delta f\rangle . \tag{1.2}$$

Consider first the case where the spectrum of A is purely discrete (and nondegenerate for simplicity). In this case, there exists a sequence of real numbers $\{a_n : n \in N\}$ and a sequence of orthonormal vectors $\{\phi_n : n \in N\}$, such that

$$A\phi_n = a_n \phi_n, \quad \text{for all } n \in N, \tag{1.3a}$$

$$\langle \phi_n, \phi_{n'}\rangle = \delta_n^{n'}, \quad \text{for all } n,n' \in N. \tag{1.3b}$$

If $\hat{\mathcal{H}}$ is the dual space of $\mathcal{H}$ (space of continuous linear forms on $\mathcal{H}$), then, for every $f \in \mathcal{H}$, a continuous linear form L_f may be defined such that for all $\varphi \in \mathcal{H}$, $L_f(\varphi) = \langle f, \varphi\rangle$. Conversely, by the well-known Riesz theorem,[3] any element of $\hat{\mathcal{H}}$ is of the form L_f for some $f \in \mathcal{H}$.

If $\varphi \in \mathcal{H}$ and $\xi \in \hat{\mathcal{H}}$, then the tensor product $\varphi \otimes \xi$ is defined to be the bounded linear operator on $\mathcal{H}$ which sends $f \in \mathcal{H}$ into

$$(\varphi \otimes \xi)(f) = \xi(f)\,\varphi.$$

By using these definitions, we can show that the spectral measure corresponding to A [as defined by (1.3a) and (1.3b)] is

$$E_\Delta = \sum_{a_n \in \Delta} \varphi_n \otimes L\varphi_n \quad \text{for all Borel sets } \Delta \subset R. \tag{1.3c}$$

In case Δ contains an infinite number of eigenvalues of A (but still a countable number), the convergence in (1.3c) is the strong pointwise convergence.

By taking $\Delta = \{a_n\}$, Eq. (1.2) reduces to

$$\begin{aligned}\text{prob}(A,f,\{a_n\}) &= \langle f, E_{\{a_n\}} f\rangle \\ &= \langle f, (\varphi_n \otimes L\varphi_n)(f)\rangle \\ &= \langle f, \varphi_n \langle \varphi_n, f\rangle\rangle \cdot \\ &= \langle f, \varphi_n\rangle\langle \varphi_n, f\rangle \\ &= |\langle \varphi_n, f\rangle|^2 .\end{aligned} \tag{1.4}$$

Equation (1.4) can be interpreted as follows:[4]

I_1: The result of any measurement of an observable can only be one of the eigenvalues of the corresponding operator. As a result of the measurements, the physical system finds itself in the state represented by the corresponding eigenvector.

I_2: If the system is known to be in the state f, then the probability that a measurement of an observable A on the state f yields the value a_n is given by

$$\text{prob}(A,f,\{a_n\}) = |\langle f, \phi_n\rangle|^2 .$$

Dirac tried to extend I_1 and I_2 to the continuous spectrum by considering a larger space.[5] His bra and ket formalism, when applied to the discrete spectrum is just a short hand. He writes, for example,

$$A\,|n\rangle = a_n\,|n\rangle \quad \text{for all } n \in N, \tag{1.5a}$$

$$\langle n\,|\,n'\rangle = \delta_n^{n'} \quad \text{for all } n,n' \in N, \tag{1.5b}$$

$$\sum_{n\in N} |n\rangle\langle n| = I. \tag{1.5c}$$

Equations (1.5a) and (1.5b) are just a shorthand for (1.3a) and (1.3b), respectively, while (1.5c) is a shorthand for the special case of (1.3c) obtained by setting $\Delta = R$, I being the identity operator on $\mathcal{H}$.

In going to the continuous spectrum, Dirac assumed the existence of a set of objects $\{|\xi\rangle : \xi \text{ in the spectrum of } A\}$ satisfying analogous properties:

$$A\,|\xi\rangle = \xi\,|\xi\rangle \quad \text{for all } \xi \text{ in the spectrum of } A, \tag{1.6a}$$

$$\langle \xi\,|\,\xi'\rangle = \delta(\xi - \xi') \quad \text{for all } \xi, \xi' \text{ in the spectrum of } A, \tag{1.6b}$$

$$\int d\xi\,|\,\xi\rangle\langle\,\xi\,| = I. \tag{1.6c}$$

Unfortunately, these equations cannot be rigorously justified in this form (Dirac never defined his larger space). We note, in particular, that Eq. (1.6b) involves the so-called δ-Dirac function, which, mathematically, fails to exist as a function.

A lot of work has been done to make these equations rigorous.[6] We mention, in particular, the extension of $\mathcal{H}$ to a rigged Hilbert space $(\Phi, \mathcal{H}, \Phi')$ introduced by Gel'fand, where $\Phi \subset \mathcal{H}$ is a dense subset endowed with a finer topology and Φ' is the dual space of Φ, equipped with the strong dual topology.

Because the topology of Φ is finer than that of $\mathcal{H}$, we have the following.

(i) The canonical embedding $\tau : \Phi \to \mathcal{H}$ is continuous.

(ii) $\hat{\mathcal{H}} \subset \Phi'$ (sometimes written as $\mathcal{H} \subset \Phi'$).

The following conditions are required to be satisfied.

(1) Φ is contained in the domain of the self-adjoint operator A under consideration, and is stable under it [i.e., $A(\Phi) \subset \Phi$].

(2) The restriction of A to Φ is a continuous map from Φ into itself, relative to the topology of Φ.

(3) For every λ in the spectrum of A, there exists an element $\xi_\lambda \in \Phi'$ such that $\xi_\lambda(A\varphi) = \lambda \xi_\lambda(\varphi)$ for all $\varphi \in \Phi$.

The operator A can easily be extended to Φ' using the definition

$$(A'\xi)(\varphi) = \xi(A\varphi) \quad \text{for all } \xi \in \Phi' \quad \text{and} \quad \varphi \in \Phi.$$

In particular,

$$A'\xi_\lambda = \lambda \xi_\lambda \quad \text{for all } \lambda \text{ in the spectrum of } A. \tag{1.7}$$

It can be shown that there exist a measure μ on the real line and a system of elements of Φ' denoted by $\{\xi_{\lambda,n} : n \in N \text{ and } \lambda \text{ in the spectrum of } A\}$, such that any element $\xi \in \Phi'$ satisfying $A'\xi = \lambda\xi$ can be written as $\xi = \sum_{n\in N} \alpha_n \xi_{\lambda,n}$ $(\alpha_n \in C)$ and such that, for all $\varphi, \psi \in \Phi$, we have

$$\langle \varphi, \psi \rangle = \int d\mu(\lambda) \sum_{n\in N} (\xi_{\lambda,n}(\varphi))^* \xi_{\lambda,n}(\psi). \tag{1.8a}$$

By defining $\xi^+_{\lambda,n}(\varphi) = (\xi_{\lambda,n}(\varphi))^*$ and $\xi^+_{\lambda,n} \otimes \xi_{\lambda,n}(\varphi,\psi) = \xi^+_{\lambda,n}(\varphi)\,\xi_{\lambda,n}(\psi)$ for the pair (φ,ψ), Eq. (1.8a) can be written as

$$\langle \varphi, \psi \rangle = \int d\mu(\lambda) \sum_{n\in N} \xi^+_{\lambda,n} \otimes \xi_{\lambda,n}(\varphi,\psi). \tag{1.8b}$$

Now, if Δ is a Borel subset of the real line, there exists a projection operator E_Δ on $\mathcal{H}$ such that, for all $\varphi, \psi \in \Phi$, we have

$$\langle \varphi, E_\Delta \psi \rangle = \int_{\lambda\in\Delta} d\mu(\lambda) \sum_{n\in N} \xi^+_{\lambda,n} \otimes \xi_{\lambda,n}(\varphi,\psi). \tag{1.8c}$$

This motivates the introduction of the symbolic definition

$$E_\Delta = \int_{\lambda\in\Delta} d\mu(\lambda) \sum_{n\in N} \xi^+_{\lambda,n} \otimes \xi_{\lambda,n} \tag{1.9}$$

using which, Eqs. (1.6a) and (1.6c) can be put into the rigorous forms (1.7) and (1.9), respectively. However, Eq. (1.6b) remains unjustified.

The physical interpretation of the rigged Hilbert space is as follows: Physical states are represented by unit rays in Φ (and not $\mathcal{H}$!). Observables are represented by self-adjoint operators densely defined in $\mathcal{H}$, with their domains containing Φ. Experiments are represented by elements of Φ'. The result of any measurement of an observable can be one of its eigenvalues, when the eigenvalue equation is solved in Φ'; but the result, if it does not belong to the spectrum of the corresponding operator in $\mathcal{H}$, must be rejected!

The probabilistic interpretation of the rigged Hilbert space structure is as follows: Let a physical system be in a state $\varphi \in \Phi$ $(\|\varphi\| = 1)$. Let λ be an admissible eigenvalue of a self-adjoint operator A. Let $f \in \Phi'$ be such that $A'f = \lambda f$.

Now, if $f \in \hat{\mathcal{H}}$, then by Riesz theorem,[3] there exists an element $h \in \mathcal{H}$ such that $f = L_h$. The probability that the measurement of A yields the value λ in the state φ is $|\langle h, \varphi\rangle|^2$ $(\|h\| = 1)$. But if $f \notin \hat{\mathcal{H}}$ (λ belongs to the continuous spectrum of A), then $|f(\varphi)|^2$ is the probability density that the measurement of A yields the value λ. The absolute probability that the measurement of A yields a value in $\Delta \subset R$ (a Borel subset) is $\int_{\lambda\in\Delta} d\mu(\lambda)\,|\xi_\lambda(\varphi)|^2$, where $A'\xi_\lambda = \lambda\xi_\lambda$ and μ is the measure used in (1.8a).

The difficulty with this formalism is that, instead of enlarging the space $\mathcal{H}$ of physical states to include eigenstates of the continuous spectrum, it reduces it to $\Phi \subset \mathcal{H}$. Thus, we, still cannot talk about transition to eigenstates induced by the measuring process itself. The role of $\mathcal{H}$ is reduced to a minimum and seems artificial. Observables, which are linear operators on Φ, continuous with respect to its fine topology, and symmetric with respect to the scalar product in Φ, must admit self-adjoint extensions to $\mathcal{H}$. Also, the eigenvalue equation when solved in Φ' gives an admissible result if and only if it corresponds to a value in the spectrum of the operator under consideration in $\mathcal{H}$. The space Φ', which is in a sense, an enlargement of $\mathcal{H}$, fails to replace $\mathcal{H}$, because it does not have a scalar product.

In this paper, we are going to extend the Hilbert space $\mathcal{H}$ to a nonstandard Hilbert space ${}^*\mathcal{H}$[7,8] in order to construct a formalism which has none of the disadvantages above. The additional properties of ${}^*\mathcal{H}$ are gained at no price, because, except for linguistic modifications, all the properties of $\mathcal{H}$ are carried into ${}^*\mathcal{H}$, due to the transfer theorem.[9] The basic idea is the nonstandard extension of a well-known property of standard Hilbert spaces. In the standard theory, if λ belongs to the continuous spectrum of a self-adjoint operator A, densely defined on $\mathcal{H}$, then for any $\epsilon > 0$ there exists a vector $f \in \mathcal{H}$, $\|f\| = 1$, such that $\|Af - \lambda f\| < \epsilon$. The vector f is said to be "almost" an eigenvector[10] with eigenvalue λ and "error" not larger than ϵ. This concept, in the standard theory, is not very helpful, because, due to the absence of the notion of absolutely small quantities, any vector $f \in \mathcal{H}$, $\|f\| = 1$ is almost an eigenvector of any operator A, with eigenvalue being any number λ and error $\|Af - \lambda f\|$. Even if we put an upperbound for errors, the determination of the eigenvalue corresponding to an almost eigenvector of a given operator is not possible.

In the nonstandard theory, the situation is different. The nonstandard real line ${}^*R$, which replaces the standard real line R and contains it as a subset (more precisely, contains a subset isomorphic to R), has the additional property of containing nonzero numbers (known as infinitesimal numbers), which are less in magnitude than any positive standard number (i.e., element of R). By allowing the error to be an infinitesimal number only, the concept of an almost eigenvector can be replaced by the concept of an ultra eigenvector. The crucial point is that for a given vector f, $\|f\| = 1$, there exists at most one standard number λ, such that $\|Af - \lambda f\|$ is infinitesimal, A being a given operator. However, if λ belongs to the spectrum of A (whether discrete or continuous), then there exists at least one vector f, $\|f\| = 1$, such that $\|Af - \lambda f\|$ is infinitesimal. The existence of f for a given λ (in the spectrum of A)

and the uniqueness of λ (if restricted to be standard) when f is given, make the concept of an ultra eigenvector very useful. Hence, by using ultra eigenstates of dynamical observables instead of eigenstates, quantum mechanics can be formulated in such a way that no distinction between the discrete and the continuous spectra is required.

In Sec. 2, we state the main facts about nonstandard analysis, and show how to construct a nonstandard Hilbert space. Section 3 is concerned with the study of the spectrum of internal operators (mainly self-adjoint) in a nonstandard Hilbert space. While in Sec. 4, we construct probability functions associated with standard self-adjoint operators and defined on a ring of Borel subsets of the real line. They will serve as probability functions for values obtained during the measuring process. In Sec. 5 we use all the preceding results to formalize quantum mechanics using a nonstandard Hilbert space. In particular, we demonstrate the manner in which both the continuous and discrete spectra can be dealt with on an equal footing. Finally, Sec. 6 is concerned with the construction of wavefunctions of nonrelativistic free particles as an illustration of the method, using a nonstandard function similar to the "delta Dirac function".

2. NONSTANDARD THEORIES

Nonstandard analysis is a set theoretic tool which allows us to embed any given set D into a larger set ${}^*D$ in such a way that any structure on D can be extended to a corresponding "internal" structure on ${}^*D$. The procedure is very simple and universal. An equivalence relation is defined on the class of all sequences. If D is a set, then ${}^*D$ is the set of equivalence classes of sequences of elements of D. If E is a structure on D (e.g., E is a subset of D, or E is a function from $D \times D$ into D, etc.), then ${}^*E$ (constructed in the same way) is not a structure on ${}^*D$, but is naturally isomorphic to a similar structure on ${}^*D$. The natural isomorphism depends on the particular case concerned, and because we meet it very frequently, the procedure seems to lose some of its simplicity, but only superficially.

The set D is naturally isomorphic to a subset of ${}^*D$ which we denote by $\hat{D}$. If E is a structure on D, then the restriction to $\hat{D}$ of the structure on ${}^*D$ isomorphic to ${}^*E$ is isomorphic to E itself. This allows us to identify D with $\hat{D}$ and consider ${}^*D$ as an extension of D. The crucial point is that the structure on ${}^*D$ isomorphic to ${}^*E$ satisfies, apart from some linguistic modifications to be stated later, the same properties satisfied by E. Hence, nothing is practically lost in going from D to ${}^*D$. What we gain is a wider concept of "convergence." Sequences which diverge in D with respect to its topology can be represented in ${}^*D$ by elements which are independent of any topology on D. The algebra applied to these elements is the same algebra applied to the "standard" elements. It appears as if we could make all sequences in D convergent.

The universal equivalence relation on the class of all sequences mentioned above can be defined by using the concept of a free ultra filter on the set of natural numbers. The existence of at least one free ultra filter is guaranteed by the axiom of choice; but the choice of a specific free ultra filter seems to be irrelevant to the problems with which we are concerned.

A. Ultra powers

Definition (2.1): A class $\mathfrak{F}$ of subsets of the set of natural numbers N (nonnegative integers) is called a free ultra filter, if an only if it satisfies:

(F1) $\phi \notin \mathfrak{F}$, but $N \in \mathfrak{F}$.

(F2) If $A \in \mathfrak{F}$ and $B \in \mathfrak{F}$, then $A \cap B \in \mathfrak{F}$.

(F3) If $A \in \mathfrak{F}$ and $A \subset B$ and $B \subset N$, then $B \in \mathfrak{F}$.

(F4) If $A \subset N$ then either $A \in \mathfrak{F}$ or $N \setminus A \in \mathfrak{F}$.

(F5) If $A \subset N$ and A is a finite set, then $N \setminus A \in \mathfrak{F}$.

Let D be a nonempty set and denote by $\Gamma(D)$ the set of all sequences of elements of D. If $a \in D$, then the constant sequence in D having the value a is denoted by $\hat{a}$. Let $\mathfrak{F}$ be a free ultra filter on N. We write $x \overset{D,\mathfrak{F}}{\sim} x'$ for $x \in \Gamma(D)$, $x' \in \Gamma(D)$ and $\{n : n \in N \text{ and } x_n = x'_n\} \in \mathfrak{F}$. It can be easily seen that $\overset{D,\mathfrak{F}}{\sim}$ is an equivalence relation on $\Gamma(D)$, induced by the free ultra filter $\mathfrak{F}$. Hence the equivalence class $[x]_{D,\mathfrak{F}}$ of $x \in \Gamma(D)$ is defined to be the set $\{x' : x \overset{D,\mathfrak{F}}{\sim} x'\}$. The quotient set $\Gamma(D)/\overset{D,\mathfrak{F}}{\sim}$ is the set of equivalence classes of elements of $\Gamma(D)$, i.e., the set $\{[x]_{D,\mathfrak{F}} : x \in \Gamma(D)\}$. We define $I_{\mathfrak{F}}(D)$ as the subset of the quotient set $\Gamma(D)/\overset{D,\mathfrak{F}}{\sim}$ whose elements are equivalence classes of constant sequences in D, i.e., the set $\{[\hat{x}]_{D,\mathfrak{F}} : x \in D\}$.

We shall adopt one free ultra filter (but leave it unspecified!) and write ${}^*D$ for $\Gamma(D)/\overset{D,\mathfrak{F}}{\sim}$ and $\hat{D}$ for $I_{\mathfrak{F}}(D)$, $\mathfrak{F}$ being understood. It is very clear that $\hat{D} \subset {}^*D$.

Theorem 2.1: The function

$$\tau_D : \begin{cases} D \to \hat{D} \\ x \to \tau_D(x) = [\hat{x}]_{D,\mathfrak{F}} \end{cases}$$

is a bijection.

Proof: τ_D, by definition of $\hat{D}$, is surjective. To show that it is injective, let $x, y \in D$ such that $x \neq y$. Now for all $n \in N$, $\hat{x}_n = x$ and $\hat{y}_n = y$. Hence, for all $n \in N$, $\hat{x}_n \neq \hat{y}_n$, i.e., $\{n : n \in N \text{ and } \hat{x}_n = \hat{y}_n\} = \phi$. But by (F1) $\phi \notin \mathfrak{F}$. Hence $[\hat{x}]_{D,\mathfrak{F}} \neq [\hat{y}]_{D,\mathfrak{F}}$, or $\tau_D(x) \neq \tau_D(y)$. ■

We shall identify D with $\hat{D}$, and call ${}^*D$ the nonstandard extension of D.

Definition 2.2: we say that a statement $P(n)$ is true for almost every $n \in N$ [and write it for short $\nabla_n P(n)$] if and only if it is true for a subset of N belonging to $\mathfrak{F}$.

Using Def. 2.2 we show how structures on D may be extended to ${}^*D$. We notice first that if x and y are two sequences of elements of D which coincide almost everywhere (i.e., for almost every $n \in N$, $x_n = y_n$), then x and y define one and the same element of ${}^*D$. Let us be given a class T of subsets of $A_1 \times A_2 \times \cdots \times A_n$, where $(A_1, A_2, \ldots, A_n)$ is an n-tuple of nonempty sets. ${}^*T$ is not a class of subsets of ${}^*A_1 \times {}^*A_2 \times \cdots \times {}^*A_n$. However, ${}^*T$ is naturally isomorphic to a class of subsets of ${}^*A_1 \times {}^*A_2 \times \cdots \times {}^*A_n$. This natural isomorphism may be visualized as follows: Let $S \in {}^*T$. If $R, R' \in S$, then R and R' are two sequences of subsets of $A_1 \times A_2 \cdots \times A_n$ which are equal almost everywhere (i.e., for almost every $j \in N, R_j = R'_j$). Let now $(y^1, \ldots, y^n) \in {}^*A_1 \times \cdots \times {}^*A_n$ and let $x^1, x'^i \in y^i$ $(i = 1, \ldots, n)$. Now x^i and x'^i are equal almost everywhere (i.e., for almost every $j \in N, x^i_j = x'^i_j$). We can verify that for almost every $j \in N$, $(x'^1_j, x'^2_j, \ldots, x'^n_j) \in R'_j$ if and only if, for almost every $j \in N$, $(x^1_j, x^2_j, \ldots, x^n_j) \in R_j$. Hence, given an element $S \in {}^*T$, we can construct an element $\check{S} \subset {}^*A_1 \times \cdots \times {}^*A_n$ in such a way that

$(y^1, \ldots, y^n) \in \check{S}$ if and only if for any $R \in S$ and $x^i \in y^i (i = 1, \ldots, n)$ the following statement holds:

For almost every $j \in N$, $(x_j^1, x_j^2, \ldots, x_j^n) \in R_j$.

Mainly, because of (F4), we may verify that if $S_1 \neq S_2$, $S_1, S_2 \in {}^*T$ then $\check{S}_1 \neq \check{S}_2$ (according to the above construction). These considerations motivate the following definition:

Definition 2.3: Let T be a class of subsets of $A_1 \times A_2 \times \cdots \times A_n$, and let $S \in {}^*T$. Then $\sigma^T_{(A_1, A_2, \ldots, A_n)}(S) \subset {}^*A_1 \times {}^*A_2 \times \cdots \times {}^*A_n$ is defined as follows:

$$(x_1, x_2, \ldots, x_n) \in \sigma^T_{(A_1, A_2, \ldots, A_n)}(S) \iff \exists_R \exists_{y^1} \exists_{y^2} \cdots \exists_{y^n} S = [R]_{T, \mathfrak{F}},$$
$$x_1 = [y^1]_{A_1, \mathfrak{F}}, x_2 = [y^2]_{A_2, \mathfrak{F}}, \ldots, x_n = [y^n]_{A_n, \mathfrak{F}},$$
$$\text{and } \forall_j (y_j^1, y_j^2, \ldots, y_j^n) \in R_j.$$

Now, if $R \in T$ (i.e., $R \subset A_1 \times A_2 \times \cdots \times A_n$) then $\tau_T(R) \in {}^*T$ (Theorem 2.1) and so $\sigma^T_{(A_1, A_2, \ldots, A_n)}(\tau_T(R)) \subset {}^*A_1 \times {}^*A_2 \times \cdots \times {}^*A_n$ (Def. 2.3). Write $\sigma_{(A_1, A_2, \ldots, A_n)}$ for $\sigma^{T_0}_{(A_1, A_2, \ldots, A_n)} \circ \tau_{T_0}$ with T_0 being the class of all subsets of $A_1 \times A_2 \times \cdots \times A_n$ [denoted by $P(A_1 \times A_2 \times \cdots \times A_n)$.].

In the following definition, take $T = P(A_1 \times A_2 \times \cdots \times A_n)$.

Definition 2.4: A relation $R \subset {}^*A_1 \times {}^*A_2 \times \cdots \times {}^*A_n$ is called:

(i) *Internal* if and only if there exists an element $S \in {}^*T$ such that $R = \sigma^T_{(A_1, A_2, \ldots, A_n)}(S)$.

(ii) *Standard* if and only if there exists an element $S \in T$ such that $R = \sigma_{(A_1, A_2, \ldots, A_n)}(S)$.

(iii) *External* if and only if it is not internal.

It is very clear that standard relations are internal. The following examples may illustrate the terminology: Consider ${}^*N$, the nonstandard extension of N (identifying N with $\hat{N}$). Now:

(a) N is an external subset of ${}^*N$. (No sequence of subsets of N generates N.)
ϕ and ${}^*N$ are standard internal subsets of ${}^*N$ (generated by ϕ and N, respectively).
If $n \in {}^*N \setminus N$ (which is not empty), then $\{n\}$ is an internal subset of ${}^*N$ which is not standard. (If $K \in n$, then the sequence $\{\{K_j\} : j \in N\}$ generates $\{n\}$.)

(b) The function

$$f: \begin{cases} {}^*N \to {}^*N \\ n \to f(n) = n \end{cases}$$

is standard internal (generated by the identity map on N). The function

$$f_0: \begin{cases} {}^*N \to {}^*N \\ n \to f(n) \end{cases} = \begin{cases} n & \text{if } n \in N \\ 0 & \text{if } n \in {}^*N \setminus N \end{cases}$$

is external. (No sequence of functions from N into N generates f_0.)

The function

$$g_K: \begin{cases} {}^*N \to {}^*N \\ n \to g_K(n) = K, \quad K \in {}^*N \end{cases}$$

is internal. It is standard if $K \in N$, but not standard if $K \in {}^*N \setminus N$.

We state now the most important theorem in nonstandard analysis known as the transfer theorem.[9] It is in fact a meta-theorem whose proof needs a lot of mathematical logic. Chapter II in Robinson's book[7] is devoted to this assertion.

Theorem 2.2 (The transfer theorem): Every abstract theory possesses a nonstandard extension in which abstract sets go into their nonstandard extensions, and relations on these sets go into *internal* relations on the extended sets. Constant relations go into their corresponding *standard* relations. No statement about *external* relations is given *a priori*.

We now apply these ideas to construct those nonstandard theories which are needed for a full understanding of the notion of a nonstandard Hilbert space.

B. Nonstandard real line

A standard theory of the real line is a second order theory given by

(1) an abstract set of individuals R,

(2) two ternary relations on R; namely S and P $(S, P \subset R \times R \times R)$.

(3) A binary relation on R; namely $L (L \subset R \times R)$, satisfying the well-known axioms which make (R, S, P, L) a complete ordered field.[11]

The interpretation of this is as follows: R is the set of real numbers, which we call individuals, where $(a, b, c) \in S$ means $a + b = c$, $(a, b, c) \in P$ $a.\, b = c$, and $(a, b) \in L$ $a \leq b$.

Nonstandard Extension: This is the second-order theory given by

(1) ${}^*R$, a nonstandard extension of R;

(2) $\check{S} = \sigma_{(R,R,R)}(S)$ and $\check{P} = \sigma_{(R,R,R)}(P)$, the standard trinary relations on ${}^*R$ corresponding to S and P, respectively;

(3) $\check{L} = \sigma_{(R,R)}(L)$, the standard binary relation on ${}^*R$ corresponding to L, and satisfying the corresponding axioms according to the transfer theorem (Theorem 2.2):

(a) $({}^*R, \check{S}, \check{P}, \check{L})$ is an ordered field, because all statements about ordered fields are first order. (Domains of variables in the statements are sets of individuals.) We shall identify R with $\hat{R}$, and write again $a + b = c$ for $(a, b, c) \in \check{S}$, $a.\, b = c$ for $(a, b, c) \in \check{P}$, and $a \leq b$ for $(a, b) \in \check{L}$.

(b) The statement concerning the completeness of R which asserts: "Every nonempty lower bounded subset of R (unary relation on R) has a greatest lower bound," being of the second order (domain of its variable is a class of relations on R) becomes: "Every nonempty lower bounded *internal* subset of ${}^*R$ has a greatest lower bound."

Definition 2.5: A nonstandard real number $x \in {}^*R$ is said to be

(i) *finite:* if and only if their exists a standard positive real number $r \in R$, $r > 0$ such that $|x| < r$;

(ii) *infinitesimal:* if and only if for every positive standard real number $r \in R$, $r > 0$, we have $|x| < r$;

(iii) *infinite:* if and only if it is not finite.

We note that all standard real numbers are finite. We note also that all infinitesimal numbers are finite. The number zero is infinitesimal. In fact, it is the only standard number which is infinitesimal. Nonzero infinitesimal numbers do exist. As a consequence, infinite numbers, being the multiplicative inverses of nonzero infinitesimal numbers do exist as well. We denote the set of finite numbers by $M_0(R)$ and infinitesimal numbers by $M_1(R)$, and prove in the appendix.

Theorem 2.3: Let $x \in M_0(R)$, then there exists a unique element $r \in R$ such that $x - r \in M_1(R)$.

Definition 2.6: The standard part of a finite number $x \in M_0(R)$ [written as st(x)] is the standard number $r \in R$ for which $x - r \in M_1(R)$.

Definition 2.7: A number $r_1 \in {}^*R$ is said to be infinitesimally close to $r_2 \in {}^*R$ (written as $r_1 \approx r_2$) if and only if $r_1 - r_2 \in M_1(R)$.

Topology

Let U be the usual topology on R (i.e., U is the class of open subsets of R). Let $\breve{U}$ be the image of ${}^*U$ under the isomorphism σ_R^U [i.e., $\breve{U} = \sigma_R^U({}^*U)$]. The elements of $\breve{U}$ are called the internal open subsets of ${}^*R$. $\breve{U}$ is closed under finite union and intersection operations, which means that if $A, B \in \breve{U}$, then $A \cup B \in \breve{U}$ and $A \cap B \in \breve{U}$. But the union of an arbitrary subclass of $\breve{U}$ need not belong to $\breve{U}$ unless we restrict ourselves to internal subclasses.

However, since $\breve{U}$ is closed under finite intersections, it serves as a basis for a topology on ${}^*R$. This topology coincides with the order topology on ${}^*R$.

We state now a property of nonstandard extensions of general topological spaces.

Let (X, T) be a topological space. Let ${}^*X$ be a nonstandard extension of X and $\breve{T}$ be the image of ${}^*T$ under the isomorphism σ_X^T. Let $\mathcal{T}$ be the topology on ${}^*X$ generated by $\breve{T}$ as a basis. It can be shown that the T-closure operator on the class of subsets of X, extended to the class of internal subsets of ${}^*X$, coincides with the $\mathcal{T}$-closure operator on the class of subsets of ${}^*X$ restricted to the class of internal subsets of ${}^*X$. This property applies in particular to R. If A is an internal subset of ${}^*R$ generated by a sequence $\{A_n: n \in N\}$ of subsets of R, then the closure of A relative to the order topology on ${}^*R$ (denoted by $\bar{A}$) is the internal subset of ${}^*R$ generated by the sequence $\{\bar{A}_n: n \in N\}$, where $\bar{A}_n$ is the closure of A_n relative to the order topology on R.

Sequences

In the standard theory, a sequence s of elements of R is a function $s: N \to R$. It is well known that every Cauchy sequence in R converges to some element in R.

Going to the nonstandard theory, a sequence s of elements of ${}^*R$, is defined to be a function $s: {}^*N \to {}^*R$, and, being by definition a subset of ${}^*N \times {}^*R$, it may be internal, standard, or external. It can be shown that a sequence $s: {}^*N \to {}^*R$ can converge at most to one point in ${}^*R$, which, when it exists, we denote by $\lim s_n$ $n \in N$. A convergent sequence in ${}^*R$ is Cauchy; but the converse is *not* true in general. However, an *internal* Cauchy sequence in ${}^*R$ is convergent.

There exists a unique *standard* function, which assigns to every *internal* sequence $s: {}^*N \to {}^*R$, and every number $n \in {}^*N$, a number called the partial sum of s up to and including n, and denoted by $\sum_{K=0}^{n} s_K$, satisfying the following conditions:

(i) $\sum_{K=0}^{0} S_K = S_0$

(ii) $\sum_{K=0}^{n+1} S_K = \sum_{K=0}^{n} S_K + S_{n+1}$

If the sequence of partial sums $\{\sum_{K=0}^{n} S_K: n \in {}^*N\}$ is convergent, we say that the internal sequence $\{S_n: n \in {}^*N\}$ is summable, and its sum is denoted by $\sum_{n \in {}^*N} S_n$ and defined to be

$$\sum_{n \in {}^*N} S_n = \lim_{n \in {}^*N} \sum_{K=0}^{n} S_K.$$

C. Nonstandard measure theory

Let (M, S, μ) be a measure space.[12] Here, M is the space, S is the σ-Ring of subsets of M, and μ is the extended real-valued measure defined on S. ("Extended real-valued" means ranging in $\bar{R} = R \cup \{-\infty, +\infty\}$). According to the general scheme outlined above, the nonstandard theory becomes:

(1) ${}^*M$, the nonstandard extension of M, is a nonempty set.

(2) $\breve{S}$, the image of ${}^*S$ under the isomorphism σ_M^S, is a nonempty class of *internal* subsets of ${}^*M$, satisfying the following conditions:

(i) $A, B \in \breve{S}$ implies $A \setminus B \in \breve{S}$.

(ii) If $\{A_n: n \in {}^*N\}$ is an *internal* sequence of elements of $\breve{S}$, then $\cup_{n \in {}^*N} A_n \in \breve{S}$.

(3) $\breve{\mu}: \breve{S} \to {}^*\bar{R}$, given by $\breve{\mu} = \sigma_{(S, \bar{R})}(\mu) \circ (\sigma_M^S)^{-1}$. It satisfies

(i) $\breve{\mu}(A) \geq 0$, for all $A \in \breve{S}$;
(ii) $\breve{\mu}(\phi) = 0$;
(iii) if $\{A_n: n \in {}^*N\}$ is an internal disjoint sequence of elements of $\breve{S}$, then

$$\breve{\mu}\Big(\bigcup_{n \in {}^*N} A_n\Big) = \sum_{n \in {}^*N} \breve{\mu}(A_n).$$

Definition 2.8: Let B be the class of Borel subsets of the real line. Let $\breve{B}$ be image of B under the isomorphism σ_R^B. A function $f: {}^*M \to {}^*R$ is said to be a measurable function if and only if for every set $\Delta \in \breve{B}$, we have $N(f) \cap f^{-1}(\Delta) \in \breve{S}$, where $N(f) = \{x: x \in {}^*M$ and $f(x) \neq 0\}$. External measurable functions may exist. However, we shall adopt a definition for simple functions which makes all simple functions internal.

Definition 2.9: Let D be a nonempty set, and let $n \in {}^*N$ be some nonstandard natural number (finite or infinite). A function $t: [n] \to {}^*D$, where $[n] = \{K: K \in {}^*N$ and $K < n\}$, is said to be an n-tuple of elements of ${}^*D$ if and only if their exists an internal sequence $s: {}^*N \to {}^*D$ such that $S_K = t_K$ for $K < n$.

Definition 2.10: A function $f: {}^*M \to {}^*R$ is said to be simple if and only if there exists an n-tuple $\{A_K: K < n\}$ of elements of $\breve{S}$ and an n-tuple $\{a_K: K < n\}$ of elements of ${}^*R$ such that $f = \sum_{K<n} a_K X_{A_K}$, where X_A is the characteristic function of A.

Definition 2.11: A simple function $f: {}^*M \to {}^*R$ is

said to be integrable if and only if $\check{\mu}(N(f)) < +\infty$ where $N(f) = \{x: x \in {}^*M \text{ and } f(x) \neq 0\}$. The integral of f is defined to be

$$\int d\check{\mu} f = \sum_{K<n} a_K \check{\mu}(A_K),$$

where $f = \sum_{K<n} a_K X_{A_K}$.

Definition 2.12: An internal measurable function $f: {}^*M \to {}^*R$ is said to be integrable if and only if there exists an internal mean fundamental sequence $\{f_n : n \in {}^*N\}$ of integrable simple functions which converges in measure to f. The integral of f is defined to be[12]

$$\int d\check{\mu} f = \lim_{n \in {}^*N} \int d\check{\mu} f_n.$$

D. Nonstandard Hilbert space

Let $(\mathcal{H}, \sigma, x, \gamma)$ be a separable complex Hilbert space. Here, $\mathcal{H}$ is the set of vectors, $\sigma \subset \mathcal{H} \times \mathcal{H} \times \mathcal{H}$ is the vector addition function, $\pi \subset C \times \mathcal{H} \times \mathcal{H}$ is the scalar vector multiplication function, and $\gamma \subset \mathcal{H} \times \mathcal{H} \times C$ is the scalar product function. We shall write $f + g = h$ for $(f,g,h) \in \sigma$, $\lambda . f = g$ for $(\lambda, f, g) \in \pi$, and $\langle f, g \rangle = \lambda$ for $(f,g,\lambda) \in \gamma$. The scalar product in particular is taken to be antilinear in the first argument, but linear in the second one. We define the norm $\| f \| = \sqrt{\langle f, f \rangle}$ for all $f \in \mathcal{H}$. The nonstandard version of the theory is given by

(1) ${}^*\mathcal{H}$, the nonstandard extension of $\mathcal{H}$;

(2) $\check{\sigma}: {}^*\mathcal{H} \times {}^*\mathcal{H} \to {}^*\mathcal{H}$, the standard function corresponding to σ under the isomorphism $\sigma(\mathcal{H}, \mathcal{H}, \mathcal{H})$;

(3) $\check{\pi}: {}^*C \times {}^*\mathcal{H} \to {}^*\mathcal{H}$, the standard function corresponding to π under the isomorphism $\sigma(C, \mathcal{H}, \mathcal{H})$;

(4) $\check{\gamma}: {}^*\mathcal{H} \times {}^*\mathcal{H} \to {}^*C$, the standard function corresponding to γ under the isomorphism $\sigma(\mathcal{H}, \mathcal{H}, C)$.

We shall write again $f + g = h$ for $(f,g,h) \in \check{\sigma}$, $\lambda . f = g$ for $(\lambda, f, g) \in \check{\pi}$, and $\langle f, g \rangle = \lambda$ for $(f,g,\lambda) \in \check{\gamma}$. The norm is defined again as $\| f \| = \sqrt{\langle f, f \rangle}$ for all $f \in {}^*\mathcal{H}$. The norm on ${}^*\mathcal{H}$ induces a topology called the strong topology.

Definition 2.13: An element $f \in {}^*\mathcal{H}$ is called

(i) *finite*, if and only if $\| f \|$ is finite;

(ii) *infinitesimal*, if and only if $\| f \|$ is infinitesimal;

(iii) *infinite* if and only if $\| f \|$ is infinite;

(iv) *near standard* if and only if there exists an element $g \in \mathcal{H}$ such that $\| f - g \| \approx 0$ [being unique, we denote g by $\mathrm{st}(f)$].

Definition 2.14:

(i) ${}^*\hat{\mathcal{H}}$ is the class of all internal bounded linear forms defined on ${}^*\mathcal{H}$ and ranging in ${}^*C$.

(ii) $L: {}^*\mathcal{H} \to {}^*\mathcal{H}$ is defined such that for all $f \in {}^*\mathcal{H}$ and $\varphi \in {}^*\mathcal{H}$, we have $L_f \in {}^*\mathcal{H}$ such that $L_f(\varphi) = \langle f, \varphi \rangle$.

By the Ricsz theorem[3] extended to ${}^*\mathcal{H}$, L is a standard anti isomorphism.

Definition 2.15: Let $f \in {}^*\mathcal{H}$ and $\xi \in {}^*\hat{\mathcal{H}}$, then their tensor product $f \otimes \xi$ is defined to be the bounded internal linear operator on ${}^*\mathcal{H}$ which sends $\varphi \in {}^*\mathcal{H}$ into $(f \otimes \xi)(\varphi) = \xi(\varphi) f \in {}^*\mathcal{H}$.

Theorem 2.4:

(i) If $\{ f_n : n \in {}^*N\}$ is an internal Cauchy sequence of elements of ${}^*\mathcal{H}$, then it converges to an element in ${}^*\mathcal{H}$.

(ii) There exists an internal sequence $\{\varphi_n : n \in {}^*N\}$ of elements of ${}^*\mathcal{H}$ such that

(a) $\langle \varphi_n, \varphi_{n'} \rangle = \delta_n^{n'}$ for all $n, n' \in {}^*N$,

(b) $\sum_{n \in {}^*N} \varphi_n \otimes L\varphi_n = I$ (the identity operator on ${}^*\mathcal{H}$).

The convergence in (b) is the pointwise strong convergence.

Proof: Since $\mathcal{H}$ is complete and separable, the above statements hold in ${}^*\mathcal{H}$ as a direct application of the transfer theorem 2. 2. ∎

Definition 2.16: Let A be a linear operator defined on a subspace $D \subset {}^*\mathcal{H}$. A pair (λ, f) where $\lambda \in {}^*C$ and $f \in D(\| f \| \neq 0)$ is called

(i) An eigen pair of A, if and only if $\| Af - \lambda f \| = 0$.

(ii) An ultra eigen pair of A, if and only if $\| Af - \lambda f \| / \| f \| \approx 0$.

It is clear that an eigen pair of A is an ultra pair of A.

1. Related standard linear forms

If $f \in {}^*\mathcal{H}$, then L_f, defined in Def. 2. 14 is a bounded internal linear form on ${}^*\mathcal{H}$; but, in general, it is not standard. We would like to construct a (not necessarily bounded) standard linear form ${}^0L_f$, defined on a standard subset of ${}^*\mathcal{H}$, denoted by D_f, and such that, for all standard $\varphi \in D_f$, we have

$${}^0L_f(\varphi) = \mathrm{st}(L_f(\varphi)).$$

Since ${}^0D_f$ (the set of standard elements of D_f) generates D_f, and ${}^0L_f$ is standard, it is sufficient to define ${}^0L_f$ on ${}^0D_f$, then extend it to D_f.

Now, for every $\varphi \in {}^0D_f$, ${}^0L_f(\varphi) = \mathrm{st}(L_f(\varphi))$, which implies that $L_f(\varphi) \in M_0(C)$. But $L_f(\varphi) = \langle f, \varphi \rangle$, hence

$${}^0D_f = \{\varphi: \varphi \in \mathcal{H} \text{ and } \langle f, \varphi \rangle \subset M_0(C)\}.$$

It is easily verified that D_f is a linear subspace of ${}^*\mathcal{H}$ over ${}^*C$. If f is a finite vector, then ${}^0D_f = \mathcal{H}$, and ${}^0L_f$ is a bounded standard linear form on ${}^*\mathcal{H}$. Now, by the Ricsz theorem[3] extended to ${}^*\mathcal{H}$, there exists a standard vector $g_f \in \mathcal{H}$ such that ${}^0L_f = L_{g_f}$.

In particular, if f is near standard, then $g_f = \mathrm{st}(f)$ and ${}^0L_f = L_{\mathrm{st}(f)}$. However, if f is standard, then ${}^0L_f = L_f$. The linear form (or functional) ${}^0L_f$ is called the related standard linear form (relative to $f \in {}^*\mathcal{H}$).

3. SPECTRAL PROPERTIES OF INTERNAL OPERATORS IN NONSTANDARD HILBERT SPACES

In this section we study the spectral properties of internal operators in a nonstandard Hilbert space in detail, and show that every point in the discrete spectrum of an internal self-adjoint operator is an eigenvalue, and that every point in the continuous spectrum of an internal self-adjoint operator is an ultra eigenvalue. At the end of this section, we relate the new concept of an ultra eigenvector which we introduce, to that of an eigenfunctional, used in the Rigged Hilbert space approach.

A. Analysis of the spectrum

Let A be a linear operator defined on a subspace $D \subset {}^*\mathcal{H}$. Let $\lambda \in {}^*C$ be any nonstandard complex number. If I is the identity operator on ${}^*\mathcal{H}$, define $\Delta(\lambda) = (A - \lambda I)D$. Let $\overline{\Delta(\lambda)}$ be the closure of $\Delta(\lambda)$ with respect to the strong topology on ${}^*\mathcal{H}$.

Definition 3.1:

(i) The resolvent set of A is $\rho(A) = \{\lambda: \Delta(\lambda) = {}^*\mathcal{H}\}$.

(ii) The discrete spectrum of A is $\sigma_d(A) = \{\lambda: \overline{\Delta(\lambda)} \neq {}^*\mathcal{H}\}$.

(iii) The continuous spectrum of A is $\sigma_c(A) = \{\lambda: \Delta(\lambda) \neq \overline{\Delta(\lambda)} = {}^*\mathcal{H}\}$.

We notice that the three sets $\rho(A), \sigma_d(A), \sigma_c(A)$ are disjoint and their union is ${}^*C$. In the following theorem we relate the spectrum of internal operators in ${}^*\mathcal{H}$ to the spectrum of the generating operators in $\mathcal{H}$.

Theorem 3.1: Let A be an internal linear operator defined on an internal subspace $D \subset {}^*\mathcal{H}$, and generated by a sequence $\{A_n: n \in N\}$ of linear operators defined on a sequence of subspace $\{D_n: n \in N\}$. Let λ be a nonstandard complex number generated by a sequence $\{\lambda_n: n \in N\}$. Then, λ belongs to the resolvent set of A, discrete spectrum of A or continuous spectrum of A if and only if for almost every $n \in N$, λ_n belongs to the resolvent set of A_n, discrete spectrum of A_n of continuous spectrum of A_n, respectively.

Proof: Let I_0 be the identity operator on $\mathcal{H}$, and define

$$\Delta(\lambda) = (A - \lambda I)D,$$
$$\Delta_n(\lambda_n) = (A_n - \lambda_n I_0) D_n .$$

Since the sequence $\{D_n: n \in N\}$ generates D and the sequence $\{A_n - \lambda_n I_0: n \in N\}$ generates $A - \lambda I$, then the sequence $\{\Delta_n(\lambda_n): n \in N\}$ generates $\Delta(\lambda)$. The three sets $\{n: \Delta_n(\lambda_n) = \mathcal{H}\}$, $\{n: \Delta_n(\lambda_n) \neq \mathcal{H}\}$ and $\{n: \Delta_n(\lambda_n) \neq \overline{\Delta_n(\lambda_n)} = \mathcal{H}\}$ are disjoint and their union is N. Hence, by (F1), at most one of them can belong to $\mathfrak{F}$, and by (F4), (F2), and (F1) (in this order), at least one of them should belong to $\mathfrak{F}$. Hence, one and only one of them does belong to $\mathfrak{F}$. We then have the following:

(i) If for almost every $n \in N$, $\Delta_n(\lambda_n) = \mathcal{H}$, then $\Delta(\lambda) = {}^*\mathcal{H}$.

(ii) If for almost every $n \in N$, $\overline{\Delta_n(\lambda_n)} \neq \mathcal{H}$, then $\overline{\Delta(\lambda)} \neq {}^*\mathcal{H}$.

(iii) If for almost every $n \in N$, $\Delta_n(\lambda_n) \neq \overline{\Delta_n(\lambda_n)} = \mathcal{H}$, then $\Delta(\lambda) \neq \overline{\Delta(\lambda)} = {}^*\mathcal{H}$.

The converse is also true by the tautology (see the Appendix):

$$\neg \nabla_n P(n) \Longleftrightarrow \nabla_n \neg P(n).$$

By applying Def. 3.1, the theorem follows. ∎

If A is an internal operator defined on a dense subspace $D \subset {}^*\mathcal{H}$, then we define a set $D' = \{f: f \in {}^*\mathcal{H}$ and $\exists_{\alpha \in {}^*R}$ such that $\forall_{\varphi \in D} |\langle f, A\varphi\rangle| \leq \alpha \| \varphi \|\}$. If $f \in D'$, then the linear form

$${}^0\xi_f: \begin{cases} D \to {}^*C \\ \varphi \to \xi_f(\varphi) = \langle f, A\varphi\rangle \end{cases}$$

is bounded and internal, hence it can be extended to $\bar{D} = {}^*\mathcal{H}$. Let ξ_f be the extension of ${}^0\xi_f$ to ${}^*\mathcal{H}$. ξ_f is still bounded and internal, hence by the Riesz theorem[3] extended to ${}^*\mathcal{H}$, there exists an element $g_f \in {}^*\mathcal{H}$ such that $\xi_f = L_{g_f}$, i.e.,

$$\xi_f(\varphi) = L_{g_f}(\varphi) \qquad \text{for all } \varphi \in {}^*\mathcal{H},$$

or

$$\langle f, A\varphi\rangle = \langle g_f, \varphi\rangle \qquad \text{for all } \varphi \in D.$$

Put $g_f = A^+ f$. It can be shown that A^+ is an internal closed operator defined on a subspace $D' \subset {}^*\mathcal{H}$. It is called the adjoint operator. An internal operator A is called self-adjoint, if and only if it satisfies the relation $A = A^+$.

Theorem 3.2: Let A be an internal self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let $\lambda \in \sigma_d(A)$, then:

(i) There exists a vector $f \in {}^*\mathcal{H}$ such that $\| f \| = 1$ and $Af = \lambda f$.

(ii) If $\lambda' \neq \lambda$ and $Af' = \lambda' f'$, then $\langle f, f'\rangle = 0$.

Proof: Since this theorem holds for standard Hilbert spaces, it holds for nonstandard Hilbert spaces (A is internal), by direct application of the transfer theorem (2. 2). ∎

Theorem 3.3: Let A be an internal self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let $\lambda \in \sigma_c(A)$, then:

(i) There exists a vector $f \in {}^*\mathcal{H}$ such that $\| f \| = 1$ and $\| Af - \lambda f \| \approx 0$.

(ii) If $\lambda' \in \sigma_c(A)$, $\lambda' \not\approx \lambda$, $\| f' \| = 1$, and $\| Af' - \lambda' f' \| \approx 0$, then $\langle f, f'\rangle \approx 0$.

Proof: It is well known[10] that if λ belongs to the continuous spectrum of a self-adjoint operator in a standard Hilbert space, then for every $\alpha > 0$ ($\alpha \in R$), there exists a unit vector $f \in \mathcal{H}$ such that $\| Af - \lambda f \| \leq \alpha$. Now, by the transfer theorem (2. 2), this is true for internal self-adjoint operators in nonstandard Hilbert spaces. By taking α to be infinitesimal we get

$$\| f \| = 1 \text{ and } \| Af - \lambda f \| \approx 0. \tag{3.1}$$

This proves (i). To prove (ii), define

$$\begin{aligned} d &= Af - \lambda f, \\ d' &= Af' - \lambda' f', \end{aligned} \tag{3.2}$$

i.e.,

$$\begin{aligned} Af &= \lambda f + d, \qquad \| d \| \approx 0, \\ Af' &= \lambda' f' + d', \qquad \| d' \| \approx 0, \end{aligned} \tag{3.3}$$

hence

$$\begin{aligned} \langle f', Af\rangle &= \lambda \langle f', f\rangle + \langle f', d\rangle, \\ \langle f, Af'\rangle &= \lambda' \langle f, f'\rangle + \langle f, d'\rangle. \end{aligned} \tag{3.4}$$

By taking the complex conjugate of the first equation in (3. 4), and by using the fact that A is self-adjoint (λ, λ' are real), we get

$$\langle f, Af'\rangle = \lambda \langle f, f'\rangle + \langle d, f'\rangle. \tag{3.5}$$

By subtracting (3. 5) from the second equation in (3. 4) we get

$$(\lambda' - \lambda)\langle f, f'\rangle = \langle d, f'\rangle - \langle f, d'\rangle, \tag{3.6}$$

since $\| d \| \approx 0$, $\| d' \| \approx 0$, $\| f' \| = \| f \| = 1$ and

$|\langle d, f'\rangle| \leq \| d \| \cdot \| f' \|$, $|\langle f, d'\rangle| \leq \| f \| \cdot \| d' \|$, then

$$\langle d, f'\rangle - \langle f, d'\rangle \approx 0,$$

i.e.,

$$(\lambda' - \lambda)\langle f, f'\rangle \approx 0, \tag{3.7}$$

but

$$\lambda' \not\approx \lambda \quad \text{(given)},$$

hence

$$\langle f, f'\rangle \approx 0 \tag{3.8}$$

and the theorem is proved. ■

We show now how to write the orthonormality condition for the standard part of the full spectrum. We shall talk about ultra eigenvectors only (because eigenvectors are a special case).

Theorem 3.4: Let ${}^0\sigma(A)$ be the standard part of the spectrum of a self-adjoint operator A (i.e., the set of standard real numbers which belong to the spectrum of A, whether discrete or continuous). Let $\{f_\lambda : \lambda \in {}^0\sigma(A)\}$ be a family of ultra eigenvectors of A, indexed by the corresponding ultra eigenvalues and normalized to unity (i.e., $\| f_\lambda \| = 1$, $\| Af_\lambda - \lambda f_\lambda \| \approx 0$). Then, for all $\lambda, \lambda' \in {}^0\sigma(A)$, $\mathrm{st}(\langle f_\lambda, f_{\lambda'}\rangle) = \delta_\lambda^{\lambda'}$.

Proof: We notice first that

$$\langle f_\lambda, f_\lambda\rangle = \| f_\lambda \|^2 = 1. \tag{3.9}$$

Now, if $\lambda \neq \lambda'$, $\lambda, \lambda' \in {}^0\sigma(A)$, then $\lambda \not\approx \lambda'$ [otherwise, if $\lambda \approx \lambda'$, then $\lambda - \lambda' \in M_1(R) \cap R = \{0\}$, hence $\lambda = \lambda'$ contrary to the assumption]. By using now Theorem 3.3, we get

$$\langle f_\lambda, f_{\lambda'}\rangle \approx 0; \tag{3.10}$$

(1) becomes: $\lambda = \lambda'$ implies $\mathrm{st}\langle f_\lambda, f_{\lambda'}\rangle = 1$; (3.11)

(2) becomes: $\lambda \neq \lambda'$ implies $\mathrm{st}\langle f_\lambda, f_{\lambda'}\rangle = 0$; (3.12)

(3) and (4) can be written as

$$\lambda, \lambda' \in {}^0\sigma(A) \Longrightarrow \mathrm{st}(\langle f_\lambda, f_{\lambda'}\rangle) = \delta_\lambda^{\lambda'} \quad \square \tag{3.13}$$

The orthonormality condition proved in the preceeding theorem cannot be extended to that part of the spectrum which is not standard, because if $\lambda \approx \lambda', \lambda \neq \lambda'$, then all what we know is $(\lambda - \lambda')\langle f_\lambda, f_{\lambda'}\rangle \approx 0$ (see Theorem 3.4). By taking the standard parts of both sides we get $0 \times \mathrm{st}\langle f_\lambda, f_{\lambda'}\rangle = 0$ which gives no information about $\mathrm{st}\langle f_\lambda, f_{\lambda'}\rangle$. However, if $f_\lambda, f_{\lambda'}$ are eigenvectors of A with eigenvalues λ and λ', then $\langle f_\lambda, f_{\lambda'}\rangle = \delta_\lambda^{\lambda'}$ exactly, without any restriction.

B. Relation with eigenfunctionals

Let A be a standard self-adjoint operator, densely defined on ${}^*\mathcal{H}$. We assume for the time that there exists a standard subspace of ${}^*\mathcal{H}$ (which we denote by Φ) satisfying the following properties:

(i) Φ is dense in ${}^*\mathcal{H}$ (i.e., $\bar{\Phi} = {}^*\mathcal{H}$);

(ii) Φ is contained in the domain of A;

(iii) Φ is stable under A [i.e., $A(\Phi) \subset \Phi$].

The existence of such a subspace for any densely defined self-adjoint operator is a well-known property of standard Hilbert spaces. By the transfer theorem 2.2, it holds for any nonstandard extension.

We assume here that Φ is subject to the extra property:

(iv) There exists a system of ultra eigenvectors whose eigenvalues belong to the standard part of the spectrum of A, which we denote by $\{f_\lambda : \lambda \in {}^0\sigma(A)\}$ such that, for all $\lambda \in {}^0\sigma(A)$,

(a) $\Phi \subset D_{f_\lambda}$

(b) ${}^0L_{f_\lambda} \neq 0$

(c) ${}^0L_{d_\lambda} = 0$ where $d_\lambda = Af_\lambda - \lambda f_\lambda$ (see Sec. 2D)

The system $\{f_\lambda : \lambda \in {}^0\sigma(A)\}$ is a system of ultra eigenvectors of A, not necessarily normalized to unity.

Property (iv) holds for the discrete part of the standard spectrum of any standard self-adjoint operator A. If $\lambda \in {}^0\sigma_d(A)$, then there exists an element $f_\lambda \in \mathcal{H}$ such that

$$Af_\lambda = \lambda f_\lambda, \qquad \| f_\lambda \| = 1.$$

Hence, $D_{f_\lambda} = {}^*\mathcal{H}$ and $d_\lambda = 0$, where $d_\lambda = Af_\lambda - \lambda f_\lambda$, leading to ${}^0L_{d_\lambda} = 0$; and ${}^0L_{f_\lambda} \neq 0$ because ${}^0L_{f_\lambda}(f_\lambda) = 1$. This proves that (iv) holds in the discrete case. But (iv) does not necessarily hold for the continuous part of the standard spectrum of an arbitrary standard self-adjoint operator. However, it holds for the absolutely continuous part[13] of the standard spectrum of any standard self-adjoint operator A, because in the A space, the subspace of ${}^*\mathcal{H}$ corresponding to that part of the spectrum of A which is absolutely continuous with respect to the Lebesgue measure can be represented by equivalence classes of nonstandard functions which are square integrable with respect to the Lebesgue measure (extended to ${}^*R$). Apart from minor modifications, due to the fact that the absolutely continuous part of the standard spectrum of A may not be the whole real line, one can construct ultra eigenvectors of A following the same procedure applied in Sec. 6 of this paper to construct ultra eigenvectors of the momentum operators. As shown in Sec. 6, these ultra eigenvectors would automatically satisfy properly (iv).

Now, since "apart from the discrete case, measures which are not absolutely continuous do not seem to occur in physical problems,"[13] we assume (iv) to hold in any case of interest.

We introduce now the definition:

Definition 3.2:

(i) If $\lambda \in {}^0\sigma(A)$ (standard part of the spectrum of A), then ξ_λ is the restriction of ${}^0L_{f_\lambda}$ to Φ.

(ii) If $\lambda \in \sigma(A) \setminus {}^0\sigma(A)$ (the part of the spectrum of A which is not standard), then ξ_λ is the internal linear form generated by $\{{}^0\xi_{\lambda_n} : n \in N\}$, where $\{\lambda_n : n \in N\}$ is a sequence generating λ and ${}^0\xi_{\lambda_n}$ is the restriction to ${}^0\Phi$ (the set of standard vectors in Φ) of the linear form ξ_{λ_n}, λ_n being standard.

The possibility of the second part of this definition follows from the fact that ξ_λ (when λ is standard) is a standard linear form.

Theorem 3.5: Let A be a standard self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let Φ be a standard subspace of ${}^*\mathcal{H}$ subject to the conditions (i)–(iv) stated above. Let $\{\xi_\lambda : \lambda \in \sigma(A)\}$ be as in Def. 3.2, then, for all $\lambda \in \sigma(A)$. $\varphi \in \Phi$, we have $\xi_\lambda(A\varphi) = \lambda\xi_\lambda(\varphi)$.

Proof: Let $\lambda \in {}^0\sigma(A)$ and $\varphi \in {}^0\Phi$ [the standard parts of $\sigma(A)$ and Φ, respectively], then

$$\begin{aligned} \xi_\lambda(\varphi) &= {}^0L_{f_\lambda}(\varphi) && \text{(Def. 3.2)} \\ &= \mathrm{st}\langle f_\lambda, \varphi\rangle && \text{(Sec. 2D)} \end{aligned} \tag{3.14}$$

$$\begin{aligned}\xi_\lambda(A\varphi) &= {}^0L_{f_\lambda}(A\varphi)\\ &= \mathrm{st}\langle f_\lambda, A\varphi\rangle \quad (A \text{ is a standard operator})\\ &= \mathrm{st}\langle Af_\lambda, \varphi\rangle \quad (A \text{ is self-adjoint})\\ &= \mathrm{st}\langle \lambda f_\lambda + d_\lambda, \varphi\rangle \quad [\text{see (iv): (c) above}]\\ &= \lambda\ \mathrm{st}\langle f_\lambda, \varphi\rangle + \mathrm{st}\langle d_\lambda, \varphi\rangle \quad (\lambda \text{ is standard})\\ &= \lambda\ \mathrm{st}\langle f_\lambda, \varphi\rangle \quad [\text{see (iv): (c) above}]\end{aligned} \tag{3.15}$$

Now, from (3. 14) and (3. 15) we have

$$\xi_\lambda(A\varphi) = \lambda\xi_\lambda(\varphi). \tag{3.16}$$

Since (3. 16) holds for any standard $\lambda \in {}^0\sigma(A)$ and $\varphi \in {}^0\Phi$, it holds for any $\lambda \in \sigma(A)$ and $\varphi \in \Phi$, because

$$\xi: \begin{cases} \sigma(A) \times \Phi \to {}^*C \\ (\lambda, \varphi) \to \xi_\lambda(\varphi) \end{cases}$$

is a standard function. This completes the proof. ■

Let T be the weakest topology for which the restriction of A to Φ is a continuous linear operator, and all the elements of the set $\{\xi_\lambda : \lambda \in \sigma(A)\}$ are continuous linear forms on Φ. Let Φ' (the dual space of Φ) be the set of all T continuous internal linear forms defined on Φ. Clearly, for all $\lambda \in \sigma(A)$, $\xi_\lambda \in \Phi'$.

Let $A' : \Phi' \to \Phi'$ be the linear operator defined as follows: If $\xi \in \Phi'$ then $A'\xi$ is the linear forms which assigns to every $\varphi \in \Phi$ the complex number $(A'\xi)(\varphi) = \xi(A\varphi) \in {}^*C$. A' is a standard linear operator. If Φ' is embedded with strong dual topology T', we see that A' is continuous with respect to T'.

Now, since $\xi_\lambda(A\varphi) = \lambda\xi_\lambda(\varphi)$ for all $\varphi \in \Phi$, $\lambda \in \sigma(A)$ (see Theorem 3. 5), we have $A'\xi_\lambda = \lambda\xi_\lambda$ for all $\lambda \in \sigma(A)$. This shows that the related standard linear forms associated with ultra eigenvectors, are precisely the eigenfunctionals used in the Rigged Hilbert space approach.

The relation between these eigenfunctionals and the spectral measure has been established in the introduction in connection with the Rigged Hilbert space approach to quantum mechanics, and there is no point in repeating it here. It is worth emphasizing that our approach has the advantage over that using Rigged Hilbert spaces, of the existence of "genuine vectors" defining the eigenfunctionals. Theorem (3. 4) which states the orthogonality condition for the ultra eigenvectors corresponding to the standard points of the full spectrum of a self-adjoint operator is remarkable. Instead of the two separate orthogonality conditions stated by Dirac, which read: $\langle n|n'\rangle = \delta_n^{n'}$ for the discrete spectrum and $\langle \xi \mid \xi'\rangle = \delta(\xi - \xi')$ for the continuous spectrum (the latter being not rigorous), we have a rigorous statement for both, stating that

$$\mathrm{st}\langle f_\lambda, f_{\lambda'}\rangle = \delta_\lambda^{\lambda'} \text{ for all } \lambda, \lambda' \in {}^0\sigma(A).$$

This statement has no analog in the rigged Hilbert space approach.

4. PROBABILITY FUNCTIONS

In this section, we try to define a probability function associated with a given standard self-adjoint operator A densely defined on ${}^*\mathcal{H}$, and a unit vector $f \in {}^*\mathcal{H}$. In the standard theory, if A is a self-adjoint operator densely defined on a Hilbert space $\mathcal{H}$ and $E(A)$ is the spectral measure[2] associated with A, then, for every unit vector $f \in \mathcal{H}$, the function

$$\mathrm{prob}(A, f, \Delta) = \langle f, E_\Delta(A)f\rangle,$$

where Δ is a Borel subset of R, is a probability measure on the Borel subsets of the real line. According to the transfer theorem (2. 2) this can be readily extended to nonstandard Hilbert spaces. If A is an internal self-adjoint operator densely defined on a nonstandard Hilbert space ${}^*\mathcal{H}$ and $E(A)$ is the nonstandard spectral measure associated with A, then for every unit vector $f \in {}^*\mathcal{H}$, the function

$$\mathrm{ns-prob}(A, f, \Delta) = \langle f, E_\Delta(A)f\rangle,$$

where Δ is a Borel internal subset of ${}^*R$, is a nonstandard probability measure on the Borel internal subsets of the nonstandard real line. However, this is *not* a probability measure. The result of a measurement can only be a *standard* real number. So, for a physical interpretation of the nonstandard theory, we need to have a *standard* probability function, defined on a ring of Borel subsets of the *standard* real line, associated with a given *standard* self-adjoint operator, densely defined on a *nonstandard* Hilbert space. We shall see that the function we are going to define is not generally σ-additive, and, hence, cannot be extended to the σ-ring of Borel subsets of the real line. However, the probability functions associated with standard and near standard vectors are σ-additive. In this case, the probability measure defined here coincides with the corresponding one in the standard Hilbert space.

A. Construction of the measure $(\mu_f A)$

Let f be a nondecreasing standard real-valued function defined on the standard real line. If $\lambda \in R$, then $f(\lambda)$ is an upper bound for the set $\{f(\lambda - \epsilon): \epsilon > 0\}$. Since this set is not empty, it has a supremum which we denote by $\mathbf{f}(\lambda)$. Hence:

Definition 4.1: If f is a nondecreasing standard real-valued function defined on the standard real line, then we define $\mathbf{f}(\lambda) = \sup_{\epsilon>0} f(\lambda - \epsilon)$.

Theorem 4.1: Let $f: R \to R$ be nondecreasing; then $\mathbf{f}: R \to R$ is nondecreasing and continuous from below.

Proof: Let $\lambda < \lambda'$, $\lambda, \lambda' \in R$. By definition $\mathbf{f}(\lambda') = \sup_{\epsilon>0} f(\lambda' - \epsilon)$, hence

$$f(\lambda) \leq \mathbf{f}(\lambda'). \tag{4.1}$$

Now

$$\mathbf{f}(\lambda) \leq f(\lambda) \quad [f(\lambda) \text{ is an upperbound of } \{f(\lambda-\epsilon): \epsilon > 0\}]. \tag{4.2}$$

Hence, from (4. 1) and (4. 2) we get

$$\lambda < \lambda' \Longrightarrow \mathbf{f}(\lambda) \leq \mathbf{f}(\lambda') \tag{4.3}$$

which shows that $\mathbf{f}$ is nondecreasing.

To show that $\mathbf{f}$ is continuous from below, let $\epsilon > 0$. Then, by definition, there exists a number $\delta > 0$, such that

$$f(\lambda - 2\delta) \geq \mathbf{f}(\lambda) - \epsilon. \tag{4.4}$$

Since $\lambda - \delta > \lambda - 2\delta$, we have (by definition)

$$\mathbf{f}(\lambda - \delta) \geq f(\lambda - 2\delta) \tag{4.5}$$

and so, from (4. 4) and (4. 5) we get

$$\mathbf{f}(\lambda - \delta) \geq \mathbf{f}(\lambda) - \epsilon. \tag{4.6}$$

Now $\lambda' > \lambda - \delta \Rightarrow \mathbf{f}(\lambda') \geq \mathbf{f}(\lambda - \delta)$ [by (4. 3)], and thus, for every $\epsilon > 0$, there exists a number $\delta > 0$ such that

$$\lambda - \delta \leq \lambda' \leq \lambda \quad \text{implies} \quad \mathbf{f}(\lambda) - \epsilon \leq \mathbf{f}(\lambda') \leq \mathbf{f}(\lambda). \tag{4. 7}$$

Hence, $\mathbf{f}$ is continuous from below. ■

Let A be a standard self-adjoint operator densely defined on ${}^*\mathcal{H}$. By the spectral theory[2] extended to ${}^*\mathcal{H}$, there exists a nonstandard spectral measure $E(A)$ defined on the Borel internal subsets of the nonstandard real line ${}^*R$, and assigning to every Borel internal subset $\Delta \subset {}^*R$ a projection operator $E_\Delta(A)$, such that, if $E_\lambda(A) \equiv E_{(-\infty, \lambda)}(A)$ then

$$A = \int_{\lambda \in {}^*R} \lambda \, dE_\lambda(A).$$

If $f \in {}^*\mathcal{H}$ and $\| f \| = 1$, then the function $\breve{\mu}_{f,A}$ defined as follows:

$$\breve{\mu}_{f,A}(\Delta) = \langle f, E_\Delta(A) f \rangle$$

for every Borel internal subset $\Delta \subset {}^*R$ is a nonstandard probability measure. In particular,

$0 \leq \breve{\mu}_{f,A}(\Delta) \leq 1$ for every Borel internal subset $\Delta \subset {}^*R$.

We now introduce the definition:

Definition 4.2: Let $\lambda \in R$, $f \in {}^*\mathcal{H}$, $\| f \| = 1$, and A be a standard self-adjoint operator densely defined on ${}^*\mathcal{H}$. If $E(A)$ is the nonstandard spectral measure associated with A, then define: $\varphi_{f,A}(\lambda) = \mathrm{st}\langle f, E_\lambda(A) f \rangle$.

Theorem 4.2: The real-valued function $\varphi_{f,A}$ defined in Def. 4. 2 is nondecreasing and lies between zero and one.

Proof: Let $\lambda < \lambda'$; then $\langle f, E_\lambda(A) f \rangle \leq \langle f, E_{\lambda'}(A) f \rangle$ (from the standard theory extended to ${}^*\mathcal{H}$). Now:

$$\varphi_{f,A}(\lambda) = \mathrm{st}\langle f, E_\lambda(A) f \rangle \leq \mathrm{st}\langle f, E_{\lambda'}(A) f \rangle = \varphi_{f,A}(\lambda');$$

therefore $\lambda < \lambda' \Rightarrow \varphi_{f,A}(\lambda) \leq \varphi_{f,A}(\lambda')$. Also, for every $\lambda \in R$, since $0 \leq \langle f, F_\lambda(A) f \rangle \leq 1$ (extension of the standard theory), we have $0 \leq \mathrm{st}\langle f, E_\lambda(A) f \rangle \leq 1$, i.e., for all $\lambda \in R$, $0 \leq \varphi_{f,A}(\lambda) \leq 1$. □

Theorem 4.3: The function $\boldsymbol{\varphi}_{f,A}$ is nondecreasing, continuous from below and lies between zero and one.

Proof: By Theorem 4. 2, $\varphi_{f,A}$ is nondecreasing. Hence, by Theorem 4. 1, $\boldsymbol{\varphi}_{f,A}$ is nondecreasing and continuous from below. Now, if $\lambda \in R$, then for every $\epsilon > 0$ we have

$$0 \leq \varphi(\lambda - \epsilon) \leq \boldsymbol{\varphi}(\lambda) \leq \varphi(\lambda) \leq 1.$$

thus, for every $\lambda \in R$, $0 \leq \boldsymbol{\varphi}(\lambda) \leq 1$. ■

Theorem 4.4: There exists a measure $\mu_{f,A}$ on the Borel subsets of the real line R, such that $\mu_{f,A}([a, b)) = \boldsymbol{\varphi}_{f,A}(b) - \boldsymbol{\varphi}_{f,A}(a)$ for $a \leq b$ and such that $\mu_{f,A}(R) \leq 1$.

Proof: We shall give here an outline of the proof only.

If $g: R \to R$ is a nondecreasing function which is continuous from below, and $\mathcal{P}$ is the class of all semi-closed semi-open intervals on the real line (i.e., $\mathcal{P} = \{[a, b): a \leq b, a, b \in R\}$), then the set function μ_0: $\mathcal{P} \to R$ defined according to

$$\mu_0([a, b)) = g(b) - g(a)$$

is finite, positive, and σ additive.[10]

If $\mathcal{S}$ is the ring generated by $\mathcal{P}$, then there exists a unique finite measure μ on $\mathcal{S}$, such that, whenever $E \in \mathcal{P}$, $\mu(E) = \mu_0(E)$. If $\mathcal{B}$ is the σ ring (of Borel subsets) generated by $\mathcal{S}$, then μ has a unique extension $\bar{\mu}$ to $\mathcal{B}$ such that $\bar{\mu}$ is a σ-finite measure on $\mathcal{B}$, and whenever $E \in \mathcal{S}$, $\bar{\mu}(E) = \mu(E)$.

Now, in particular,

$$\bar{\mu}([a, b)) = \mu([a, b)) = \mu_0([a, b)) = g(b) - g(a).$$

Since our function $\varphi_{f,A}$ is nondecreasing and continuous from below, there exists a σ-finite measure $\mu_{f,A}$ on the Borel subsets of R, such that

$$\mu_{f,A}([a, b)) = \varphi_{f,A}(b) - \varphi_{f,A}(a).$$

To show that $\mu_{f,A}$ is finite (and has 1 as an upper bound), let $\Delta_n = [-n, n)$.

The sequence $\{\Delta_n: n \in N\}$ is increasing with $\cup_{n \in N} \Delta_n = R$. Hence,

$$\mu_{f,A}(R) = \mu_{f,A}\Big(\bigcup_{n \in N} \Delta_n\Big) = \sup_{n \in N} \mu_{f,A}(\Delta_n)$$
$$= \sup_{n \in N} [\boldsymbol{\varphi}_{f,A}(n) - \boldsymbol{\varphi}_{f,A}(n)].$$

But, for all $\lambda \in R$, $0 \leq \boldsymbol{\varphi}_{f,A}(\lambda) \leq 1$ (Theorem 4. 3), thus $\boldsymbol{\varphi}_{f,A}(-n) \geq 0$, $\boldsymbol{\varphi}_{f,A}(n) \leq 1$ or $\boldsymbol{\varphi}_{f,A}(n) - \boldsymbol{\varphi}_{f,A}(-n) \leq 1$,

Since for all $n \in N$, $\mu_{f,A}(\Delta_n) \leq 1$, we have

$$\sup_{n \in N} \mu_{f,A}(\Delta_n) \leq 1$$

and, therefore, $\mu_{f,A}(R) \leq 1$.

This completes the proof. □

Definition 4.3: A unit vector $f \in {}^*\mathcal{H}$ is called a probability measure inducing vector with respect to a standard self-adjoint operator A densely defined on ${}^*\mathcal{H}$, if and only if the function $\boldsymbol{\varphi}_{f,A}$ satisfies the properties:

(i) If $\{\lambda_n: n \in N\}$ is an increasing sequence of real numbers tending to $+\infty$, then $\lim \boldsymbol{\varphi}_{f,A}(\lambda_n) = 1$, $n \in N$.

(ii) If $\{\lambda_n: n \in N\}$ is a decreasing sequence of real numbers tending to $-\infty$, then $\lim \boldsymbol{\varphi}_{f,A}(\lambda_n) = 1$, $n \in N$.

It can be easily verified that we may replace $\boldsymbol{\varphi}_{f,A}$ in Def. 4. 3 by $\varphi_{f,A}$ without affecting it [i.e., conditions (i) and (ii) in the definition hold for $\varphi_{f,A}$ if and only if they hold for $\boldsymbol{\varphi}_{f,A}$].

Theorem 4.5: Let f be a probability measure inducing vector with respect to a standard self-adjoint operator A densely defined on ${}^*\mathcal{H}$, then, $\mu_{f,A}$ is a probability measure satisfying

$$\mu_{f,A}((-\infty, a)) = \boldsymbol{\varphi}_{f,A}(a) \quad \text{for every } a \in R.$$

Proof: Let $a \in R$, and define $\lambda_n = a - n$ for every $n \in N$. Now, $\{\lambda_n: n \in N\}$ is a decreasing sequence converging to $-\infty$, and so

$$\lim_{n \in N} \boldsymbol{\varphi}_{f,A}(\lambda_n) = 0. \tag{4. 8}$$

Let $\Delta_n = [\lambda_n, a)$. By Theorem 4. 4

$$\mu_{f,A}(\Delta_n) = \boldsymbol{\varphi}_{f,A}(a) - \boldsymbol{\varphi}_{f,A}(\lambda_n) \tag{4. 9}$$

which, since $\{\Delta_n: n \in N\}$ is an increasing sequence and $\cup_{n \in N} \Delta_n = (-\infty, a)$ implies

$$\mu_{f,A}((-\infty, a)) = \mu_{f,A}\Big(\bigcup_{n \in N} \Delta_n\Big) = \sup_{n \in N} \mu_{f,A}(\Delta_n)$$
$$= \sup_{n \in N} [\boldsymbol{\varphi}_{f,A}(a) - \boldsymbol{\varphi}_{f,A}(\lambda_n)] \text{ [by Eq. (4. 9)]}$$
$$= \boldsymbol{\varphi}_{f,A}(a) - \inf_{n \in N} \boldsymbol{\varphi}_{f,A}(\lambda_n)$$
$$= \boldsymbol{\varphi}_{f,A}(a) - \lim_{n \in N} \boldsymbol{\varphi}_{f,A}(\lambda_n) = \boldsymbol{\varphi}_{f,A}(a)$$
[by Eq. (4. 8)]

i.e.,

$$\mu_{f,A}((-\infty, a)) = \boldsymbol{\varphi}_{f,A}(a). \tag{4. 10}$$

To show that $\mu_{f,A}$ is a probability measure, we have to show that $\mu_{f,A}(R) = 1$.

Let $\lambda'_n = n$ for every $n \in N$, hence $\{\lambda'_n : n \in N\}$ is an increasing sequence of real numbers converging to $+\infty$, and therefore

$$\lim_{n\in N} \varphi_{f,A}(\lambda'_n) = 1. \tag{4.11}$$

Let $\Delta'_n = (-\infty, \lambda'_n)$ and by (4.10), we have

$$\mu_{f,A}(\Delta'_n) = \varphi_{f,A}(\lambda'_n) \tag{4.12}$$

and since $\{\Delta'_n : n \in N\}$ is an increasing sequence with $R = \cup_{n\in N} \Delta'_n$, we find

$$\begin{aligned}\mu_{f,A}(R) &= \mu_{f,A}\Big(\bigcup_{n\in N} \Delta'_n\Big) = \sup_{n\in N} \mu_{f,A}(\Delta'_n) \\ &= \sup_{n\in N} \varphi_{f,A}(\lambda'_n) \qquad \text{[by Eq. (4.12)]} \\ &= \lim_{n\in N} \varphi_{f,A}(\lambda'_n) = 1 \quad \text{[by Eq. (4.11)]};\end{aligned}$$

thus

$$\mu_{f,A}(R) = 1. \qquad \blacksquare \tag{4.13}$$

Definition 4.4: Let A be a standard self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let $f \in {}^*\mathcal{H}$ be a probability measure inducing vector with respect to A. Define

$${}^0\text{prob}(A, f, \Delta) = \mu_{f,A}(\Delta)$$

for every Borel subset $\Delta \subset R$.

We relate now the probability measure ${}^0\text{prob}(A, f, \Delta)$ with the probability measure $\text{prob}({}^0A, \text{st}(f), \Delta)$ where ${}^0A$ is the restriction of A to $\mathcal{H}$ and f is a near standard vector. The measure $\text{prob}({}^0A, \text{st}(f), \Delta)$ is defined as $\langle \text{st}(f), E_\Delta({}^0A)\,\text{st}(f)\rangle$, where $\text{st}(f) \in \mathcal{H}$ and $E({}^0A)$ is the *standard* spectral measure associated with ${}^0A$. Δ is a Borel subset of the real line R.

Theorem 4.6: Let A be a standard self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let $f \in {}^*\mathcal{H}$ be a unit near standard vector. Then f is a probability measure inducing vector relative to A, and, for every Borel subset $\Delta \subset R$,

$${}^0\text{prob}(A, f, \Delta) = \text{prob}({}^0A, \text{st}(f), \Delta),$$

where ${}^0A$ is the restriction of A to $\mathcal{H}$.

Proof: Since f is near standard, there exists a standard vector $g \in \mathcal{H}$ and an infinitesimal vector $d \in {}^*\mathcal{H}$ such that

$$f = g + d. \tag{4.14}$$

Now

$$\langle f, f\rangle = \langle g, g\rangle + \langle g, d\rangle + \langle d, g\rangle + \langle d, d\rangle.$$

But

$$\langle d, d\rangle = \|d\|^2 \approx 0,$$

$$|\langle d, g\rangle| = |\langle g, d\rangle| \leq \|g\|\ \|d\| \approx 0;$$

hence

$$\langle g, d\rangle + \langle d, g\rangle + \langle d, d\rangle \approx 0,$$

$$\text{st}\langle f, f\rangle = \text{st}\langle g, g\rangle,$$

i.e., $\text{st}\,\|f\|^2 = \text{st}\,\|g\|^2$.

But $\|f\| = 1$ and $\|g\| \in R$ (g is a standard vector), hence

$$\|g\| = 1. \tag{4.15}$$

Now let $\lambda \in R$ and $E(A)$ be the nonstandard spectral measure associated with A, then

$$\begin{aligned}\langle f, E_\lambda(A) f\rangle &= \langle g + d, E_\lambda(A)(g + d)\rangle \\ &= \langle g, E_\lambda(A) g\rangle + \langle d, E_\lambda(A) d\rangle \\ &\quad + \langle g, E_\lambda(A) d\rangle + \langle d, E_\lambda(A) g\rangle.\end{aligned} \tag{4.16}$$

Now

$$|\langle g, E_\lambda(A) d\rangle| = |\langle d, E_\lambda(A) g\rangle| \leq \|d\| \cdot \|E_\lambda(A) g\| \leq \|d\| \cdot \|g\| = \|d\| \approx 0,$$

$$\therefore\ |\langle g, E_\lambda(A) d\rangle| = |\langle d, E_\lambda(A) g\rangle| \approx 0;$$

also

$$\langle d, E_\lambda(A) d\rangle = \|E_\lambda(A) d\|^2 \leq \|d\|^2 \approx 0$$

hence

$$\langle d, E_\lambda(A) d\rangle + \langle g, E_\lambda(A) d\rangle + \langle d, E_\lambda(A) g\rangle \approx 0$$

or

$$\text{st}\langle f, E_\lambda(A) f\rangle = \text{st}\langle g, E_\lambda(A) g\rangle. \tag{4.17}$$

But, since $g \in \mathcal{H}$ and $\lambda \in R$, then $\langle g, E_\lambda(A) g\rangle = \langle g, E_\lambda({}^0A) g\rangle$, where ${}^0A$ is the restriction of A to $\mathcal{H}$.

Equation (4.17) can be written in the form

$$\varphi_{f,A}(\lambda) = \langle g, E_\lambda({}^0A) g\rangle. \tag{4.18}$$

Now, since the function $\lambda \to \langle g, E_\lambda({}^0A) g\rangle$ is continuous from below, we get

$$\varphi_{f,A}(\lambda) = \varphi_{f,A}(\lambda) = \langle g, E_\lambda({}^0A) g\rangle. \tag{4.19}$$

Now, if $\{\lambda_n : n \in N\}$ is an increasing (decreasing) sequence of real numbers converging to $+\infty$ ($-\infty$), then $\lim_{n\in N} \langle g, E_{\lambda_n}({}^0A) g\rangle = 1$ ($= 0$). Hence f is a probability measure inducing vector with respect to A.

Using this fact, Eq. (4.19) yields to

$${}^0\text{prob}(A, f, \Delta) = \text{prob}({}^0A, g, \Delta)$$

or, for all Borel subsets $\Delta \subset R$, $f \in {}^*\mathcal{H}$ is near standard;

$${}^0\text{prob}(A, f, \Delta) = \text{prob}({}^0A, \text{st}(f), \Delta). \qquad \blacksquare \tag{4.20}$$

In the special case where f is standard, we have $\text{st}(f) = f$, and so

$${}^0\text{prob}(A, f, \Delta) = \text{prob}({}^0A, f, \Delta).$$

However, vectors which are not near standard have interesting properties, some of which are dealt with in the following subsection.

B. Measure of point sets

The probability measure $\text{prob}(A, f, \Delta)$ in a standard Hilbert space gives a zero measure for a point set $\{\lambda\}$ if λ does not belong to the discrete spectrum of A. The probability measure $\mu_{f,A}$ in a nonstandard Hilbert space has different properties. We shall show that if f is an ultra eigenvector of A with ultra eigenvalue $\lambda \in R$, then $\mu_{f,A}(\{\lambda\}) = 1$ or, equivalently,

$$\mu_{f,A}(\Delta) = \begin{cases} 1 & \text{if } \lambda \in \Delta \\ 0 & \text{if } \lambda \notin \Delta \end{cases} \text{ for every Borel subset } \Delta \subset R.$$

Theorem 4.7: Let A be a standard self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let $f \in {}^*\mathcal{H}$, $\|f\| = 1$ be an ultra eigenvector of A with ultra eigenvalue $\lambda \in R$, then

$$\varphi_{f,A}(\lambda') = \begin{cases} 0 & \text{if } \lambda' < \lambda \\ 1 & \text{if } \lambda' > \lambda . \end{cases}$$

Proof: Let $d = Af - \lambda f$, $\|d\| \approx 0$, since $\|Af - \lambda f\| / \|f\| \approx 0$ and $\|f\| = 1$. Now

$$\|d\|^2 = \langle d, d\rangle = \langle (A-\lambda)f, (A-\lambda)f\rangle$$
$$= \langle f, (A-\lambda)^2 f\rangle \ (A \text{ is self-adjoint}).$$

But, from standard spectral theory extended to ${}^*\mathcal{H}$

$$\langle f, (A-\lambda)^2 f\rangle = \int (\lambda' - \lambda)^2 d\langle f, E_{\lambda'}(A) f\rangle;$$

therefore

$$\int (\lambda' - \lambda)^2 d\langle f, E_{\lambda'}(A) f\rangle \approx 0. \tag{4.21}$$

Let $\delta > 0$, $\delta \in R$ then (4.21) becomes

$$\int_{-\infty}^{\lambda-\delta} (\lambda'-\lambda)^2 \, d\langle f, E_{\lambda'}(A) f\rangle + \int_{\lambda-\delta}^{\lambda+\delta} (\lambda' - \lambda)^2 d\langle f, E_{\lambda'}(A) f\rangle$$
$$+ \int_{\lambda+\delta}^{+\infty} (\lambda' - \lambda)^2 d\langle f, E_{\lambda'}(A) f\rangle \approx 0. \tag{4.22}$$

From the fact that the integrand in each of the three integrals in (4.22) is nonnegative, we deduce that each of these integrals is infinitesimal. In particular,

$$\int_{-\infty}^{\lambda-\delta} (\lambda' - \lambda)^2 \, d\langle f, E_{\lambda'}(A) f\rangle \approx 0, \tag{4.23}$$

$$\int_{\lambda+\delta}^{+\infty} (\lambda' - \lambda)^2 d\langle f, E_{\lambda'}(A) f\rangle \approx 0. \tag{4.24}$$

But

$$\lambda' < \lambda - \delta \text{ implies that } (\lambda' - \lambda)^2 \geq \delta^2; \tag{4.25}$$

also

$$\lambda + \delta \leq \lambda' \text{ implies that } (\lambda' - \lambda)^2 \geq \delta^2, \tag{4.26}$$

Hence, from (4.23) and (4.25) we get

$$0 \approx \int_{-\infty}^{\lambda-\delta} (\lambda' - \lambda)^2 d\langle f, E_{\lambda}(A) f\rangle \geq \delta^2 \int_{-\infty}^{\lambda-\delta} d\langle f, E_{\lambda'}(A) f\rangle$$
$$= \delta^2 \langle f, E_{\lambda-\delta}(A) f\rangle$$

i.e.,

$$\langle f, E_{\lambda-\delta}(A) f\rangle \approx 0 \quad \text{for all } \delta \in R, \ \delta > 0 \tag{4.27}$$

Also, from (4.24) and (4.26) we get

$$0 \approx \int_{\lambda+\delta}^{+\infty} (\lambda' - \lambda)^2 d\langle f, E_{\lambda'}(A) f\rangle \geq \delta^2 \int_{\lambda+\delta}^{+\infty} d\langle f, E_{\lambda'}(A) f\rangle$$
$$= \delta^2 [1 - \langle f, E_{\lambda+\delta}(A) f\rangle]$$

i.e.,

$$\langle f, E_{\lambda+\delta}(A) f\rangle \approx 1 \quad \text{for all } \delta \in R, \ \delta > 0. \tag{4.28}$$

From (4.27) and (4.28) we finally conclude:

$$\varphi_{f,A}(\lambda') = \text{st}\langle f, E_{\lambda'}(A) f\rangle = \begin{cases} 0 & \text{if } \lambda' < \lambda \\ 1 & \text{if } \lambda' > \lambda . \end{cases} \quad \blacksquare$$

Theorem 4.8: Let A be a self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let $f \in {}^*\mathcal{H}$, $\|f\| = 1$ be an ultra eigenvector of A with ultra eigenvalue $\lambda \in R$. Then, f is a probability measure inducing vector relative to A, and satisfies $\mu_{f,A}(\{\lambda\}) = 1$.

Proof: By Theorem 4.7 we have

$$\varphi_{f,A}(\lambda') = \begin{cases} 0 & \text{if } \lambda' < \lambda \\ 1 & \text{if } \lambda' > \lambda . \end{cases} \tag{4.29}$$

It can be directly verified (using Def. 4.1 that

$$\varphi_{f,A}(\lambda') = \begin{cases} 0 & \text{if } \lambda' \leq \lambda \\ 1 & \text{if } \lambda' > \lambda . \end{cases} \tag{4.30}$$

Equation (4.30) may be obtained from (4.31) without any knowledge about the specific value of $\varphi_{f,A}(\lambda)$, except, of course, the fact that $0 \leq \varphi_{f,A}(\lambda) \leq 1$ which is a consequence of Theorem 4.2. It can also be directly verified that $\varphi_{f,A}$ satisfies the conditions of Def. 4.3 which make f a probability measure inducing vector relative to A.

Now define $\Delta_n = [\lambda, \lambda + (1/n))$, then $\{\Delta_n : n \in N\}$ is a decreasing sequence of Borel subsets with $\{\lambda\} = \cap_{n\in N} \Delta_n$. Hence,

$$\mu_{f,A}(\{\lambda\}) = \mu_{f,A}(\cap_{n\in N} \Delta_n) = \inf_{n\in N} \mu_{f,A}(\Delta_n). \tag{4.31}$$

But

$$\mu_{f,A}(\Delta_n) = \varphi_{f,A}[\lambda + (1/n)] - \varphi_{f,A}(\lambda) \quad \text{(Theorem 4.4)}$$
$$= 1 \quad \text{[by Eq. (4.30)]}. \tag{4.32}$$

From (4.31) and (4.32) we deduce that

$$\mu_{f,A}(\{\lambda\}) = 1. \quad \blacksquare \tag{4.33}$$

C. General probability set functions

We study here the case where a unit vector $f \in {}^*\mathcal{H}$ does not satisfy the conditions of Def. 4.3. In this case, the measure $\mu_{f,A}$ defined on the Borel subsets of the real line is not a probability measure [because $\mu_{f,A}(R) < +1$]. We substitute it by a set function $\nu_{f,A}$ defined on a ring $G(R)$ of Borel subsets of the real line, and satisfying:

(i) $\nu_{f,A}(\phi) = 0$;

(ii) $\nu_{f,A}(R) = 1$;

(iii) for all $E \in G(R)$, $\nu_{f,A}(E) \geq 0$;

(iv) If $E_1, E_2 \in G(R)$, $E_1 \cap E_2 = \phi$, then $\nu_{f,A}(E_1 \cup E_2) = \nu_{f,A}(E_2) + \nu_{f,A}(E_2)$.

We notice that if $f: R \to R$ is a nondecreasing function, then for every $\lambda \in R$ the set $\{f(\lambda + \epsilon) : \epsilon > 0\}$ is not empty and has $f(\lambda)$ as a lower bound. Hence it has an infimum.

Definition 4.5: Let $f: R \to R$ be a nondecreasing function. Define $\bar{f}(\lambda) = \inf_{\epsilon>0} f(\lambda + \epsilon)$.

Theorem 4.9: Let $f: R \to R$ is nondecreasing, then $\bar{f}: R \to R$ is nondecreasing and continuous from above

Proof: If $\lambda < \lambda'$, then $\bar{f}(\lambda) \leq f(\lambda')$ (by definition). But $f(\lambda') \leq \bar{f}(\lambda')$, hence $\bar{f}(\lambda) \leq \bar{f}(\lambda')$, and $\bar{f}$ is nondecreasing.

Also for every $\epsilon > 0$ there exists a $\delta > 0$ such that $f(\lambda + 2\delta) \leq \bar{f}(\lambda) + \epsilon$.

But $\bar{f}(\lambda + \delta) \leq f(\lambda + 2\delta)$ (because $\lambda + \delta < \lambda + 2\delta$), hence $\bar{f}(\lambda + \delta) \leq \bar{f}(\lambda) + \epsilon$.

Now $\bar{f}(\lambda) \leq \bar{f}(\lambda + \delta)$ (proved) hence, for every $\epsilon > 0$ there exists a $\delta > 0$ such that for every $\lambda < \lambda' < \lambda + \delta$,

we have $\bar{f}(\lambda) \leq \bar{f}(\lambda') \leq \bar{f}(\lambda) + \epsilon$. Hence $\bar{f}$ is continuous from above. ■

Theorem 4.10: Let $f: R \to R$ be a nondecreasing function, then, for every $\lambda \in R$,

(i) $\inf_{\epsilon<0} \mathbf{f}(\lambda + \epsilon) = \bar{f}(\lambda)$;

(ii) $\sup_{\epsilon>0} \bar{f}(\lambda - \epsilon) = \mathbf{f}(\lambda)$.

Proof: Since $\mathbf{f}(\lambda) \leq f(\lambda)$ for every $\lambda \in R$, then

$$\inf_{\epsilon>0} \mathbf{f}(\lambda + \epsilon) \leq \inf_{\epsilon>0} f(\lambda + \epsilon) = \bar{f}(\lambda). \tag{4.34}$$

Let $g(\lambda) = \inf_{\epsilon>0} \mathbf{f}(\lambda + \epsilon)$, hence for every $\epsilon > 0$ there exists a $\delta > 0$ such that $\mathbf{f}(\lambda + 2\delta) \leq g(\lambda) + \epsilon$.

But $\mathbf{f}(\lambda + 2\delta) \geq f(\lambda + \delta) \geq \bar{f}(\lambda)$ (by definition) hence

$$\bar{f}(\lambda) \leq g(\lambda) + \epsilon. \tag{4.35}$$

Since $\epsilon > 0$, but otherwise arbitrary, (4.35) implies that

$$\bar{f}(\lambda) \leq g(\lambda) = \inf_{\epsilon>0} \mathbf{f}(\lambda + \epsilon). \tag{4.36}$$

From (4.34) and (4.36) we conclude that $\bar{f}(\lambda) = \inf_{\epsilon>0} \mathbf{f}(\lambda + \epsilon)$. Similarly, we can prove that $\sup_{\epsilon>0} \bar{f}(\lambda - \epsilon) = \mathbf{f}(\lambda)$. ■

Definition 4.6:

(i) $\mathcal{T}_- = \{(-\infty, a): a \in R\}$;

(ii) $\mathcal{T}_+ = \{(a, +\infty): a \in R\}$;

(iii) $\mathcal{T} = \mathcal{T}_- \cup \mathcal{T}_+ \cup \{\phi\}$.

Definition 4.7: $G(R)$ is the minimal ring containing $\mathcal{T}$ (defined in Def. 4.6).

Definition 4.8: Let A be a standard self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let $f \in {}^*\mathcal{H}$, $\|f\| = 1$, and define $\nu^0_{f,A}: \mathcal{T} \to R$ as follows:

(i) $\nu^0_{f,A}((-\infty, a)) = \varphi_{f,A}(a)$;

(ii) $\nu^0_{f,A}((a, +\infty)) = 1 - \bar{\varphi}_{f,A}(a)$;

(iii) $\nu^0_{f,A}(\phi) = 0$;

where $\varphi_{f,A}$ is the function defined in Def. 4.2.

Theorem 4.22: Let A be a standard self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let $f \in {}^*\mathcal{H}$, $\|f\| = 1$, then the function $\nu^0_{f,A}$ has a unique extension $\nu_{f,A}$ to $G(R)$, such that:

(i) $\nu_{f,A}(\phi) = 0$;

(ii) $\nu_{f,A}(R) = 1$;

(iii) for every $E \in G(R)$, $\nu_{f,A}(E) \geq 0$;

(iv) if $E_1, E_2 \in G(R)$, $E_1 \cap E_2 = \phi$, then $\nu_{f,A}(E_1 \cup E_2) = \nu_{f,A}(E_1) + \nu_{f,A}(E_2)$.

Proof: Let $a < b$ then $[a, b) = (-\infty, b) \setminus (-\infty, a)$ which means that $G(R)$ contains the ring $\mathcal{E}_1$ generated by semi-closed semi-open intervals of the real line.

By Theorem 4.4 there exists a measure $m^{(1)}_{f,A}$ on $\mathcal{E}_1$ such that

$$m^{(1)}_{f,A}([a, b)) = \varphi_{f,A}(b) - \varphi_{f,A}(a). \tag{4.37}$$

We have also $(a, b] = (a, +\infty) \setminus (b, +\infty)$ which means that $G(R)$ contains the ring $\mathcal{E}_2$ generated by semi-open semi-closed intervals of the real line.

By a similar argument, there exists a measure $m^{(2)}_{f,A}$ on $\mathcal{E}_2$ such that

$$m^{(2)}_{f,A}((a, b]) = \bar{\varphi}_{f,A}(b) - \bar{\varphi}_{f,A}(a). \tag{4.38}$$

Let $G_0(R)$ be the minimal ring generated by $\mathcal{E}_1 \cup \mathcal{E}_2$. Hence $G_0(R) \subset G(R)$ [by definition of $G_0(R)$].

As a result of Theorem 4.10, it can be shown that there exists a measure $m_{f,A}$ on $G_0(R)$, such that:

(i) if $\Delta \in \mathcal{E}_1$, then $m_{f,A}(\Delta) = m^{(1)}_{f,A}(\Delta)$;

(ii) if $\Delta \in \mathcal{E}_2$, then $m_{f,A}(\Delta) = m^{(2)}_{f,A}(\Delta)$.

Now, a general element $E \in G(R) \setminus \{R\}$ can be written as

$$E = \Delta_- \cup \Delta_0 \cup \Delta_+, \quad \Delta_- \cap \Delta_0 = \Delta_0 \cap \Delta_+ = \phi,$$

where

$$\Delta_- \in \mathcal{T}_- \cup \{\phi\}, \quad \Delta_+ \in \mathcal{T}_+ \cup \{\phi\}, \quad \Delta_0 \in G_0(R). \tag{4.39}$$

Since $E \neq R$, we can easily see that $\Delta_- \cap \Delta_+ = \phi$.

Assume now that

$$E = \Delta'_- \cup \Delta'_0 \cup \Delta'_+, \quad \Delta'_- \cap \Delta'_0 = \Delta'_0 \cap \Delta_+ = \phi$$

with

$$\Delta'_- \in \mathcal{T}_- \cup \{\phi\}, \quad \Delta'_+ \in \mathcal{T}_+ \cup \{\phi\}, \quad \Delta'_0 \in G_0(R).$$

If

$$\left.\begin{aligned} \Delta_- &\subset \Delta'_- \\ \Delta_+ &\subset \Delta'_+ \end{aligned}\right\} \text{ then } \left\{\begin{aligned} \Delta'_- &= \Delta_- \cup (\Delta'_- \setminus \Delta_-) \\ \Delta'_+ &= \Delta_+ \cup (\Delta'_+ \setminus \Delta_+) \end{aligned}\right. \tag{4.40}$$

and hence, $\Delta_0 = \Delta'_0 \cup (\Delta'_- \setminus \Delta_-) \cup (\Delta'_+ \setminus \Delta_+)$

with

$$\Delta'_0 \cap (\Delta'_- \setminus \Delta_-) = (\Delta'_- \setminus \Delta_-) \cap (\Delta'_+ \setminus \Delta_+) = (\Delta'_+ \setminus \Delta_+) \cap \Delta'_0 = \phi. \tag{4.41}$$

Now

$$\nu^0_{f,A}(\Delta'_-) = \nu^0_{f,A}(\Delta_-) + m^{(1)}_{f,A}(\Delta'_- \setminus \Delta_-), \tag{4.42a}$$

$$\nu^0_{f,A}(\Delta'_+) = \nu^0_{f,A}(\Delta_+) + m^{(2)}_{f,A}(\Delta'_+ \setminus \Delta_+), \tag{4.42b}$$

$$m_{f,A}(\Delta_0) = m_{f,A}(\Delta'_0) + m_{f,A}(\Delta'_- \setminus \Delta_-) + m_{f,A}(\Delta'_+ \setminus \Delta_+). \tag{4.42c}$$

From Eqs. (4.42a)–(4.42c) we get

$$\nu^0_{f,A}(\Delta'_-) + \nu^0_{f,A}(\Delta'_+) + m_{f,A}(\Delta'_0) = \nu^0_{f,A}(\Delta_-) + \nu^0_{f,A}(\Delta_+) + m_{f,A}(\Delta_0). \tag{4.43}$$

Result (4.43) which holds for case (4.41), holds also for other cases. Hence we may introduce the definition

$$\nu_{f,A}(E) = \nu^0_{f,A}(\Delta_-) + m_{f,A}(\Delta_0) + \nu^0_{f,A}(\Delta_+). \tag{4.44}$$

This function is unambiguously defined, irrespective of the decomposition of E used. We define also

$$\nu_{f,A}(R) = 1. \tag{4.45}$$

The set function $\nu_{f,A}$ thus defined, is an extension of both $\nu^0_{f,A}$ and $m_{f,A}$. It satisfies conditions (i), (ii), and (iii) required by the theorem. To show that condition (iv) is satisfied as well, let $E_1, E_2 \in G(R)$ with $E_1 \cap E_2 = \phi$.

Now, if $E_1 = R$ then $E_2 = \phi$ and

$$\nu_{f,A}(E_1 \cup E_2) = \nu_{f,A}(E_1) + \nu_{f,A}(E_2). \tag{4.46}$$

Similarly, (4.46) holds for the case $E_2 = R$. Let now $E_1 \neq R$ and $E_2 \neq R$ then,

$$E_1 = \Delta_-^{(1)} \cup \Delta_0^{(1)} \cup \Delta_+^{(1)}, \quad \Delta_-^{(1)} \cap \Delta_0^{(1)} = \Delta_0^{(1)} \cap \Delta_+^{(1)} = \phi,$$
$$E_2 = \Delta_-^{(2)} \cup \Delta_0^{(2)} \cup \Delta_+^{(2)}, \quad \Delta_-^{(2)} \cap \Delta_0^{(2)} = \Delta_0^{(2)} \cap \Delta_+^{(2)} = \phi$$

with

$$\Delta_-^{(1)}, \Delta_-^{(2)} \in J_- \cup \{\phi\}, \quad \Delta_+^{(1)}, \Delta_+^{(2)} \in J_+ \cup \{\phi\},$$
$$\Delta_0^{(1)}, \Delta_0^{(2)} \in G_0(R).$$

Now

$$E_1 \cup E_2 = (\Delta_-^{(1)} \cup \Delta_-^{(2)}) \cup (\Delta_0^{(1)} \cup \Delta_0^{(2)}) \cup (\Delta_+^{(1)} \cup \Delta_+^{(2)}). \tag{4.47}$$

But, since $E_1 \cap E_2 = \phi$, then

(a) $\Delta_-^{(1)} = \phi$ or $\Delta_-^{(2)} = \phi$,

(b) $\Delta_+^{(1)} = \phi$ or $\Delta_+^{(2)} = \phi$,

(c) $\Delta_0^{(1)} \cap \Delta_0^{(2)} = \phi$,

(d) $(\Delta_-^{(1)} \cup \Delta_-^{(2)}) \cap (\Delta_0^{(1)} \cup \Delta_0^{(2)}) = (\Delta_0^{(1)} \cup \Delta_0^{(2)}) \cap (\Delta_+^{(1)} \cup \Delta_+^{(2)}) = \phi.$

From (a) we get

$$\nu_{f,A}^0(\Delta_-^{(1)} \cup \Delta_-^{(2)}) = \nu_{f,A}^0(\Delta^{(1)}) + \nu_{f,A}^0(\Delta_-^{(2)}). \tag{4.48a}$$

From (b) we get

$$\nu_{f,A}^0(\Delta_+^{(1)} \cup \Delta_+^{(2)}) = \nu_{f,A}^0(\Delta_+^{(1)}) + \nu_{f,A}^0(\Delta_+^{(2)}). \tag{4.48b}$$

From (c) we get

$$m_{f,A}(\Delta_0^{(1)} \cup \Delta_0^{(2)}) = m_{f,A}(\Delta_0^{(1)}) + m_{f,A}(\Delta_0^{(2)}). \tag{4.48c}$$

From (d) we get

$$\nu_{f,A}(E_1 \cup E_2) = \nu_{f,A}^0(\Delta_-^{(1)} \cup \Delta_-^{(2)}) + m_{f,A}(\Delta_0^{(1)} \cup \Delta_0^{(2)}) + \nu_{f,A}^0(\Delta_+^{(1)} \cup \Delta_+^{(2)}). \tag{4.48d}$$

By using (4.48a)–(4.48d) we get finally

$$\nu_{f,A}(E_1 \cup E_2) = \nu_{f,A}(E_1) + \nu_{f,A}(E_2). \tag{4.49}$$

Thus, the theorem is proved. □

We have now to link the probability set function $\nu_{f,A}$ (which need not be a measure) with the probability measure $\mu_{f,A}$.

Theorem 4.12: Let A be a standard self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let $f \in {}^*\mathcal{H}$, $\|f\| = 1$ be a probability measure inducing vector, then for every $E \in G(R)$, $\mu_{f,A}(E) = \nu_{f,A}(E)$.

Proof: By Theorem (4.5)

$$\mu_{f,A}((-\infty, a)) = \varphi_{f,A}(a) = \nu_{f,A}((-\infty, a)),$$
$$\mu_{f,A}(R) = 1. \tag{4.50}$$

Let $\Delta_n = [a + (1/n), +\infty)$, then $\{\Delta_n : n \in N\}$ is an increasing sequence with $\cup_{n\in N}\ \Delta_n = (a, +\infty)$. Hence

$$\mu_{f,A}((a, +\infty)) = \mu_{f,A}(\cup_{n\in N}\ \Delta_n) = \sup_{n\in N}\ \mu_{f,A}(\Delta_n)$$
$$= \sup_{n\in N}\ \mu_{f,A}(R \setminus (-\infty, a + (1/n))$$
$$= \sup_{n\in N}\ [1 - \varphi_{f,A}(a + (1/n))] = 1 - \inf_{n\in N}\ \varphi_{f,A}(a + (1/n))$$
$$= 1 - \bar{\varphi}_{f,A}(a) = \nu_{f,A}^0((a, +\infty)) \tag{4.51}$$

Now,

$$\mu_{f,A}(\phi) = 0; \tag{4.52}$$

hence for all $E \in \mathcal{T}$, $\mu_{f,A}(E) = \nu_{f,A}^0(E)$.

The extension of $\nu_{f,A}^0$ to an additive function $\nu_{f,A}$ on $G(R)$ is unique. Since $\mu_{f,A}$ is also additive on $G(R)$ and coincides with $\nu_{f,A}^0$ on $\mathcal{T}$, we have for every $E \in G(R)$.

$$\mu_{f,A}(E) = \nu_{f,A}(E). \qquad \square$$

Definition 4.9: Let A be a standard self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let $f \in {}^*\mathcal{H}$, $\|f\| = 1$. Define the probability function $q - \text{prob}(A, f, \Delta) = \nu_{f,A}(\Delta)$ for every $\Delta \in G(R)$.

The function: $\Delta \to q - \text{prob}(A, f, \Delta)$ is a quasi probability measure. The only condition which is not satisfied in general is the σ-additivity.

5. FORMULATION OF QUANTUM MECHANICS

In this section we reformulate quantum mechanics using the concept of a nonstandard separable Hilbert space. There are two methods by which the problem may be solved. In the inductive method, one starts with a formalized physical structure (e.g., the lattice of propositions associated with a set of yes-no experiments[1]), and proves that it is isomorphic to a mathematical structure (e.g., a complex Hilbert space). In the deductive method, one starts with a mathematical structure and deduces the physical results.

We shall follow here the deductive method. The reason for this choice is that is is very difficult (if possible at all) to prove that a given structure is isomorphic to a nonstandard Hilbert space, because we cannot construct any specific nonstandard Hilbert space. All that we know is that the existence of nonstandard Hilbert spaces is a consequence of the axiom of choice. Although it is usually postulated that the dynamical variables of a quantal system satisfy some Lie algebra, we are not going to assume anything of this sort here. Our approach is quite general. It may be applied to relativistic as well as nonrelativistic quantum mechanics.

A. Kinematics

In standard quantum mechanics, states (or pure states) are represented by unit rays in a separable complex Hilbert space. If $f \in \mathcal{H}$ is a unit vector, then f and $e^{i\varphi} f (\varphi \in R)$ represent the same state. Hence, if we define the relation: $f \overset{\mathcal{H}}{\longleftrightarrow} g$ if and only if there exists a $\varphi \in R$ such that $f = e^{i\varphi} g$, we immediately see that $\overset{\mathcal{H}}{\longleftrightarrow}$ is an equivalence relation on $\mathcal{H}$. If $U(\mathcal{H})$ is the set of unit vectors, then the set of physical states is the quotient set $S(\mathcal{H}) = U(\mathcal{H})/\overset{\mathcal{H}}{\longleftrightarrow}$. What is significant about this definition is that the expectation value of any operator is a property of the state and not the specific vector representing the state. Thus, if $f \in \mathcal{H}$, $\|f\| = 1$ belongs to the domain of an operator A on $\mathcal{H}$, then $\langle f, Af\rangle = \langle e^{i\varphi} f, Ae^{i\varphi} f\rangle$.

In the nonstandard theory, we meet another possible degree of freedom (other than the phase). Infinitesimal vectors should not be distinguished from the zero vector. This means that if $d \in {}^*\mathcal{H}$, $f \in {}^*\mathcal{H}$, $\|d\| \approx 0$, $\|f\| = \|f + d\| = 1$, then f and $f + d$ should represent the same state. Hence we introduce the definition:

Definition 5.1: $f \overset{{}^*\mathcal{H}}{\longleftrightarrow} g$ if and only if $f, g \in {}^*\mathcal{H}$, $\|f\| = \|g\| = 1$, and there exists a $\varphi \in {}^*R$ such that $\|e^{i\varphi} f - g\| \approx 0$.

Theorem 5.1: Let $U(^*\mathcal{H})$ be the set of unit vectors in $^*\mathcal{H}$, then "$\overset{^*\mathcal{H}}{\longleftrightarrow}$" defined in Def. 5. 1 is an equivalence relation on $U(^*\mathcal{H})$.

Proof: Let $f \in U(^*\mathcal{H})$, then $\| f\| = 1$ and $\|f - f\| = 0$, hence $f \overset{^*\mathcal{H}}{\longleftrightarrow} f$ and "$\overset{^*\mathcal{H}}{\longleftrightarrow}$" is reflexive.

Assume now that $f \overset{^*\mathcal{H}}{\longleftrightarrow} g$. It follows that $\|f\| = \|g\| = 1$ and that there exists a number $\varphi \in {}^*R$ such that $\| e^{i\varphi} f - g\| \approx 0$. But $\| e^{-i\varphi} g - f\| = \| e^{i\varphi} f - g\|$. Hence $g \overset{^*\mathcal{H}}{\longleftrightarrow} f$ and "$\overset{^*\mathcal{H}}{\longleftrightarrow}$" is symmetric.

Assume, finally, that $f \overset{^*\mathcal{H}}{\longleftrightarrow} g$ and $g \overset{^*\mathcal{H}}{\longleftrightarrow} h$. Hence $\| f\| = \| g\| = \| h\| = 1$ and there exist two numbers $\varphi, \psi \in {}^*R$ such that $\| e^{i\varphi} f - g\| \approx 0$ and $\| e^{i\psi} g - h\| \approx 0$. But $\| e^{i(\varphi+\psi)} f - e^{i\psi} g\| = \| e^{i\varphi} f - g\|$ and $\| e^{i(\varphi+\psi)} f - h\| \leq \| e^{i(\varphi+\psi)} f - e^{i\psi} g\| + \| e^{i\psi} g - h\|$, thus $\| e^{i(\varphi+\psi)} f - h\| \approx 0$. Hence $f \overset{^*\mathcal{H}}{\longleftrightarrow} h$ and "$\overset{^*\mathcal{H}}{\longleftrightarrow}$" is transitive. This completes the proof. ■

Definition 5.2: The set of physical states $S(^*\mathcal{H})$ is defined as $S(^*\mathcal{H}) = U(^*\mathcal{H})/\overset{^*\mathcal{H}}{\longleftrightarrow}$, where $U(^*\mathcal{H}) = \{f: f \in {}^*\mathcal{H} \text{ and } \|f\| = 1\}$.

We prove now that all vectors corresponding to the same physical state define one and the same probability function $\nu_{f,A}$ with respect to any standard self-adjoint operator A densely defined on $^*\mathcal{H}$.

Theorem 5.2: Let A be a standard self-adjoint operator densely defined on $^*\mathcal{H}$. Let $f \overset{^*\mathcal{H}}{\longleftrightarrow} g$. Then $\nu_{f,A} = \nu_{g,A}$, where $\nu_{f,A}$ is the probability function introduced in Theorem 4. 11.

Proof: By definition of the relation $f \overset{^*\mathcal{H}}{\longleftrightarrow} g$, there exists a number $\varphi \in {}^*\mathcal{H}$ and a vector $d \in {}^*\mathcal{H}$, $\| d\| \approx 0$, such that $g = e^{i\varphi} f + d$, $\| f\| = \| g\| = 1$. Let $E(A)$ be the spectral measure associated with A. Let $\lambda \in R$. Now

$$\begin{aligned}\langle g, E_\lambda(A) g\rangle &= \langle e^{i\varphi} f + d, E_\lambda(A)(e^{i\varphi} f + d)\rangle \\ &= \langle f, E_\lambda(A) f\rangle + \langle d, E_\lambda(A) d\rangle \\ &\quad + e^{i\varphi}\langle d, E_\lambda(A) f\rangle + e^{-i\varphi}\langle f, E_\lambda(A) d\rangle; \quad (5.1)\end{aligned}$$

but

$$\begin{aligned}&|\langle d, E_\lambda(A) d\rangle| \leq \| d\|^2 \approx 0, \\ &| e^{i\varphi}\langle d, E_\lambda(A) f\rangle| = | e^{-i\varphi}\langle f, E_\lambda(A) d\rangle| \leq \| d\| \approx 0. \quad (5.2)\end{aligned}$$

Hence, from (5. 1) and (5. 2) we conclude that

$$\mathrm{st}\langle g, E_\lambda(A) g\rangle = \mathrm{st}\langle f, E_\lambda(A) f\rangle. \quad (5.3)$$

However, (5. 3) is equivalent to

$$\varphi_{g,A}(\lambda) = \varphi_{f,A}(\lambda) \quad (5.4)$$

and (5. 4), finally, leads to $\nu_{g,A} = \nu_{f,A}$. ■

We introduce now the following axioms:

Axiom 1: Physical states are represented by classes of equivalent unit vectors in a nonstandard Hilbert space $^*\mathcal{H}$. The equivalence relation is defined in Def. 5. 1.

Axiom 2: Observables (dynamical variables) are represented by *standard* self-adjoint operators densely defined on $^*\mathcal{H}$.

Axiom 3: The result of any measurement of an observable can only be one of the *standard spectral values* of the corresponding operator. As a result of the measurement, the physical system finds itself in a state represented by an ultra eigenvector of the operator representing the measured observable, corresponding to the measured spectral value.

Axiom 4: If a system is known to be in the state represented by the vector f, then the probability that a measurement of an observable A on the state represented by f yields a value in a Borel set $\Delta \in G(R)$ is given by

$$q - \mathrm{prob}(A, f, \Delta) = \nu_{f,A}(\Delta).$$

Axiom 5: If a system makes a transition between the state represented by the vector f_1 and the state represented by the vector f_2, then the transition probability is given by

$$\mathrm{tran\ prob}(f_1 \to f_2) = \mathrm{st}|\langle f_2, f_1\rangle|^2.$$

In these axioms we have two definitions for probability. We want to show here that no contradiction may arise between these two definitions.

Let a system be in a state represented by the vector $f \in {}^*\mathcal{H}$. Let A be an observable, and let λ belong to the standard spectrum of A. Hence, λ is a possible outcome of a measurement of A carried out on the system in the state represented by f. The probability of getting this value is

$$q - \mathrm{prob}(A, f, \{\lambda\}) = \nu_{f,A}(\{\lambda\}) = \overline{\varphi}_{f,A}(\lambda) - \varphi_{f,A}(\lambda). \quad (5.5)$$

As a result of this experiment, the system finds itself in a state represented by an ultra eigenvector of A with ultra eigenvalue λ. Let $\{g_i: i \in I\}$ be a system of vectors satisfying

$$\| Ag_i - \lambda g_i\| \approx 0 \quad \text{for all } i \in I, \quad (5.6)$$

$$\mathrm{st}\langle g_i, g_j\rangle = \delta_i^j \quad \text{for all } i, j \in I. \quad (5.7)$$

The probability that the system undergoes a transition to the state represented by g_i is

$$\mathrm{tran\ prob}(f \to g_i) = \mathrm{st}|\langle g_i, f\rangle|^2. \quad (5.8)$$

By (5. 7), there is no probability that the system may go from state g_i to state g_j, if $i \neq j$. Hence, the probability of transition to one of the states belonging to a countable subclass $\{g_i: i \in J\}$ is

$$\sum_{i \in J} \mathrm{st}|\langle g_i, f\rangle|^2. \quad (5.9)$$

However, by (5. 6), the probability calculated in (5. 9) is the probability of getting the value λ as a result of a measurement of A carried out on the system in the state represented by f and subject to the condition that the system undergoes a transition to one of the states $\{g_i: i \in J\}$. Obviously, this probability should not exceed the unconditioned probability of getting the value λ as a result of the same measurement. Mathematically, it is required that

$$\sum_{i \in J} \mathrm{st}|\langle g_i, f\rangle|^2 \leq \overline{\varphi}_{f,A}(\lambda) - \varphi_{f,A}(\lambda). \quad (5.10)$$

We show now that this condition is satisfied.

Theorem 5.3: Let A be a standard self-adjoint operator densely defined on $^*\mathcal{H}$. Let λ belong to the standard spectrum of A. Let $\{g_i: i \in J\}$ be a countable set satisfying (i) $\| Ag_i - \lambda g_i\| \approx 0$, $\| g_i\| = 1$ for all $i \in J$; (ii) $\mathrm{st}\langle g_i, g_j\rangle = \delta_i^j$ for all $i, j \in J$.

Let $f \in {}^*\mathcal{H}$, $\|f\| = 1$, then

$$\sum_{i\in J} \mathrm{st}|\langle g_i, f\rangle|^2 \leq \bar{\varphi}_{f,A}(\lambda) - \varphi_{f,A}(\lambda).$$

Proof: Let $E(A)$ be the spectral measure associated with A and write $E_{\lambda'}$ for $E^{(A)}_{(-\infty,\lambda')}$. Let J_0 be any finite subset of J. Define the operator

$$B = E_{\lambda+\delta} - E_{\lambda-\delta'} - \sum_{i\in J_0} g_i \otimes L_{g_i}, \tag{5.11}$$

where $\delta, \delta' \in R$, $\delta, \delta' > 0$.

We can easily verify that $B^+ = B$. Hence,

$$\langle f, B^2 f\rangle = \langle f, B^+ Bf\rangle = \langle Bf, Bf\rangle \geq 0. \tag{5.12}$$

Now,

$$\begin{aligned} B^2 &= (E_{\lambda+\delta} - E_{\lambda-\delta'})^2 + \sum_{i\in J_0}\sum_{j\in J_0} (g_i \otimes L_{g_i})(g_j \otimes L_{g_j}) \\ &\quad - (E_{\lambda+\delta} - E_{\lambda-\delta'}) \sum_{i\in J_0} g_i \otimes L_{g_i} \\ &\quad - \sum_{i\in J_0} g_i \otimes L_{g_i}(E_{\lambda+\delta} - E_{\lambda-\delta'}). \end{aligned} \tag{5.13}$$

However,

$$(E_{\lambda+\delta} - E_{\lambda-\delta'})^2 = E_{\lambda+\delta} - E_{\lambda-\delta'}, \tag{5.14}$$

$$\begin{aligned} &\sum_{i\in J_0}\sum_{j\in J_0} (g_i \otimes L_{g_i})(g_j \otimes L_{g_i}) \\ &\quad = \sum_{i\in J_0}\sum_{j\in J_0} g_i \otimes L_{g_j}\langle g_i, g_j\rangle \\ &\quad = \sum_{i\in J_0} g_i \otimes L_{g_i} + \sum_{i\neq j\in J_0} g_i \otimes L_{g_j}\langle g_i, g_j\rangle. \end{aligned} \tag{5.15}$$

Let

$$\begin{aligned} C &= (1 - E_{\lambda+\delta} + E_{\lambda-\delta'}) \sum_{i\in J_0} g_i \otimes L_{g_i} \\ &\quad + \sum_{i\in J_0} g_i \otimes L_{g_i}(1 - E_{\lambda+\delta} + E_{\lambda-\delta'}) \\ &\quad + \sum_{i\neq j\in J_0} g_i \otimes L_{g_j}\langle g_i, g_j\rangle \end{aligned} \tag{5.16}$$

using which, we conclude

$$B^2 = B + C. \tag{5.17}$$

(5.12) and (5.17) imply that

$$\langle f, Bf\rangle + \langle f, Cf\rangle \geq 0. \tag{5.18}$$

But

$$\begin{aligned} \langle f, Cf\rangle &= \sum_{i\in J_0} \langle f, (1 - E_{\lambda+\delta} + E_{\lambda-\delta'})g_i\rangle\langle g_i, f\rangle \\ &\quad + \sum_{i\in J_0} \langle f, g_i\rangle\langle g_i, (1 - E_{\lambda+\delta} + E_{\lambda-\delta'})f\rangle \\ &\quad + \sum_{i\neq j\in J_0} \langle f, g_i\rangle\langle g_i, g_j\rangle\langle g_j, f\rangle. \end{aligned} \tag{5.19}$$

By using Theorem 4.7 we can show that $\|(1 - E_{\lambda+\delta} + E_{\lambda-\delta'})g_i\| \approx 0$ for all $i \in J_0$, from which it follows that

$$\begin{aligned} &\sum_{i\in J_0} \langle f, (1 - E_{\lambda+\delta} + E_{\lambda-\delta'})g_i\rangle\langle g_i, f\rangle \\ &\quad + \sum_{i\in J_0} \langle f, g_i\rangle\langle g_i, (1 - E_{\lambda+\delta} + E_{\lambda-\delta'})f\rangle \approx 0. \end{aligned} \tag{5.20}$$

By using condition (ii), given in the theorem we conclude that

$$\sum_{i\neq j\in J_0} \langle f, g_i\rangle\langle g_i, g_j\rangle\langle g_j, f\rangle \approx 0. \tag{5.21}$$

By inserting (5.20) and (5.21) in (5.19), one gets $\langle f, Cf\rangle \approx 0$, from which it follows that

$$\mathrm{st}\langle f, Bf\rangle \geq 0. \tag{5.22}$$

But, by using (5.11) and (5.22) we get

$$\mathrm{st}\langle f, E_{\lambda+\delta} f\rangle - \mathrm{st}\langle f, E_{\lambda-\delta'} f\rangle - \sum_{i\in J_0} \mathrm{st}\,|\langle g_i, f\rangle|^2 \geq 0,$$

which is equivalent to

$$\sum_{i\in J_0} \mathrm{st}|\langle g_i, f\rangle|^2 \leq \varphi_{f,A}(\lambda + \delta) - \varphi_{f,A}(\lambda - \delta'). \tag{5.23}$$

Since (5.23) is true for every $\delta, \delta' > 0$, $\delta, \delta' \in R$ we get

$$\begin{aligned} \sum_{i\in J_0} \mathrm{st}|\langle g_i, f\rangle|^2 &\leq \inf_{\delta,\delta'>0} [\varphi_{f,A}(\lambda + \delta) - \varphi_{f,A}(\lambda - \delta')] \\ &= \inf_{\delta>0} \varphi_{f,A}(\lambda + \delta) - \sup_{\delta'>0} \varphi_{f,A}(\lambda - \delta') \\ &= \bar{\varphi}_{f,A}(\lambda) - \varphi_{f,A}(\lambda). \end{aligned} \tag{5.24}$$

Since for every partial sum (over a finite index set $J_0 \subset J$) we have

$$\sum_{i\in J_0} \mathrm{st}|\langle g_i, f\rangle|^2 \leq \bar{\varphi}_{f,A}(\lambda) - \varphi_{f,A}(\lambda),$$

we conclude that the supremum of these partial sums (which is the total sum) satisfies the same relation.

Hence

$$\sum_{i\in J} \mathrm{st}|\langle g_i, f\rangle|^2 \leq \bar{\varphi}_{f,A}(\lambda) - \varphi_{f,A}(\lambda). \tag{5.25}$$

This completes the proof of the theorem. ■

Expectation values

In standard quantum mechanics, the expectation value of an observable A when the system is in a state represented by a vector $f \in \mathcal{H}$ is given by $\langle f, Af\rangle = \int \lambda d\langle f, E_\lambda(A) f\rangle$.

In the nonstandard theory, the expression $\langle f, Af\rangle$ cannot serve as an expectation value. The reason is as follows: Let $\{\varphi_n : n \in {}^*N\}$ be a standard orthonormal basis for ${}^*\mathcal{H}$. Let A be the linear operator whose domain is the set of vectors $D = \{f : f \in {}^*\mathcal{H}$ and $\sum_{n\in{}^*N} n^2|\langle\varphi_n, f\rangle|^2 < +\infty\}$, and which maps $f \in D$ into the vector $Af = \sum_{n\in{}^*N} n\varphi_n\langle\varphi_n, f\rangle$. A is a standard self-adjoint operator.

Let $\epsilon \in {}^*R$ be any infinitesimal number and let $\omega \in {}^*N \setminus N$ be infinite. Let $n \in N$ be any natural number. The two vectors φ_n and $f = (1 + \epsilon^2)^{-1/2}(\varphi_n + \epsilon\varphi_\omega)$ satisfy the relation $\varphi_n \overset{*\mathcal{H}}{\Longleftrightarrow} f$. But $\langle f, Af\rangle = (1 + \epsilon^2)^{-1}(n + \epsilon^2\omega) = (1 + \epsilon^2)^{-1}(\langle\varphi_n, A\varphi_n\rangle + \epsilon^2\omega)$. Since ϵ and ω are arbitrary, we may choose $\omega \geq 1/\epsilon^3$ which leads to an infinite difference between the two values $\langle f, Af\rangle$ and $\langle\varphi_n, A\varphi_n\rangle$, though φ_n and f define the same physical state. We introduce now the definitions:

Definition 5.3: Let A be a standard self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let $f \in {}^*\mathcal{H}$, $\|f\| = 1$. Define:

(i) $\bar{\rho}_{f,A} \equiv 1 - \lim_{\lambda\to+\infty} \varphi_{f,A}(\lambda)$;

(ii) $\rho_{f,A} \equiv \lim_{\lambda\to-\infty} \varphi_{f,A}(\lambda)$.

Definition 5.4: Let A and f be as in the preceding definition. A measurable function $u: R \to R$ is said to be integrable with respect to the set function $\nu_{f,A}$ if and

only if it is integrable with respect to the measure $\mu_{f,A}$, and the two limits

$$\lim\ u(\lambda)\bar{\rho}_{f,A} \text{ as } \lambda \to +\infty,\ \lim\ u(\lambda)\rho_{f,A} \text{ as } \lambda \to -\infty$$

exist. The integral of u with respect to the set function $\nu_{f,A}$ is defined to be

$$\int u(\lambda)\,d\nu_{f,A}(\lambda) \equiv \int u(\lambda)\,d\mu_{f,A}(\lambda) + \lim_{\lambda\to+\infty} u(\lambda)\bar{\rho}_{f,A} + \lim_{\lambda\to-\infty} u(\lambda)\rho_{f,A}.$$

It is not difficult to verify that $\int d\nu_{f,A}(\lambda) = 1$ knowing that $\mu_{f,A}(R) = 1 - (\bar{\rho}_{f,A} + \rho_{f,A})$.

Definition 5.5: An observable A is said to have a strong expectation value in the state defined by the vector $f \in {}^*\mathcal{H}$, $\| f \| = 1$ if and only if the identity map $I: R \to R$ defined as $I(\lambda) = \lambda$ for all $\lambda \in R$ is integrable with respect to the set function $\nu_{f,A}$. The strong expectation value of A is then defined to be

$$\overline{\langle A\rangle}_f \equiv \int \lambda\, d\nu_{f,A}(\lambda).$$

Definition 5.6: An observable A is said to have a weak expectation value in the state defined by the vector $f \in {}^*\mathcal{H}$, $\| f \| = 1$ if and only if the identity map $I: R \to R$ defined as $I(\lambda) = \lambda$ for all $\lambda \in R$ is integrable with respect to the measure $\mu_{f,A}$. The weak expectation value of A is then defined to be

$$\langle A\rangle_f \equiv \int \lambda\, d\mu_{f,A}(\lambda).$$

If f is a near standard unit vector, then the following relation is true for every standard self-adjoint operator A:

$$\overline{\langle A\rangle}_f = \langle A\rangle_f = \langle \mathrm{st}(f), A\ \mathrm{st}(f)\rangle.$$

If f is a unit ultra eigenvector of A with ultra eigenvalue $\lambda \in R$, then $\overline{\langle A\rangle}_f = \langle A\rangle_f = \lambda$.

B. Dynamics

The laws of dynamics depend to a great deal on the underlying symmetry. In nonrelativistic quantum mechanics, a physical system develops with time, i.e., a state $\psi(t)$ is given at every instance of time t. The relation between states of the same system at different times is established through a one-parameter unitary group: For every time interval τ there exists a unitary operator U_τ such that $\psi(t+\tau) = U_\tau \psi(t)$. In relativistic quantum mechanics, a state should be defined on a spacelike surface σ. The development of the system takes place as we pass from one spacelike surface to another. In any case, a concept of initial and final states is given. Final states are connected with initial ones through a unitary transformation.

In our approach we postulate the existence of such transformation:

Axiom 6: There exists a standard unitary operator S defined on ${}^*\mathcal{H}$ which maps "initial" states into "final" states.

The S operator introduced in Axiom 6 above, is nothing but the nonstandard extension of the corresponding S operator in the standard theory. Axiom 6 asserts that the dynamics of the nonstandard theory is the nonstandard extension of the dynamics of the standard theory. The nature of this S operator is left unspecified. It may have some labels (denoting a one-parameter group of transformations) or it may be label free, connecting asymptotic states (as in a classical scattering problem). However, we are only concerned with the unitarity of S, from which the following theorems follow.

Theorem 5.4: If $f \overset{{}^*\mathcal{H}}{\longleftrightarrow} g$, then $Sf \overset{{}^*\mathcal{H}}{\longleftrightarrow} Sg$.

Proof: Let $f \overset{{}^*\mathcal{H}}{\longleftrightarrow} g$, then, there exists a number $\varphi \in {}^*R$ and an infinitesimal vector $d \in {}^*\mathcal{H}$ such that $g = e^{i\varphi} f + d$.

From linearity of S we have

$$Sg = e^{i\varphi} Sf + Sd. \tag{5.26}$$

But $\| Sd \|^2 = \langle Sd, Sd\rangle = \langle d, S^* Sd\rangle = \langle d, d\rangle = \| d \|^2$ (due to unitarity of S), hence

$$\| Sd \| \approx 0. \tag{5.27}$$

In general, S (due to unitarity) does not change the normalization of any vector, hence,

$$\| Sf \| = \| Sg \| = 1. \tag{5.28}$$

From (5.26)–(5.28) it follows that $Sf \overset{{}^*\mathcal{H}}{\longleftrightarrow} Sg$. ∎

The significance of this theorem lies in the fact that dynamics is independent of the specific vector chosen to represent a given state. An infinitesimal variation in the initial conditions does not lead to more than an infinitesimal variation in the final conditions. Hence, the solution of any dynamical problem is stable. We prove now the conservation of the whole spectrum, of a self-adjoint operator densely defined on ${}^*\mathcal{H}$ and commuting with S.

Theorem 5.5: Let A be a self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let $[S, A] = 0$. Let $f \in {}^*\mathcal{H}$ be a unit ultra eigenvector of A with ultra eigenvalue $\lambda \in {}^*R$. Then, Sf is a unit ultra eigenvector of A with ultra eigenvalue $\lambda \in {}^*R$.

Proof: Let $d = Af - \lambda f$, hence $\| d \| \approx 0$ by definition ($\| f \| = 1$). Now

$$ASf = SAf = S(\lambda f + d) = \lambda Sf + Sd \tag{5.29}$$

But $\| Sf \| = 1$, $\| Sd \| \approx 0$ (Theorem 5.4), hence Sf is a unit ultra eigenvector of A with ultra eigenvalue $\lambda \in {}^*R$. ∎

In the special case where f is a unit ultra eigenvector of A with ultra eigenvalue $\lambda \in R$, we see that $\mu_{f,A}(\{\lambda\}) = 1$ as a result of Theorem 4.8. Since Sf is also a unit ultra eigenvector of A with ultra eigenvalue $\lambda \in R$ (Theorem 5.5), it follows that $\mu_{Sf,A}(\{\lambda\}) = 1$. Hence, the probability measure is conserved.

Theorem 5.6: Let A be a (standard) self-adjoint operator densely defined on ${}^*\mathcal{H}$. Let $\{f_\lambda : \lambda \in {}^0\sigma(A)\}$ be a family of unit ultra eigenvectors of A indexed by their standard ultra eigenvalues. If $[S, A] = 0$, then: For every $\lambda, \lambda' \in {}^0\sigma(A)$, $\mathrm{st}\langle f_\lambda, Sf_{\lambda'}\rangle = \delta^\lambda_{\lambda'}\ \mathrm{st}\langle f_\lambda, Sf_\lambda\rangle$.

Proof: By Theorem 5.5, $Sf_{\lambda'}$ is an ultra eigenvector of A with ultra eigenvalue $\lambda' \in R$. If $\lambda \neq \lambda', \lambda, \lambda' \in R$, then by Theorem 3.4

$$\mathrm{st}\langle f_\lambda, Sf_{\lambda'}\rangle = 0. \tag{5.30}$$

However, when $\lambda = \lambda'$, we have

$$\mathrm{st}\langle f_\lambda, Sf_{\lambda'}\rangle = \mathrm{st}\langle f_\lambda, Sf_\lambda\rangle . \tag{5.31}$$

(5.30) and (5.31) combine to: For every

$$\lambda, \lambda' \in {}^0\sigma(A),\ \mathrm{st}\langle f_\lambda, Sf_{\lambda'}\rangle = \delta_\lambda^{\lambda'}\ \mathrm{st}\langle f_\lambda, Sf_\lambda\rangle . \qquad \blacksquare$$

Theorem (5.6) proves the conservation of the standard spectrum of a self-adjoint operator commuting with S. Our results are equally valid for both the discrete and the continuous spectra.

6. APPLICATION TO WAVE MECHANICS

In this section we try to construct wavefunctions representing a nonrelativistic free particle. In the standard theory, the wavefunction of a free particle of definite momentum turns out to be a function which is not square integrable in the position representations or a sort of delta function in the momentum representation. As a way out of this situation, Dirac suggested the use of a function $\delta^{(n)}(x)$ defined on an n-dimensional Euclidean space, such that, for every complex valued function ϕ, defined on the n-dimensional Euclidean space R^n, and belonging to the class Φ of good functions, the following conditions are satisfied:

(D1) $\int d^n x\, \delta^{(n)}(x)\, \varphi(x) = \varphi(0)$;

(D2) $$\int d^n x \left(\frac{\partial^{P_1}}{\partial x_1^{P_1}} \frac{\partial^{P_2}}{\partial x_2^{P_2}} \cdots \frac{\partial^{P_n}}{\partial x_n^{P_n}} \delta^{(n)}(x)\right) \varphi(x) = (-)^{P_1+P_2+\cdots+P_n} \int d^n x\, \delta^{(n)}(x) \frac{\partial^{P_1}}{\partial x_1^{P_1}} \frac{\partial^{P_2}}{\partial x_1^{P_1}} \cdots \frac{\partial^{P_n}}{\partial x_n^{P_n}} \varphi(x).$$

(A good function $\varphi: R^n \to C$ is a C^∞-function such that it and all its derivative vanish faster than $\|x\|^{-m}$ for every $m \in N$ as $\|x\| \to +\infty$, $\|x\| \equiv \sqrt{(x_1^2 + x_1^2 + \cdots + x_n^2)}$ $x \in R^n$)

It is well known that no such function exists. The distributions theory proves the existence of a linear functional δ defined on Φ and continuous with respect to an appropriate topology on Φ, such that

$$\delta(\varphi) = \varphi(0), \tag{6.1}$$

and the differentiation of a distribution T is defined in general as

$$(D_1^{P_1} D_2^{P_2} \cdots D_n^{P_n} T)(\varphi) = (-1)^{P_1+P_2+\cdots+P_n} T(D_1^{P_1} D_2^{P_2} \cdots D_n^{P_n} \varphi), \tag{6.2}$$

where

$$D_i^{P_i} = \frac{\partial^{P_i}}{\partial x_i^{P_i}}, \qquad i = 1, 2, \ldots, n.$$

But the δ distribution defined above is not a complex valued function defined on R^n. Hence the expression $\delta^{(n)}(x - x')$ which is used extensively in quantum mechanics and quantum field theory is meaningless (no one can tell what $\delta^{(n)}_{(0)}$, for example, means). The expression

$$\int d^n x'\, \delta^{(n)}(x - x')\, \varphi(x')$$

is just a "bad" way of writing $\delta_x^{(n)}(\varphi)$. It is even worse when the integration sign is not there.

We construct first a nonstandard function defined on the nonstandard n-dimensional Euclidean space ${}^*R^n$. We shall give it the symbol $\delta^{(n)}$. If φ is a complex-valued standard function defined on ${}^*R^n$ and continuous at the origin, then $\delta^{(n)}$ has the property that

$$\mathrm{st} \int d^n x\, \delta^{(n)}(x)\, \varphi(x) = \varphi(0). \tag{6.3}$$

Furthermore, if φ has all the derivatives required in D2, then D2 holds exactly.

We shall use this function to construct the momentum wavefunctions of a nonrelativistic particle, and show how we can reproduce a great deal of Dirac's notation for the continuous spectrum.

A. Construction of $\delta^{(n)}$

It is known that if $\{f_n : n \in N\}$ is a sequence of functions defined on R^n and ranging in C, such that for every good function $\varphi: R^n \to C$ (where a good function is a C^∞ function, such that it and all its derivatives vanish faster than $\|x\|^{-m}$ for every $m \in N$ as $\|x\| \to +\infty$, $\|x\| = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$, $x \in R^n$), the following condition holds:

$$\lim_{m \in N} \int d^n x\, f_m(x)\, \varphi(x) = \varphi(0) \tag{6.4}$$

and, if δ is the internal function, defined on ${}^*R^n$ and ranging in ${}^*C$, which is generated by the sequence $\{f_n : n \in N\}$, then for every standard good function φ: ${}^*R^n \to {}^*C$ the following condition holds:

$$\mathrm{st} \int d^n x\, \delta(x)\, \varphi(x) = \varphi(0). \tag{6.5}$$

We want to construct here a function $\delta^{(n)}$ such that (6.5) holds for a wider class of functions (not necessarily good functions only). In addition to the mathematical requirement that $\delta^{(n)}$ be at least a fairly good function in order to justify (D2) (infinite differentiability) and (6.5) (integrability with good functions), $\delta^{(n)}$, from a physical point of view, must be a good function, in order to justify taking its Fourier transform (when going from X space to P space, etc). It is very useful to impose the extra condition on $\delta^{(n)}$ of having a compact (infinitesimal) support (in the internal sense) in the neighborhood of the origin, a fact which is used in proving many theorems in this section. Any good (real nonnegative even) function with a compact infinitesimal support (in the internal sense) in the neighborhood of the origin, and normalized such that its integral is unity, would serve the purpose. However, we construct here a specific one, but we emphasize that our choice has no effect on the proofs of the following theorems. All theorems remain valid if we replace the δ function adopted here by any other member of its equivalence class mentioned above.

Definition 6.1: Let $x \in {}^*R^n$, $x = (x_1, x_2, \ldots, x_n)$, $n \in N$ (a finite natural number). Define $\|x\|_n = \sum_{i=1}^n |x_i|^2)^{1/2}$.

Definition 6.2: Let $\lambda(x) = \theta(x)\, e^{-1/x}$, $x \in {}^*R$, where $\theta(x)$ is the Heaviside step function extended to ${}^*R$ [$\theta(x) = 1$ if $x > 0$, $\theta(x) = 0$ if $x < 0$]. Define:

(i) $$I_n(c) = \int_0^c \frac{e^{-(c-x)^{-1}}}{e^{-(c-x)^{-1}} + e^{-1/x}}\, x^{n-1}\, dx$$

for $n \in N$, $c \in {}^*R$, $c > 0$;

(ii) $$\Delta_c^{(n)}(x) = \frac{\Gamma(\frac{1}{2}n)}{2[\Gamma(\frac{1}{2})]^n I_n(c)} \cdot \frac{\lambda(c - \|x\|_n)}{\lambda(c - \|x\|_n) + \lambda(\|x\|_n)}$$

for $n \in N$, $c \in {}^*R$, $c > 0$, $x \in {}^*R^n$.

Theorem 6.1: The function $\Delta_c^{(n)}$ defined in Def. 6.2 has the properties:

(i) $\| x \|_n \geqslant c \Rightarrow \Delta_c^{(n)}(x) = 0$;

(ii) it is a C^∞ function;

(iii) $\int_{x \in {}^*R^n} d^n x \, \Delta_c^{(n)}(x) = 1$.

Proof: To prove (i), we notice that $\| x \|_n > c$ implies that $\theta(c - \| x \|_n) = 0$, from which it follows that $\lambda(c - \| x \|_n) = 0$, hence

$$\| x \|_n > c \Rightarrow \Delta_c^{(n)}(x) = 0 \tag{6.6}$$

Also $\lim e^{-(c-\| x \|_n)^{-1}} = 0$ as $\| x \|_n \to c$ and $\| x \|_n < c$, hence $\lim \lambda(c - \| x \|_n) = 0$ as $\| x \|_n \to c$ and $\| x \|_n < c$, and

$$\| x \|_n = c \Rightarrow \Delta_c^{(n)}(x) = 0. \tag{6.7}$$

(6.6) and (6.7) proves (i).

To prove (ii) we notice that $\Delta_c^{(n)}$ is analytic in the domain $D_1 = \{x: x \in {}^*R^n \text{ and } 0 < \| x \|_n < c\}$ and the domain $D_2 = \{x: x \in {}^*R^n \text{ and } \| x \|_n > c\}$. Hence it is C^∞ in D_1 and D_2.

Let x be a point on the surface $E = \{x': x' \in {}^*R^n \text{ and } \| x' \|_n = c\}$. We have shown that

$$\lim_{\substack{x' \to x \\ x' \in D_1}} \Delta_c^{(n)}(x') = \lim_{\substack{x' \to x \\ x' \in D_2}} \Delta_c^{(n)}(x') = 0, \tag{6.8}$$

which guarantees continuity across E.

Let $(P_1, P_2, \ldots, P_n) \in {}^*N^n$ with $P = P_1 + P_2 + \cdots + P_n$. It can be shown by mathematical induction that for every $x \in D_1$,

$$\frac{\partial^{P_1}}{\partial x_1^{P_1}} \frac{\partial^{P_2}}{\partial x_2^{P_2}} \cdots \frac{\partial^{P_n}}{\partial x_n^{P_n}} \Delta_c^{(n)}(x) = \sum_{\alpha,\beta=1}^{P} \sum_{j=2}^{2P} \sum_{K=3}^{3P} \sum_{r=2}^{P+1} \times \frac{Q^{\alpha,\beta,j,K,r}_{(P_1,P_2,\ldots,P_n)}(x')}{c^K e^{-(r-\beta)/c}} \frac{e^{-\alpha/(c-\| x \|_n)} e^{-\beta/(\| x \|_n)}}{(e^{-(c-\| x \|_n)^{-1}} + e^{-(\| x \|_n)^{-1}})^r} \tag{6.9}$$

where $Q^{\alpha,\beta,j,K,r}_{(P_1,P_2,\ldots,P_n)}(x)$ are polynomials in the components $(x_1, x_2, \ldots, x_n)$. Now, let $x' \in E$, hence

$$\lim_{\substack{x \to x' \\ x \in D_1}} \frac{\partial^{P_1}}{\partial x_1^{P_1}} \frac{\partial^{P_2}}{\partial x_2^{P_2}} \cdots \frac{\partial^{P_n}}{\partial x_n^{P_n}} \Delta_c^{(n)}(x) = \sum_{\alpha,\beta=1}^{P} \sum_{j=2}^{2P} \sum_{K=3}^{3P} \sum_{r=2}^{P+1} \frac{Q^{\alpha,\beta,j,K,r}_{(P_1,P_2,\ldots,P_n)}(x')}{c^K e^{-(r-\beta)/c}} \lim_{\substack{\| x \|_n \to c \\ \| x \|_n < c}} \frac{e^{-\alpha/(c-\| x \|_n)}}{(c - \| x \|_n)^j} = 0.$$

Hence

$$\lim_{\substack{x \to x' \\ x \in D_1}} \frac{\partial^{P_1}}{\partial x_1^{P_1}} \frac{\partial^{P_2}}{\partial x_2^{P_2}} \cdots \frac{\partial^{P_n}}{\partial x_n^{P_n}} \Delta_c^{(n)}(x) = \lim_{\substack{x \to x' \\ x \in D_2}} \times \frac{\partial^{P_1}}{\partial x_1^{P_1}} \frac{\partial^{P_2}}{\partial x_2^{P_2}} \cdots \frac{\partial^{P_n}}{\partial x_n^{P_n}} \Delta_c^{(n)}(x) = 0. \tag{6.10}$$

(6.8) and (6.10) prove that $\Delta_c^{(n)}$ is C^∞ across E. At the origin

$$\lim_{\substack{x \to 0 \\ x \in D_1}} \Delta_c^{(n)}(x) = \frac{\Gamma(\frac{1}{2}n)}{2[\Gamma(\frac{1}{2})]^n I_n(c)} \lim_{\substack{x \to 0 \\ x \in D_1}} \frac{e^{-(c-\| x \|_n)^{-1}}}{e^{-(c-\| x \|_n)^{-1}} + e^{-(\| x \|_n)^{-1}}}$$

$$= \frac{\Gamma(\frac{1}{2}n)}{2[\Gamma(\frac{1}{2})]^n I_n(c)}. \tag{6.11}$$

Hence $\Delta_c^{(n)}(x)$ is continuous at the origin. Also

$$\lim_{\substack{x \to 0 \\ x \in D_1}} \frac{\partial^{P_1}}{\partial x_1^{P_1}} \frac{\partial^{P_2}}{\partial x_2^{P_2}} \cdots \frac{\partial^{P_n}}{\partial x_n^{P_n}} \Delta_c^{(n)}(x) = \sum_{\alpha,\beta=1}^{P} \sum_{j=2}^{2P} \sum_{K=3}^{3P} \sum_{r=2}^{P+1} \frac{Q^{\alpha,\beta,j,K,r}_{(P_1,P_2,\ldots,P_n)}(0)}{c^j e^{-r-x/c}} \lim_{\substack{\| x \|_n \to 0 \\ \| x \|_n > 0}} \frac{e^{-\beta/\| x \|_n}}{(\| x \|_n)^K} = 0.$$

Hence

$$\lim_{\substack{x \to 0 \\ x \in D_1}} \frac{\partial^{P_1}}{\partial x_1^{P_1}} \frac{\partial^{P_2}}{\partial x_2^{P_2}} \cdots \frac{\partial^{P_n}}{\partial x_n^{P_n}} \Delta_c^{(n)}(x) = 0. \tag{6.12}$$

(6.11) and (6.12) prove that $\Delta_c^{(n)}$ is C^∞ at the origin. This completes the proof of (ii).

To prove (iii) we have

$$\int_{x \in {}^*R^n} d^n x \, \Delta_c^{(n)}(x) = \int_{\| x \|_n \leq c} d^n x \, \Delta_c^{(n)}(x)$$

$$= \frac{\Gamma(\frac{1}{2}n)}{2[\Gamma(\frac{1}{2})]^n I_n(c)} \int_{\| x \|_n \leq c} d^n x \, \frac{e^{-(c-\| x \|_n)^{-1}}}{e^{-(c-\| x \|_n)^{-1}} + e^{-(\| x \|_n)^{-1}}}$$

$$= \frac{\Gamma(\frac{1}{2}n)}{2[\Gamma(\frac{1}{2})]^n I_n(c)} \cdot \frac{2[\Gamma(\frac{1}{2})]^n}{\Gamma(\frac{1}{2}n)} \int_0^c \frac{e^{-(c-x)^{-1}}}{e^{-(c-x)^{-1}} + e^{-x^{-1}}} x^{n-1} \, dx.$$

$$= 1.$$

Hence the theorem is proved. ■

The number c in $\Delta_c^{(n)}$ need only be positive, but, otherwise, arbitrary. We choose some infinitesimal number $\epsilon > 0$ and set $c = \epsilon$ (but leave ϵ unspecified.)

Definition 6.3: The function $\delta^{(n)}: {}^*R^n \to {}^*R$ is defined as $\delta^{(n)} = \Delta_\epsilon^{(n)}$ for some $\epsilon > 0$, $\epsilon \approx 0$.

Theorem 6.2: Let $\varphi: {}^*R^n \to {}^*C$ be a standard function continuous at the origin. Then

$$\text{st} \int_{x \in {}^*R^n} d^n x \, \delta^{(n)}(x) \, \varphi(x) = \varphi(0).$$

Proof:

$$\left| \int_{x \in {}^*R^n} d^n x \, \delta^{(n)}(x) [\varphi(x) - \varphi(0)] \right|$$

$$= \left| \int_{\| x \|_n \leq \epsilon} d^n x \, \delta^{(n)}(x) [\varphi(x) - \varphi(0)] \right|$$

$$\leqslant \left| \int_{\| x \|_n < \epsilon} d^n x \, \delta^{(n)}(x) \, | \, \varphi(x) - \varphi(0) \, | \right|. \tag{6.13}$$

But if φ is standard and continuous at $a \in R^n$ (a standard point), then for every $x \in {}^*R^n$, $\| x - a \|_n \approx 0$ implies that[7]

$$\varphi(x) - \varphi(a) \approx 0.$$

Hence, setting $a = 0$ we get

$$\| x \|_n \leqslant \epsilon \quad \text{implies} \quad \varphi(x) - \varphi(0) \approx 0 \tag{6.14}$$

because ϵ is infinitesimal.

Let $\alpha > 0$, $\alpha \in R$ (a standard positive number), then, from (6.14)

$$\| x \|_n \leqslant \epsilon \quad \text{implies} \quad | \, \varphi(x) - \varphi(0) \, | < \alpha. \tag{6.15}$$

Now, (6.13) and (6.15) imply

$$\left| \int_{x \in {}^*R^n} d^n x \, \delta^{(n)}(x) [\varphi(x) - \varphi(0)] \right| \leqslant \alpha \int_{\| x \|_n \leq \epsilon} \times d^n x \, \delta^{(n)}(x) = \alpha.$$

Hence, for all $\alpha > 0$, $\alpha \in R$, we have

$$\left| \int_{x \in {}^*R^n} d^n x\, \delta^{(n)}(x)\,[\varphi(x) - \varphi(0)] \right| < \alpha ,$$

i.e.,

$$\int_{x \in {}^*R^n} d^n x\, \delta^{(n)}(x)\,[\varphi(x) - \varphi(0)] \in M_1(C) \qquad (6.16)$$

or

$$\int_{x \in {}^*R^n} d^n x\, \delta^{(n)}(x)\,\varphi(x) - \varphi(0) \in M_1(C).$$

But $\varphi(0) \in C$, hence (from the definition of the standard part of a number) we get

$$\mathrm{st} \int_{x \in {}^*R^n} d^n x\, \delta^{(n)}(x)\,\varphi(x) = \varphi(0). \quad \blacksquare$$

Theorem 6.3: Let φ: ${}^*R^n \to {}^*C$ be a standard function such that all its derivatives up to $D_1^{P_1} D_2^{P_2} \cdots D_n^{P_n} \varphi$ exist and are continuous ($D_i^{P_i} = \partial^{P_i} / \partial x_i^{P_i}$, $i = 1, 2, \ldots, n$), then

$$\mathrm{st} \int_{x \in {}^*R^n} d^n x \left(\frac{\partial^{P_1}}{\partial x_1^{P_1}} \frac{\partial^{P_2}}{\partial x_2^{P_2}} \cdots \frac{\partial^{P_n}}{\partial x_n^{P_n}} \delta_{(x)}^{(n)} \right) \varphi(x)$$

$$= (-1)^{P_1 + P_2 + \cdots + P_n} D_1^{P_1} D_2^{P_2} \cdots D_n^{P_n}\, \varphi(0).$$

Proof: First, notice that if $P_1' < P_1, P_2' \leq P_2, \ldots, P_n' \leq P_n$, and that if for every ψ: ${}^*R^n \to {}^*C$, for which $D_1^{P_1'} D_2^{P_2'} \cdots D_n^{P_n'} \psi$ exists and is continuous, we have

$$\int_{x \in {}^*R^n} d^n x \left(\frac{\partial^{P_1'}}{\partial x_1^{P_1'}} \frac{\partial^{P_2'}}{\partial x_2^{P_2'}} \cdots \frac{\partial^{P_n'}}{\partial x_n^{P_n'}} \delta_{(x)}^{(n)} \right)$$

$$\times \psi(x) = (-1)^{P_1' + P_2' + \cdots + P_n'}$$

$$\times \int_{x \in {}^*R^n} d^n x\, \delta^{(n)}(x) \frac{\partial^{P_1'}}{\partial x_1^{P_1'}} \frac{\partial^{P_2'}}{\partial x_2^{P_2'}} \cdots \frac{\partial^{P_n'}}{\partial x_n^{P_n'}} \psi(x), \qquad (6.17)$$

then

$$\int_{x \in {}^*R^n} d^n x \left(\frac{\partial^{P_1+1}}{\partial x_1^{P_1'+1}} \frac{\partial^{P_2'}}{\partial x_2^{P_2'}} \cdots \frac{\partial^{P_2'}}{\partial x^{P_n'}} \right) \delta^{(n)}(x) \quad \varphi(x)$$

$$= \int_{x \in {}^*R^n} d^n x \frac{\partial}{\partial x_1} \left(\frac{\partial^{P_1'}}{\partial x_1^{P_1'}} \frac{\partial^{P_2'}}{\partial x_2^{P_2'}} \cdots \frac{\partial^{P_n'}}{\partial x_n^{P_n'}} \delta^{(n)}(x) \right) \varphi(x)$$

$$= \int_{-\infty}^{+\infty} dx_2 \cdots \int_{-\infty}^{+\infty} dx_n \int_{-\infty}^{+\infty} dx_1 \frac{\partial}{\partial x_1}$$

$$\times \left(\frac{\partial^{P_1'}}{\partial x_1^{P_1'}} \frac{\partial^{P_2'}}{\partial x_2^{P_2'}} \cdots \frac{\partial^{P_n'}}{\partial x_n^{P_2'}} \delta^{(n)}(x) \right) \varphi(x)$$

$$= \int_{-\infty}^{+\infty} dx_2 \cdots \int_{-\infty}^{+\infty} dx_n \left[\left(\frac{\partial^{P_1'}}{\partial x_1^{P_1'}} \frac{\partial^{P_2'}}{\partial x_2^{P_2'}} \cdots \frac{\partial^{P_n'}}{\partial x_n^{P_2'}} \delta^{(n)}(x) \right) \varphi(x) \right]_{-\infty}^{+\infty}$$

$$- \int_{-\infty}^{+\infty} dx_2 \cdots \int_{-\infty}^{+\infty} dx_n \int_{-\infty}^{+\infty} dx_1$$

$$\times \left(\frac{\partial^{P_1'}}{\partial x_1^{P_1'}} \frac{\partial^{P_2'}}{\partial x_2^{P_2'}} \cdots \frac{\partial^{P_n'}}{\partial x_n^{P_n'}} \delta^{(n)}(x) \right) \frac{\partial \varphi(x)}{\partial x_1}$$

$$= - \int_{x \in {}^*R^n} d^n x \left(\frac{\partial^{P_1'}}{\partial x_1^{P_1'}} \frac{\partial^{P_2'}}{\partial x_2^{P_2'}} \cdots \frac{\partial^{P_n'}}{\partial x_n^{P'}} \delta^{(n)}(x) \right)$$

$$\times \frac{\partial \varphi(x)}{\partial x_1}. \qquad (6.18)$$

But $D_1^{P_1'} D_2^{P_2'} \cdots D_n^{P_n'}(D, \varphi)$ exists and is continuous (given), hence, by applying (6.17) to (6.18) we get

$$\int_{x \in {}^*R^n} d^n x \left(\frac{\partial^{P_1'+1}}{\partial x_1^{P_1'+1}} \frac{\partial^{P_2'}}{\partial x_2^{P'}} \cdots \frac{\partial^{P_n'}}{\partial x_n^{P_n'}} \delta_{(x)}^{(n)} \right) \varphi(x)$$

$$= - (-1)^{P_1' + P_2' + \cdots + P_n'} \int_{x \in {}^*R^n} d^n x\, \delta^{(n)}(x)$$

$$\times \frac{\partial^{P_1'+1}}{\partial x_1^{P_1'+1}} \frac{\partial^{P_2'}}{\partial x_2^{P_2'}} \cdots \frac{\partial^{P_n'}}{\partial x_n^{P_n'}} \varphi(x)$$

$$= (-1)^{(P_1'+1) + P_2' + \cdots + P_n'} \int_{x \in {}^*R^n} d^n x\, \delta^{(n)}(x)$$

$$\times \frac{\partial^{P_1'+1}}{\partial x_1^{P_1'+1}} \frac{\partial^{P_2'}}{\partial x_2^{P_2'}} \cdots \frac{\partial^{P_n'}}{\partial x_n^{P_n'}} \varphi(x). \qquad (6.19)$$

Equation (6.17) is true if we set $\psi = \varphi$ and $P_1' = P_2' = \cdots = P_n' = 0$. Now, setting $\psi = \varphi$ in Eq. (6.17) and assuming that it is true, Eq. (6.19) follows. So, by mathematical induction we get

$$\int_{x \in {}^*R^n} d^n x \left(\frac{\partial^{P_1}}{\partial x_1} \frac{\partial^{P_2}}{\partial x_2} \cdots \frac{\partial^{P_n}}{\partial x_n} \delta_{(x)}^{(n)} \right) \varphi(x) = (-)^{P_1 + P_2 + \cdots + P_n}$$

$$\times \int_{x \in {}^*R^n} d^n x\, \delta_{(x)}^{(n)} \left(\frac{\partial^{P_1}}{\partial x_1^{P_1}} \frac{\partial^{P_2}}{\partial x_2^{P_2}} \cdots \frac{\partial^{P_n}}{\partial x_n^{P_n}} \varphi(x) \right). \qquad (6.20)$$

By applying Theorem 6.2 to (6.20), we get finally

$$\mathrm{st} \int_{x \in {}^*R^n} d^n x \left(\frac{\partial^{P_1}}{\partial x_1^{P_1}} \frac{\partial^{P_2}}{\partial x_2^{P_2}} \cdots \frac{\partial^{P_n}}{\partial x_n^{P_n}} \delta_{(x)}^{(n)} \right)$$

$$\times \varphi(x) = (-)^{P_1 + P_2 + \cdots + P_n} D_1^{P_1} D_2^{P_2} \cdots D_n^{P_n}\, \varphi(0). \quad \blacksquare$$

Higher orders of $\delta^{(n)}$

We shall call $\delta^{(n)}$, the n-dimensional δ function of the first order. We define now n-dimensional δ functions of higher orders.

Definition 6.4: The n-dimensional δ function of the p order is the function $\delta_p^{(n)}$: ${}^*R^n \to {}^*R$ defined as follows:

$$\delta_p^{(n)}(x) \equiv \int_{x_1 \in {}^*R^n} d^n x_1 \int_{x_2 \in {}^*R^n} dx_2^n \cdots \int_{x_{p-1} \in {}^*R^n} d^n x_{p-1}$$

$$\times \delta^{(n)}(x - x_1)\, \delta^{(n)}(x_1 - x_2) \cdots \delta^{(n)}(x_{p-2} - x_{p-1})$$

$$\times \delta^{(n)}(x_{p-1}).$$

Theorem 6.4: The n-dimensional δ function of the p order satisfies the following properties:

(i) If $x \in {}^*R^n$, $\| x \|_n \geq p\epsilon$, then $\delta_p^{(n)}(x) = 0$;

(ii) $\int_{x \in {}^*R^n} d^n x\, \delta_p^{(n)}(x) = 1$;

(iii) If φ: ${}^*R^n \to {}^*C$ is a standard function which is continuous at the origin, then

$$\mathrm{st} \int_{x \in {}^*R^n} d^n x\, \delta_p^{(n)}(x)\,\varphi(x) = \varphi(0).$$

Proof: To prove property (i), consider the function

$$f(x, x_1, \ldots, x_{p-1}) = \delta^{(n)}(x - x_1)\, \delta^{(n)}(x_1 - x_2)$$

$$\cdots \delta^{(n)}(x_{p-2} - x_{p-1})\, \delta^{(n)}(x_{p-1}). \qquad (6.21)$$

Let $\| x \|_n \geq p\epsilon$. If $\| x_1 \|_n \leq (p-1)\epsilon$, then $\| x - x_1 \|_n \geq | \| x \|_n - \| x_1 \|_n | \geq \epsilon$ which leads to $\delta^{(n)}(x - x_1) = 0$ and, hence, $f(x, x_1, \ldots, x_{p-1}) = 0$. So, in order that $f(x, x_1, \ldots, x_{p-1})$ does not vanish we should have $\| x_1 \|_n > (p-1)\epsilon$. If we continue in this fashion, we find that a necessary condition for the function f to be different from zero is that

$$\| x_1 \|_n > (p-1)\epsilon, \| x_2 \|_n > (p-2)\epsilon, \ldots, \| x_{p-1} \|_n > \epsilon. \tag{6.22}$$

But, if $\| x_{p-1} \|_n > \epsilon$, then $\delta^{(n)}(x_{p-1}) = 0$ and f vanishes again. Hence $\| x \|_n \geq p\epsilon$ implies that

$$f(x, x_1, \ldots, x_{p-1}) = 0. \tag{6.23}$$

But

$$\delta_p^{(n)}(x) = \int_{x_1 \in {}^*R^n} d^n x_1 \int_{x_2 \in {}^*R^n} d^n x_2 \cdots \int_{x_{p-1} \in {}^*R^n} \times d^n x_{p-1} f(x, x_1, x_2, \ldots, x_{p-1}). \tag{6.24}$$

Therefore, by using (6.23) in (6.24) we get

$$\| x \|_n \geq p\epsilon \quad \text{implies that} \quad \delta_p^{(n)}(x) = 0. \tag{6.25}$$

To prove properly (ii) we notice that

$$\delta_{p+1}^{(n)}(x) = \int_{x' \in {}^*R^n} d^n x' \, \delta^{(n)}(x - x') \, \delta_p^{(n)}(x'), \tag{6.26}$$

thus

$$\int_{x \in {}^*R^n} d^n x \, \delta_{p+1}^{(n)}(x) = \int_{x \in {}^*R^n} d^n x \int_{x' \in {}^*R^n} d^n x' \delta^{(n)}(x - x') \, \delta_p^{(n)}(x') = \int_{x \in {}^*R^n} d^n x \, \delta_p^{(n)}(x). \tag{6.27}$$

Equation (6.2), together with the fact that $\int_{x \in {}^*R^n} d^n x \, \delta^{(n)}(x) = 1$ ($\delta_1^{(n)} = \delta^{(n)}$) implies property (ii).

Finally, to prove (iii), we use (i) and (ii), and apply Theorem 6.3. This completes the proof of the theorem. ∎

Although the n-dimensional δ functions of all orders satisfy the "definition" of the δ function in the conventional sense, yet they are different functions. They coincide on the whole nonstandard real line, save for an infinitesimal neighborhood of zero.

B. Nonrelativistic particles

The nonstandard Hilbert space of a nonrelativistic particle (in the momentum representation) is the set ${}^*\mathcal{H}$ of equivalence classes of square integrable internal complex-valued functions defined on ${}^*R^3$. If $f, g \in {}^*\mathcal{H}$, then, their scalar product is defined to be

$$\langle f, g \rangle = \int_{p \in {}^*R^3} d^3 p \, f^*(p) g(p) = (2\pi)^{-3} \int_{p \in {}^*R^3} d^3 p \, f^*(p) g(p).$$

The momentum operator P_i in the ith direction is the standard operator whose domain is the set of vectors

$$D(P_i) \equiv \{ f : f \in {}^*\mathcal{H} \text{ and } \int d^3 p \, P_i^2 \, | f(p) |^2 < +\infty \}$$

and which maps $f \in D(P_i)$ into $P_i f$ defined as follows:

$$(P_i f)(p') = p_i' f(p').$$

The position operator X_i in the ith direction is the standard operator whose domain is the set of vectors

$$D(X_i) \equiv \{ f : f \in {}^*\mathcal{H} \text{ and } D_i f \text{ exists and } \int d^3 p \times | \partial / \partial p_i \, f(p) |^2 < +\infty \}$$

and which maps $f \in D(X_i)$ into $X_i f$ defined as follows:

$$(X_i f)(p') = i \frac{\partial}{\partial p_i'} f(p').$$

Each of the six operators $\{P_1, P_2, P_3, X_1, X_2, X_3\}$ is self-adjoint. One can easily verify that on a common dense domain:

$$[X_i, X_j] = 0, \quad [P_i, P_j] = 0, \quad [X_i, P_j] = i\delta_{ij}.$$

Definition 6.5: Let $\Delta \subset {}^*R$ be a Borel subset. The operator $E_\Delta(P_i)$ $(i = 1, 2, 3)$ is defined on ${}^*\mathcal{H}$ as follows: For every $f \in {}^*\mathcal{H}$,

$$(E_\Delta(P_i) f)(p') = \chi_\Delta(p_i') f(p'),$$

where χ_Δ is the characteristic function of Δ.

Definition 6.6: Let $\Delta \subset {}^*R$ be a Borel subset. The operator $E_\Delta(X_i)$ $(i = 1, 2, 3)$ is defined on a dense domain in ${}^*\mathcal{H}$ as follows: For every $f \in {}^*\mathcal{H}$ for which the function

$$\tilde{f}_p(x) = \int_{p' \in {}^*R^3} d^3 p' \, e^{i(\mathbf{p}' - \mathbf{p}) \cdot \mathbf{x}} f(p') \quad \text{for every } p \in {}^*R^3$$

exists and is integrable, we define

$$(E_\Delta(X_i) f)(p) = \int_{x_i \in \Delta} dx_i \prod_{j \neq i} \int_{x_j \in {}^*R} dx_j \int_{p' \in {}^*R^3} d^3 p' e^{i(\mathbf{p}' - \mathbf{p}) \cdot \mathbf{x}} f(p').$$

Then $E_\Delta(X_i)$ is extended to ${}^*\mathcal{H}$ by continuity.

It is easy to verify that $E(P_i)$, $E(X_i)$ $(i = 1, 2, 3)$ are the nonstandard spectral measures associated with P_i and X_i $(i = 1, 2, 3)$.

Definition 6.7: Define the vectors $f_p(p'), f_x(x') \in {}^*\mathcal{H}$ as follows:

$$(f_p(p'))(p'') = \delta^{(3)}(p'' - p') = (2\pi)^3 \delta^{(3)}(p'' - p'),$$

$$(f_x(x'))(p'') = \int_{x'' \in {}^*R^3} d^3 x'' \, e^{-i\mathbf{p}'' \cdot \mathbf{x}''} \delta^{(3)}(x'' - x').$$

Theorem 6.5:

(i) $f_p(p')$ is an ultra eigenvector of (P_1, P_2, P_3) with ultra eigenvalues (p_1', p_2', p_3').

(ii) $f_x(x')$ is an ultra eigenvector of (X_1, X_2, X_3) with ultra eigenvalues (x_1', x_2', x_3').

Proof:

$$\| (P_i - p_i') f_p(p') \|^2 = \int_{p'' \in {}^*R^3} d^3 p'' (p_i'' - p_i')^2 \, | \delta^{(3)}(p'' - p') |^2; \tag{6.28}$$

but if $(p_i'' - p_i')^2 \geq \epsilon^2 (\epsilon \simeq 0)$, then $\delta^{(3)}(p'' - p') = 0$. Hence Eq. (6.28) becomes

$$\| (P_i - p_i') f_p(p') \|^2 \leq \epsilon^2 \int_{p'' \in {}^*R^3} d^3 p'' \, | \delta^{(3)}(p'' - p') |^2 = \epsilon^2 \| f_p(p') \|^2 \tag{6.29}$$

i.e., $\| (P_i - p_i') f_p(p') \| \leq \epsilon \| f_p(p') \|$ or

$$\| (P_i - p_i') f_p(p') \| / \| f_p(p') \| \approx 0. \tag{6.30}$$

This proves (i).

To prove (ii), we proceed as follows:

$$\begin{aligned} \| (X_i - x_i') f_x(x') \|^2 &= \int_{p'' \in {}^*R^3} d^3 p'' \int_{x''' \in {}^*R^3} d^3 x''' (x_i''' - x_i') e^{i\mathbf{p}'' \cdot x'''} \\ &\quad \times \delta^{(3)}(x''' - x') \int_{x'' \in {}^*R^3} d^3 x'' (x_i'' - x_i') e^{i\mathbf{p}'' \cdot \mathbf{x}''} \delta^{(3)}(x'' - x') \\ &= \int_{p'' \in {}^*R^3} d^3 p'' \int_{x''' \in {}^*R^3} d^3 x''' \int_{x'' \in {}^*R^3} d^3 x'' \, e^{i\mathbf{p}'' \cdot (\mathbf{x}''' - \mathbf{x}'')} \\ &\quad \times (x_i''' - x_i') (x_i'' - x_i') \, \delta^{(3)}_{(x''' - x')} \, \delta^{(3)}_{(x'' - x')} \\ &= \int_{x'' \in {}^*R^3} d^3 x'' (x_i'' - x_i')^2 \, | \, \delta^{(3)}(x'' - x') \, |^2. \end{aligned} \tag{6.31}$$

Now

$$\| f_X(x')\|^2 = \int_{x''\in {}^*R^3} d^2x'' \mid \delta^{(3)}(x''-x') \mid^2. \quad (6.32)$$

In the same way as we concluded (6.30) from (6.28) and (6.29), we get from (6.31) and (6.32):

$$\|(X_i - x_i')f_X(x')\| \;/\; \| f_X(x')\| \approx 0. \quad (6.33)$$

This completes the proof of the theorem. ■

Eigenfunctionals

(See Sec. 2D; related standard linear forms, and Sec. 3B; relation with eigenfunctionals).

Let ${}^0D$ be the set of standard good functions (${}^0D \subset \mathcal{K}$) and let D be the standard subset of ${}^*\mathcal{K}$ generated by ${}^0D$. It is well known that D is dense in ${}^*\mathcal{K}$, because, from the standard theory, ${}^0D$ is dense in $\mathcal{K}$. Let ξ_P, ξ_X: ${}^*R^3 \times D \to {}^*C$ be the *standard* functions defined as follows: If $\lambda \in R^3$ (a standard triple) and $g \in {}^0D$ (a standard vector), then

$$\xi_P(\lambda)(g) = \mathrm{st}\langle f_P(\lambda), g\rangle, \qquad \xi_X(\lambda)(g) = \mathrm{st}\langle f_X(\lambda), g\rangle$$

the existence of ξ_P follows from the fact that

$$\mathrm{st}\langle f_P(\lambda), g\rangle = \mathrm{st}\int d^3p''\, \check{\delta}^{(3)}(p''-\lambda) g(p'') = g(\lambda)$$

since g is continuous at $\lambda \in R^3$ (Theorem 6.2). The existence of ξ_X follows from the fact that

$$\begin{aligned}\mathrm{st}\langle f_X(\lambda), g\rangle &= \mathrm{st}\int d^3p'' \int d^3x''\, e^{i\mathbf{p}''\cdot\mathbf{x}''}\delta^{(3)}(x''-\lambda) g(p'') \\ &= \mathrm{st}\int d^3x''\, \delta^{(3)}(x''-\lambda)\int d^3p''\, e^{i\mathbf{p}''\cdot\mathbf{x}''} g(p'') \\ &= \int d^3p''\, e^{i\mathbf{p}''\cdot\mathbf{x}}\, g(p'')\end{aligned}$$

since g is a good function.

The uniqueness of ξ_P and ξ_X follows from the fact that they are standard.

Theorem 6.6: Let $f \in D$ (where D is the subspace defined above). Then

$$\xi_P(p')(P_i f) = p_i' \xi_P(p')(f) \quad \text{and} \quad \xi_X(x')(X_i f) = x_i' \xi_X(x')(f).$$

Proof: Let $p' \in R^3$ (a standard triple) and $f \in {}^0D$ (a standard vector). Now:

$$\begin{aligned}\xi_P(p')(P_i f) &= \mathrm{st}\langle f_P(p'), P_i f\rangle \\ &= \mathrm{st}\int_{p''\in {}^*R^3} d^3p''\, \check{\delta}^{(3)}(p''-p') P_i'' f(p'') \\ &= p_i' f(p') \\ &= p_i'\, \mathrm{st}\int_{p''\in {}^*R^3} d^3p''\, \check{\delta}^{(3)}(p''-p') f(p'') \\ &= p_i'\, \mathrm{st}\langle f_P(p'), f\rangle \\ &= p_i' \xi_P(p')(f). \end{aligned} \quad (6.34)$$

Since ξ_P and P_i are standard functions and (6.34) is true for every standard $p' \in R^3$ and standard $f \in {}^0D$, Eq. (6.34) is also true for all $p \in {}^*R^3$ and $f \in D$.

Let again $x' \in R^3$ (a standard triple) and $f \in {}^0D$ (a standard vector). Now:

$$\begin{aligned}\xi_X(x')(X_i f) &= \mathrm{st}\langle f_X(x'), X_i f\rangle \\ &= \mathrm{st}\int_{p''\in {}^*R^3} d^3p'' \int_{x''\in {}^*R^3} d^3x'' e^{i\mathbf{p}''\cdot\mathbf{x}''}\delta^{(3)}(x''-x') i\frac{\partial}{\partial p_i''} f(p'') \\ &= \mathrm{st}\int_{x''\in {}^*R^3} d^3x''\delta^{(3)}(x''-x') x_i'' \int_{p''\in {}^*R^3} d^3p''\, e^{i\mathbf{p}''\cdot\mathbf{x}''} f(p'') \\ &= x_i' \int_{p''\in {}^*R^3} d^3p''\, e^{i\mathbf{p}''\cdot\mathbf{x}'} f(p'') \\ &= x_i'\, \mathrm{st}\int_{x''\in {}^*R^3} d^3x''\delta^{(3)}(x''-x')\int_{p''\in {}^*R^3} d^3p''\, e^{i\mathbf{p}''\cdot\mathbf{x}''} f(p'') \\ &= x_i' \int_{p''\in {}^*R^3} d^3p'' \int_{d^3x''} d^3x'' e^{i\mathbf{p}''\cdot\mathbf{x}''}\delta^{(3)}(x''-x') f(p'') \\ &= x_i'\, \mathrm{st}\langle f_X(x'), f\rangle \\ &= x_i' \xi_X(x')(f). \end{aligned} \quad (6.35)$$

By a similar argument, (6.35) holds for all $x' \in {}^*R^3$ and $f \in D$. Hence the theorem is proved. ■

Definition 6.8:

(i) $\xi_{\mathbf{p}}^+(p') \otimes \xi_{\mathbf{p}}(p')$: $\begin{cases} D \times D \to {}^*C \\ (f,g) \to (\xi_{\mathbf{p}}^+(p') \otimes \xi_{\mathbf{p}}(p'))(f,g) = (\xi_{\mathbf{p}}(p')(f))^*(\xi_{\mathbf{p}}(p')(g)) \end{cases}$

for every $p' \in {}^*R^3$.

(ii) $\xi_X^+(x') \otimes \xi_X(x')$: $\begin{cases} D \times D \to {}^*C \\ (f,g) \to (\xi_X^+(x') \otimes \xi_X(x'))(f,g) = (\xi_X(x')(f))^*(\xi_X(x')(g)) \end{cases}$

for every $x' \in {}^*R^3$.

Theorem 6.7: Let $f,g \in D$ and $\Delta \subset {}^*R$ be a Borel subset. Then:

(i) $\langle f, E_\Delta(P_i) g\rangle = \int_{P_i'\in\Delta} dp_i' \prod_{j\neq i} \int_{P_j'\in {}^*R} dP_j' \times (\xi_{\mathbf{p}}^+(p') \otimes \xi_{\mathbf{p}}(p'))(f,g);$

(ii) $\langle f, E_\Delta(X_i) g\rangle = \int_{x_i'\in\Delta} dx_i' \prod_{j\neq i} \int_{x_j'\in {}^*R} dx_j' \times (\xi_X^+(x') \otimes \xi_X(x'))(f,g).$

Proof:

$$\begin{aligned}\langle f, E_\Delta(P_i) g\rangle &= \int_{p'\in {}^*R^3} d^3p'\, f^*(p')\, x_\Delta(p_i') g(p') \\ &= \int_{P_i'\in\Delta} dp_i' \prod_{j\neq i} \int_{P_j'\in {}^*R} dp_j'\, f^*(p') g(p') \\ &= \int_{P_i'\in\Delta} dp_i' \prod_{j\neq i} \int_{P_j'\in {}^*R} dp_j' (\xi_{\mathbf{p}}(p')(f))^*(\xi_{\mathbf{p}}(p') g) \\ &= \int_{P_i'\in\Delta} dp_i' \prod_{j\neq i} \int_{P_j'\in {}^*R} dp_j' (\xi_{\mathbf{p}}^+(p') \otimes \xi_{\mathbf{p}}(p'))(f,g), \end{aligned} \quad (6.36)$$

$$\begin{aligned}\langle f, E_\Delta(X_i) g\rangle &= \int_{P''\in {}^*R^3} d^3p''\, f^*(p'') \int_{x_i'\in\Delta} dx_i' \prod_{j\neq i} \times \int_{x_j'\in {}^*R} dx_j' \int_{P'\in {}^*R^3} d^3p'\, e^{i(\mathbf{p}'-\mathbf{p}'')\cdot\mathbf{x}'} g(p') \\ &= \int_{x_i'\in\Delta} dx_i' \prod_{j\neq i} \int_{x_j'\in {}^*R} dx_j' \int_{P''\in {}^*R^3} \times d^3p'' (e^{i\mathbf{p}''\cdot\mathbf{x}'} f(p''))^* \int_{P'\in {}^*R^3} d^3p'\, e^{i\mathbf{p}'\cdot\mathbf{x}'} g(p') \\ &= \int_{x_i'\in\Delta} dx_i' \prod_{j\neq i} \int_{x_j'\in {}^*R} dx_j' (\xi_X(x')(f))^*(\xi_X(x')(g)) \\ &= \int_{x_i'\in\Delta} dx_i' \prod_{j\neq i} \int_{x_j'\in {}^*R} dx_j' (\xi_X^+(x') \otimes \xi_X(x'))(f,g). \end{aligned}$$

■ (6.37)

Since D is dense in ${}^*\mathcal{K}$, Theorem 6.7 allows us to introduce the symbolic definition.

Definition 6.9: Let $\Delta \subset {}^*R$ be a Borel subset. Define:

(i) $\int_{P_i\in\Delta} dp_i' \prod_{j\neq i} \int_{P_j'\in {}^*R} dp_j'\, \xi_{\mathbf{p}}^+(p') \otimes \xi_{\mathbf{p}}(p') \equiv E_\Delta(P_i);$

(ii) $\int_{x_i' \in \Delta} dx_i' \prod_{j \neq i} \int_{x_j' \in {}^*R} dx_j' \; \xi_X^+(x') \otimes \xi_X(x') \equiv E_\Delta(X_i)$.

Theorem 6. 6 together with Def. 6. 9 reproduces the results obtained in the rigged Hilbert space approach to quantum mechanics.

Uncertainty principle

The states $f_p(p'), f_X(x')$ satisfy the properties

$$\langle f_p(p'), f_p(p'') \rangle = \delta_2^{(3)}(p' - p''),$$

$$\langle f_X(x'), f_X(x') \rangle = \delta_2^{(3)}(x' - x'')$$

as may be easily verified. $\delta_2^{(3)}$ is the three-dimensional δ function of the second order.

To normalize the states to unity, define the following.

Definition 6.10:

(i) $\Pi_{p'} \equiv [\delta_2^{(3)}(0)]^{-1/2} f_p(p')$ for any $p' \in {}^*R^3$.

(ii) $Y_{x'} \equiv [\delta_2^{(3)}(0)]^{-1/2} f_X(x')$ for any $x' \in {}^*R^3$.

As a result of Theorem 4. 8 together with Theorem 6. 5,

$\mu_{\Pi_{p'}, P_i}(\{p_i'\}) = 1$ for any $p_i' \in R$ and

$\mu_{Y_{x'}, X_i}(\{x_i'\}) = 1$ for any $x_i' \in R$.

Hence if a particle is in the state represented by the vector $\Pi_{p'}$, $p' \in R^3$, it is certain that it has momentum p', and if a particle is in the state represented by the vector $Y_{x'}$, $x' \in R^3$, it is certain that it is at the point x' in space. We wish to prove now that a complete certainty in the momentum of a particle gives rise to complete uncertainty in the position of that particle, and vice versa (i.e., complete certainty in the position gives rise to complete uncertainty in the momentum).

Theorem 6.8: Let $p \in R^3$, then for every $\lambda \in R$,

$$\mathrm{st}\langle \Pi_p, E_\lambda(X_i) \Pi_p \rangle = \tfrac{1}{2}, \quad i = 1, 2, 3$$

Proof:

$$\begin{aligned}
&\langle \Pi_p, E_\lambda(X_i) \Pi_p \rangle \\
&= \frac{1}{\delta_2^{(3)}(0)} \int_{p'' \in {}^*R^3} d^3p'' \, \delta^{(3)}(p'' - p) \int_{-\infty}^{\lambda} dx_i \prod_{j \neq i} \\
&\quad \times \int_{x_j \in {}^*R} dx_j \int_{p' \in {}^*R^3} d^3p' \, e^{i(p' - p'') \cdot x} \, \delta^{(3)}(p' - p) \\
&= \frac{1}{\delta_2^{(3)}(0)} \int_{-\infty}^{\lambda} dx_i \prod_{j \neq i} \int_{x_j \in {}^*R} dx_j \left(\int_{p'' \in {}^*R^3} d^3p'' e^{-ip'' \cdot x} \right. \\
&\quad \left. \times \, \delta^{(3)}(p'') \right) \left(\int_{p' \in {}^*R^3} d^3p' e^{ip' \cdot x} \, \delta^{(3)}(p') \right) \\
&= \frac{1}{\delta_2^{(3)}(0)} \int_{-\infty}^{0} dx_i \prod_{j \neq i} \int_{x_j \in {}^*R} dx_j \\
&\quad \times \left(\int_{p'' \in {}^*R^3} d^3p'' \, e^{-ip'' \cdot x} \, \delta^{(3)}(p'') \right) \\
&\quad \times \left(\int_{p' \in {}^*R^3} d^3p' \, e^{ip' \cdot x} \, \delta^{(3)}(p') \right) \\
&\quad + \frac{1}{\delta_2^{(3)}(0)} \int_0^{\lambda} dx_i \prod_{j \neq i} \int_{x_j \in {}^*R} dx_j \\
&\quad \times \left(\int_{p'' \in {}^*R^3} d^3p'' \, e^{-ip'' \cdot x} \, \delta^{(3)}(p'') \right) \\
&\quad \times \left(\int_{p' \in {}^*R^3} d^3p' \, e^{ip' \cdot x} \, \delta^{(3)}(p') \right) \\
&= \tfrac{1}{2} \frac{1}{\delta_2^{(3)}(0)} \int_{x \in {}^*R^3} d^3x \left(\int_{p'' \in {}^*R^3} d^3p'' \, e^{-ip'' \cdot x} \, \delta^{(3)}(p'') \right) \\
&\quad \times \left(\int_{p' \in {}^*R^3} d^3p' \, e^{ip' \cdot x} \, \delta^{(3)}(p') \right) \\
&\quad + \frac{1}{\delta_2^{(3)}(0)} \int_0^{\lambda} dx_i \prod_{j \neq i} \int_{x_j \in {}^*R} dx_j \left(\int_{p'' \in {}^*R^3} d^3p'' \right. \\
&\quad \left. \times \, e^{-ip' \cdot x'} \, \delta^{(3)}(p'') \right) \left(\int_{p' \in {}^*R^3} d^3p' \, e^{ip' \cdot x'} \, \delta^{(3)}(p') \right) \\
&= \tfrac{1}{2} + \frac{1}{\delta_2^{(3)}(0)} \int_0^{\lambda} dx_i \prod_{j \neq i} \int_{x_j \in {}^*R} dx_j \\
&\quad \times \int_{p'' \in {}^*R^3} d^3p'' e^{-ip'' \cdot x} \int_{p' \in {}^*R} d^3p' \delta^{(3)}(p') \delta^{(3)}(p'' - p') \\
&= \tfrac{1}{2} + \frac{1}{\delta_2^{(3)}(0)} \int_0^{\lambda} dx_i \prod_{j \neq i} \int_{x_j \in {}^*R} dx_j \int_{p'' \in {}^*R^3} \\
&\quad \times d^3p'' \, e^{-ip'' \cdot x} \, \delta_2^{(3)}(p'') \\
&= \tfrac{1}{2} + \frac{1}{\delta_2^{(3)}(0)} \int_0^{\lambda} dx_i \int_{p_i' \in {}^*R} dp_i' \, e^{-ip_i' x_i} \prod_{j \neq i} \\
&\quad \times \int_{x_j \in {}^*R} dx_j \int_{p_j \in {}^*R} e^{-ip_j' x_j} \, \delta_2^{(3)}(p') dp_j' \\
&= \tfrac{1}{2} + \frac{1}{\delta_2^{(3)}(0)} \int_0^{\lambda} dx_i \int_{|p_i'| \leq 2\epsilon} dp_i' e^{-ip_i' x_i} \delta_2^{(3)}(p') \Big|_{p_j' = 0 \text{ for } j \neq i}.
\end{aligned} \tag{6.38}$$

But

$$\begin{aligned}
\delta_2^{(3)}(p) &= \int d^3p' \, \delta^{(3)}(p - p') \delta^{(3)}(p') \\
&= \left| \int d^3p' \, \delta^{(3)}(p - p') \delta^{(3)}(p') \right| \\
&\leq \left(\int d^3p' \, |\delta^{(3)}(p - p')|^2 \right)^{1/2} \left(\int d^3p' \, |\delta^{(3)}(p')|^2 \right)^{1/2} \\
&= \delta_2^{(3)}(0)
\end{aligned}$$

or

$$\max_{p \in {}^*R^3} \delta_2^{(3)}(p) = \delta_2^{(3)}(0),$$

hence

$$\begin{aligned}
&\left| \int_0^{\lambda} dx_i \int_{|p_i'| \leq 2\epsilon} dp_i' \, e^{-ip_i' x_i} \, \delta^{(3)}(p') \Big|_{p_j' = 0 \text{ for } j \neq i} \right| \\
&\quad \leq \int_{\min(0,\lambda)}^{\max(0,\lambda)} dx_i \int_{|p_i'| \leq 2\epsilon} dp_i' \, \delta_2^{(3)}(0) = \frac{2\epsilon |\lambda|}{\pi} \delta_2^{(3)}(0).
\end{aligned} \tag{6.39}$$

Hence

$$|\langle \Pi_p, E_\lambda(X_i) \Pi_p \rangle - \tfrac{1}{2}| \leq 2\epsilon |\lambda| / \pi \approx 0 \tag{6.40}$$

or

$$\mathrm{st}\langle \Pi_p, E_\lambda(X_i) \Pi_p \rangle = \tfrac{1}{2}.$$

Theorem 6. 8 shows that Π_p is *not* a probability measure inducing vector relative to X_i. The probability of finding the particle in any interval $[a, b]$ is zero. However, the probability of finding the particle in the interval $(-\infty, a)$ or $(a, +\infty)$ is $\frac{1}{2}$. X_i has no strong expectation value in the state Π_p (because $\bar{\rho}_{\Pi_p, X_i} = \rho_{\Pi_p, X_i} = \frac{1}{2}$, and $\frac{1}{2}\lambda$ does not tend to any limit as $\lambda \to \pm\infty$). But X_i has a weak expectation value in the state Π_p:

$$\langle X_i \rangle_{\Pi_p} = 0$$

because the measure $\mu_{\Pi_p, X_i} = 0$.

Theorem 6. 8 still holds if we replace the momentum by the position and vice versa.

Dirac's notation

Let us write the following shorthand:

$|p'\rangle$ for $f_p(p')$, $|x'\rangle$ for $f_X(x')$, $|p')$ for $\xi_p^+(p')$, $|x')$ for $\xi_X^+(x')$, $(p'|$ for $\xi_p(p')$, $(x'|$ for $\xi_X(x')$, $\langle p'|p''\rangle$ for $\langle f_{\mathbf{p}}(p'), f_{\mathbf{p}}(p'')\rangle$, $\langle x'|x''\rangle$ for $\langle f_X(x'), f_X(x'')\rangle$, $\langle x'|p'\rangle$ for $\langle f_X(x'), f_p(p')\rangle$, $|p')(p'|$ for $\xi_{\mathbf{p}}^+(p')\otimes\xi_{\mathbf{p}}(p')$, $|x')(x'|$ for $\xi_X^+(x')\otimes\xi_X(x')$.

By using these notations we may transform a lot of expressions proved above into Dirac's language.

Normalization:

$$\langle p'|p''\rangle = \delta_2^{(3)}(p'-p''), \quad \langle x'|x''\rangle = \delta_2^{(3)}(x'-x').$$

Fourier transform: We find

$$\mathrm{st}\langle x'|p'\rangle = \mathrm{st}\int_{p''\in{}^*R^3} d^3p'' \int_{x''\in{}^*R^3} d^3x'' e^{i\mathbf{p}''\cdot\mathbf{x}''}\delta^{(3)}(x''-x')\delta^{(3)}(p''-p') = e^{i\mathbf{p}'\cdot\mathbf{x}'},$$

hence

$$\mathrm{st}\langle x'|p'\rangle = e^{i\mathbf{p}'\cdot\mathbf{x}'}.$$

Eigen functionals: The results of Theorem 6. 6 may be written in the form

$$(p'|P_i|f\rangle = p_i'(p'|f\rangle, \quad (x'|X_i|f\rangle = x_i'(x'|f\rangle,$$

which allow the introduction of the definition[6]

$$P_i'|p') = p_i'|p'), \quad X_i'|x') = x_i'|x').$$

Finally, Def. 6. 9 can be written in the form

$$E_\Delta(P_i) = \int_{p_i'\in\Delta} dp_i' \prod_{j\neq i} \int_{p_j'\in{}^*R} dp_j'|p')(p'|,$$

$$E_\Delta(X_i) = \int_{x_i'\in\Delta} dx_i \prod_{j\neq i} \int_{x_i'\in{}^*R} dx_j'|x')(x'|.$$

As a special case where $\Delta = R$ we get

$$\int_{p'\in{}^*R^3} d^3p'|p')(p'| = I, \quad \int_{x'\in{}^*R^3} d^3x'|x')(x'| = I.$$

ACKNOWLEDGMENTS

I would like to thank my supervisor, Dr. C. J. Isham, for his enormous help and encouragement and for the most valuable discussions and suggestions, without which I could not have completed this work. I would like, also, to thank the Lebanese Government for a maintenance grant.

APPENDIX

Theorem A: The quantifier $\forall$ satisfies the tautology

$$\neg\forall_n P(n) \Longleftrightarrow \forall_n \neg P(n)$$

($\neg P$ reads not "p").

Proof:

(i) Assume $\neg\forall_n P(n)$ i.e., $\{n: n\in N \text{ and } P(n)\}\notin\mathfrak{F}$ hence, by (F4) $N\setminus\{n: n\in N \text{ and } P(n)\}\in\mathfrak{F}$, i.e., $\{n: n\in N \text{ and } \neg P(n)\}\in\mathfrak{F}$ or $\forall_n\neg P(n)$. Hence

$$\neg\forall_n P(n) \Longrightarrow \forall_n \neg P(n). \tag{A1}$$

(ii) Assume $\forall_n\neg P(n)$, i.e., $\{n: n\in N \text{ and } \neg P(n)\}\in\mathfrak{F}$ hence, by (F1) and (F2) $N\setminus\{n: n\in N \text{ and } \neg P(n)\}\notin\mathfrak{F}$, i.e., $\{n: n\in N \text{ and } P(n)\}\notin\mathfrak{F}$ or $\neg\forall_n P(n)$. Hence

$$\forall_n \neg P(n) \Longrightarrow \neg\forall_n P(n) \tag{A2}$$

by (A1) and (A2) we get

$$\neg\forall_n P(n) \Longleftrightarrow \forall_n \neg P(n). \quad \blacksquare$$

Theorem B: Let $x\in M_0(R)$, then there exists a unique element $r\in R$ such that $x-r\in M_1(R)$.

Proof: By definition, there exists a positive $r'\in R$ such that $|x|<r'$ i.e.,

$$-r' < x < r'. \tag{A3}$$

Let $A(x)=\{r'': x\leq r'' \text{ and } r''\in R\}$. By (A3), $r'\in A(x)$ and $-r'$ is a lower bound for $A(x)$. Hence $A(x)$ has an infimum in R. Let $r=\inf A(x)$.

To show that $x-r\in M_1(R)$, assume the contrary, i.e., $x-r\notin M_1(R)$, which means that

$$\exists_{\delta\in R,\,\delta>0} \text{ such that } \delta<|x-r| \tag{A4}$$

if $x-r>0$, then

$$\delta < x-r \quad \text{and} \quad r+\delta < x. \tag{A5}$$

But (A5) means that $r+\delta$ is a lower bound of $A(x)$. Hence

$r+\delta\leq r$ by definition of r,

$\therefore$ therefore $\delta\leq 0$ contradicting the definition of δ.

If $x-r<0$, then

$$\delta < r-x \quad \text{and} \quad x < r-\delta \tag{A6}$$

But (A6) means that $r-\delta\in A(x)$ and so $r\leq r-\delta$, leading to the same contradiction $\delta\leq 0$.

The case $x-r=0$ gives an immediate contradiction to (A4). Hence

$$x-r\in M_1(R). \tag{A7}$$

To prove uniqueness, let $x-r'\in M_1(R)$. It can be shown that $M_1(R)$ is a commutative ring. Hence $(x-r)-(x-r')\in M_1(R)$ i.e.,

$$r'-r\in M_1(R). \tag{A8}$$

Now, if $r'\neq r$, then $\delta=\frac{1}{2}|r'-r|$ satisfies $\delta\in R$ and $\delta<|r'-r|$ contradicting (A8), hence $r=r'$. $\blacksquare$

Theorem C: The sets ${}^*R\setminus M_0(R)$ and $M_1(R)\setminus\{0\}$ are not empty.

Proof: Since R is an infinite set, it follows that[7] ${}^*R\setminus R\neq\phi$. Let $x\in{}^*R\setminus R$.

(i) If $x\in{}^*R\setminus M_0(R)$, then $x\neq 0$ [$0\in M_0(R)$] and $1/x$ exists. Now, for all $r\in R, r>0$ we have $1/r<|x|$ (by definition of x); hence $1/|x|<r$, which means that $1/x\in M_1(R)$ but $1/x\neq 0$, hence $1/x\in M_1(R)\setminus\{0\}$. i.e., ${}^*R\setminus M_1(R)\neq\phi$ and $M_1(R)\setminus\{0\}\neq\phi$.

(ii) If $x\notin{}^*R\setminus M_0(R)$, then $x\in M_0(R)$. Now $\mathrm{st}(x)\in R$ and $x-\mathrm{st}(x)\in M_1(R)$. But $\mathrm{st}(x)\neq x$ (because $x\notin R$). Hence $x-\mathrm{st}(x)\neq 0$ and $x-\mathrm{st}(x)\in M_1(R)\setminus\{0\}$. Also $1/[x-\mathrm{st}(x)]\in{}^*R\setminus M_0(R)$. Hence ${}^*R\setminus M_0(R)\neq\phi$ and $M_1(R)\setminus\{0\}\neq\phi$. Hence the theorem is proved. $\blacksquare$

[1] J. M. Jauch, *Foundations of Quantum Mechanics* (Addison-Wesley, 1968).
[2] E. R. Lorch, *Spectral Theory* (Oxford U. P., 1962).
[3] F. Riesz and B. Sz. Nagy, *Functional Analysis* (Ungar, New York, 1955).
[4] D. K. Sen, *Fields and/or Particles* (Ryerson, Toronto, 1968).
[5] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford U. P., 1958).
[6] J-P. Antoine, J. Math. Phys. **10**, 53 (1969).
[7] A. Robinson, *Non Standard Analysis* (North-Holland, Amsterdam, 1970).
[8] A. Voros, "Introduction to Non Standard Analysis," DPh-T/72/16, Centre d'Etudes Nucleaires de Saclay, France.
[9] P. G. Kelemen and A. Robinson, J. Math. Phys. 13, 1870 (1972).
[10] I. M. Gel'fand & G. E. Shilov, *Generalized Functions* (Academic, New York, (1967), Vol. 3.
[11] R. A. Dean, *Elements of Abstract Algebra* (Wiley, New York, 1966).
[12] P. R. Halmos, *Measure Theory* (Van Nostrand, Princeton, N. J. 1962).
[13] J. M. Jauch and B. Misra, Helv. Phys. Acta 38, 30 (1965).

An algebraic representation of continuous superselection rules

Renzo Cirelli and Franco Gallone

Istituto di Scienze Fisiche dell'Università, 20133 Milano, Italy

Istituto Nazionale di Fisica Nucleare, Sezione di Milano

Bella Gubbay

Istituto di Scienze Fisiche dell'Università, 20133 Milano, Italy

(Received 6 September 1973)

From the logic approach to quantum and classical mechanics, the W^*-algebraic approach is deduced in dependence of a suitable "prestate." An algebraic representation of the logic description is in fact constructed in a framework in which continuous superselection rules can be present. Logic propositions, observables, and states are represented by decomposable projections, decomposable self-adjoint operators, and normal states in a direct integral of Hilbert spaces. In this representation each algebraic term becomes the representative of a homologous logic one and the expectation values as well as the superselection rules are conserved. When a principle of "undistinguishability" is taken into account, the representation is faithful. In the classical case, the representation results in Koopman's formalism.

INTRODUCTION

Recently, much interest has been devoted to the logic and algebraic axiomatic approaches to the foundations of physical theories. Both approaches describe as particular cases the usual models of quantum and classical mechanics.

The logic (or "lattice" or "events") approach has been developed in recent years[1] according to an idea advanced by Birkhoff and von Neumann.[2] It encompasses the essential features and conceptual problems of other axiomatic approaches and, owing to its semplicity and direct adherence to the phenomenological interpretation of the physical "experiments", has made possible a better understanding of some of the physical concepts, such as the definition of symmetries and their representations,[3] the nonoccurrence of hidden variables in quantum theory,[4] the nature of superselection rules.[5]

The interest in the algebraic approach is mainly a consequence of the work of Segal,[6] who, generalizing an idea of Jordan, von Neumann, and Wigner,[7] assumes that the observables of a physical system are properly represented by the self-adjoint elements of a C^*-algebra and the states by the normalized positive linear functionals on it. The principal aims of this approach are: to show that some of the difficulties encountered by quantum mechanics when applied to situations involving an infinite number of degrees of freedom are due to some intrinsic limitations of the traditional framework; to provide a scheme in which these difficulties could be by-passed. Moreover, by GNS construction,[8] it brings to physical theory the entirely new concept that the representations of the algebraic picture depend on the states of the physical system which we want to describe. Segal's algebraic approach has been widely applied to statistical mechanics and quantum field theory. An important standpoint related to this approach has been expressed by Haag and Kastler.[9]

In a previous paper,[10] developing an idea of Jauch,[11] we examined the possibility of a complete description of a quantum system (including superselection rules) by a formalism based on W^*-algebras and normal states. The choice of these algebras and states instead of the more general C^*-algebras and positive linear functionals is almost unavoidable if one wants an algebraic model deducible from a logic one, since the set of the projections of a W^*-algebra forms a σ-complete orthocomplemented lattice; moreover, normalcy of states amounts, for a separable Hilbert space, to σ-additivity on the "logic" of the projections. Furthermore, von Neumann's model of quantum mechanics[12] is in fact a model with a W^*-algebra and normal states and, under fairly general conditions, the Gibbs state of a statistical system is "locally normal."[13] Finally, also a classical system can be described with W^*-algebras and normal states by means of the following construction. Take for a classical system an Abelian C^*-algebra $\mathfrak{A}$ (to describe the observables) and the set of normalized positive linear functionals on it (to describe the states). Then, by Gel'fand–Naimark isomorphism, $\mathfrak{A}$ is isomorphic with the C^*-algebra of the continuous complex functions vanishing at infinity on a locally compact space Λ (the phase space of the system) and, by a representation theorem of Riesz, the states are represented by probability measures on Λ. Hence, fixing a distinguished state μ, by Koopman's formalism[14] we get the C^*-algebra D_c of continuously decomposable operators on $L^2(\Lambda, \mu)$,[15] in which only the probability measures which are absolutely continuous with respect to μ can be faithfully represented. If compatibility of this description with the logic approach is now required, we have to replace D_c with the W^*-algebra D generated by D_c (which contains decomposable operators of general kind)[15] because Ludwig's axiom of sensitivity increase[16] in a sense holds for a C^*-algebra if and only if the algebra is in fact a W^*-algebra.[17] Moreover the states selected in the previously described way (namely, absolutely continous with respect to μ) can be easily shown to be normal states on D.

For the important and partially complementary roles played by the logic and algebraic approaches, an important task is to investigate the correlations between them. Plymen has shown how it is possible to deduce from an algebraic picture a logic one.[18] In the aforementioned paper[10] we succeeded in the converse deduction for a physical system in which just discrete superselection rules[19] were allowed. This was indeed a rather serious limitation which, for instance, excluded the possibility of encompassing in our investigation also classical systems.

In the present work we solve the problem of deducing the algebraic picture based on W^*-algebras and normal

states from a logic description "à la Jauch"[20] in which superobservables with purely continuous spectrum (namely continuous superselection rules) can be present. The algebraic representation that we get is ruled by a "prestate" which selects a part of the physical system described in the logic picture in such a way that the different superselected sectors are linked in a measurable way. The representation we get is therefore dependent on a state, in accordance with the spirit of GNS construction, and it furnishes a faithful representation for propositions, observables and states when "undistinguishable" elements are identified; moreover, this representation sets up a complete correspondence between the algebraic and logic pictures since each algebraic object is the representative of a homologous logic one. The representation we construct sets up also a complete and faithful correspondence between logic and algebraic superobservables. Finally, it is worthwhile to mention that, seeking what our construction results in, if the physical system is a classical one, we are led to recover Koopman's formalism.

For the definitions and the results of the algebraic picture that we use (expecially for superselection rules, sectors, decomposable operators) see our aforementioned paper.[10] Most of the mathematical tools we use in connection with W^*-algebras, direct integrals, and normal functionals are the content of book of Dixmier,[15] to which we refer also for the notations and the terminology on these subjects.

1. THE LOGIC PICTURE

The *propositions* of a physical system are represented by the elements of an orthocomplemented, weakly modular and σ-complete lattice $\mathfrak{L}$.[20] Such a structure will be called a logic.[21] Assuming $\mathfrak{L}$ atomic and complete, it can be decomposed[22] into a direct union[23] of nontrivial irreducible logics $\mathfrak{L}^{(\lambda)}$ (the superselected sectors): $\mathfrak{L} = \vee^{\oplus}_{\lambda\in\Lambda}\mathfrak{L}^{(\lambda)}$. The center $\mathcal{C}(\mathfrak{L})$ of $\mathfrak{L}$ is a Boolean logic; hence[24] it can be considered to represent a classical system (namely, the classical part of the physical system represented by $\mathfrak{L}$). Therefore Λ can be considered as the phase space of a classical system, since $\mathcal{C}(\mathfrak{L})$ is isomorphic[25] (in an obvious way) to the power set $\mathbb{P}(\Lambda)$. For this reason we may reasonably make the assumption that a σ-algebra $\mathfrak{B}$ exists on Λ such that Λ results in a standard Borel space (for instance, Λ could be a locally compact topological space with a countable open basis or a complete separable metric space).

A *state* ψ is a mapping

$$\psi : \mathfrak{D}_\psi \to [0, 1]$$

such that

(a) $\mathfrak{D}_\psi$ is a sublogic[26] of $\mathfrak{L}$,

(b) $\psi(\emptyset_{\mathfrak{D}_\psi}) = 0,\ \psi(I_{\mathfrak{D}_\psi}) = 1$,

(c) $\psi(\vee_n p_n) = \sum_n \psi(p_n)$ for any sequence $\{p_n\}$ of mutually orthogonal elements of $\mathfrak{D}_\psi$.

While $\mathfrak{L}$ represents the set of all the "elementary detectors" for the physical system, the set S of the states represents in a sense the set of all the "preparation procedures". We have not defined the states *on* $\mathfrak{L}$; this amounts to assume that not any pair "source detector" makes sense as a physical arrangement to get measures (an analogous standpoint can be found in Sec. 1. 5 of Ludwig's book quoted in Ref. 1). If, for instance, $\mathfrak{D}_\psi$ is isomorphic to the Borel σ-algebra $\mathfrak{B}$ existing by assumption on Λ, then through this isomorphism ψ reduces to a probability measure on Λ, and we have in fact a classical description for the classical part of the system represented by $\mathfrak{L}$. If at the opposite side $\mathfrak{D}_\psi$ is isomorphic to $\mathfrak{L}^{(\lambda)}$ for some $\lambda \in \Lambda$, then $\mathfrak{D}_\psi$ is irreducible and ψ provides a purely quantum description, since $\mathfrak{D}_\psi$ is centerless and no classical part is then allowed.

For a state ψ, the set O_ψ of *ψ-observables* is the set of $\alpha \in \hom(\mathfrak{B}_{\mathbb{R}}, \mathfrak{D}_\psi)$ such that $\alpha(\mathbb{R}) = I_{\mathfrak{D}_\psi}$ ($\mathfrak{B}_{\mathbb{R}}$ stands for the Borel σ-algebra on the real line $\mathbb{R}$). The spectrum of an observable, bounded and constant observables, functions of an observable, are defined in the usual way.[21] In particular, if k is an element of $\mathbb{R}$ different from zero and α is a ψ-observable, by $k\alpha$ we denote the ψ-observable defined by $(k\alpha)(E) = \alpha(k^{-1}E), \forall E \in \mathfrak{B}_{\mathbb{R}}$, while for $k = 0$ we define, for $E \in \mathfrak{B}_{\mathbb{R}}$,

$$(0\alpha)(E) = \begin{cases} I_{\mathfrak{D}_\psi} & \text{if } 0 \in E, \\ \emptyset_{\mathfrak{D}_\psi} & \text{if } 0 \notin E. \end{cases}$$

For any $k \in \mathbb{R}$, the observable $k\alpha$ just defined coincides in fact with $f_k(\alpha)$, where $f_k : \mathbb{R} \to \mathbb{R}, f_k(t) = kt$. The ψ-questions are ψ-observables of considerable importance. A ψ-question α_p (with $p \in \mathfrak{D}_\psi$) is the ψ-observable such that $\alpha_p(\{0, 1\}) = I_{\mathfrak{D}_\psi}$ and $\alpha_p(\{1\}) = p$. A ψ-superobservable is a nonconstant element of O_ψ, with range in the center of $\mathfrak{D}_\psi$. If α is a bounded ψ-observable, $\psi\circ\alpha$ is a bounded probability measure on $\mathbb{R}$ and

$$\langle\alpha\rangle_\psi := \int_{\mathbb{R}} t d(\psi\circ\alpha)(t)$$

is the expectation value of α in ψ. The set of bounded ψ-observables will be denoted by $O_\psi{}^0$. We point out that for a ψ-question α_p the expectation-value $\langle\alpha_p\rangle_\psi$ can be interpreted as the probability that a sample of the physical system represented by $\mathfrak{L}$ possesses the "property" p when it is prepared in the state ψ, since $\langle\alpha_p\rangle_\psi = \psi(p)$.

2. THE PRESTATE

To get an algebraic picture from the logic one outlined above, the first step is to take into account a representation theorem[22] along with other results,[27] from which it follows that each $\mathfrak{L}^{(\lambda)}$ can be identified with the logic $\mathcal{P}(\mathcal{H}(\lambda))$ of the projections of a real, complex or quaternionic (depending on λ) nontrivial Hilbert space $\mathcal{H}(\lambda)$. We shall assume $\mathcal{H}(\lambda)$ to be complex for any λ; this seems in fact to entail no loss of generality.[28] In each sector $\mathfrak{L}^{(\lambda)}$ it is then possible to obtain an algebraic picture from the logic one in a standard way based on Gleason's theorem and spectral theory.[29] It seems sensible that it will be possible to achieve the same result for the overall picture carried by $\mathfrak{L}$ if the sectors $\mathfrak{L}^{(\lambda)}$ will be linked in a suitable way. In our picture, such a link is in fact established by a state in the way which will be expounded in the assumption below. First we need a rather technical result.

Proposition 2.1: Let $\{x_n\}$ be a sequence of elements of $\Pi_{\lambda\in\Lambda}\mathcal{H}(\lambda)$ (namely of fields of vectors) such that $\sum_{n=1}^{\infty}\|x_n(\lambda)\|^2 = 1, \forall\lambda \in \Lambda$, and μ a measure on Λ such that the function $\Lambda \ni \lambda \to (x_n(\lambda)|x_m(\lambda))$ is μ-measurable, $\forall n, m$; let $p_0(\lambda)$ be the element of $\mathcal{P}(\mathcal{H}(\lambda))$ with range the subspace $\mathcal{H}_0(\lambda)$ of $\mathcal{H}(\lambda)$ spanned by $\{x_n(\lambda)\}$. Then[30]

$$\mathfrak{D} := \{p = \vee^{\oplus}_{\lambda\in\Lambda} p(\lambda) \in \mathfrak{L};\ p(\lambda) \leq p_0(\lambda), \forall\lambda \in \Lambda \text{ and } (x_n(\lambda)|p(\lambda)x_m(\lambda)) \ \mu\text{-measurable}, \forall n, m\}$$

is a sublogic of $\mathfrak{L}$ with unit element $I_{\mathfrak{D}} = \vee^{\oplus}_{\lambda\in\Lambda} p_0(\lambda)$.

Two elements p and q of $\mathfrak{D}$ are orthogonal iff $p(\lambda)q(\lambda) = \mathbb{O}_\lambda$ [the null projection on $\mathcal{K}(\lambda)$], $\forall \lambda \in \Lambda$, and compatible iff $[p(\lambda), q(\lambda)] = \mathbb{O}_\lambda, \forall \lambda \in \Lambda$. The center of $\mathfrak{D}$ is $\mathcal{C}(\mathfrak{D}) = \{p \in \mathfrak{L}; p = \vee^{\oplus}_{\lambda\in\Lambda} \chi_E(\lambda)p_0(\lambda), E \in \mathcal{B}_\mu\}$,[31] where $\mathcal{B}_\mu$ is the μ-completion of the σ-algebra $\mathcal{B}$.

Proof: Let p_1 and p_2 be two elements of $\mathfrak{D}$; then $(p_1 \wedge p_2)(\lambda) = p_1(\lambda) \wedge p_2(\lambda) \leq p_0(\lambda)$ trivially holds, $\forall \lambda \in \Lambda$; next, since

$$(x_n(\lambda)|(p_1 \wedge p_2)(\lambda)x_m(\lambda)) = \lim_{k\to\infty} (x_n(\lambda)|(p_1(\lambda)p_2(\lambda))^k x_m(\lambda)),$$[32]

from μ-measurability of $(x_n(\lambda)|(p_1(\lambda)p_2(\lambda))^k x_m(\lambda))$ for any k (II. 2. 1 in Ref. 15), μ-measurability of $(x_n(\lambda)|(p_1 \wedge p_2)(\lambda)x_m(\lambda))$ follows for any n and m, whence $p_1 \wedge p_2 \in \mathfrak{D}$. We get the same result for $p_1 \vee p_2$, since $(p_1 \vee p_2)(\lambda) = p_1(\lambda) \vee p_2(\lambda) \leq p_0(\lambda)$ trivially holds, $p_1 \vee p_2 = (p_1' \wedge p_2')'$ ($'$ means orthocomplementation in $\mathfrak{L}$) is true and, if $(x_n(\lambda)|p(\lambda)x_m(\lambda))$ is μ-measurable for an element $p = \vee^{\oplus}_{\lambda\in\Lambda} p(\lambda)$ of $\mathfrak{L}$, the same holds for p', since

$$(x_n(\lambda)|p'(\lambda)x_m(\lambda)) = (x_n(\lambda)|x_m(\lambda)) - (x_n(\lambda)|p(\lambda)x_m(\lambda)).$$

In this way we have shown that $\mathfrak{D}$ is closed with respect to finite join and meet. Moreover $\mathfrak{D}$ has least element $\tilde{\mathcal{O}}_\mathfrak{D} = \mathcal{O}_\mathfrak{L} = \vee^{\oplus}_{\lambda\in\Lambda} \mathbb{O}_\lambda$ and greatest element $I_\mathfrak{D} = \vee^{\oplus}_{\lambda\in\Lambda} p_0(\lambda)$. Take now a sequence $\{p_n\}$ of elements of $\mathfrak{D}$. Setting $q_i := \vee_{n=1}^{i} p_n$, we get $\vee_n p_n = \vee_i q_i$, since for any element p of $\mathfrak{L}$ we have $p \geq q_i, \forall_i$ iff $p \geq p_n, \forall_n$. Moreover $\vee_i q_i \in \mathfrak{D}$, as $(\vee_i q_i)(\lambda) = \vee_i q_i(\lambda) \leq p_0(\lambda)$ trivially holds $\forall \lambda \in \Lambda$, and for μ-measurability we have

$$(x_n(\lambda)|\vee_i q_i(\lambda)x_m(\lambda)) = \lim_{i\to\infty} (x_n(\lambda)|q_i(\lambda)x_m(\lambda)),$$

since $\{q_i(\lambda)\}$ is an increasing sequence of projections in the Hilbert space $\mathcal{K}(\lambda)$.[33] Therefore $\vee_n p_n \in \mathfrak{D}$. In the same way we can prove $\wedge_n p_n \in \mathfrak{D}$. Finally, for any $p \in \mathfrak{D}$ we get $(p' \wedge I_\mathfrak{D})(\lambda) = (\mathbb{1}_\lambda - p(\lambda)) \wedge p_0(\lambda) = p_0(\lambda) - p(\lambda)$ [where $\mathbb{1}_\lambda$ is the unit operator on $\mathcal{K}(\lambda)$], $\forall \lambda \in \Lambda$, whence $p' \wedge I_\mathfrak{D} \in \mathfrak{D}$ easily follows. Hence $\mathfrak{D}$ is a sublogic of $\mathfrak{L}$; the orthocomplement in $\mathfrak{D}$ of an element p of $\mathfrak{D}$ will be denoted by $\tilde{p}$ (hence $\tilde{p} = p' \wedge I_\mathfrak{D}$).

Two elements p_1 and p_2 of $\mathfrak{D}$ are said orthogonal if $p_1 \leq \tilde{p}_2$ (Ref. 21, Chap. VI, Sec. 1), namely if $p_1(\lambda) = p_1(\lambda)\tilde{p}_2(\lambda) = p_1(\lambda)(p_0(\lambda) - p_2(\lambda)) = p_1(\lambda) - p_1(\lambda)p_2(\lambda)$, which is equivalent to $p_1(\lambda)p_2(\lambda) = \mathbb{O}_\lambda, \forall \lambda \in \Lambda$.

Two elements p and q of $\mathfrak{D}$ are said compatible if there are

$$p_1 = \vee^{\oplus}_{\lambda\subset\Lambda} p_1(\lambda), \qquad q_1 = \vee^{\oplus}_{\lambda\in\Lambda} q_1(\lambda), \qquad r = \vee^{\oplus}_{\lambda\in\Lambda} r(\lambda)$$

mutually orthogonal in $\mathfrak{D}$ such that $p = p_1 \vee r$ and $q = q_1 \vee r$ (Ref. 21, Chap, VI, Sec. 5). If this is the case, $p(\lambda) = p_1(\lambda) + r(\lambda)$ and $q(\lambda) = q_1(\lambda) + r(\lambda)$ with $p_1(\lambda)q_1(\lambda) = p_1(\lambda)r(\lambda) = q_1(\lambda)r(\lambda) = \mathbb{O}_\lambda$, whence $p(\lambda)q(\lambda) = r(\lambda) = q(\lambda)p(\lambda), \forall \lambda \in \Lambda$. If conversely for $p, q \in \mathfrak{D}$ we have $[p(\lambda)q(\lambda)] = \mathbb{O}_\lambda$, $\forall \lambda \in \Lambda$, we can construct r, p_1, q_1 in $\mathfrak{D}$ setting $r(\lambda) := p(\lambda)q(\lambda), p_1(\lambda) := p(\lambda) - r(\lambda), q_1(\lambda) := q(\lambda) - r(\lambda)$; these are easily shown to be mutually orthogonal elements of $\mathfrak{D}$ such that $p = p_1 \vee r$ and $q = q_1 \vee r$.

If $E \in \mathcal{B}_\mu$, $\vee^{\oplus}_{\lambda\in\Lambda} \chi_E(\lambda)p_0(\lambda)$ is obviously an element of $\mathfrak{D}$ compatible with any element of $\mathfrak{D}$, namely it is an element of $\mathcal{C}(\mathfrak{D})$. If conversely $p \in \mathcal{C}(\mathfrak{D})$, then $[p(\lambda), q(\lambda)] = \mathbb{O}_\lambda, \forall q \in \mathfrak{D}, \forall \lambda \in \Lambda$, and this condition can be shown (easily by absurd) to imply $p(\lambda) \in \{\mathbb{O}_\lambda, p_0(\lambda)\}$, $\forall \lambda \in \Lambda$. Hence a subset E of Λ exists such that

$$p = \vee^{\oplus}_{\lambda\in\Lambda} \chi_E(\lambda)p_0(\lambda).$$

The function

$$\chi_E(\lambda) = \sum_{n=1}^{\infty} \chi_E(\lambda)\|x_n(\lambda)\|^2 = \sum_{n=1}^{\infty} (x_n(\lambda)|\chi_E(\lambda)p_0(\lambda)x_n(\lambda))$$

is μ-measurable since p is an element of $\mathfrak{D}$, whence $E \in \mathcal{B}_\mu$. QED

Assumption: The passage from the logic description of Sec. 1 to an algebraic one is ruled by a state φ which can be attained in the following way.

Let a "purely classical" situation be given in $\mathfrak{L}$ (namely a probability measure μ on $\mathcal{B}$) along with a "purely quantum" situation in each superselected sector $\mathfrak{L}^{(\lambda)}$ (namely a state φ_λ defined on $\mathfrak{L}^{(\lambda)}$) in such a way that if $\{x_n(\lambda)\}$ is the sequence of vectors in $\mathcal{K}(\lambda)$ related to φ_λ by Gleason's theorem, the function $\Lambda \ni \lambda \to (x_n(\lambda)|x_m(\lambda))$ is μ-measurable, $\forall n, m$ (this is in fact the most important point, because this is the only link that we require among the $\mathfrak{L}^{(\lambda)}$'s). Denoting by $p_0(\lambda)$ the projection in $\mathcal{K}(\lambda)$ onto the subspace $\mathcal{K}_0(\lambda)$ spanned by $\{x_n(\lambda)\}$, define

$$\mathfrak{D}_\varphi := \{p = \vee^{\oplus}_{\lambda\in\Lambda} p(\lambda) \in \mathfrak{L}; p(\lambda) \leq p_0(\lambda), \forall \lambda \in \Lambda, \text{ and}$$

$$\Lambda \ni \lambda \to (x_n(\lambda)|p(\lambda)x_m(\lambda)) \ \mu\text{-measurable}, \forall n, m\}$$

and

$$\varphi : \mathfrak{D}_\varphi \to [0, 1], \qquad \varphi(p) = \int_\Lambda \varphi_\lambda(p(\lambda))d\mu(\lambda)$$

(the definition of φ is consistent, namely the integral exists, just by the definition of $\mathfrak{D}_\varphi$). The mapping φ is a state [$\mathfrak{D}_\varphi$ is a sublogic after Prop. 2. 1 and properties (b) and (c) of Sec. 1 are easily verified]. It is the state which rules the picture which we are going to construct.

We notice that if $\alpha^{(E)}$ is the φ-question related to the proposition $\vee^{\oplus}_{\lambda\in\Lambda} \chi_E(\lambda)p_0(\lambda)$, where $E \in \mathcal{B}$, then we can easily show that $\mu(E) = \langle \alpha^{(E)} \rangle_\varphi$. Hence $\mu(E)$ is the probability that if a sample of the system represented by $\mathfrak{L}$ is prepared by the preparation procedure φ, it will be found in the sectors $\mathfrak{L}^{(\lambda)}$ with $\lambda \in E$. Therefore, φ can be thought of as a "continuous" classical superposition of states concentrated in the superselected sectors, with weights which are the probabilities in φ of the sectors. In the description which we will now construct in dependence of the assumption above, φ will play the role of a "prestate." It is in a sense a "coarse" state, while the φ-states which will be soon introduced represent a refinement of φ. This process is particularly transparent for the conditional states we are going to define, since in this case the "filtering" subsequent to φ is performed observing a "property" of the system.

From φ we can get in fact other states in the following way. From a component φ_λ of φ [which is a state on the sector $\mathcal{P}(\mathcal{K}(\lambda))$] and a component $p(\lambda)$ of an element p of $\mathfrak{D}_\varphi$ such that $\varphi_\lambda(p(\lambda)) \neq 0$, we get in a standard way[34] the new state on $\mathcal{P}(\mathcal{K}(\lambda))$

$$\tilde{\varphi}_\lambda^{(p)} : \mathcal{P}(\mathcal{K}(\lambda)) \to [0, 1], \qquad \tilde{\varphi}_\lambda^{(p)}(q_\lambda) = [\varphi_\lambda(p^{(\lambda)})]^{-1}\varphi_\lambda^{(p)}(q_\lambda),$$

where

$$\varphi_\lambda^{(p)}(q_\lambda) = \sum_{n=1}^{\infty} (p(\lambda)x_n(\lambda)|q_\lambda p(\lambda)x_n(\lambda)).$$

The state $\tilde{\varphi}_\lambda^{(p)}$ is interpreted as the preparation procedure described by φ_λ and conditioned by $p(\lambda)$ [namely, after φ_λ the occurrence or nonoccurrence of the "property" $p(\lambda)$ is determined and the resulting physical system is accepted as a sample physical system if and only if $p(\lambda)$ occurred; $\varphi_\lambda(p(\lambda))$ is a normalization factor which represents the frequency of accepted cases]. Then, from φ and an element p of $\mathfrak{D}_\varphi$ such that $\varphi(p) \neq 0$ we can get a

new state φ_p defined on $\mathfrak{D}_\varphi$ composing the conditioned states $\tilde{\varphi}_p^{(\lambda)}$ resulting in each $\mathfrak{L}^{(\lambda)}$ from the components φ_λ and $p(\lambda)$, each with its own "frequency of transmission," and normalizing after the composition (since after φ we observe the "property" p as a whole); namely,

$$\varphi_p : \mathfrak{D}_\varphi \to [0,1],$$

$$\varphi_p(q) = \varphi_p\left(\overset{\oplus}{\underset{\lambda\in\Lambda}{\vee}} q(\lambda)\right) = \frac{1}{\varphi(p)} \int_\Lambda \varphi_\lambda^{(p)}(q(\lambda))d\mu(\lambda),$$

where $\varphi_\lambda^{(p)}(q(\lambda)) := 0$ if $\varphi_\lambda(p(\lambda)) = 0$ [hence, in any case,

$$\varphi_\lambda^{(p)}(q(\lambda)) = \sum_{n=1}^{\infty} (p(\lambda)x_n(\lambda) \mid q(\lambda)p(\lambda)x_n(\lambda))]$$

It is a matter of direct inspection to control that the definition of φ_p is consistent and that φ_p is really a state. We notice that φ_p includes also a classical description, since

$$\mathfrak{A} := \{p \in \mathfrak{D}_\varphi; p = \overset{\oplus}{\underset{\lambda\in\Lambda}{\vee}} \chi_E(\lambda)p_0(\lambda), E \in \mathfrak{B}\}$$

is a sub-σ-algebra of the center of $\mathfrak{D}_\varphi$ isomorphic (in an obvious way) with $\mathfrak{B}$ and, through this isomorphism, φ_p defines on $\mathfrak{B}$ the probability measure

$$\mu_p(E) := \varphi_p\left(\overset{\oplus}{\underset{\lambda\in\Lambda}{\vee}} \chi_E(\lambda)p_0(\lambda)\right), \quad \forall E \in \mathfrak{B}.$$

The state φ_p can be interpreted as "the conditional state with respect to p and φ".[35] We shall denote by $\mathfrak{D}_\varphi^0$ the set of the elements p of $\mathfrak{D}_\varphi$ which give rise in this way to states, namely for which $\varphi(p) \neq 0$.

We shall now define as φ-states the states defined on $\mathfrak{D}_\varphi$ which preserve the "sum" of two observables whenever such a sum exists with respect to the states which can be constructed from φ and $\mathfrak{D}_\varphi^0$ in the aforementioned way.

Definition 2.1: The elements of the set

$$S_\varphi := \{\psi \in S; \mathfrak{D}_\psi = \mathfrak{D}_\varphi \text{ and } \langle\alpha\rangle_\psi + \langle\beta\rangle_\psi = \langle\gamma\rangle_\psi \text{ whenever}$$

$$\alpha, \beta, \gamma \in O_\varphi^0 \text{ are such that } \langle\alpha\rangle_{\varphi_p} + \langle\beta\rangle_{\varphi_p} = \langle\gamma\rangle_{\varphi_p}, \forall p \in \mathfrak{D}_\varphi^0\}$$

are called φ-states.

Obviously $\varphi_p \in S_\varphi, \forall p \in \mathfrak{D}_\varphi^0$. It is not in general true that such states exhaust S_φ, as it is shown by the counterexample of Sec. 4, in which elements of S_φ can be found different from $\varphi_p, \forall p \in \mathfrak{D}_\varphi^0$.

Remark 2.1: Let α_1 and α_2 be two elements of O_φ^0 such that $\langle\alpha_1\rangle_{\varphi_p} = \langle\alpha_2\rangle_{\varphi_p}, \forall p \in \mathfrak{D}_\varphi^0$. Since for any $\psi \in S$ and $\alpha \in O_\varphi^0$ we have $\langle k\alpha\rangle_\psi = \int_R td(\psi\circ\alpha)(k^{-1}t) = \int_R ktd(\psi\circ\alpha)(t') = k\langle\alpha\rangle_\psi, \forall k \in \mathbb{R} - \{0\}$, then $\langle\alpha_1\rangle_{\varphi_p} = \frac{1}{2}(\langle\alpha_2\rangle_{\varphi_p} + \langle\alpha_2\rangle_{\varphi_p}) = \langle\frac{1}{2}\alpha_2\rangle_{\varphi_p} + \langle\frac{1}{2}\alpha_2\rangle_{\varphi_p}, \forall p \in \mathfrak{D}_\varphi^0$, whence $\langle\alpha_1\rangle_\psi = \langle\frac{1}{2}\alpha_2\rangle_\psi + \langle\frac{1}{2}\alpha_2\rangle_\psi = \langle\alpha_2\rangle_\psi, \forall\psi \in S_\varphi$.

The next defintion amounts to consider, in the picture ruled by the prestate φ of the assumption, as states only the elements of S_φ. In that case two φ-observables result in fact to be undistinguishable when their expectation values are the same in each element of S_φ.

Definition 2.2: Two elements $\alpha, \beta \in O_\varphi^0$ are said to be equivalent if $\langle\alpha\rangle_\psi = \langle\beta\rangle_\psi, \forall_\psi \in S_\varphi$. In that case we write α R β and R is easily shown to be an equivalence relation on O_φ^0.

Remark 2.2: Particular elements of O_φ^0 are the φ-questions. For two φ-questions α_{p_1} and α_{p_2} ($p_1, p_2 \in \mathfrak{D}_\varphi$), $\alpha_{p_1} R \alpha_{p_2}$ is equivalent to $\psi(p_1) = \langle\alpha_{p_1}\rangle_\psi = \langle\alpha_{p_2}\rangle_\psi = \psi(p_2), \forall_\psi \in S_\varphi$. Hence the equivalence relation on O_φ^0 involves the equivalence relation $\sim$ on $\mathfrak{D}_\varphi$ defined in the following way: for

$$p_1, p_2 \in \mathfrak{D}_\varphi, \quad p_1 \sim p_2 \text{ iff } \psi(p_1) = \psi(p_2), \quad \forall_\psi \in S_\varphi.$$

We could try to define, for φ-states, an equivalence relation which is the dual of that introduced by Def. 2.2, namely to call equivalent two elements ψ_1, ψ_2 of S_φ if $\langle\alpha\rangle_{\psi_1} = \langle\alpha\rangle_{\psi_2}, \forall\alpha \in O_\varphi^0$. However, this is a trivial relation indeed, since for two such elements of S_φ we get $\langle\alpha_p\rangle_{\psi_1} = \langle\alpha_p\rangle_{\psi_2}$, whence $\psi_1(p) = \psi_2(p), \forall p \in \mathfrak{D}_\varphi$, namely $\psi_1 = \psi_2$.

3. THE ALGEBRAIC REPRESENTATION

We shall now construct an algebraic representation of the logic concepts previously introduced in dependence of the prestate φ, namely of φ-propositions (the elements of $\mathfrak{D}_\varphi$), φ-observables (the elements of O_φ^0) and φ-states (the elements of S_φ). We have at our disposal the field $\Pi_{\lambda\in\Lambda} \mathcal{K}_0(\lambda)$ of the complex and separable (by construction) Hilbert spaces $\mathcal{K}_0(\lambda)$, wherein a sequence $\{x_n\}$ exists of fields of vectors (x_n takes up in $\lambda \in \Lambda$ the value $x_n(\lambda)$ which appears in the assumption of Sec. 2) which fulfils the conditions of II. 1. 4 Prop. 4 in Ref. 15. Hence we can define a measurable structure for $\Pi_{\lambda\in\Lambda} \mathcal{K}_0(\lambda)$ and the Hilbert space

$$\mathcal{K}_\varphi := \int_\Lambda^{\oplus} \mathcal{K}_0(\lambda)d\mu(\lambda).$$

We point out that since Λ has been assumed to be a standard Borel space, $\mathcal{K}_\varphi$ results in a separable Hilbert space (II. 1. 6 Coroll. in Ref. 15). We shall denote by $R(\mathcal{K}_\varphi)$ and $D(\mathcal{K}_\varphi)$ ($\mathfrak{Z}$ in the notation of Ref. 15) the W^*-algebras of decomposable and diagonal operators on $\mathcal{K}_\varphi$ and by $\mathfrak{L}_\varphi$ the set of the projections of $R(\mathcal{K}_\varphi)$ (for the content of this paragraph, see Chap. II in Ref. 15).

The set $\mathfrak{L}_\varphi$ is easily shown to be a logic under the usual partial ordering and orthocomplementation for projections. We will prove in fact that $\mathfrak{L}_\varphi$ gives rise to a suitable algebraic representation of φ-propositions.

Proposition 3.1: Let i be the mapping

$$i : \mathfrak{D}_\varphi \to \mathfrak{L}_\varphi, \quad i(p) = \int_\Lambda^{\oplus} p(\lambda)d\mu(\lambda) \quad \text{if } p = \overset{\oplus}{\underset{\lambda\in\Lambda}{\vee}} p(\lambda).$$

Then i is a homomorphism of $\mathfrak{D}_\varphi$ onto $\mathfrak{L}_\varphi$.

In $\Delta_\varphi := \mathfrak{D}_\varphi/\sim$ (where $\sim$ is the relation of Remark 2. 2) a partial ordering and an orthocomplementation can be defined such that Δ_φ results in a logic and the natural mapping ϕ of $\mathfrak{D}_\varphi$ onto Δ_φ in a homomorphism. An isomorphism j of Δ_φ with $\mathfrak{L}_\varphi$ exists such that i can be decomposed into $i = j\circ\phi$.

Proof: First we notice that i is well defined. If $p = \vee_{\lambda\in\Lambda}^{\oplus} p(\lambda)$ is an element of $\mathfrak{D}_\varphi$, then $\Lambda \ni \lambda \to p(\lambda)$ is a μ-measurable field of projections in the μ-measurable field of Hilbert spaces $\Pi_{\lambda\in\Lambda} \mathcal{K}_0(\lambda)$, by definition of $\mathfrak{D}_\varphi$ and by II. 2. 1 Prop. 1 in Ref. 15.

Next we show that i is a homomorphism onto. If $\{p_n\}$ is a sequence of elements of $\mathfrak{D}_\varphi$, both $\wedge_n i(p_n)$ and $i(\wedge_n p_n)$ are elements of $\mathfrak{L}_\varphi$. Moreover, they have the same range, since, for an element $x = \int_\Lambda^{\oplus} x(\lambda)d\mu(\lambda)$ of $\mathcal{K}_\varphi$, we get

$$i(\wedge_n p_n)x = \int_\Lambda^\oplus (\wedge_n p_n(\lambda))x(\lambda)d\mu(\lambda) = x$$

iff $(\wedge_n p_n(\lambda))x(\lambda) = x(\lambda)$, μ a.e.,

iff $p_n(\lambda)x(\lambda) = x(\lambda)$, $\forall n, \mu$ a.e.,

iff $p_n(\lambda)x(\lambda) = x(\lambda)$, μ a.e., $\forall n$,

iff $i(p_n)x = x$, $\forall n$,

iff $\wedge_n i(p_n)x = x$.

Hence $i(\wedge_n p_n) = \wedge_n i(p_n)$. Denoting by $\tilde{p}$ the orthocomplement $p' \wedge I_{\mathfrak{D}_\varphi}$ of an element p of $\mathfrak{D}_\varphi$, from the proof of Prop. 2. 1 we have $\tilde{p}(\lambda) = p_0(\lambda) - p(\lambda)$, whence $i(\tilde{p}) = \int_\Lambda^\oplus p_0(\lambda)d\mu(\lambda) - \int_\Lambda^\oplus p(\lambda)d\mu(\lambda) = 1_{\mathcal{H}_\varphi} - i(p)$ (where $1_{\mathcal{H}_\varphi}$ denotes the unit operator on $\mathcal{H}_\varphi$). Hence i fulfils condition (b) of Footnote 25. Since i preserves the meet and the orthocomplementation, it preserves also the join. Therefore, it is a homomorphism. It is onto by definition of $\mathcal{L}_\varphi$.

Let $p_1 = \vee_{\lambda\in\Lambda}^\oplus p_1(\lambda)$ and $p_2 = \vee_{\lambda\in\Lambda}^\oplus p_2(\lambda)$ be two elements of $\mathfrak{D}_\varphi$. Because of II. 2. 3 Coroll. in Ref. 15, the condition $i(p_1) = i(p_2)$ is equivalent to $p_1(\lambda) = p_2(\lambda)$ μ a.e. Before introducing in Δ_φ a structure of orthocomplemented partially ordered set, we need to prove that these two conditions are in turn equivalent to $p_1 \sim p_2$. Suppose in fact $p_1(\lambda) = p_2(\lambda)$ μ a.e. Then

$$\varphi_p(p_1) = [1/\varphi(p)] \int_\Lambda \varphi_\lambda^{(p)}(p_1(\lambda))d\mu(\lambda)$$
$$= [1/\varphi(p)] \int_\Lambda \varphi_\lambda^{(p)}(p_2(\lambda))d\mu(\lambda) = \varphi_p(p_2)$$

holds for any $p \in \mathfrak{D}_\varphi^0$, whence

$$\langle \alpha_{p_1} \rangle_{\varphi_p} = \varphi_p(p_1) = \varphi_p(p_2) = \langle \alpha_{p_2} \rangle_{\varphi_p}, \quad \forall p \in \mathfrak{D}_\varphi^0,$$

and this implies (Remark 2. 1) $\langle \alpha_{p_1} \rangle_\psi = \langle \alpha_{p_2} \rangle_\psi, \forall \psi \in S_\varphi$, namely (Remark 2. 2) $p_1 \sim p_2$. If conversely $p_1 \sim p_2$, then in particular

$$\varphi_p(p_1) = [1/\varphi(p)] \int_\Lambda \sum_n (p(\lambda)x_n(\lambda) | p_1(\lambda)p(\lambda)x_n(\lambda))d\mu(\lambda)$$
$$= [1/\varphi(p)] \int_\Lambda \sum_n (p(\lambda)x_n(\lambda) | p_2(\lambda)p(\lambda)x_n(\lambda))d\mu(\lambda)$$
$$= \varphi_p(p_2), \quad \forall p \in \mathfrak{D}_\varphi^0,$$

whence, setting $q = \int_\Lambda^\oplus q(\lambda)d\mu(\lambda) := i(p_1) - i(p_2)$, we get

$$[1/\varphi(p)] \int_\Lambda \sum_n (p(\lambda)x_n(\lambda) | q(\lambda)p(\lambda)x_n(\lambda))d\mu(\lambda) = 0, \quad \forall p \in \mathfrak{D}_\varphi^0,$$

and this in turn implies

$$\int_\Lambda \sum_n (p(\lambda)x_n(\lambda) | q(\lambda)p(\lambda)x_n(\lambda))d\mu(\lambda) = 0, \quad \forall p \in \mathfrak{D}_\varphi$$

[if $\varphi(p) = 0$, then $p(\lambda)x_n(\lambda) = 0$, μ a.e.]. Let now

$$X := \{x \in \prod_{\lambda\in\Lambda} \mathcal{H}_0(\lambda); x \text{ square integrable and } \|x(\lambda)\| \in \{0, 1\}, \forall \lambda \in \Lambda\}$$

and, for $x \in X$,

$$p_x(\lambda): \mathcal{H}(\lambda) \to \mathcal{H}(\lambda), \quad p_x(\lambda)y_0 = (x(\lambda)|y_0)x(\lambda)$$

[$x(\lambda)$ denotes the λ-component of x]. Then $\vee_{\lambda\in\Lambda}^\oplus p_x(\lambda) \in \mathfrak{D}_\varphi$ for any $x \in X$, since $p_x(\lambda) \leq p_0(\lambda)$, $\forall \lambda \in \Lambda$ [$x(\lambda) \in \mathcal{H}_0(\lambda), \forall \lambda \in \Lambda$] and $(x_n(\lambda)|p_x(\lambda)x_m(\lambda)) = (x_n(\lambda)|x(\lambda))$ $(x(\lambda)|x_m(\lambda))$ is a μ-measurable function, $\forall n, m$ (II. 1. 4 Prop. 2 in Ref. 15). Writing the last equality to 0 above with p_x in the place of p, we get

$$0 = \int_\Lambda \sum_n (x_n(\lambda)|x(\lambda))(x(\lambda)|x_n(\lambda))\,(x(\lambda)|q(\lambda)x(\lambda))d\mu(\lambda)$$
$$= \int_\Lambda \sum_n (x_n(\lambda)|p_x(\lambda)x_n(\lambda))\,(x(\lambda)|q(\lambda)x(\lambda))d\mu(\lambda)$$
$$= \int_\Lambda \varphi_\lambda(p_x(\lambda))(x(\lambda)|q(\lambda)x(\lambda))\,d\mu(\lambda), \quad \forall x \in X.$$

If $x \in X$ and $E \in \mathfrak{B}_\mu$ then $\chi_E x \in X$, whence

$$0 = \int_\Lambda \varphi_\lambda(p_{\chi_E x(\lambda)})\chi_E(\lambda)(x(\lambda)|q(\lambda)x(\lambda))d\mu(\lambda)$$
$$= \int_\Lambda \chi_E(\lambda)\varphi_\lambda(p_x(\lambda))(x(\lambda)|q(\lambda)x(\lambda))\,d\mu(\lambda),$$
$$\forall E \in \mathfrak{B}_\mu, \quad \forall x \in X.$$

Hence $\varphi_\lambda(p_x(\lambda))(x(\lambda)|q(\lambda)x(\lambda)) = 0$, μ a.e., $\forall x \in X$.[36] If $\lambda \in \Lambda$ is such that $\varphi_\lambda(p_x(\lambda)) = 0$, then $p_x(\lambda)x_n(\lambda) = 0$, $\forall n$, whence $p_x(\lambda) = O_\lambda$ [since $x(\lambda) \in \mathcal{H}_0(\lambda)$ and $\mathcal{H}_0(\lambda)$ is spanned by $\{x_n(\lambda)\}$], and this implies $x(\lambda) = 0$. Therefore $(x(\lambda)|q(\lambda)x(\lambda)) = 0$, μ a.e., $\forall x \in X$. If $\hat{y}$ is now any element of $\mathcal{H}_\varphi$, let y be a square integrable field of vectors such that $\hat{y} = \int_\Lambda^\oplus y(\lambda)d\mu(\lambda)$ and x be the field of vectors

$$x(\lambda) := \begin{cases} (\|y(\lambda)\|)^{-1}y(\lambda) & \text{if } y(\lambda) \neq 0, \\ 0 & \text{if } y(\lambda) = 0. \end{cases}$$

The field of vectors x is an element of X, whence

$$(\hat{y}|q\hat{y}) = \int_\Lambda (y(\lambda)|q(\lambda)y(\lambda))d\mu(\lambda)$$
$$= \int_\Lambda \|y(\lambda)\|^2(x(\lambda)|q(\lambda)x(\lambda))d\mu(\lambda) = 0,$$

since $(x(\lambda)|q(\lambda)x(\lambda)) = 0$, μ a.e. Therefore, $q = O_{\mathcal{H}_\varphi}$ (the null operator on $\mathcal{H}_\varphi$). In this way we have shown that $i(p_1) = i(p_2)$ follows from $p_1 \sim p_2$.

Now we define in Δ_φ a partial ordering and an orthocomplementation. If $\phi(p_1)$ and $\phi(p_2)$ are two elements of Δ_φ, we define: $\phi(p_1) \leq \phi(p_2)$ iff $p_1(\lambda) \leq p_2(\lambda)$, μ a.e. (this definition is consistent since, for p and q in $\mathfrak{D}_\varphi$, $p \sim q$ is equivalent to $p(\lambda) = q(\lambda)$, μ a.e.). The relation now introduced in Δ_φ is easily shown to be a partial ordering, with least element $\emptyset_{\Delta_\varphi} = \phi(\emptyset_{\mathfrak{D}_\varphi})$ and greatest element $I_{\Delta_\varphi} = \phi(I_{\mathfrak{D}_\varphi})$. The mapping $\Delta_\varphi \ni \phi(p) \to (\phi(p))' := \phi(\tilde{p}) \in \Delta_\varphi$ is well defined [since, for p and q in $\mathfrak{D}_\varphi$, $p \sim q$ is equivalent to $p(\lambda) = q(\lambda)$, μ a.e., namely to $p_0(\lambda) - p(\lambda) = p_0(\lambda) - q(\lambda)$, μ a.e., namely to $\tilde{p} \sim \tilde{q}$] and it is easily shown to be an orthocomplementation.

We will now show that Δ_φ is a σ-complete lattice. If $\{\phi(p_n)\}$ is a sequence of elements of Δ_φ, then for an element $\phi(p)$ of Δ_φ we have $\phi(p) \geq \phi(\vee_n p_n)$ iff $p(\lambda) \geq \vee_n p_n(\lambda)$, μ a.e., iff $p(\lambda) \geq p_n(\lambda)$, $\forall n$, μ a.e., iff $p(\lambda) \geq p_n(\lambda)$, μ a.e., $\forall n$, iff $\phi(p) \geq \phi(p_n)$, $\forall n$. Therefore, $\vee_n \phi(p_n)$ exists and $\vee_n \phi(p_n) = \phi(\vee_n p_n)$. In a similar way the existence of $\wedge_n \phi(p_n)$ can be proved along with $\wedge_n \phi(p_n) = \phi(\wedge_n p_n)$. To prove weak modularity in Δ_φ, take $\phi(p)$ and $\phi(q)$ in Δ_φ such that $\phi(p) \leq \phi(q)$. Since $p(\lambda) \leq q(\lambda)$, μ a.e., and $\mathcal{P}(\mathcal{H}_0(\lambda))$ is a logic for any $\lambda \in \Lambda$, then, by property 2, ii, Chap. VI, Sec. 1 in Ref. 21, we get

$$q(\lambda) = p(\lambda) \vee ((p_0(\lambda) - p(\lambda)) \wedge q(\lambda))$$
$$= (p \vee (\tilde{p} \wedge q))(\lambda), \quad \mu \text{ a.e.}$$

Hence $\phi(q) = \phi(p \vee (\tilde{p} \wedge q)) = \phi(p) \vee ((\phi(p))' \wedge \phi(q))$. Therefore, weak modularity (property 2, ii, Chap. VI, Sec. 1 in Ref. 21) has been proved in Δ_φ, which results in a logic.

The mapping ϕ results in a homomorphism of $\mathfrak{D}_\varphi$ onto Δ_φ, since we have $\phi(\vee_n p_n) = \vee_n \phi(p_n)$ and $\phi(\wedge_n p_n) = \wedge_n \phi(p_n)$ for any sequence $\{p_n\}$ of elements of $\mathfrak{D}_\varphi$ along with $\phi(\tilde{p}) = (\phi(p))'$ for any element p of $\mathfrak{D}_\varphi$ (by the very definition of the orthocomplementation in Δ_φ).

Let j be the mapping

$$j: \Delta_\varphi \to \mathcal{L}_\varphi, \qquad j(\phi(p)) = i(p)$$

[this definition is consistent, since we have shown that, for p_1 and p_2 in $\mathfrak{D}_\varphi$, $p_1 \sim p_2$ is equivalent to $i(p_1) = i(p_2)$]. It is a homomorphism since

$$j(\vee_n \phi(p_n)) = j(\phi(\vee_n p_n)) = i(\vee_n p_n) = \vee_n i(p_n) = \vee_n j(\phi(p_n))$$

for any sequence $\{\phi(p_n)\}$ in Δ_φ, the analogous relation holds for the meet and

$$j(\phi(p))') = i(\tilde{p}) = 1_{\mathcal{K}_\varphi} - i(p) = 1_{\mathcal{K}_\varphi} - j(\phi(p)), \qquad \forall \phi(p) \in \Delta_\varphi.$$

The mapping j is onto since i is onto. Moreover, if $\phi(p) \in \Delta_\varphi$ is such that $j(\phi(p)) = 0_{\mathcal{K}_\varphi}$, then $i(p) = 0_{\mathcal{K}_\varphi}$, whence $p(\lambda) = 0_\lambda$, μ a.e. and this in turn implies $p \sim \emptyset_{\mathfrak{D}_\varphi}$, namely $\phi(p) = \emptyset_{\Delta_\varphi}$. Hence j is one-to-one (Prop. 1.5 in Ref. 23). Therefore j is an isomorphism of Δ_φ with $\mathcal{L}_\varphi$ and, by the definition of j, $i = j \circ \phi$. QED

We next come to the problem of finding an algebraic representation of φ-observables. To avoid unessential complications, we will limit ourselves to bounded φ-observables. If also unbounded φ-observables were taken into account, the same results that we are going to expose would hold, if in Prop. 3.2 the class of the Hermitian elements of $R(\mathcal{K}_\varphi)$ were enlarged to include also the self-adjoint operators affiliated with $R(\mathcal{K}_\varphi)$ [namely self-adjoint operators whose spectral projections are elements of $R(\mathcal{K}_\varphi)$] and in the discussion after Prop. 3.3 care were taken for the domains of the operators and the existence of expectation values. We remark that if α is a φ-observable, then, because of Prop. 3.1, $i \circ \alpha$ is a projection-valued measure in $\mathcal{K}_\varphi$ which defines a self-adjoint operator—usually denoted by $\int_R t d(i\circ\alpha)(t)$—on a suitable domain.[37] We shall denote by R^h the set of the Hermitian elements of $R(\mathcal{K}_\varphi)$.

Proposition 3.2: Let ω be the mapping

$$\omega: O_\varphi^0 \to R^h, \qquad \omega(\alpha) = \int_R t d(i\circ\alpha)(t).$$

(a) If R is the equivalence relation on O_φ^0 of Def. 2.2, then ω can be decomposed into $\omega = \omega_1 \circ \omega_2$, where ω_2 is the natural mapping of O_φ^0 onto O_φ^0/R and

$$\omega_1: O_\varphi^0/\mathrm{R} \to R^h, \ \omega_1(\omega_2(\alpha)) = \omega(\alpha)$$

is a one-to-one mapping.

(b) Let

$$\Omega_\varphi^0 := \{\hat{\alpha} \in \hom(\mathfrak{B}_\mathrm{R}, \Delta_\varphi);\ \hat{\alpha} \text{ bounded and } \hat{\alpha}(\mathbb{R}) = I_{\Delta_\varphi}\};$$

then

$$\rho: O_\varphi^0/\mathrm{R} \to \Omega_\varphi^0, \qquad \rho(\omega_2(\alpha)) = \phi\circ\alpha$$

is a bijection and ω can be decomposed into $\omega = \omega_1' \circ \omega_2'$, where

$$\omega_2' := \rho\circ\omega_2 \quad [\omega_2'(\alpha) = \phi\circ\alpha, \forall \alpha \in O_\varphi^0]$$

is a mapping of O_φ^0 onto Ω_φ^0 and $\omega_1' := \omega_1 \circ \rho^{-1}$ [$\omega_1'(\hat{\alpha}) = \omega(\alpha), \forall \hat{\alpha} \in \Omega_\varphi^0$, if $\alpha \in O_\varphi^0$ is such that $\phi\circ\alpha = \hat{\alpha}$] is a bijection of Ω_φ^0 with R^h.

(c) Along with ω_1', also ω and ω_1 are mappings onto R^h, whence ω_1 is a bijection of O_φ^0/R with R^h.

Proof: First we show that the definition of ω is consistent. If $\alpha \in O_\varphi^0$ then $i\circ\alpha$ is a bounded projection-valued measure and $\int_R t d(i\circ\alpha)(t)$ is a Hermitian (namely bounded and self-adjoint) operator. Its spectral projections $(i\circ\alpha)(E)$, $E \in \mathfrak{B}_\mathrm{R}$, are elements of the W^*-algebra $R(\mathcal{K}_\varphi)$ on account of the definition of i. Hence $\int_R t d(i\circ\alpha)(t)$ is an element of R^h.[38]

Next we establish a fact which turns out to be useful in the sequel. If $\alpha \in O_\varphi^0$, then

$$\begin{aligned}\langle\alpha\rangle_{\varphi_p} &= \int_R t d(\varphi_p\circ\alpha)(t) \overset{(0)}{=} \lim_{s\to\infty} \sum_{k=-s!}^{s!} t_k^{(s)} \varphi_p(\alpha(E_k^{(s)})) \\ &= \lim_{s\to\infty} \sum_{k=-s!}^{s!} t_k^{(s)} \left(\sum_{n=1}^{\infty} (y_n^{(p)} | i(\alpha(E_k^{(s)})) y_n^{(p)})_{\mathcal{K}_\varphi} \right) \\ &\overset{(1)}{=} \sum_{n=1}^{\infty} \left(y_n^{(p)} \middle| \lim_{s\to\infty} \sum_{k=-s!}^{s!} t_k^{(s)} i(\alpha(E_k^{(s)})) y_n^{(p)} \right)_{\mathcal{K}_\varphi} \\ &\overset{(2)}{=} \sum_{n=1}^{\infty} (y_n^{(p)} | \omega(\alpha) y_n^{(p)})_{\mathcal{K}_\varphi}, \qquad \forall p \in \mathfrak{D}_\varphi^0,\end{aligned}$$

where

$$y_n^{(p)} := \frac{1}{[\varphi(p)]^{1/2}} \int_\Lambda^{\oplus} p(\lambda) x_n(\lambda) d\mu(\lambda)$$

[$\lambda \to p(\lambda)x_n(\lambda)$ is a square integrable field of vectors since $p \in \mathfrak{D}_\varphi^0$], $E_k^{(s)} = [k/2^s, (k+1)/2^s)$ and $t_k^{(s)} := k/2^s$ if $k \geq 0$, $:= (k+1)/2^s$ if $k < 0$; setting

$$f_s := \sum_{k=-s!}^{s!} t_k^{(s)} \chi_{E_k^{(s)}}$$

we have that $\{f_s\}$ is a sequence of simple Borel functions on $\mathbb{R}$ such that $|f_s(t)| \leq |t|$ and

$$\lim_{s\to\infty} f_s(t) = t, \qquad \forall t \in \mathbb{R},$$

whence (0) holds[39];

restricting now the integral $\int_R t d(\varphi_p\circ\alpha)(t)$ and the functions f_s to the spectrum $\sigma(\alpha)$ of α, which is a bounded Borel set of $\mathbb{R}$ (see p. 110 of Ref. 21), we notice that

$$\sup_{t\in\sigma(\alpha)} |t - f_s(t)| \to 0$$

and from this it follows that

$$\begin{aligned}\omega(\alpha) = \int_{\sigma(\alpha)} t d(i\circ\alpha)(t) &= \lim_{s\to\infty} \sum_{k=-s!}^{s!} t_k^{(s)} i(\alpha(E_k^{(s)}) \cap \sigma(\alpha)) \\ &= \lim_{s\to\infty} \sum_{k=-s!}^{s!} t_k^{(s)} i(\alpha(E_k^{(s)}))\end{aligned}$$

where convergence is in the uniform topology of bounded operators on $\mathcal{K}_\varphi$[40]; from this argument the equality[2] follows as well as[1] if we take into account that the mapping

$$\mathfrak{B}(\mathcal{K}_\varphi) \ni A \to \sum_{n=1}^{\infty} (y_n^{(p)} | A y_n^{(p)}) \in \mathbb{C}$$

[where $\mathfrak{B}(\mathcal{K}_\varphi)$ stands for the algebra of bounded operators on $\mathcal{K}_\varphi$ and $\mathbb{C}$ for the complex field] is continuous with respect to the uniform topology of operators (it is in fact a normal state). In this way we have proved that

$$(*) \qquad \langle\alpha\rangle_{\varphi_p} = \sum_{n=1}^{\infty} (y_n^{(p)} | \omega(\alpha) y_n^{(p)}), \qquad \forall \alpha \in O_\varphi^0, \forall p \in \mathfrak{D}_\varphi^0.$$

(a) Let α_1 and α_2 be two elements of O_φ^0 such that $\omega(\alpha_1) = \omega(\alpha_2)$. From (*) then we get $\langle\alpha_1\rangle_{\varphi_p} = \langle\alpha_2\rangle_{\varphi_p}$,

$\forall p \in \mathfrak{D}^0_\varphi$, whence (Remark 2.1) $\alpha_1 \, R \, \alpha_2$. If conversely $\alpha_1 \, R \, \alpha_2$, then in particular $\langle \alpha_1 \rangle_{\varphi_p} = \langle \alpha_2 \rangle_{\varphi_p}$, $\forall p \in \mathfrak{D}^0_\varphi$, and by (*)

$$\sum_{n=1}^{\infty} (y_n^{(p)} | \omega(\alpha_1) y_n^{(p)}) = \sum_{n=1}^{\infty} (y_n^{(p)} | \omega(\alpha_2) y_n^{(p)}), \qquad \forall p \in \mathfrak{D}^0_\varphi,$$

whence, setting

$$A = \int_\Lambda^{\oplus} A(\lambda) d\mu(\lambda) := \omega(\alpha_1) - \omega(\alpha_2),$$

$$\frac{1}{\varphi(p)} \int_\Lambda \sum_{n=1}^{\infty} (p_x(\lambda) x_n(\lambda) | A(\lambda) p_x(\lambda) x_n(\lambda)) d\mu(\lambda) = 0, \qquad \forall x \in X$$

[where X and $p_x(\lambda)$ have the same meaning as in the proof of Prop. 3.1]. Then we have (by the same arguments as in the proof of Prop. 3.1) $A = \mathbb{O}_{R(\mathcal{H}_\varphi)}$, namely $\omega(\alpha_1) = \omega(\alpha_2)$. Thus, for any two elements α_1 and α_2 of O^0_φ, $\omega(\alpha_1) = \omega(\alpha_2)$ is equivalent to $\alpha_1 \, R \, \alpha_2$. Therefore the definition of ω_1 is consistent and ω_1 is one-to-one. The equality $\omega = \omega_1 \circ \omega_2$ holds by the definition of ω_1.

(b) The definition of ρ is consistent and ρ is one-to-one since, for two elements α_1 and α_2 of O^0_φ, $\phi \circ \alpha_1 = \phi \circ \alpha_2$ is equivalent to $\alpha_1 \, R \, \alpha_2$. We have in fact $\phi \circ \alpha_1 = \phi \circ \alpha_2$ iff (the mapping j of Prop. 3.1 is one-to-one) $i \circ \alpha_1 = j \circ \phi \circ \alpha_1 = j \circ \phi \circ \alpha_2 = i \circ \alpha_2$ iff (by the definition of ω, since the correspondence between self-adjoint operators and projection valued measures is one-to-one, see Theorem VIII. 6 in Ref. 37) $\omega(\alpha_1) = \omega(\alpha_2)$ iff [see the proof of part (a)] $\alpha_1 \, R \, \alpha_2$. Besides, ρ is onto, since if $\hat{\alpha}$ is an element of Ω^0_φ, an $\alpha_0 \in O^0_\varphi$ exists such that $\hat{\alpha} = \phi \circ \alpha_0$ by the Appendix (with $\mathfrak{D}_\varphi$ and Δ_φ in the place of $\mathfrak{L}$ and $\hat{\mathfrak{L}}$ and ϕ in the place of $\hat{h}$).

The mapping ω'_2 is from O^0_φ onto Ω^0_φ, since both ω_2 and ρ are onto. The mapping ω'_1 is one-to-one since both ρ^{-1} and ω_1 are one-to-one. We shall now show that it is onto. Let in fact A be an element of R^h; then[38] $P \in$ hom $(\mathfrak{B}_R, \mathfrak{L}_\varphi)$ with $P(\mathbb{R}) = 1_{\mathcal{H}_\varphi}$ exists such that $A = \int_R t dP(t)$. The mapping $j^{-1} \circ P$ is an element of Ω^0_φ and α in O^0_φ exists such that $\phi \circ \alpha = j^{-1} \circ P$ (since ω'_2 is onto), whence

$$\omega'_1(j^{-1} \circ P) = \omega(\alpha) = \int_R t d(i \circ \alpha)(t) = \int_R t d(j \circ \phi \circ \alpha)(t) = \int_R t dP(t) = A.$$

The decomposition $\omega = \omega'_1 \circ \omega'_2$ follows from $\omega'_1 \circ \omega'_2 = \omega_1 \circ \omega_2$.

(c) The mapping ω is from O^0_φ onto R^h since both ω'_1 and ω'_2 are onto. Therefore, ω_1 as well is onto R^h and thus it results in a bijection of O^0_φ / R onto R^h. QED

The task of obtaining an algebraic representation of the logic concepts introduced in Sec. 2 will be accomplished if a suitable representation for φ-states will be found. Because of the important role they will play with respect to this problem, we need to define properly the logic states related to Δ_φ. First we notice that, for $\hat{p} \in \Delta^0_\varphi := \phi(\mathfrak{D}^0_\varphi)$ and $p \in \mathfrak{D}^0_\varphi$ such that $\hat{p} = \phi(p)$, the mapping

$$\hat{\varphi}_{\hat{p}} : \Delta_\varphi \to [0, 1], \qquad \hat{\varphi}_{\hat{p}}(\phi(q)) = \varphi_p(q)$$

is defined in a consistent way just by the definition of the relation $\sim$, and

(a) $\hat{\varphi}_{\hat{p}}(\mathcal{O}_{\Delta_\varphi}) = \varphi_p(\mathcal{O}_{\mathfrak{D}_\varphi}) = 0$, $\hat{\varphi}_{\hat{p}}(I_{\Delta_\varphi}) = \varphi_p(I_{\mathfrak{D}_\varphi}) = 1$,

(b) if $\{\phi(p_n)\}$ is a sequence of mutually orthogonal elements of Δ_φ, then $p_n(\lambda) \leq p_0(\lambda) - p_m(\lambda)$, μ a.e., $\forall n, m$, whence [since the family $\{\phi(p_n)\}$ is countable] $p_n(\lambda) \leq p_0(\lambda) - p_m(\lambda), \forall n, m, \mu$ a.e. ($\forall \lambda \notin E, E \in \mathfrak{B}$, $\mu(E) = 0$). We can choose other representatives for $\phi(p_n)$ (for instance turning $p_n(\lambda)$ into $\mathbb{O}_\lambda$ for $\lambda \in E$), which we denote still by p_n, such that $\{p_n\}$ is a sequence of mutually orthogonal elements of $\mathfrak{D}_\varphi$. Thus we get

$$\hat{\varphi}_{\hat{p}}(\vee_n \phi(p_n)) = \varphi_p(\vee_n p_n) = \sum_n \varphi_p(p_n) = \sum_n \hat{\varphi}_{\hat{p}}(\phi(p_n)).$$

Definition 3.1: Let Σ_φ be the set of the mappings $\hat{\psi} : \Delta_\varphi \to [0, 1]$ such that

(a) $\hat{\psi}(\mathcal{O}_{\Delta_\varphi}) = 0$, $\hat{\psi}(I_{\Delta_\varphi}) = 1$,

(b) $\hat{\psi}(\vee_n \phi(p_n)) = \sum_n \hat{\psi}(\phi(p_n))$ for any sequence $\{\phi(p_n)\}$ of mutually orthogonal elements of Δ_φ,

(c) $\langle \hat{\alpha} \rangle_{\hat{\psi}} + \langle \hat{\beta} \rangle_{\hat{\psi}} = \langle \hat{\gamma} \rangle_{\hat{\psi}}$ whenever $\hat{\alpha}, \hat{\beta}, \hat{\gamma} \in \Omega^0_\varphi$ are such that $\langle \hat{\alpha} \rangle_{\hat{\varphi}_{\hat{p}}} + \langle \hat{\beta} \rangle_{\hat{\varphi}_{\hat{p}}} = \langle \hat{\gamma} \rangle_{\hat{\varphi}_{\hat{p}}}, \forall \hat{p} \in \Delta^0_\varphi$, where

$$\langle \hat{\alpha} \rangle_{\hat{\psi}} := \int_R t d(\hat{\psi} \circ \hat{\alpha})(t).$$

The elements of Σ_φ are the equivalents in Δ_φ to the elements of S_φ in $\mathfrak{D}_\varphi$, since $\hat{\varphi}_{\hat{p}}$ is nothing else than φ_p "transferred" on Δ_φ. We point out in particular that $\mathcal{C}(\Delta_\varphi)$ could be shown to be isomorphic with $\mathfrak{B}/\mathfrak{N}$, where $\mathfrak{N}$ stands for the class of μ-negligible elements of $\mathfrak{B}$ (see at the end of this section) as well as the restriction of $\hat{\varphi}_{\hat{p}}$ to $\mathfrak{B}/\mathfrak{N}$ could be shown to be the measure defined on $\mathfrak{B}/\mathfrak{N}$ by μ_p.

Proposition 3.3: Let $R_{*,1}$ be the set of normal functionals on $R(\mathcal{H}_\varphi)$ of unit norm.

(a) If $\psi \in S_\varphi$, then

$$f_\psi : R(\mathcal{H}_\varphi) \to \mathbb{C}, \qquad f_\psi(A) = \langle \alpha_1 \rangle_\psi + i \langle \alpha_2 \rangle_\psi \text{ if } \alpha_1, \alpha_2 \in O^0_\varphi$$

are such that

$$\omega(\alpha_1) = A_1 := \tfrac{1}{2}(A + A^\dagger), \omega(\alpha_2) = A_2 := \tfrac{1}{2} i(-A + A^\dagger)$$

is an element of $R_{*,1}$. The mapping

$$f : S_\varphi \to R_{*,1}, \qquad f(\psi) = f_\psi$$

is a bijection.

(b) If $\psi \in S_\varphi$, then

$$\sigma_\psi : \Delta_\varphi \to [0, 1], \qquad \sigma_\psi(\phi(p)) = \psi(p),$$

is an element of Σ_φ. The mapping

$$\sigma : S_\varphi \to \Sigma_\varphi, \qquad \sigma(\psi) = \sigma_\psi,$$

is a bijection and $f' := f \circ \sigma^{-1}$ [$f'(\hat{\psi}) = f_\psi, \forall \hat{\psi} \in \Sigma_\varphi$, if $\psi \in S_\varphi$ is such that $\sigma_\psi = \hat{\psi}$] is a bijection of Σ_φ with $R_{*,1}$.

Proof: First we show that the definition of f_ψ is consistent. Elements α_1 and α_2 like those involved in the definition of f_ψ always exist in O^0_φ since ω maps O^0_φ onto R^h. Moreover, for $\alpha'_1, \alpha'_2 \in O^0_\varphi$ such that $\omega(\alpha'_1) = A_1$ and $\omega(\alpha'_2) = A_2$, we get $\omega(\alpha_1) = \omega(\alpha'_1)$ and $\omega(\alpha_2) = \omega(\alpha'_2)$, whence [see the proof of part (a) of Prop. 3.2] $\langle \alpha_1 \rangle_\psi = \langle \alpha'_1 \rangle_\psi$ and $\langle \alpha_2 \rangle_\psi = \langle \alpha'_2 \rangle_\psi, \forall \psi \in S_\varphi$.

Next we prove that f_ψ is indeed an element of $R_{*,1}$. If $A, B, C \in R(\mathcal{H}_\varphi)$ are such that $A + B = C$, then $C_1 = A_1 + B_1$ and $C_2 = A_2 + B_2$. Let $\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2 \in O^0_\varphi$ be such that $\omega(\alpha_1) = A_1, \ldots, \omega(\gamma_2) = C_2$. In view of (*) in the proof of Prop. 3.2 we have

$$\langle \gamma_k \rangle_{\varphi_p} = \sum_{n=1}^{\infty} (y_n^{(p)} | C_k y_n^{(p)}) = \sum_{n=1}^{\infty} (y_n^{(p)} | A_k y_n^{(p)}) + \sum_{n=1}^{\infty} (y_n^{(p)} | B_k y_n^{(p)}) = \langle \alpha_k \rangle_{\varphi_p} + \langle \beta_k \rangle_{\varphi_p}$$

for $k = 1, 2, \forall p \in \mathfrak{D}_\varphi^0$. Therefore, by the definition of S_φ,

$$\langle \gamma_k \rangle_\psi = \langle \alpha_k \rangle_\psi + \langle \beta_k \rangle_\psi \qquad (k = 1, 2)$$

whence

$$f_\psi(C) = \langle \gamma_1 \rangle_\psi + i\langle \gamma_2 \rangle_\psi$$
$$= (\langle \alpha_1 \rangle_\psi + i\langle \alpha_2 \rangle_\psi) + (\langle \beta_1 \rangle_\psi + i\langle \beta_2 \rangle_\psi) = f_\psi(A) + f_\psi(B).$$

Having proved in this way the additivity of f_ψ, we will now show that $f_\psi(kA) = kf_\psi(A), \forall' k \in \mathbb{C}, \forall A \in R(\mathcal{K}_\varphi)$. First notice that, for $k \in \mathbb{R}$ and $\alpha \in O_\varphi^0$, $\omega(k\alpha) = k\omega(\alpha)$ holds, since

$$\omega(k\alpha) = \int_{\mathbb{R}} t d(i \circ k\alpha)(t) = \int_{\mathbb{R}} t d(i \circ \alpha)(k^{-1}t)$$
$$= k \int_{\mathbb{R}} t' d(i \circ \alpha)(t') = k\omega(\alpha),$$

if $k \neq 0$, and $\omega(0\alpha) = \mathbb{0}_{R(\mathcal{K}_\varphi)} = 0\omega(\alpha)$ by the definition of 0α. For $k \in \mathbb{C}$ let k_1 and k_2 be its real and imaginary parts. For any $A \in R(\mathcal{K}_\varphi)$, from $kA = k_1A_1 + ik_2A_1 + ik_1A_2 - k_2A_2$, and from the additivity of f_ψ we get $f_\psi(kA) = f_\psi(k_1A_1) + f_\psi(ik_2A_1) + f_\psi(ik_1A_2) + f_\psi(-k_2A_2)$. If α_1, $\alpha_2 \in O_\varphi^0$ are such that $\omega(\alpha_n) = A_n (n = 1, 2)$, then, by the remark above, by the definition of f_ψ, and by Remark 2.1, we obtain

$$f_\psi(kA) = k_1\langle \alpha_1 \rangle_\psi + ik_2\langle \alpha_1 \rangle_\psi + ik_1\langle \alpha_2 \rangle_\psi - k_2\langle \alpha_2 \rangle_\psi$$
$$= k_1 f_\psi(A_1) + ik_2 f_\psi(A_1) + k_1 f_\psi(iA_2) + ik_2 f_\psi(iA_2)$$
$$= (k_1 + ik_2)(f_\psi(A_1) + f_\psi(iA_2)),$$

whence, by the additivity of f_ψ,

$$f_\psi(kA) = kf_\psi(A_1 + iA_2) = kf_\psi(A).$$

Thus f_ψ is a linear functional on $R(\mathcal{K}_\varphi)$. If A is a positive element of R^h, then $\alpha_0 \in O_\varphi^0$ exists such that $\omega(\alpha_0) = A$ and $\alpha_0(E_0) = I_{\mathfrak{D}_\varphi}$ with $E_0 \in \mathfrak{B}_{\mathbb{R}}$ and $E_0 \subset [0, \infty)$. Let in fact $\alpha \in O_\varphi^0$ be such that $\omega(\alpha) = A$. Since $i \circ \alpha$ is the projection valued measure of A and A is positive, $E_0 \in \mathfrak{B}_{\mathbb{R}}$ exists such that $E_0 \subset [0, \infty)$, $E_0 \neq \emptyset$, and $i \circ \alpha(E_0) = \mathbb{1}_{\mathcal{K}_\varphi}$. Take $t_0 \in E_0$ and define the mapping

$$\alpha_0 : \mathfrak{B}_{\mathbb{R}} \to \mathfrak{D}_\varphi,$$

$$\alpha_0(E) = \begin{cases} (\alpha(E) \wedge \alpha(E_0)) \vee (\alpha(E_0))^\sim, & \text{if } t_0 \in E_0, \\ \alpha(E) \wedge \alpha(E_0), & \text{if } t_0 \notin E_0. \end{cases}$$

The mapping α_0 is easily seen to be a φ-observable and $\omega(\alpha_0) = A$ since $i \circ \alpha_0(E) = i(\alpha(E)) \wedge i(\alpha(E_0)) = i(\alpha(E))$, $\forall E \in \mathfrak{B}_{\mathbb{R}}$. Moreover, $\alpha_0(E_0) = \alpha(E_0) \vee (\alpha(E_0))^\sim = I_{\mathfrak{D}_\varphi}$. Then

$$f_\psi(A) = \langle \alpha_0 \rangle_\psi = \int_{\mathbb{R}} t d(\psi \circ \alpha_0)(t) \geq 0$$

since the measure $\psi \circ \alpha_0$ has the support contained in $E_0 \subset [0, \infty)$. In this way we have shown that f_ψ is a positive linear functional on $R(\mathcal{K}_\varphi)$. Let now $\{P_n\}$ be a sequence of mutually orthogonal projections of $R(\mathcal{K}_\varphi)$. In the same way as in (b) of $\hat{\varphi}_{\hat{p}}$ before Def. 3.1, we can find a sequence $\{p_n\}$ of mutually orthogonal elements of $\mathfrak{D}_\varphi$ such that $i(p_n) = P_n, \forall_n$. From the definitions of $\omega(\alpha)$ and $\langle \alpha \rangle_\psi$ for $\alpha \in O_\varphi^0$, we get $\omega(\alpha_{p_n}) = i(p_n) = P_n$ and $\langle \alpha_{p_n} \rangle_\psi = \psi(p_n)$ and (since i is a homomorphism and from Ref. 41)

$$\omega(\alpha_{\vee_n p_n}) = i(\vee_n p_n) = \vee_n i(p_n) = \sum_n P_n$$

along with (since ψ is σ-additive)

$$\langle \alpha_{\vee_n p_n} \rangle_\psi = \psi(\vee_n p_n) = \sum_n \psi(p_n).$$

Hence, we get, from the definition of f_ψ,

$$f_\psi\left(\sum_n P_n\right) = \langle \alpha_{\vee_n p_n} \rangle_\psi = \sum_n \psi(p_n) = \sum_n \langle \alpha_{p_n} \rangle_\psi = \sum_n f_\psi(P_n).$$

Because of the separability of $\mathcal{K}_\varphi$, and family of mutually orthogonal and different from zero projections of $R(\mathcal{K}_\varphi)$ is at most countable. Therefore $f_\psi \in R_*$ (I. 4, exercise 9, in Ref. 15). Finally we notice that

$$\omega\left(\alpha_{I_{\mathfrak{D}_\varphi}}\right) = \mathbb{1}_{\mathcal{K}_\varphi},$$

whence

$$f_\psi(\mathbb{1}_{\mathcal{K}_\varphi}) = \langle \alpha_{I_{\mathfrak{D}_\varphi}} \rangle_\psi = \psi(I_{\mathfrak{D}_\varphi}) = 1.$$

Hence f_ψ has unit norm[42] and this completes the proof that $f_\psi \in R_{*,1}$.

Next we shall establish that f is a mapping from S_φ onto $R_{*,1}$. Let f_0 be an element of $R_{*,1}$. Then the mapping

$$\psi_0 : \mathfrak{D}_\varphi \to [0, 1], \qquad \psi_0(p) = f_0(i(p)),$$

is an element of S, since properties (b) and (c) of Sec. 1 hold as i is a homomorphism and $f_0(\mathbb{1}_{\mathcal{K}_\varphi}) = 1$. Take now $\alpha, \beta, \gamma \in O_\varphi^0$ such that $\langle \alpha \rangle_{\varphi_p} + \langle \beta \rangle_{\varphi_p} = \langle \gamma \rangle_{\varphi_p}, \forall p \in \mathfrak{D}_\varphi^0$. Then with the procedure used in the proof of part (a) of Prop. 3.2 it can be shown that $\omega(\alpha) + \omega(\beta) = \omega(\gamma)$. We observe finally that, $\forall \alpha \in O_\varphi^0$,

$$\langle \alpha \rangle_{\psi_0} = \lim_{s \to \infty} \sum_{k=-s!}^{s!} t_k^{(s)} \psi_0(\alpha(E_k^{(s)})) = \lim_{s \to \infty} \sum_{k=-s!}^{s!} t_k^{(s)} f_0(i(\alpha(E_k^{(s)})))$$
$$= f_0\left(\lim_{s \to \infty} \sum_{k=-s!}^{s!} t_k^{(s)} i(\alpha(E_k^{(s)}))\right) = f_0(\omega(\alpha)),$$

where the symbols have the same meaning and the equalities hold for the same reasons as in establishing (*) in the proof of Prop. 3.2 (since f_0 is a normal state). Hence we get

$$\langle \alpha \rangle_{\psi_0} + \langle \beta \rangle_{\psi_0} = f_0(\omega(\alpha) + \omega(\beta)) = f_0(\omega(\gamma)) = \langle \gamma \rangle_{\psi_0}$$

and this shows that $\psi_0 \in S_\varphi$. Construct now f_{ψ_0}. For any $A \in R(\mathcal{K}_\varphi)$ we have $f_{\psi_0}(A) = \langle \alpha_1 \rangle_{\psi_0} + i\langle \alpha_2 \rangle_{\psi_0}$, if α_1, $\alpha_2 \in O_\varphi^0$ are such that $\omega(\alpha_k) = A_k (k = 1, 2)$. Since $\langle \alpha_k \rangle_{\psi_0} = f_0(\omega(\alpha_k))$ $(k = 1, 2)$, from the linearity of f_0 we have $f_{\psi_0}(A) = f_0(A_1) + if_0(A_2) = f_0(A)$. Therefore $f_{\psi_0} = f_0$ and f is onto.

Finally f is one-to-one since, for $\psi_1, \psi_2 \in S_\varphi$, $f(\psi_1) = f(\psi_2)$ implies $f_{\psi_1}(A) = f_{\psi_2}(A), \forall A \in R(\mathcal{K}_\varphi)$, whence $\langle \alpha \rangle_{\psi_1} = \langle \alpha \rangle_{\psi_2}, \forall \alpha \in O_\varphi^0$, which in turn implies $\psi_1(p) = \langle \alpha_p \rangle_{\psi_1} = \langle \alpha_p \rangle_{\psi_2} = \psi_2(p), \forall p \in \mathfrak{D}_\varphi$, namely $\psi_1 = \psi_2$. This completes the proof that f is a bijection.

(b) The definition of σ_ψ is consistent since, if $p_1, p_2 \in \mathfrak{D}_\varphi$ are such that $p_1 \sim p_2$, then $\psi(p_1) = \psi(p_2)$ holds just by the definition of the relation $\sim$. We have now to show that $\sigma_\psi \in \sum_\varphi$. Properties (a) and (b) of Def. 3.1 can be shown for σ_ψ in exactly the same way as they have been shown for $\hat{\varphi}_{\hat{p}}$ before Def. 3.1. Besides, for $\hat{\alpha} \in \Omega_\varphi^0$ and $\alpha \in O_\varphi^0$ such that $\phi \circ \alpha = \hat{\alpha}$ (α exists for any $\hat{\alpha}$ since ω_2' is onto), we get $\langle \hat{\alpha} \rangle_{\hat{\varphi}_{\hat{p}}} = \int_{\mathbb{R}} t d(\hat{\varphi}_{\hat{p}} \circ \hat{\alpha})(t) = \int_{\mathbb{R}} t d(\varphi_p \circ \alpha)(t) = \langle \alpha \rangle_{\varphi_p}, \forall \hat{p} \in \Delta_\varphi^0$ and $p \in \mathfrak{D}_\varphi^0$ such that $\phi(p) = \hat{p}$, as well as $\langle \hat{\alpha} \rangle_{\sigma_\psi} = \langle \alpha \rangle_\psi$, whence property (c) of Def. 3.1 for σ_ψ since $\psi \in S_\varphi$

The mapping σ will now be shown to be a bijection. It is one-to-one, since from $\sigma_{\psi_1}(\hat{p}) = \sigma_{\psi_2}(\hat{p}), \forall \hat{p} \in \Delta_\varphi$, it follows that $\psi_1(p) = \psi_2(p), \forall p \in \mathfrak{D}_\varphi$, namely $\psi_1 = \psi_2$. Moreover it is onto. In fact, let $\hat{\psi}$ be an element of Σ_φ.

The mapping

$$\psi: \mathfrak{D}_\varphi \to [0, 1], \quad \psi(p) = \hat{\psi}(\phi(p)).$$

is an element of S_φ since:

(i) $\psi(\emptyset_{\mathfrak{D}_\varphi}) = \hat{\psi}(\emptyset_{\Delta_\varphi}) = 0$ and $\psi(I_{\mathfrak{D}_\varphi}) = \hat{\psi}(I_{\Delta_\varphi}) = 1$;

(ii) If $\{p_n\}$ is a sequence of mutually orthogonal elements of $\mathfrak{D}_\varphi$ the same holds for $\{\phi(p_n)\}$ in Δ_φ since ϕ is a homomorphism, whence

$$\psi(\vee_n p_n) = \hat{\psi}(\phi(\vee_n p_n)) = \hat{\psi}(\vee_n \phi(p_n)) = \sum_n \hat{\psi}(\phi(p_n)) = \sum_n \psi(p_n);$$

(iii) we have shown above that $\langle\hat{\alpha}\rangle_{\hat{\varphi}_{\hat{p}}} = \langle\alpha\rangle_{\varphi_p}, \forall p \in \mathfrak{D}_\varphi^0$ and $\hat{p} = \phi(p), \forall \alpha \in O_\varphi^0$ with $\hat{\alpha} = \omega_2'(\alpha)$; the equality $\langle\hat{\alpha}\rangle_{\hat{\psi}} = \langle\alpha\rangle_\psi$ can be proved in the same way; thus $\psi \in S_\varphi$.

We have now trivially $\sigma_\psi(\phi(p)) = \psi(p) = \hat{\psi}(\phi(p))$, $\forall \phi(p) \in \Delta_\varphi$. This completes the proof that σ is a bijection.

The mapping f' is a bijection since it results from the composition of two bijections. QED

With this theorem we have accomplished the task of finding an algebraic representation of the logic concepts which are ruled by the "prestate" of the assumption of Sec. 2. For φ-propositions (the elements of the logic $\mathfrak{D}_\varphi$) we have in fact found the homomorphism i onto $\mathfrak{L}_\varphi$, for φ-observables (the elements of O_φ^0) the mapping ω onto R^h, and for φ-states (the elements of S_φ) the mapping f onto $R_{*,1}$ in such a way that, $\forall \psi \in S_\varphi$,

$$f_\psi(i(p)) = \psi(p), \quad \forall p \in \mathfrak{D}_\varphi,$$

[since $i(p) = \omega(\alpha_p)$ and $\psi(p) = \langle\alpha_p\rangle_\psi$] and

$$f_\psi(\omega(\alpha)) = \langle\alpha\rangle_\psi, \quad \forall \alpha \in O_\varphi^0$$

(by the definition of f_ψ).

Hence the algebraic representation conserves the experimentally meaningful quantities of the theory, which are probability functions on the logic of propositions and expectation values for observables.

At the logic level, indeed, it seems sensible to consider as proper φ-propositions and φ-observables the elements of Δ_φ and $O_\varphi^0/\mathbf{R}$ rather than the elements of $\mathfrak{D}_\varphi$ and O_φ^0 since this amounts to considering physically equivalent the propositions and the observables which are undistinguishable whatever the state is in which the system is prepared (this attitude is, for instance, the starting point in Ludwig's approach[43]). There is no need to construct equivalence classes in S_φ since the elements of S_φ have been shown to be already separated by the observables (see after Remark 2.2). Also with respect to the algebraic representation things work well for Δ_φ, $O_\varphi^0/\mathbf{R}$, and S_φ since for them we have constructed faithful algebraic representations (through j, ω_1, and f). We meet anyway a very serious difficulty: While Δ_φ is suitable to be interpreted as a set of logic propositions since it is indeed a logic, neither the elements of $O_\varphi^0/\mathbf{R}$ are homomorphisms from $\mathfrak{B}_\mathbf{R}$ into the "correct" logic Δ_φ (hence they are not logic observables related to Δ_φ) nor the elements of S_φ are σ-additive "measures" normalized to one on Δ_φ (hence they are not logic states related to Δ_φ). Anyway we get rid of this difficulty since we have shown that a bijection exists of $O_\varphi^0/\mathbf{R}$ with Ω_φ^0 as well as of S_φ with Σ_φ and Ω_φ^0 and Σ_φ are observables and states (in the proper logic sense) related to the propositions Δ_φ. These "correct" sets of propositions (Δ_φ), observables (Ω_φ^0), and states (Σ_φ) admit an "isomorphic" algebraic representation in $\mathcal{K}_\varphi$, since there are an isomorphism j of Δ_φ with $\mathfrak{L}_\varphi$, a bijection ω_1' of Ω_φ^0 with R^h and a bijection f' of Σ_φ with $R_{*,1}$ such that

$$(f'(\hat{\psi}))(j(\hat{p})) = \hat{\psi}(\hat{p}), \quad \forall \hat{\psi} \in \Sigma_\varphi, \quad \forall \hat{p} \in \Delta_\varphi$$

and

$$(f'(\hat{\psi}))(\omega_1'(\hat{\alpha})) = \langle\hat{\alpha}\rangle_{\hat{\psi}}, \quad \forall \hat{\psi} \in \Sigma_\varphi, \forall \hat{\alpha} \in \Omega_\varphi^0.$$

In this way we have deduced, in dependence of a suitable prestate φ, from the logic picture the usual algebraic picture based on W^*-algebras.[10,11] Indeed, we have shown that for the logic picture an algebraic isomorphic representation exists in which the experimental values are conserved. In the representation of Ω_φ^0 we can show that also the superobservables are conserved. First we need a remark about the center $\mathcal{C}(\Delta_\varphi)$ of Δ_φ.

Remark 3.1: The equality

$$\mathcal{C}(\Delta_\varphi) = \{\phi(p); p \in \mathfrak{D}_\varphi, p(\lambda) \in \{\mathbb{O}_\lambda, p_0(\lambda)\}, \mu \text{ a.e.}\}$$

holds. From Lemma 6.7 in Ref. 21 we get that the compatibility of two elements $\phi(p)$, $\phi(q)$ of Δ_φ is equivalent to $(\phi(p) \wedge \phi(q))' \wedge \phi(q) = (\phi(p))' \wedge \phi(q)$, which amounts to $(p_0(\lambda) - p(\lambda) \wedge q(\lambda)) \wedge q(\lambda) = (p_0(\lambda) - p(\lambda)) \wedge q(\lambda)$, μ a.e. namely to the compatibility of $p(\lambda)$ and $q(\lambda)$, μ a.e. Therefore, an element $\phi(p)$ of Δ_φ lies in the center of Δ_φ iff $[p(\lambda), q(\lambda)] = \mathbb{O}_\lambda$, μ a.e., $\forall q \in \mathfrak{D}_\varphi$ and this condition in turn can be shown to be equivalent to $p(\lambda) \in \{\mathbb{O}_\lambda, p_0(\lambda)\}$, μ a.e. (one side of the equivalence is trivial, for the other side proceed by absurd).

Proposition 3.4: The superobservables are conserved by ω_1', the bijection of Ω_φ^0 with R^h. If $\hat{\alpha} \in \Omega_\varphi^0$, then $\omega_1'(\hat{\alpha})$ is an algebraic superobservable [namely $\omega_1'(\hat{\alpha}) \in D(\mathcal{K}_\varphi)$, the W^*-algebra of diagonal operators on $\mathcal{K}_\varphi$, and $\omega_1'(\hat{\alpha}) \neq k\mathbf{1}_{\mathcal{K}_\varphi}, \forall k \in \mathbf{R}$] iff $\hat{\alpha}$ is a superobservable in Ω_φ^0 [namely $\hat{\alpha}(E) \in \mathcal{C}(\Delta_\varphi), \forall E \in \mathfrak{B}$, and $\hat{\alpha}$ is not constant, that is $\hat{\alpha}(\{k\}) \neq I_{\Delta_\varphi}, \forall k \in \mathbb{R}$].

Proof: For $\hat{\alpha} \in \Omega_\varphi^0$ let $\alpha \in O_\varphi^0$ be such that $\hat{\alpha} = \phi \circ \alpha$ (such an α exists because ω_1' is onto). Then $\omega_1'(\hat{\alpha}) \in D(\mathcal{K}_\varphi)$ iff $i(\alpha(E)) \in D(\mathcal{K}_\varphi), \forall E \in \mathfrak{B}_\mathbf{R}$ [$i \circ \alpha$ is the projection valued measure of $\omega_1'(\alpha)$ by the definition of ω_1', then use Ref. 38 since $D(\mathcal{K}_\varphi)$ is a W^*-algebra], namely iff $\alpha(E)(\lambda) \in \{p_0(\lambda), \mathbb{O}_\lambda\}$, μ a.e., $\forall E \in \mathfrak{B}_\mathbf{R}$ [use the definitions of i and of $D(\mathcal{K}_\varphi)$], namely iff $\phi(\alpha(E)) \in \mathcal{C}(\Delta_\varphi), \forall E \in \mathfrak{B}_\mathbf{R}$ (Remark 3.1), namely iff $\hat{\alpha}(E) \in \mathcal{C}(\Delta_\varphi), \forall E \in \mathfrak{B}_\mathbf{R}$. Moreover, we get, for $k \in \mathbf{R}$, $\omega_1'(\hat{\alpha}) = k\mathbf{1}_{\mathcal{K}_\varphi}$ iff $i(\alpha(\{k\})) = \mathbf{1}_{\mathcal{K}_\varphi}$ iff $\alpha(\{k\})(\lambda) = p_0(\lambda)$, μ a.e., iff $\phi(\alpha(\{k\})) = I_{\Delta_\varphi}$ iff $\hat{\alpha}(\{k\}) = I_{\Delta_\varphi}$. QED

We want now to point out in what sense our description can include also continuous superselection rules. Consider the quotient algebra $\mathfrak{B}/\mathfrak{N}$, namely the quotient set of $\mathfrak{B}$ with respect to the equivalence relation

$$E_1 \sim E_2 \text{ iff } E_1 \Delta E_2 \in \mathfrak{N} \quad (E_1, E_2 \in \mathfrak{B}),$$

where $E_1 \Delta E_2$ stands for the symmetric difference of E_1 and E_2 and $\mathfrak{N} := \{E \in \mathfrak{B}; \mu(E) = 0\}$; on $\mathfrak{B}/\mathfrak{N}$ it is defined the structure of σ-algebra such that the natural map from $\mathfrak{B}$ onto $\mathfrak{B}/\mathfrak{N}$ results in a homomorphism (Sec. 40 in Ref. 36). Since we have assumed that $\mathfrak{B}$ defines on Λ a structure of standard Borel space, $\mathfrak{B}$ has a countable set of generators and the same holds obviously for $\mathfrak{B}/\mathfrak{N}$.

Hence, from Sec. 40, Theorem B in Ref. 36 and Chap. 14, Sec. 2, Problem 12 in Ref. 39, we get that $\mathfrak{B}/\mathfrak{N}$ is isomorphic either to $\mathfrak{B}_{[a,b]}/m$ (where $\mathfrak{B}_{[a,b]}$ is the Borel σ-algebra of an interval $[a,b]$ and m the Lebesgue measure on it), to a discrete σ-algebra containing at most a countable number of elements, or to a σ-algebra which is the direct union of the preceding two.

Since the mapping

$$\sigma: \mathfrak{B}/\mathfrak{N} \to \mathcal{C}(\Delta_\varphi),\ \sigma([E]) = \phi\Big(\bigvee_{\lambda\in\Lambda}^{\oplus} \chi_E(\lambda)p_0(\lambda)\Big)$$

($[E]$ is the equivalence class containing the element E of $\mathfrak{B}$) can be easily seen to be an isomorphism of $\mathfrak{B}/\mathfrak{N}$ with $\mathcal{C}(\Delta_\varphi)$, then $\mathcal{C}(\Delta_\varphi)$ is isomorphic to one of the three σ-algebras listed above in dependence with Λ and φ ($\mathfrak{N}$ is in fact determined, through μ, by φ). In the first and in the third case continuous superselection rules are present, while in the second case they can be only discrete. To this statement a definite meaning will now be given. Consider for instance the first case. Then continuous superselection rules are present not only in the sense that Λ is a noncountable set in $\mathcal{L} = \bigvee_{\lambda\in\Lambda}^{\oplus} \mathcal{L}^{(\lambda)}$,[5] but in the precise sense that a superobservable α exists in Ω^0_φ with purely continuous spectrum, namely such that $\alpha(\{t\}) = 0$ for any real number t in the spectrum of α. We can prove in fact an even stronger result. Since $\mathcal{C}(\Delta_\varphi)$ is a separable sub-σ-algebra of Δ_φ ($\mathfrak{B}/\mathfrak{N}$ has a countable set of generators) then, from the theorems on p. 100 and 101 in Ref. 20, the existence of a "total" superobservable α_0 in Ω^0_φ follows such that any superobservable α in Ω^0_φ can be expressed as $\alpha = f^{(\alpha)}(\alpha_0)$ with a suitable real valued Borel function $f^{(\alpha)}$ on $\mathbb{R}$. We are now going to construct an α_0 with these properties and such that its spectrum is purely continuous (hence any other "total" superobservable has to share this property). Take the function

$$\tau: [a,b] \to \mathbb{R},\qquad \tau(t) = t;$$

then

$$\alpha_0: \mathfrak{B}_{[a,b]} \to \mathcal{C}(\Delta_\varphi),\qquad \alpha_0(E) = \sigma\circ\pi\circ\tau^{-1}(E)$$

(where π is the natural mapping of $\mathfrak{B}_{[a,b]}$ onto $\mathfrak{B}_{[a,b]}/m$, which has been identified with $\mathfrak{B}/\mathfrak{N}$ for brevity) is a superobservable in Ω^0_φ with the required properties.

In a similar way, in the second case the spectrum of a "total" superobservable would result to be discrete (this is the situation examined in Ref. 10), while in the third case it would have a continuous as well as a discrete part.

Finally, we notice that the physical interpretation of a "total" superobservable depends on the particular physical system described by $\mathcal{L}$ and φ.

4. THE CLASSICAL CASE

We will now examine what results to be the construction of the previous sections if the physical system we are dealing with is a classical one. In the classical case, $\mathcal{L}$ is assumed to be a Boolean algebra,[20,21] whence $\mathcal{K}(\lambda)$ has to be a one-dimensional Hilbert space for every λ in Λ (we have supposed in fact $\mathcal{K}(\lambda)$ nontrivial for any λ in Λ; if $\lambda_0 \subset \Lambda$ could exist such that dim $\mathcal{K}_0(\lambda) > 1$, then $\mathcal{P}(\mathcal{K}_0(\lambda))$ would not be a Boolean algebra, and this would entail nondistributivity also for $\mathcal{L}$). Then

$$\chi: \mathbb{P}(\Lambda) \to \mathcal{L},\qquad \chi(S) = \bigvee_{\lambda\in\Lambda}^{\oplus} \chi_S(\lambda)\mathbb{1}_{\mathcal{K}(\lambda)}$$

is an isomorphism of the logic $\mathbb{P}(\Lambda)$ (the power set of Λ equipped with set-theoretical join, meet and complementation) with $\mathcal{L}$. The sequence $\{x_n(\lambda)\}$ of the assumption reduces in this case to $\{x(\lambda)\}$ with $\|x(\lambda)\| = 1, \forall\lambda\in\Lambda$. Through χ the sublogic $\mathfrak{D}_\varphi$ of $\mathcal{L}$ is isomorphic with the μ-completion $\mathfrak{B}_\mu$ of the σ-algebra $\mathfrak{B}$ and any state on $\mathfrak{D}_\varphi$ defines a probability measure on $\mathfrak{B}_\mu$; in particular φ defines in this way the completion of μ, which is an extension of μ to $\mathfrak{B}_\mu$ which we will denote still by μ. No distinction will be made hereafter between $\mathfrak{D}_\varphi$ and $\mathfrak{B}_\mu$ and between a state on $\mathfrak{D}_\varphi$ and the measure it defines on $\mathfrak{B}_\mu$. If $\mathfrak{F}^0$ is the set of bounded μ-measurable real functions on Λ, then

$$\eta: \mathfrak{F}^0 \to O^0_\mu,\qquad \eta(g) = g^{-1},$$

(where g^{-1} stands for the "inverse image" of g restricted to $\mathfrak{B}_\mu$) is a bijection of $\mathfrak{F}^0$ with O^0_μ (as it can be easily deduced from Theorem 1.4 of Ref. 21) such that

$$\langle\eta(g)\rangle_\nu = \int_{\mathbb{R}} td(\nu\circ g^{-1})(t) = \int_\Lambda g(\lambda)d\nu(\lambda) = \nu(g)$$

for any $g \in \mathfrak{F}^0$ and for any state ν on $\mathfrak{B}_\mu$ (the second equality holds by Ref. 44).

If E_0 is an element of $\mathfrak{B}_\mu$ such that $\mu(E_0) \neq 0$, then the "conditional state with respect to E_0 and μ" turns out to be the probability measure μ_{E_0} on $\mathfrak{B}_\mu$ defined as

$$\mu_{E_0}(E) = [1/\mu(E_0)]\int_\Lambda \chi_{E_0}(\lambda)\chi_E(\lambda)d\mu(\lambda) = \mu(E\cap E_0)/\mu(E_0),\qquad \forall E \in \mathfrak{B}_\mu.$$

We will now prove that the set S_μ of μ-states (Def. 2.1) coincides with the set of probability measures on $\mathfrak{B}_\mu$ which are absolutely continuous with respect to μ. Let ν be an element of S_μ and E an element of $\mathfrak{B}_\mu$ such that $\mu(E) = 0$; for the elements χ_Λ and χ_E of $\mathfrak{F}^0$ we have $\langle\eta(\chi_\Lambda)\rangle_{\mu_{E_0}} = \int_\Lambda 1d\mu_{E_0}(\lambda) = \mu_{E_0}(\Lambda) = 1$ and $\langle\eta(\chi_E)\rangle_{\mu_{E_0}} = \int_\Lambda \chi_E(\lambda)d\mu_{E_0}(\lambda) = \mu(E\cap E_0)/\mu(E_0) = 0$, whence $\langle\eta(\chi_\Lambda)\rangle_{\mu_{E_0}} + \langle\eta(\chi_E)\rangle_{\mu_{E_0}} = \langle\eta(\chi_\Lambda)\rangle_{\mu_{E_0}}, \forall E_0 \in \mathfrak{B}_\mu$. Therefore, we get $\langle\eta(\chi_\Lambda)\rangle_\nu + \langle\eta(\chi_E)\rangle_\nu = \langle\eta(\chi_\Lambda)\rangle_\nu$, which easily reduces to $\nu(\Lambda) + \nu(E) = \nu(\Lambda)$, namely to $\nu(E) = 0$. Conversely, let ν be a probability measure on $\mathfrak{B}_\mu$ absolutely continuous with respect to μ. If g_a, g_b, g_c are elements of $\mathfrak{F}^0$ such that $\langle\eta(g_a)\rangle_{\mu_{E_0}} + \langle\eta(g_b)\rangle_{\mu_{E_0}} = \langle\eta(g_c)\rangle_{\mu_{E_0}}$, namely

$$[1/\mu(E_0)]\int_{E_0}(g_a(\lambda) + g_b(\lambda) - g_c(\lambda))d\mu(\lambda) = 0,\quad \forall E_0 \in \mathfrak{B}_\mu \text{ for which } \mu(E_0) \neq 0$$

{since $d\mu_{E_0}/d\mu = [1/\mu(E_0)]\chi_{E_0}$}, then we get $\int_{E_0}(g_a(\lambda) + g_b(\lambda) - g_c(\lambda))d\mu(\lambda) = 0, \forall E_0 \in \mathfrak{B}_\mu$, whence $g_a(\lambda) + g_b(\lambda) - g_c(\lambda) = 0, \mu$ a.e.,[36] and this implies $g_a(\lambda) + g_b(\lambda) - g_c(\lambda) = 0, \nu$ a.e. From this equality we get

$$\int_\Lambda g_a(\lambda)d\nu(\lambda) + \int_\Lambda g_b(\lambda)d\nu(\lambda) = \int_\Lambda g_c(\lambda)d\nu(\lambda),$$

namely $\langle\eta(g_a)\rangle_\nu + \langle\eta(g_b)\rangle_\nu = \langle\eta(g_c)\rangle_\nu$, which completes the proof.

We notice that we can easily find elements of S_μ which do not reduce to any μ_{E_0}. Take in fact a probability measure ν on $\mathfrak{B}_\mu$ different from μ but equivalent to μ. Then no $E_0 \in \mathfrak{B}_\mu$ exists such that $\mu_{E_0} = \nu$, because of the only μ_{E_0} equivalent to μ is μ_Λ, namely μ itself, since $d\mu_{E_0}/d\mu = \chi_{E_0}$.

Let us now examine μ-propositions, μ-observables, and μ-states after the reduction into equivalence classes of undistinguishable elements.

First we notice that the equivalence relation of Remark 2.2 can be identified, through χ, with the following relation in $\mathfrak{B}_\mu$:

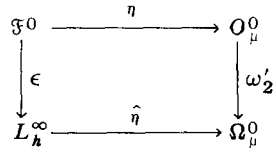
$$E_1 \sim E_2 \text{ iff } \mu(E_1 \triangle E_2) = 0, \text{ where } E_1, E_2 \in \mathfrak{B}_\mu.$$

We have in fact $\vee^{\oplus}_{\lambda\in\Lambda} \chi_{E_1}(\lambda)\mathbb{1}_{\mathcal{H}(\lambda)} \sim \vee^{\oplus}_{\lambda\in\Lambda} \chi_{E_2}(\lambda)\mathbb{1}_{\mathcal{H}(\lambda)}$ iff (see the proof of Prop. 3.1) $\chi_{E_1}(\lambda) = \chi_{E_2}(\lambda)$ μ-a.e. iff $\chi_{E_1\triangle E_2}(\lambda) = |\chi_{E_1}(\lambda) - \chi_{E_2}(\lambda)| = 0$, μ a.e., iff $\mu(E_1 \triangle E_2) = 0$. Hence Δ_μ can be identified with $\mathfrak{B}_\mu/\mathfrak{N}_\mu$, where $\mathfrak{N}_\mu = \{E \in \mathfrak{B}_\mu;\ \mu(E) = 0\}$, and the elements of Σ_μ with the measures which are defined on $\mathfrak{B}_\mu/\mathfrak{N}_\mu$ in a straightforward way by the elements of S_μ [Sec. 40, note (11), in Ref. 36].

If L^∞_h is the set of self-adjoint elements of the C^*-algebra $L^\infty(\Lambda, \mathfrak{B}_\mu, \mu)$ of equivalence classes of complex functions from $\mathcal{L}^\infty(\Lambda, \mathfrak{B}_\mu, \mu)$, then

$$\hat{\eta}: L^\infty_h \to \Omega^0_\mu, \qquad \hat{\eta}([g]) = \phi\circ g^{-1}$$

is a bijection of L^∞_h with Ω^0_μ such that the diagram



commutes, where ϵ is the natural mapping from $\mathfrak{F}^0$ onto L^∞_h. First we show that the definition of $\hat{\eta}$ is consistent. If g is a real element of $\mathcal{L}^\infty(\Lambda, \mathfrak{B}_\mu, \mu)$, then $g^{-1}\epsilon$ hom $(\mathfrak{B}_\mathbb{R}, \mathfrak{B}_\mu)$ with $g^{-1}(\mathbb{R}) = \Lambda$ and (identifying Δ_μ with $\mathfrak{B}_\mu/\mathfrak{N}_\mu$) $\phi\circ g^{-1} \in$ hom $(\mathfrak{B}_\mathbb{R}, \Delta_\mu)$ with $(\phi\circ g^{-1})(\mathbf{R}) = I_{\Delta_\mu}$; moreover, there exists a compact subset K of $\mathbb{R}$ such that $g(\lambda) \in K$, μ a.e., whence $g^{-1}(\mathbb{R} - K) \in \mathfrak{N}_\mu$ and $(\phi\circ g^{-1})(\mathbf{R} - K) = \emptyset_{\Delta_\mu}$; from this equality we get $(\phi\circ g^{-1})(K) = I_{\Delta_\mu}$, namely that $\phi\circ g^{-1}$ is bounded. Furthermore, for two real elements g_1 and g_2 of $\mathcal{L}^\infty(\Lambda, \mathfrak{B}_\mu, \mu)$, $g_1(\lambda) = g_2(\lambda)$, μ a.e., is equivalent to $\mu(g_1^{-1}(E) \triangle g_2^{-1}(E)) = 0$, namely to $\phi(g_1^{-1}(E)) = \phi(g_2^{-1}(E))$, $\forall E \in \mathfrak{B}_\mathbb{R}$. This completes the proof that the definition of $\hat{\eta}$ is consistent and also shows that $\hat{\eta}$ is one-to-one. To prove that $\hat{\eta}$ is onto, take any $\hat{\alpha} \in \Omega^0_\mu$; then, by the Appendix, an element $\alpha \in O^0_\mu$ exists such that $\hat{\alpha} = \phi\circ\alpha$, namely (since η is onto) an element g of $\mathfrak{F}^0$ exists such that $\hat{\alpha} = \phi\circ g^{-1}$; it is now sufficient to notice that $\mathfrak{F}^0$ is a subset of the real part of $\mathcal{L}^\infty(\Lambda, \mathfrak{B}_\mu, \mu)$. The commutativity of the diagram is clear from the definitions of the mappings involved.

Summing up, we have shown that if the system represented by $\mathcal{L}$ is a classical one, the logic picture described in the previous sections reduces to the special example of probability theory that one obtains when in a probability space a distinguished measure is fixed.[35] Here the probability space is $(\Lambda, \mathfrak{B}_\mu)$ and the distinguished measure is μ. Then the states can be identified with the probability measures which are absolutely continuous with respect to μ and the observables with the bounded real valued random variables in such a way that $\nu(g)$ is the expectation value of an observable which corresponds to the random variable g with respect to a state which corresponds to the measure ν. Finally, conditional states result in the usual conditional probabilities and, when undistinguishable elements of the theory are identified, we get $\mathfrak{B}_\mu/\mathfrak{N}_\mu$ and L^∞_h as the proper sets of propositions and observables. Therefore, the aforementioned model of probability theory can be completely deduced as a particular case of the logic picture of the previous sections.

We are now going to investigate what the algebraic representation of Sec. 3 becomes. The Hilbert space $\mathcal{H}_\varphi$ in which the representation is constructed is nothing else than $L^2(\Lambda, \mathfrak{B}_\mu, \mu)$. The elements $\hat{E}$ of $\mathfrak{B}_\mu/\mathfrak{N}_\mu$ are represented through the map j as $j(\hat{E}) = P_E$, where P_E is the projection on $L^2(\Lambda, \mathfrak{B}_\mu, \mu)$ defined by $(P_E x)(\lambda) = \chi_E(\lambda)x(\lambda)$, $\forall\lambda \in \Lambda$, $\forall x \in L^2(\Lambda, \mathfrak{B}_\mu, \mu)$ [in this notation we ignore the difference, in $L^2(\Lambda, \mathfrak{B}_\mu, \mu)$, between equivalence classes and representative elements thereof] and E is a representative of $\hat{E}$.

An element $[g]$ of L^∞_h {or, equivalently, $\hat{\eta}([g])$ of Ω^0_μ} is represented by the Hermitian operator on $L^2(\Lambda, \mathfrak{B}_\mu, \mu)$

$$\omega'_1(\hat{\eta}([g])) = \omega'_1(\phi\circ g^{-1}) = \omega(g^{-1}),$$

where g has been chosen in $[g]$ such that $g \in \mathfrak{F}^0$. Denoting by A_g the Hermitian operator on $L^2(\Lambda, \mathfrak{B}_\mu, \mu)$ defined by $(A_g x)(\lambda) = g(\lambda)x(\lambda)$, $\forall\lambda \in \Lambda$, $\forall x \in L^2(\Lambda, \mathfrak{B}_\mu, \mu)$, we get $\omega(g^{-1}) = A_g$, since, for any $x, y \in L^2(\Lambda, \mathfrak{B}_\mu, \mu)$,

$$\begin{aligned}(x|\omega(g^{-1})y) &\overset{(1)}{=} \int_\mathbb{R} t\,d(x|P_{g^{-1}(t)}y)\\ &\overset{(2)}{=} \int_\mathbb{R} t\,d\mu_{x,y}(g^{-1}(t))\\ &\overset{(3)}{=} \int_\Lambda g(\lambda)d\mu_{x,y}(\lambda)\\ &\overset{(4)}{=} \int_\Lambda (x(\lambda)|g(\lambda)y(\lambda))d\mu(\lambda) = (x|A_g y),\end{aligned}$$

where (1) holds by definition of ω, since $i(g^{-1}(E)) = P_{g^{-1}(E)}$, $\forall E \in \mathfrak{B}_\mathbb{R}$, (2) holds if $\mu_{x,y}$ denotes the complex measure on Λ defined by $\mu_{x,y}(E) = (x|P_E y)$, $\forall E \in \mathfrak{B}$, (3) holds by Ref. 44, (4) holds because $(d\mu_{x,y}/d\mu)(\lambda) = (x(\lambda)|y(\lambda))$, since, by definition of $\mu_{x,y}$

$$\mu_{x,y}(E) = \int_E (x(\lambda)|y(\lambda))d\mu(\lambda), \forall E \in \mathfrak{B}.$$

From Prop. 3.3 we know that any state from S_μ can be represented as a normal state on $R(L^2(\Lambda, \mathfrak{B}_\mu, \mu))$. In the present case, in correspondence with any element ν of S_μ it is possible to construct a definite vector x such that $\nu(A) = (x|Ax)$, $\forall A \in R(L^2(\Lambda, \mathfrak{B}_\mu, \mu))$. This is an easy task since now we know the structure of the Hilbert space and any element of S_μ is absolutely continuous with respect to μ. For $\nu \in S_\mu$ consider in fact the vector x of $L^2(\Lambda, \mathfrak{B}_\mu, \mu)$ given by

$$x(\lambda) = \left(\frac{d\nu}{d\mu}(\lambda)\right)^{1/2}, \qquad \forall\lambda \in \Lambda.$$

Then

$$\begin{aligned}f_\nu(A) &= \langle\alpha_1\rangle_\nu + i\langle\alpha_2\rangle_\nu = \int_\Lambda g_1(\lambda)d\nu(\lambda) + i\int_\Lambda g_2(\lambda)d\nu(\lambda)\\ &= \int_\Lambda g_1(\lambda)\frac{d\nu}{d\mu}(\lambda)d\mu(\lambda) + i\int_\Lambda g_2(\lambda)\frac{d\nu}{d\mu}(\lambda)d\mu(\lambda)\\ &= (x|(A_{g_1} + iA_{g_2})x)\\ &= (x|Ax), \qquad \forall A \in R(L^2(\Lambda, \mathfrak{B}_\mu, \mu)),\end{aligned}$$

where α_k, $k = 1, 2$, is an element of O^0_μ such that $\omega(\alpha_k) = A_k$ and g_k an element of $\mathfrak{F}^0$ such that $\alpha_k = g_k^{-1}$ (such α_k and g_k exist since ω and η are mappings onto), whence $A_{g_k} = A_k$ by the result of the preceding paragraph.

Therefore, when the physical system is a classical one, in the algebraic representation of the previous sections the observables are represented by the "multiplication operators" $A_g (g \in \mathfrak{F}^0)$ in the Hilbert space $L^2(\Lambda, \mathfrak{B}_\mu, \mu)$, where Λ can be interpreted as the phase space of the system and μ is a definite "prestate." A representation of this kind for a classical system was proposed with different motivations by Koopman[45] and has proved useful

in analyzing properties of a classical system, for instance in ergodic theory.[46] Hence Koopman's representation can be deduced as a particular case of our description.

APPENDIX

We shall state a theorem the framework of which is the functional calculus of observables. It is a generalization of Theorem 1.4 of Ref. 21, to which it reduces in the particular case when the logics involved are a Boolean σ-algebra $\mathfrak{L}$ of subset of a fixed set and a Boolean σ-algebra $\hat{\mathfrak{L}}$.

Theorem: Let $\mathfrak{L}$ and $\hat{\mathfrak{L}}$ be two logics and h a homomorphism of $\mathfrak{L}$ onto $\hat{\mathfrak{L}}$. Suppose further that $\hat{\alpha}$ is any $\hat{\mathfrak{L}}$-observable, namely $\hat{\alpha} \in \hom(\mathfrak{B}_{\mathbb{R}}, \hat{\mathfrak{L}})$ with $\hat{\alpha}(\mathbb{R}) = I_{\hat{\mathfrak{L}}}$. Then there exists an $\mathfrak{L}$-observable α, namely $\alpha \in \hom(\mathfrak{B}_{\mathbb{R}}, \mathfrak{L})$ with $\alpha(\mathbb{R}) = I_{\mathfrak{L}}$, such that $\hat{\alpha} = h \circ \alpha$.

Proof: Let $p, q \in \mathfrak{L}$ be such that $p \leq q$, whence $h(p) \leq h(q)$ since a homomorphism is order-preserving, and $\hat{r} \in \hat{\mathfrak{L}}$ be such that $h(p) \leq \hat{r} \leq h(q)$; since h is onto, $r_1 \in \mathfrak{L}$ exists such that $h(r_1) = \hat{r}$. If $r := (r_1 \wedge q) \vee p$, we have $p \leq r \leq q$ along with $h(r) = (h(r_1) \wedge h(q)) \vee h(p) = h(r_1) = \hat{r}$.

Let $r_1, r_2, \ldots, r_n, \ldots$ be an enumeration of the rational numbers in $\mathbb{R}$ and $D_i := (-\infty, r_i)$. Evidently $\hat{\alpha}(D_i) \leq \hat{\alpha}(D_j)$ whenever $r_i < r_j$. Now we shall construct a sequence $\{p_n\}$ of elements of $\mathfrak{L}$ such that

(a) $h(p_n) = \hat{\alpha}(D_n), \forall n$,

(b) $p_i \leq p_j$ whenever $r_i < r_j$.

For D_1, an element p_1 exists in $\mathfrak{L}$ such that $h(p_1) = \hat{\alpha}(D_1)$, since h is onto $\hat{\mathfrak{L}}$. Suppose $\{p_1, \ldots, p_k\}$ in $\mathfrak{L}$ have been found such that

(i) $h(p_n) = \hat{\alpha}(D_n), n = 1, \ldots, k$,

(ii) $p_i \leq p_j$ whenever $r_i < r_j, 1 \leq i, j \leq k$.

We shall construct p_{k+1} as follows. Let $(n_1, n_2, \ldots, n_k)$ be the permutation of $(1, 2, \ldots, k)$ such that $r_{n_1} < r_{n_2} < \ldots < r_{n_k}$. Then ther exists a unique i such that $r_{n_i} < r_{k+1} < r_{n_{i+1}}$ (we define $r_{n_0} = -\infty$ and $r_{n_{k+1}} = \infty$) and by the remark made in the first paragraph an element p_{k+1} exists in $\mathfrak{L}$ such that $p_{n_i} \leq p_{k+1} \leq p_{n_{i+1}}$ and $h(p_{k+1}) = \hat{\alpha}(D_{k+1})$. The collection $\{p_1, \ldots, p_k, p_{k+1}\}$ then has the same properties relative to $r_1, \ldots, r_k, r_{k+1}$ as $\{p_1, \ldots, p_k\}$ had relative to $r_1, \ldots, r_k$. Thus it follows by induction that there exists a sequence $\{p_n\}$ of elements of $\mathfrak{L}$ with the properties (a) and (b).

Let now $\mathfrak{A}(\{p_n\})$ be the sublogic of $\mathfrak{L}$ generated by $\{p_n\}$, namely the smallest sublogic of $\mathfrak{L}$ containing the family $\{p_n, p_n', \emptyset_{\mathfrak{L}}, I_{\mathfrak{L}}\}, \forall n$ ($'$ means orthocomplementation). By property (b) either $p_{n_1} \leq p_{n_2}$ or $p_{n_2} \leq p_{n_1}, \forall n_1, n_2$; hence $\mathfrak{A}(\{p_n\})$ is a Boolean sub-σ-algebra of $\mathfrak{L}$ by Corollary 6.15 in Ref. 21. By construction $\mathfrak{A}(\{p_n\})$ is separable; hence (Lemma 6.16 in Ref. 21) there exists an $\mathfrak{L}$-observable β such that $\mathfrak{A}(\{p_n\}) = \{\beta(E); E \in \mathfrak{B}_{\mathbb{R}}\}$. Since h is a homomorphism onto, $h(I_{\mathfrak{L}}) = I_{\hat{\mathfrak{L}}}$ and $h \circ \beta$ is an $\hat{\mathfrak{L}}$-observable such that

$$\{h \circ \beta(E); E \in \mathfrak{B}_{\mathbb{R}}\} = \mathfrak{A}(\{h(p_n)\}) = \mathfrak{A}(\{\hat{\alpha}(D_n)\}) = \{\hat{\alpha}(E); E \in \mathfrak{B}_{\mathbb{R}}\},$$

where property (a) has been taken into account. Therefore, a real valued Borel function u on $\mathbb{R}$ exists such that $\hat{\alpha} = \mu(h \circ \beta)$ (Theorem on p. 101 in Ref. 20). Finally we notice that

$$u(h \circ \beta)(E) = h \circ \beta(u^{-1}(E)) = h(\beta(u^{-1}(E))) = h(u(\beta)(E)) = h \circ u(\beta)(E), \quad \forall E \in \mathfrak{B}_{\mathbb{R}},$$

whence $\hat{\alpha} = h \circ \alpha$, if α is the $\mathfrak{L}$-observable $\alpha := u(\beta)$. QED

If the $\hat{\mathfrak{L}}$-observable $\hat{\alpha}$ of the theorem is bounded, then a bounded $\mathfrak{L}$-observable α_0 can be constructed such that $\hat{\alpha} = h \circ \alpha_0$. If α is the $\mathfrak{L}$-observable whose existence is asserted by the theorem and K is a compact subset of $\mathbb{R}$ such that $\hat{\alpha}(K) = I_{\hat{\mathfrak{L}}}$, take a t_0 in K and define for any $E \in \mathfrak{B}$

$$\alpha_0(E) := \begin{cases} \alpha(E \cap K) & \text{if } t_0 \notin E, \\ \alpha(E \cap K) \vee \alpha(K)' & \text{if } t_0 \in E \end{cases}$$

[$\alpha(K)'$ is the orthocomplement of $\alpha(K)$ in $\mathfrak{L}$]. The mapping α_0 can be easily shown to be an $\mathfrak{L}$-observable with the required properties.

[1] To quote just the textbooks: G. W. Mackey, *Mathematical Foundations of Quantum Mechanics* (Benjamin, New York, 1963); J. M. Jauch, *Foundations of Quantum Mechanics* (Addison-Wesley, Reading, Mass., 1968); V. S. Varadarajan, *Geometry of Quantum Theory, Vol. I* (Princeton U. P., Princeton, N. J., 1968); G. Ludwig, *Deutung des Begriffs "physikalische Theorie" und axiomatische Grundlegung der Hilbertraumstruktur durch Hauptsätze des Messens* (Springer, Berlin, 1970).

[2] G. Birkhoff and J. von Neumann, Ann. Math. **37**, 823 (1936).

[3] G. G. Emch and C. Piron, J. Math. Phys. **4**, 469 (1963); F. Gallone and A. Maniá, Ann. Inst. H. Poincaré **15 A**, 37 (1971); S. P. Gudder, Can. J. Math. **23**, 659 (1971).

[4] J. M. Jauch and C. Piron, Helv. Phys. Acta **36**, 827 (1963).

[5] C. Piron, Helv. Phys. Acta **42**, 330 (1969).

[6] I. E. Segal, Ann. Math. **48**, 930 (1947).

[7] P. Jordan, J. von Neumann, and E. Wigner, Ann. Math. **35**, 29 (1934).

[8] I. Gel'fand and M. A. Naimark, Mat. Sborn. **12**, (54), 197 (1943); Ref. 6.

[9] R. Haag and D. Kastler, J. Math. Phys. **5**, 848 (1964).

[10] R. Cirelli and F. Gallone, "Algebra of Observables and Quantum Logic," to be published on Ann. Inst. H. Poincaré.

[11] J. M. Jauch, Helv. Phys. Acta **33**, 711 (1960); J. M. Jauch and B. Misra, Helv. Phys. Acta **34**, 699 (1961).

[12] J. von Neumann, *Grundlagen der Quantenmechanik* (Springer, Berlin, 1932).

[13] G. F. Dell'Antonio, S. Doplicher, and D. Ruelle, Commun. Math. Phys. **2**, 223 (1966).

[14] see, for instance, G. G. Emch, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory* (Wiley-Interscience, New York, 1972), Chap. I, 2, d.

[15] J. Dixmier, *Les algèbres d'opérateurs dans l'espace Hilbertien (Algèbres de von Neumann)* (Gauthier-Villars, Paris, 1969), remark on p. 119 and II.2.4.

[16] G. Ludwig, Z. Phys. **181**, 223 (1964).

[17] E. Chen, J. Math Phys. **12**, 2364 (1971).

[18] R. J. Plymen, Helv. Phys. Acta **41**, 69 (1968).

[19] G. C. Wick, A. S. Wightman, and E. P. Wigner, Phys. Rev. **88**, 101 (1952).

[20] Jauch's book. Ref. 1.

[21] Varadarajan's book, Ref. 1. We shall denote by $\vee$ and $\wedge$ the join and the meet in any logic we will consider; the least and the greatest elements in a logic denoted by $\mathfrak{L}$ will be denoted by $\emptyset_{\mathfrak{L}}$ and $I_{\mathfrak{L}}$; for the orthocomplementation we shall use either the symbol $'$ or $\sim$

[22] C. Piron, Helv. Phys. Acta **37**, 439 (1964).

[23] F. Gallone and A. Maniá, Ann. Inst H. Poincaré **15 A**, 37 (1971).

[24] C. Piron, "Observables in General Quantum Theory," in *Foundations of Quantum Mechanics (Varenna)* (Academic, New York, 1971).

[25] If $\mathfrak{L}_1$ and $\mathfrak{L}_2$ are two logics, a homomorphism of $\mathfrak{L}_1$ into $\mathfrak{L}_2$ is a mapping $\tau: \mathfrak{L}_1 \to \mathfrak{L}_2$ such that

(a) $\tau(\vee_n p_n) = \vee_n \tau(p_n)$ and $\tau(\wedge_n p_n) = \wedge_n \tau(p_n)$ for an sequence $\{p_n\}$ of elements of $\mathfrak{L}_1$

(b) $\tau(p') = (\tau(p))' \wedge \tau(I_{\mathfrak{L}_1})$ for any element p of $\mathfrak{L}_1$.

An isomorphism is a homorphism which is one-to-one and onto. The set of homomorphisms from $\mathfrak{L}_1$ into $\mathfrak{L}_2$ will be denoted by $\hom(\mathfrak{L}_1, \mathfrak{L}_2)$.

[26]A sublogic $\mathfrak{D}$ of a logic $\mathfrak{L}$ is a subset of $\mathfrak{L}$ closed with respect to countable joins and meets (hence $\mathfrak{D}$ is a σ-complete lattice with respect to the restriction to $\mathfrak{D}$ of the order relation of $\mathfrak{L}$), with least element $0_{\mathfrak{D}} = 0$ and greatest element $I_{\mathfrak{D}}$ (in general different from $I_{\mathfrak{L}}$) and such that $p' \wedge I_{\mathfrak{D}} \in \mathfrak{D} \forall p \in \mathfrak{D}$ (then the mapping $\mathfrak{D} \ni p \to \tilde{p} := p' \wedge I_{\mathfrak{D}} \in \mathfrak{D}$ is an orthocomplementation on $\mathfrak{D}$ and $\mathfrak{D}$ results in a logic; see Prop. 1.2 of Ref. 23). A Boolean sublogic of $\mathfrak{L}$ will be called a sub-σ-algebra of $\mathfrak{L}$.

[27]S. Gudder and C. Piron, J. Math. Phys. 12, 1583 (1971); R. Cirelli and P. Cotta-Ramusino, "On the Isomorphism of a 'Quantum Logic' with the logic of the projections in a Hilbert space", to be published on Intern. J. Theor. Phys.

[28]E. C. G. Stuckelberg, Helv. Phys. Acta **33**, 727 (1960); G. Emch. Helv. Phys. Acta **36**, 739, 770 (1963).

[29]A. M. Gleason, J. Ratl. Mech. Anal. **6**, 885 (1957). See also Refs. 20 and 21.

[30]When in connection with an element p of $\mathfrak{D}$ (or $\mathfrak{D}_\varphi$, after the assumption) we write $p(\lambda)$, we mean $p = \vee_\delta {}^{+}{}_\lambda\, p(\lambda)$ also when this equality is not written explicitly.

[31]With χ_E we denote the characteristic function of a subset E of a set Λ

[32]P. R. Halmos, *A Hilbert Space Problem Book* (Princeton U. P., Princeton, N. J., 1967), Problem 96.

[33]G. Ludwig, *Die Grundlagen der Quantenmechanik* (Springer, Berlin, 1954), A 1, Sec. 6.

[34]A. Messiah, *Quantum Mechanics, Vol. I* (North-Holland, Amsterdam, 1961), p. 333.

[35]J. C. T. Pool, "Events, Observables and Operations and the Mathematical Approaches to Quantum Theory", Lecture notes of Marburg Advanced Study Institute, 1973.

[36]P. R. Halmos, *Measure Theory* (Princeton U. P., Princeton, N.J , 1950), Sec. 25, Theorem E.

[37]M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. I* (Academic, New York, 1972).

[38]J. Dixmier, *Les C*-algèbres et leurs représentations* (Gauthier-Villars, Paris, 1964) A 7.

[39]H. L. Royden, *Real Analysis* (Macmillan, New York, 1963), Ch. 11, Th. 14.

[40]N. Dunford and J. T. Schwartz, *Linear Operators, Part II* (Interscience, New York, 1963), p. 892.

[41]P. R. Halmos, *Introduction to Hilbert Space and the Theory of Spectral Multiplicity* (Chelsea Publ., New York, 1951), Sec. 30, Theorem 1.

[42]S. Sakai, *C*-Algebras and W*-Algebras* (Springer, Berlin, 1971), 1.5.1.

[43]A. Hartkämper, "Preparing and Measuring Procedures," Lecture notes of Marburg Advanced Study Institute, 1973.

[44]N. Dunford and J. T. Schwartz, *Linear Operators, Part I* (Interscience, New York, 1958), **III**. 10.8.

[45]B. O. Koopman, Proc. Natl. Acad. Sci. USA **17**, 315 (1931).

[46]P. R. Halmos, *Lectures on Ergodic Theory* (Chelsea, New York, 1958); V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics* (Benjamin, New York, 1968).

Scattering of a finite beam in a random medium with a nonhomogeneous background

M.J. Beran and A.M. Whitman

School of Engineering, Tel-Aviv University, Tel-Aviv, Israel
Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia, Pennsylvania 19174
(Received 25 October 1973)

We consider the scattering of a finite beam of radiation in a random medium with a nonhomogeneous background that contains linear and quadratic variation. We assume that both the fluctuations in the random medium and the background nonhomogeneity are weak. We obtain general expressions for the coherence function and intensity distribution. We present explicit solutions in the multiple scattering region. Results are compared to the case in which the background is homogeneous.

1. INTRODUCTION

It was shown in a previous paper[1] that the equation governing the propagation of a finite beam of radiation in a random medium with homogeneous statistics is

$$\frac{\partial\{\hat{\Gamma}\}}{\partial z} = \left(k^2[\bar{\sigma}(s) - \bar{\sigma}(0)] + \frac{i}{2k}(\nabla^2_{\mathbf{x}_{T1}} - \nabla^2_{\mathbf{x}_{T2}})\right)\{\hat{\Gamma}\}. \tag{1}$$

Here $\{\hat{\Gamma}\}$ is the ensemble averaged coherence function, z is the propagation direction, $\mathbf{x}_{T1}$ and $\mathbf{x}_{T2}$ are the transverse (x, y) coordinates in the plane perpendicular to z, $\nabla^2_{\mathbf{x}_T}$ is the transverse laplacian operator, $s = |\mathbf{x}_{T1} - \mathbf{x}_{T2}|$, and

$$\bar{\sigma}(s) = \tfrac{1}{4}\int_{-\infty}^{\infty} \sigma[(s^2 + z^2)^{1/2}]\,dz,$$

where σ is the correlation function associated with the index of refraction fluctuations. The fluctuations are assumed to be homogeneous and isotropic. $k = 2\pi\nu/c$ is the wavenumber of the radiation which is governed by a scalar wave equation. For convenience the mean index of refraction has been set equal to unity.

In the derivation of Eq. (1) the fluctuations were taken to be weak so that there existed a distance Δz which satisfied the conditions

$$\Delta z \gg l_M, \quad \theta_c^2 k \Delta z \ll 1, \quad k^2\bar{\sigma}(0)\Delta z \ll 1. \tag{2}$$

Here l_M is the maximum correlation distance associated with the fluctuations and θ_c is the characteristic spread of the radiation. An additional condition, $a \gg l_M$, where a is the characteristic beam dimension, was shown subsequently[2] not to be necessary.

Here we wish to include the effect of a nonhomogeneous background, $\bar{k}(y)$. That is, we consider now the case in which the mean wavenumber varies in one of the transverse directions. Physically this could result from the mean density variation with altitude in the atmosphere or a mean vertical temperature gradient in the ocean. We assume here that meanir.gful values of y_1 and y_2 are restricted by the condition

$$\frac{\bar{k}(y_1) - \bar{k}(y_2)}{\bar{k}(y_1)} \ll 1 \tag{3}$$

and that if θ_R^2 is the average angular deviation of the beam the condition

$$\theta_R^2 \bar{k}\Delta z \ll 1 \tag{4}$$

is satisfied. In addition we require the radiation to be quasimonochromatic so that

$$\Delta\nu/\bar{\nu} \ll 1. \tag{5}$$

If the above conditions are met, the effect of a nonhomogeneous background may be included in Eq. (1), by the addition of the term

$$i[\bar{k}(y_1) - \bar{k}(y_2)]\,\{\hat{\Gamma}\} \tag{6}$$

on the right-hand side.

The derivation of Eq. (1) is based on the fact that in any small interval, Δz, the scattering term (the first term on the rhs) and the diffraction term (the second term) act independently in changing $\{\hat{\Gamma}\}$ and the change resulting from each effect is small. By imposing the conditions given in Eqs. (3), (4), and (5) we require that the effect due to refraction in the interval Δz is also small. Moreover, since the condition $\theta^2\bar{k}\Delta z \ll 1$ ($\theta = \theta_c$ or θ_R) essentially allows the use of a geometric optics type approximation in the interval Δz, then the refraction term is given by Eq. (6). That is, in the interval Δz, we consider only the effect of a phase shift. The effect of refraction on the scattering and diffraction terms is neglected in Δz.

The equation we treat in this paper is

$$\frac{\partial\{\hat{\Gamma}\}}{\partial z} = \left(\bar{k}^2[\bar{\sigma}(s) - \bar{\sigma}(0)] + \frac{i}{2\bar{k}}(\nabla^2_{\mathbf{x}_{T1}} - \nabla^2_{\mathbf{x}_{T2}}) + i[\bar{k}(y_1) - \bar{k}(y_2)]\right)\{\hat{\Gamma}\}. \tag{7}$$

As a boundary condition we take on the plane $z = 0$ the coherent Gaussian form

$$\{\hat{\Gamma}(\mathbf{x}_{T1}, \mathbf{x}_{T2}, 0)\} = \hat{I}_0 \exp[-(|\mathbf{x}_{T1}|^2 + |\mathbf{x}_{T2}|^2)/2b^2]. \tag{8}$$

Other initial conditions may be used, but the Gaussian form yields particularly simple results and illustrates the nature of the results expected for finite beams in general.

We choose the following form for $\bar{k}(y)$:

$$\bar{k}(y) = \bar{k} + a_1 y - a_2 y^2. \tag{9}$$

This choice allows us to convert Eq. (7), by Fourier transformation to a first-order partial differential equation. Inclusion of a y^3 term in Eq. (9) would lead to a second-order term and make the solution intractable. We thus restrict our attention here to this quadratic form. The values of y that may be significant in our problem are restricted by the condition given in Eq. (3).

2. GENERAL SOLUTION OF EQ. (7)

In this section we solve Eq. (7) using Eq. (9). We first transform to the sum and difference coordinates

$$\begin{aligned} &s_x = x_1 - x_2, \quad s = (s_x^2 + s_y^2)^{1/2} \\ &s_y = y_1 - y_2, \\ &p_x = (x_1 + x_2)/2, \\ &p_y = (y_1 + y_2)/2 \end{aligned} \tag{10}$$

and define

$$F(s) \equiv [\bar{\sigma}(s) - \bar{\sigma}(0)].$$

We find

$$\frac{\partial\{\hat{\Gamma}\}}{\partial z} = [\bar{k}^2 F(s) + ia_1 s_y - 2ia_2 s_y p_y]\{\hat{\Gamma}\} + \frac{i}{\bar{k}}\left(\frac{\partial^2}{\partial s_x \partial p_x} + \frac{\partial^2}{\partial s_y \partial p_y}\right)\{\hat{\Gamma}\}. \tag{11}$$

Proceeding as we have in a previous paper[3] we take the Fourier transform of both sides of Eq. (7). This yields

$$\frac{\partial\{\tilde{\Gamma}\}}{\partial z} = [\bar{k}^2 F(s) + ia_1 s_y]\{\tilde{\Gamma}\} - 2a_2 s_y \frac{\partial\{\tilde{\Gamma}\}}{\partial \mu_y} + \frac{\mu_y}{\bar{k}}\frac{\partial\{\tilde{\Gamma}\}}{\partial s_y} + \frac{\mu_x}{\bar{k}}\frac{\partial\{\tilde{\Gamma}\}}{\partial s_x}, \tag{12}$$

where

$$\{\tilde{\Gamma}\} = \int_{-\infty}^{\infty}\int \exp[+i(\mu_x p_x + \mu_y p_y)]\{\hat{\Gamma}\}\, dp_x dp_y. \tag{13}$$

The initial condition is

$$\{\hat{\Gamma}\} = \hat{\Gamma}_0(s, p), \quad z = 0,$$

or

$$\{\tilde{\Gamma}\} = \tilde{\Gamma}_0(s, \mu), \quad z = 0, \quad \mu = (\mu_x^2 + \mu_y^2)^{1/2}.$$

The characteristic equations associated with Eq. (12) are

$$\begin{aligned} &\frac{dz}{dt} = 1, \quad \frac{ds_x}{dt} = -\frac{\mu_x}{\bar{k}}, \quad \frac{ds_y}{dt} = -\frac{\mu_y}{\bar{k}}, \\ &\frac{d\mu_x}{dt} = 0, \quad \frac{d\mu_y}{dt} = 2a_2 s_y, \\ &\frac{d\{\tilde{\Gamma}\}}{dt} = [\bar{k}^2 F(s) + ia_1 s_y]\{\tilde{\Gamma}\}, \end{aligned} \tag{14}$$

where t is an arbitrary parameter.

The solution of Eqs. (14) is, for $a_2 > 0$,

$$\begin{aligned} &z = t, \quad s_x = s_{x_0} - (\mu_{x_0}/\bar{k})t, \\ &s_y = s_{y_0}\cos\omega t - (\mu_{y_0}/\bar{k}\omega)\sin\omega t, \quad \omega^2 = 2a_2/\bar{k}, \\ &\mu_x = \mu_{x_0}, \quad \mu_y = \bar{k}\omega s_{y_0}\sin\omega t + \mu_{y_0}\cos\omega t, \\ &\{\tilde{\Gamma}\} = \{\tilde{\Gamma}_0\}\exp\{ia_1[(s_{y_0}/\omega)\sin\omega t + (\mu_{y_0}/\bar{k}\omega^2)(\cos\omega t - 1)] \\ &\qquad + G(s_{x_0}, s_{y_0}, \mu_{x_0}, \mu_{y_0}, t)\}, \end{aligned} \tag{15}$$

where

$$G(s_{x_0}, s_{y_0}, \mu_{x_0}, \mu_{y_0}, t) = \bar{k}^2 \int_0^t F[(s_x^2 + s_y^2)^{1/2}]dt'.$$

For $a_2 < 0$,

$$\begin{aligned} &z = t, \quad s_x = s_{x_0} - (\mu_{x_0}/\bar{k})t, \\ &s_y = s_{y_0}\cosh\Omega t - (\mu_{y_0}/\bar{k}\Omega)\sinh\Omega t, \quad \Omega^2 = -2a_2/\bar{k}, \\ &\mu_x = \mu_{x_0}, \quad \mu_y = -s_{y_0}\bar{k}\Omega\sinh\Omega t + \mu_{y_0}\cosh\Omega t, \\ &\{\tilde{\Gamma}\} = \{\tilde{\Gamma}_0\}\exp\{ia_1[(s_{y_0}/\Omega)\sinh\Omega t - (\mu_{y_0}/\bar{k}\Omega^2) \\ &\qquad \times(\cosh\Omega t - 1)] + G(s_{x_0}, s_{y_0}, \mu_{x_0}, \mu_{y_0}, t)\}. \end{aligned} \tag{16}$$

We next introduce the boundary condition given by Eq. (8). Straightforward but somewhat lengthy manipulation yields finally the following results.

(A) $a_2 > 0$:

$$\begin{aligned} \{\tilde{I}(s_x, s_y, \mu_x, \mu_y, z)\} = I_0\pi b^2 \exp\Bigg[&-\frac{s_x^2}{4b^2} - \frac{s_y^2}{4b^2} \\ &\times(\cos^2\omega z + \bar{k}^2\omega^2 b^4 \sin^2\omega z) - \frac{s_x\mu_x z}{2\bar{k}b^2} \\ &- \frac{s_y\mu_y \cos\omega z \sin\omega z}{2b^2\bar{k}\omega}(1 - \bar{k}^2\omega^2 b^4) \\ &+ is_y a_1 \frac{\sin\omega z}{\omega} - \frac{\mu_x^2 b^2}{4}\left(1 + \frac{z^2}{\bar{k}^2 b^4}\right) \\ &- \frac{\mu_y^2 b^2}{4}\left(\cos^2\omega z + \frac{\sin^2\omega z}{k^2\omega^2 b^4}\right) \\ &+ \frac{i\mu_y a_1(1 - \cos\omega z)}{\bar{k}\omega^2} + G_+(s_x, s_y, \mu_x, \mu_y)\Bigg], \end{aligned} \tag{17}$$

where

$$G_+ = \frac{\bar{k}^3}{\mu}\int_0^{\mu z/\bar{k}} F\Bigg[\Bigg\{\left(s_x + \frac{\mu_x}{\mu}w\right)^2 + \left[s_y\cos\left(\frac{\omega\bar{k}w}{\mu}\right) + \frac{\mu_y}{\bar{k}\omega}\sin\left(\frac{\omega\bar{k}w}{\mu}\right)\right]^2\Bigg\}^{1/2}\Bigg] dw.$$

(B) $a_2 < 0$:

$$\begin{aligned} \{\tilde{I}(s_x, s_y, \mu_x, \mu_y, z)\} = I_0\pi b^2 \exp\Bigg[&-\frac{s_x^2}{4b^2} - \frac{s_y^2}{4b^2} \\ &\times(\cosh^2\Omega z + \bar{k}^2\Omega^2 b^4 \sinh^2\Omega z) - \frac{s_x\mu_x z}{2b^2\bar{k}} \\ &- \frac{s_y\mu_y}{2b^2\bar{k}\Omega}\cosh\Omega z \sinh\Omega z(1 - \bar{k}^2\Omega^2 b^4) \\ &+ is_y a_1 \frac{\sinh\Omega z}{\Omega} - \frac{\mu_x^2 b^2}{4}\left(1 + \frac{z^2}{\bar{k}^2 b^4}\right) \\ &- \frac{\mu_y^2 b^2}{4}\left(\cosh^2\Omega z + \frac{\sinh^2\Omega z}{\bar{k}^2\Omega^2 b^4}\right) \\ &- i\mu_y a_1\left(\frac{1 - \cosh\Omega z}{\bar{k}\Omega^2}\right) + G_-(s_x, s_y, \mu_x, \mu_y)\Bigg], \end{aligned} \tag{18}$$

where

$$G_- = \frac{\bar{k}^3}{\mu}\int_0^{\mu z/\bar{k}} F\Bigg[\Bigg\{\left(s_x + \frac{\mu_x}{\mu}w\right)^2 + \left[s_x\cosh\left(\frac{\Omega\bar{k}w}{\mu}\right) + \frac{\mu_y}{\bar{k}\Omega}\sinh\left(\frac{\Omega\bar{k}w}{\mu}\right)\right]^2\Bigg\}^{1/2}\Bigg] dw.$$

The Fourier transform of the intensity distribution, $\{\tilde{I}(\mu_x, \mu_y, z)\}$, is obtained by setting $s_x = s_y = 0$. We find then for $a_2 > 0$

$$\{\tilde{I}(\mu_x, \mu_y, z)\} = I_0\pi b^2 \exp\left\{-\frac{\mu_x^2 b^2}{4}\left(1 + \frac{z^2}{\bar{k}^2 b^4}\right)\right.$$

$$- \frac{\mu_y^2 b^2}{4}\left(\cos^2\omega z + \frac{\sin^2\omega z}{\bar{k}^2\omega^2 b^4}\right)$$

$$+ i a_1 \frac{\mu_y(1-\cos\omega z)}{\bar{k}\omega^2}$$

$$+ \frac{\bar{k}^3}{\mu}\int_0^{\mu z/\bar{k}} F\left[\left[\frac{\mu_x^2}{\mu^2}w^2 + \frac{\mu_y^2}{\bar{k}^2\omega^2}\right.\right.$$

$$\left.\left.\times \sin^2\left(\frac{\omega\bar{k}w}{\mu}\right)\right]^{1/2}\right] dw\Bigg\} \tag{19}$$

and for $a_2 < 0$

$$\{\tilde{I}(\mu_x, \mu_y, z)\} = I_0\pi b^2 \exp\Bigg\{-\frac{\mu_x^2 b^2}{4}\left(1 + \frac{z^2}{\bar{k}b^4}\right)$$

$$- \frac{\mu_y^2 b^2}{4}\left[\cosh^2\omega z + \frac{\sinh^2\omega z}{\bar{k}^2\omega^2 b^4}\right]$$

$$- i\mu_y \frac{a_1(1-\cosh\Omega z)}{\bar{k}\Omega^2}$$

$$+ \frac{\bar{k}^3}{\mu}\int_0^{\mu z/\bar{k}} F\left[\left[\frac{\mu_x^2}{\mu^2}w^2 + \frac{\mu_y^2}{\bar{k}^2\omega^2}\right.\right.$$

$$\left.\left.\times \sinh^2\left(\frac{\omega\bar{k}w}{\mu}\right)\right]^{1/2}\right] dw\Bigg\}. \tag{20}$$

3. SPECIFIC SOLUTIONS

Equations (17)—(20) reduce to previous results when a_1 and a_2 are set equal to zero. When $a_2 = 0$ but a_1 is finite (i.e., we have only a linear variation in the non-homogeneous background), we find that intensity peak is shifted by the known amount

$$\Delta p = a_1 z^2/2\bar{k}.$$

When $a_2 \neq 0$, the shift is

$$\Delta p = a_1(1-\cos\omega z)/\bar{k}\omega^2, \qquad a_2 > 0, \tag{21}$$

$$\Delta p = -a_1(1-\cosh\Omega z)/\bar{k}\Omega^2, \qquad a_2 < 0.$$

The shift Δp is independent of the scattering since the above terms are not dependent on F and the scattering integrals in Eqs. (19) and (20) are symmetric in μ_y. It is expected that if an $a_3 y^3$ term was included in Eq. (9) this would no longer be true.

To proceed further in the solution of $\{\hat{\Gamma}\}$ and $\{\hat{I}\}$ requires an explicit knowledge of $F(s) = \bar{\sigma}(s) - \bar{\sigma}(0)$. It is, however, possible to evaluate these functions in the multiple scatter region where the only important contribution comes in the limit $s \to 0$.[2]

We note that after obtaining the solution we may obtain a solution in the absence of scattering by setting $\bar{\sigma}(s) - \bar{\sigma}(0)$ equal to zero.

We know from solutions obtained when a_1 and a_2 are equal to zero that as $z \to \infty$ the intensity distribution is determined by the form of $F(s) = \bar{\sigma}(s) - \bar{\sigma}(0)$ as $s \to 0$. In this region we may expand $\bar{\sigma}(s)$ in a Taylor series and we have

$$F(s) \approx \bar{\sigma}_0 + \bar{\sigma}_0' s + \tfrac{1}{2}\bar{\sigma}_0'' s^2 + \cdots - \bar{\sigma}_0$$

$$\approx \tfrac{1}{2}\bar{\sigma}_0'' s^2 \tag{22}$$

since $\bar{\sigma}_0' = 0$.

When a_1 and a_2 are not equal to zero, this approximation is still valid in the multiple scatter region. When $a_2 > 0$, the average spread of the beam is inhibited and we have a phenomenon that is exemplified by sound trapping in the ocean. When $a_2 < 0$, the average spread of the beam is enhanced by the variation of $\bar{k}(y)$. In both cases, however, the angular spread of the scattered radiation grows with distance, and as $z \to \infty$ Eq. (22) is valid. We shall demonstrate the consistancy of this argument after the results for $\{\hat{\Gamma}\}$ have been obtained.

If Eq. (22) is substituted into Eqs. (17)—(20), we find after manipulation:

(A) $a_2 > 0$:

$$\{\hat{\Gamma}\} = \frac{I_0 b^2}{4R_x R_{y+}} \exp\Bigg\{\frac{-s_x^2}{4b^2}(1 - 2\bar{\sigma}_0''\bar{k}^2 b^2 z)$$

$$- \frac{s_y^2}{4b^2}\left[\cos^2\omega z + \bar{k}^2\omega^2 b^4 \sin^2\omega z - \bar{\sigma}_0''\bar{k}^2 b^2\left(z + \frac{\sin 2\omega z}{2\omega}\right)\right]$$

$$+ \frac{i a_1 s_y \sin\omega z}{\omega} - \frac{1}{4R_x^2}\left(p_x - \frac{i s_x z}{2b^2\bar{k}}(1 - \bar{\sigma}_0''\bar{k}^2 b^2 z)\right)^2$$

$$- \frac{1}{4R_{y+}^2}\left[p_y - \frac{a_1(1-\cos\omega z)}{\bar{k}\omega^2} - \frac{i s_y}{2b^2\bar{k}\omega}\left(\tfrac{1}{2}\sin^2\omega z\right.\right.$$

$$\left.\left.\times(1 - \bar{k}^2\omega^2 b^4) - \frac{\bar{\sigma}_0''\bar{k}^2 b^2}{\omega}\sin^2\omega z\right)\right]^2\Bigg\}, \tag{23}$$

$$\{\hat{I}\} = \frac{I_0 b^2}{4R_x R_{y+}} \exp\left(-\frac{p_x^2}{4R_x^2} - \frac{[p_y - a_1(1-\cos\omega z)/\bar{k}\omega^2]^2}{4R_{y+}^2}\right), \tag{24}$$

where

$$R_x^2 = \frac{b^2}{4}\left(1 + \frac{z^2}{\bar{k}^2 b^4} - \frac{2}{3}\frac{\bar{\sigma}_0'' z^3}{b^2}\right),$$

$$R_{y+}^2 = \frac{b^2}{4}\left(\cos^2\omega z + \frac{\sin^2\omega z}{\bar{k}^2\omega^2 b^4} - \frac{\bar{\sigma}_0''}{\omega^2 b^2}\frac{(z - \sin 2\omega z)}{2\omega}\right). \tag{25}$$

(B) $a_2 < 0$:

$$\{\hat{\Gamma}\} = \frac{I_0 b^2}{4R_x R_{y-}} \exp\Bigg\{-\frac{s_x^2}{4b^2}(1 - 2\bar{\sigma}_0''\bar{k}^2 b^2 z)$$

$$- \frac{s_y^2}{4b^2}\left[\cosh^2\Omega z + \bar{k}^2 b^4\Omega^2 \sinh^2\Omega z\right.$$

$$\left. - \bar{\sigma}_0''\bar{k}^2 b^2\left(z + \frac{\sinh 2\Omega z}{2\Omega}\right)\right]$$

$$+ \frac{i a_1 s_y \sinh\Omega z}{\Omega} - \frac{1}{4R_x^2}\left(p_x - \frac{i s_x z}{2b^2\bar{k}}(1 - \bar{\sigma}_0''\bar{k}^2 b^2 z)\right)^2$$

$$- \frac{1}{4R_{y-}^2}\left[p_y + a_1\frac{(1-\cosh\Omega z)}{\bar{k}\Omega^2} - \frac{i s_y}{2b^2\bar{k}\Omega}\left(\frac{\sinh\Omega z}{2\Omega}\right)\right.$$

$$\left.\times\left((1 + \bar{k}^2\Omega^2 b^4) - \frac{\bar{\sigma}_0''\bar{k}^2 b^2}{\Omega}\sinh^2\Omega z\right)\right]^2\Bigg\}, \tag{26}$$

$$\{\hat{I}\} = \frac{I_0 b^2}{4R_x R_{y-}}$$

$$\times \exp\left[-\frac{p_x^2}{4R_x^2} - \left(\frac{p_y + a_1(1-\cosh\Omega z)/\bar{k}\Omega^2}{4R_{y-}^2}\right)^2\right], \tag{27}$$

where

$$R_{y-}^2 = \frac{b^2}{4}\left[\cosh^2\Omega z + \frac{\sinh^2\Omega z}{\bar{k}^2\Omega^2 b^4} + \frac{\bar{\sigma}_0''}{\Omega^2 b^2}\left(\frac{z - \sinh 2\Omega z}{2\Omega}\right)\right]. \tag{28}$$

A comparison of results to those obtained when $a_1 = a_2 = 0$ shows that the x dependence of $\{\hat{\Gamma}\}$ and $\{\hat{I}\}$ is unaffected by the variation of $\bar{k}$. The effects in the two directions are essentially uncoupled. When $\omega \to 0$ and $\Omega \to 0$, the results approach the results obtained when $a_2 = 0$ except for the shift of the peak of the radiation (when $a_1 \neq 0$). The effects of a_2 are important when $(|a_2|/\bar{k})^{1/2} z \geq O(1)$.

To demonstrate the consistency of the expansion in Eq. (22), we examine Eqs. (23) and (24) when $p_x = s_x = 0$, $p_y = a_1(1 - \cos\omega z)/\bar{k}\omega^2$. From Eq. (23) we find

$$\{\hat{\Gamma}\} = \frac{I_0 b^2}{4R_x R_{y+}} \exp\left\{-\frac{s_y^2}{4b^2}\left[\cos^2\omega z + \bar{k}^2\omega^2 b^4 \sin^2\omega z - \bar{\sigma}_0''\bar{k}^2 b^2\left(z + \frac{\sin 2\omega z}{2\omega}\right)\right] - \frac{s_y^2}{4R_{y+}^2}\frac{1}{4b^2\bar{k}^2\omega^2}\left(\frac{\sin 2\omega z}{2}(1 - \bar{k}^2\omega^2 b^4) - \frac{\bar{\sigma}_0''\bar{k}^2 b^2}{\omega}\sin^2\omega z\right)^2 + \frac{i a_1 s_y \sin\omega z}{\omega}\right\}. \tag{29}$$

As $z \to \infty$, we find for $|\{\hat{\Gamma}\}|$ (we are not here interested in the phase of $\{\hat{\Gamma}\}$)

$$|\{\hat{\Gamma}\}| \to \frac{I_0 b^2}{4R_x R_{y+}} \exp\left[-s_y^2\left(-\frac{\bar{\sigma}_0''\bar{k}^2 z}{4}\right)\right] \quad z \to \infty, \tag{30}$$

where we note that $\bar{\sigma}_0''$ is a negative quantity. Thus we see that the characteristic distance in which $|\{\hat{\Gamma}\}|$ decays is proportional to $z^{1/2}$ as $z \to \infty$. Hence the angular spread of the radiation grows as $z^{1/2}$ as $z \to \infty$, and Eq. (22) is a valid expansion as $z \to \infty$. It is important to emphasize that although the angular spectrum grows as $z^{1/2}$, as it does when $a_2 = 0$, here ($a_2 > 0$) the characterisitc spread of the intensity distribution in the y direction grows only at a rate proportional to $z^{1/2}$. When $a_2 = 0$, the intensity distribution grows at a rate proportional to $z^{3/2}$.

When $a_2 < 0$ the angular spectrum grows exponentionally as $z \to \infty$ ($\sinh x$ and $\cosh x$ are proportional to e^x as $x \to \infty$). This is true for both the intensity distribution and the angular spectrum. It is thus consistent here too to use the approximation given in Eq. (22).

The solution given in Eqs. (23)—(26), valid in the multiple scatter region must still meet the condition that $(a_1 y + a_2 y^2)/\bar{k} \ll 1$. When $a_2 > 0$, this condition reduces to

$$a_1^2/\bar{k}a_2 \ll 1, \quad \bar{\sigma}_0'' z/2 \ll 1, \tag{31}$$

if we require $\omega z \gg 1$.

In Ref. 3 we defined the multiple scatter region by two conditions. The important condition was

$$z\,\bar{k}^2 l_m^2\ \bar{\sigma}_0''/6 \gg 1. \tag{32}$$

When $a_2 < 0$, $\Omega z \gg 1$, the conditions are

$$a_1^2 \cosh^2(\Omega z)/\bar{k}\,a_2 \ll 1, \quad a_1 R_{y-}/\bar{k} \text{ and } a_2 R_{y-}^2/\bar{k} \ll 1. \tag{33}$$

The utility of Eqs. (23)—(28) in any physical problem thus depends on the relative values of the parameters $a_1, a_2, \bar{\sigma}_0'', \bar{k}, l_m$ and the propagation distance z. When the solution is valid, the intensity distribution [Eqs. (24) and (27)] has the particularly simple Gaussian form with a shifted peak determined by a_1, a_2, and $\bar{k}$. The characteristic spread of the distribution is determined by diffraction and scattering as presented in Eqs. (25) and (28). In Eq. (32) is satisfied for a value of z such that $\omega z \ll 1$ or $\Omega z \ll 1$, then Eqs. (31) and (33) need not be satisfied.

4. SUMMARY

In this paper we determined expressions for the coherence function $\{\hat{\Gamma}\}$ and the average intensity distribution $\{\hat{I}\}$ under conditions for which the nonhomogeneous background varied slowly according to Eq. (9) and the radiation was forward scattered. The conditions for validity of the general solutions [Eqs. (17)—(20)] are given in Eqs. (2), (3), and (4). An explicit solution in the multiple scatter region is given in Eqs. (23)—(28). The additional conditions for the validity of these latter solutions are given in Eqs. (31), (32), and (33).

When a_2 is positive, we find the familiar channeling effect. In this case we have determined explicit expressions for the coherence function (which is directly related to angular spectrum of the scattered radiation) and for the intensity distribution. When a_2 is negative, the angular spectrum and intensity distribution grow rapidly when $\Omega z > 1$. Explicit expressions for $\{\hat{\Gamma}\}$ and $\{\hat{I}\}$ are also given in this case.

[1] M.J. Beran, J. Opt. Soc. Am. 60, 518 (1970).
[2] M.J. Beran and A.M. Whitman, J. Opt. Soc. Am. 61, 1044 (1971).
[3] A.M. Whitman and M.J. Beran, J. Opt. Soc. Am. 60, 1595 (1970).

Riemann–Green's functions for solving electromagnetic problems exhibiting rotational symmetry in media moving with superluminal velocities*

Salvatore Solimeno

Istituto Elettrotecnico, Università di Napoli, Naples, Italy
(Received 8 August 1974)

A new method of investigation of the electromagnetic field configuration in media moving with superluminal velocities is illustrated. The basic assumption is the rotational symmetry of the field. It is shown that a Volterra integral equation can be written for the magnetic field on the metallic boundary surface. This equation rests on the knowledge of the Riemann–Green's functions of two differential operators. Some expressions for these functions are obtained.

1. INTRODUCTION

In recent years there has been much interest in the study of electromagnetic radiation in moving media.[1] In this context, the innovation of operating in the frame moving together with the sources has shed new light on these problems. In fact, the introduction of Maxwell–Minkowski constitutive relations[1,2] has allowed transformation of time-dependent problems in time-independent ones.

In the wake of this simplification, it has been further shown[3] that, by a suitable extension of the Clemmow[4] scaling procedure, the radiation problem in moving media can be solved by referring to an equivalent isotropic and nondispersive medium with simple boundary conditions. In particular, it has been proved that this holds true for sub- and super-luminal velocities. With superluminal velocities, the search for a solution can be conducted by a suitable analytic continuation of the solution obtained for a stationary simple medium. In practice, this implies looking for the analytic solution of the EM problem, thus excluding numerical techniques. It, therefore, is useful when analytic solutions are available, e.g., when mode expansion technique is applicable.

In more complicated problems, where numerical techniques are more powerful, the above line of search looses its efficacy.

Fortunately a large class of boundary value problems in moving media exhibit a rotational symmetry around the velocity direction. Surprisingly, this circumstance does not seem to have drawn the attention of scientists. Nevertheless, it emerges as an essential factor for applying integral approaches.

As will be shown in the following, the search for the field reduces to solving a Volterra integral equation, whose kernel is a well behaved function devoid of the typical singularity of the usual Green's functions.

2. WAVE EQUATION FOR THE ELECTROMAGNETIC FIELD IN A SIMPLE MEDIUM MOVING AT SUPERLUMINAL VELOCITY

Let us consider a uniform, isotropic and nondispersive medium moving with uniform velocity $\mathbf{v}=v\hat{v}$ with respect to the sources and the boundaries.

By introducing the quantities[1]

$$a=(1-\beta^2)/(n^2\beta^2-1),$$

$$\mathbf{A}=a(\mathbf{I}-\mathbf{V})+\mathbf{V}$$

where $\beta=v/c$, c being the light velocity in vacuum, $n=\sqrt{\epsilon}$ is the refractive index, I is the unit dyadic, and $\mathbf{V}=\hat{v}\hat{v}$, a time-harmonic ($\exp(i\omega t)$) electromagnetic field $\Phi=\mathbf{E}$(or H) satisfies the equation[3]

$$\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\Phi^2}-\frac{\partial^2}{\partial z^2}-K^2\right)\mathbf{\Phi}'=-\mathbf{S} \tag{1}$$

where r, Φ, z are cylindrical coordinates with the z axis parallel to $\hat{v}$, $K^2=\omega^2a^2\mu\epsilon$, $\mathbf{\Phi}'=\mathbf{A}^{1/2}\cdot\mathbf{\Phi}$ and S is the source term.[3]

When $\underline{S}$ and the boundaries exhibit rotational symmetry with respect to the z axis, (1) simplifies as

$$\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}-\frac{\partial^2}{\partial z^2}-K^2\ \Phi_z'\equiv\hat{L}_z\Phi_z'=-S_z, \tag{2}$$

$$\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}-\frac{1}{r^2}-\frac{\partial^2}{\partial z^2}-K^2\right)\Phi_t'=\hat{L}_t\mathbf{\Phi}_t'=-\mathbf{S}_t \tag{3}$$

for the z and the transverse component of $\underline{\Phi}'$, respectively. In addition, as $(\partial/\partial r)\hat{r}=(\partial/\partial r)\hat{\Phi}=0$, the scalars Φ_r' and Φ_ϕ' also satisfy (3).

3. RIEMANN-GREEN'S FUNCTIONS

A. Introduction

The first general solution of the problem of Cauchy for an extensive class of partial differential equations was given by Riemann[5] almost a century ago in his well-known paper on the propagation of sound waves of finite amplitude. Although stated only for certain special equations, it is applicable to any linear equation of hyperbolic type of the second order in two independent variables: it depends ultimately on finding a certain subsidiary function, usually called the Riemann–Green's function, which is the solution of a characteristic boundary value problem for the adjoint equation.[6] Riemann gave explicit formulas for this subsidiary function in two cases of special importance in gas dynamics. Successively, Chaundy,[7] Titchmarsh,[8] Henrici,[9] and Copson[10] were able to find the Riemann–Green's functions for many other equations.

B. A brief description of Riemann's method

Let us consider the hyperbolic equation

$$\frac{\partial^2 u}{\partial r^2}-\frac{\partial^2 u}{\partial z^2}+2a\frac{\partial u}{\partial r}-2b\frac{\partial u}{\partial z}+cu\equiv\hat{L}_{(r,z)}[u]=f \tag{4}$$

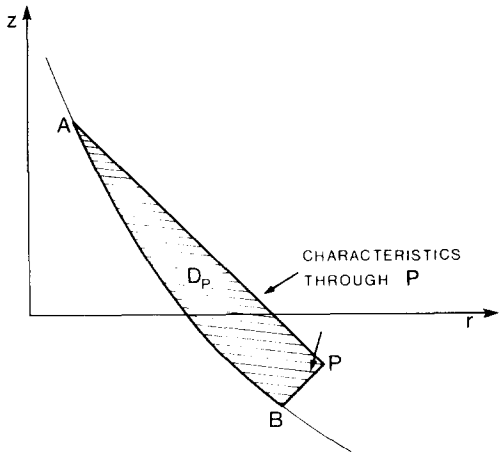


FIG. 1. Domain of dependence of P.

where a, b, c, and f are functions of r, $z \in R$, R being a region of the (r, z) plane. We prescribe the initial values $u \equiv \phi_1$, and a free curve[6] C and, say, the normal derivative $\partial u/\partial n \equiv \phi_2$. We consider the solution of the boundary value problem at all points P for which the corresponding domain of dependence D_P is contained in R. By resorting to a general technique developed by Riemann, we can write

$$u(P) = \tfrac{1}{2}u(A)R(A, P) + \tfrac{1}{2}u(B)R(B, P) + \tfrac{1}{2}\int_{AB}\left[\left(2buR + R\frac{\partial u}{\partial z} - u\frac{\partial R}{\partial z}\right)dr + \left(2auR + R\frac{\partial u}{\partial r} - u\frac{\partial R}{\partial r}\right)dz\right] + \int_{D_P} Rf\,dr\,dz, \tag{5}$$

where R is the Riemann–Green's function of the operator $\hat{L}$, A and B are the points where the characteristics through P cut the curve C (see Fig. 1).

The Riemann–Green's function $R(Q, P)$ is defined imposing the following conditions:

(a) $\hat{L}^*_{(Q)}[R] = \hat{L}_{(P)}[R] = 0$,

(b) $\dfrac{\partial R}{\partial Q} = (a - b)R$ along AP and $\dfrac{\partial R}{\partial Q} = (a + b)R$ along BP,

(c) $R(P, P) = 1$. (6)

By simple algebra Eq. (5) can be recast as

$$u(P) = \tfrac{1}{2}u(A)\,R(A, P) + \tfrac{1}{2}u(B)R(B, P) + \tfrac{1}{2}\int_{AB}\left[R\left(\frac{\partial u}{\partial s}\sin 2\widehat{rs} - \frac{\partial u}{\partial n}\cos 2\widehat{rs}\right) - u\left(\frac{\partial R}{\partial s}\sin 2\widehat{sr} - \frac{\partial R}{\partial n}\cos 2\widehat{sr}\right) + 2uR(b\cos\widehat{sr} + a\cos\widehat{sz})\right]ds + \int_{D_P} Rf\,dr\,dz \tag{7}$$

where s is the curvilinear coordinate along C and $\widehat{sr}$ in the angle formed by $\hat{r}$ with $\hat{s}$.

C. Application of Riemann's method to the radiation problem

For $\hat{L} = \hat{L}_z$ or $\hat{L}_t$, Eq. (7) reads

$$u(P) = \tfrac{1}{2}u(A)R(A, P) + \tfrac{1}{2}u(B)R(B, P) + \tfrac{1}{2}\int_{AB}\left[R\left(\frac{\partial u}{\partial s}\sin 2\widehat{sr} - \frac{\partial u}{\partial n}\cos 2\widehat{sr}\right) - u\left(\frac{\partial R}{\partial s}\sin 2\widehat{sr} - \frac{\partial R}{\partial n}\cos 2\widehat{sr}\right) + u\frac{R}{r}\cos\widehat{sz}\right]ds + \int_{D_P} RS\,dr\,dz. \tag{8}$$

Now, it is convenient to introduce the following quantity:

$$F = \cos\widehat{nr}\left(\frac{\partial}{\partial s}(ru)\cos\widehat{rs} + \frac{\partial}{\partial n}(ru)\cos\widehat{nr}\right) - \cos\widehat{nz}\left(\frac{\partial}{\partial s}(ru)\cos\widehat{sz} + \frac{\partial}{\partial n}(ru)\cos\widehat{nz}\right) = \frac{\partial}{\partial s}(ru)\sin 2\widehat{sr} - \frac{\partial}{\partial n}(ru)\cos 2\widehat{sr}. \tag{9}$$

Since

$$\frac{\partial u}{\partial s}\sin 2\widehat{sr} - \frac{\partial u}{\partial n}\cos 2\widehat{sr} = \frac{F}{r} - \frac{u}{r}\cos\widehat{zs}, \tag{10}$$

plugging (10) into (8) we get

$$u(P) = \tfrac{1}{2}u(A)R(A, P) + \tfrac{1}{2}u(B)R(B, P) + \int_{AB}\frac{FR}{r}\,ds - \tfrac{1}{2}\int_{AB} u\left(\frac{\partial R}{\partial s}\sin 2\widehat{sr} - \frac{\partial R}{\partial n}\cos 2\widehat{sr}\right)ds + \int_{D_P} RS\,dr\,dz. \tag{11}$$

4. INTEGRAL EQUATION FOR THE MAGNETIC FIELD

Let us consider a finite source distribution S contained in a region R with some metallic boundaries ∂R. In our two-dimensional representation, we assume C coinciding with the boundaries.

In Ref. 3 we have shown that the EM field satisfies the equation

$$(-\mathbf{I} + 2\mathbf{V}) \cdot \nabla \times \mathbf{H}^f = i\omega\epsilon a \mathbf{E}^f, \tag{12}$$

where $\mathbf{H}^f$ and $\mathbf{E}^f$ are fictitious fields simply related to $\mathbf{E}$ and $\mathbf{H}$. For a class of problems where $\mathbf{H}^f = u\hat{\phi}$, (12) reads

$$(-\mathbf{I} + 2\mathbf{V}) \cdot \left(-\frac{\partial u}{\partial z}\hat{r} + \frac{1}{r}\frac{\partial}{\partial r}(ru)\hat{z}\right) = \frac{\partial u}{\partial z}\hat{r} + \frac{1}{r}\frac{\partial}{\partial r}(ru)\hat{z} = i\omega\epsilon a \mathbf{E}^f.$$

Projecting on C, we obtain

$$i\omega\epsilon a \mathbf{E}^f \cdot \hat{s} = \cos\widehat{rs}\,\frac{1}{r}\frac{\partial}{\partial z}(ur) + \cos\widehat{sz}\,\frac{1}{r}\frac{\partial}{\partial r}(ru) = F.$$

Even though F vanishes identically on metallic boundaries, it is worth assuming $F \neq 0$ in order to include in the following discussion the radiation from slotted metallic antennas. Then we assume F be an assigned function on C.

Let us now consider a straight line K parallel to the z axis, such that the half-plane containing the point $z = -\infty$ and limited by K does not contain the sources.

Now, at each point $P \in R$ lying on K, the field vanishes

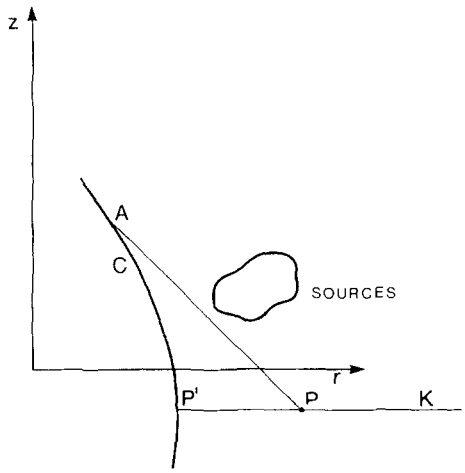


FIG. 2. The free curve C is assumed coincident with the metallic boundary of the region R where the electromagnetic field is confined. The integral appearing in the integral equation solving the boundary value problem is calculated between P' and A.

identically (see Ref. 3). Therefore, (11) simplifies as

$$\tfrac{1}{2}u(A)R(A,P)=\frac{1}{2}\int_{AP'}u\left(\frac{\partial R}{\partial s}\sin 2\widehat{sr}-\frac{\partial R}{\partial n}\cos 2\widehat{sr}\right)ds$$

$$-\frac{1}{2}\int_{AP'}\frac{FR}{r}ds-\int_{D_P}RS\,dr\,dz \tag{13}$$

where P' is the projection of P on C (see Fig. 2).

On writing Eq. (13), the search for the H^f_Φ field on C rests on the solution of a Volterra integral equation of the second kind, whose kernel $\partial R/\partial s\sin 2\widehat{sr}-\partial R/\partial n\times\cos 2\widehat{sr}$ is a finite continuous function of its argument. Once H_Φ is known on C, H^f_Φ can be calculated at any point P by using (11).

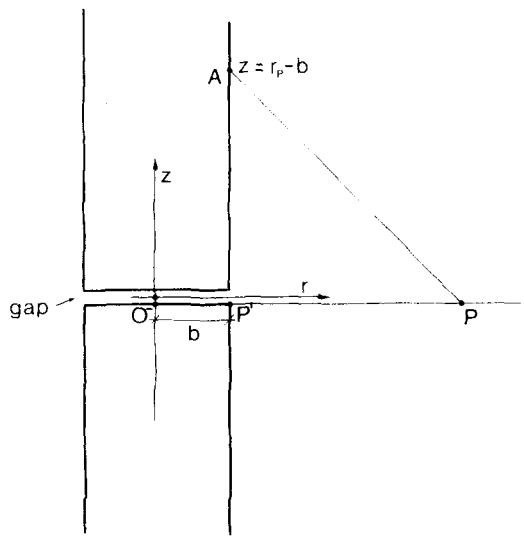


FIG. 3. Geometry of the infinite cylindrical antenna problem. The excitation is applied through the gap. No other sources are supposed present.

5. INFINITE CYLINDRICAL ANTENNA

As an example, let us consider an infinite cylindrical antenna of radius b oriented in such a way that its axis coincides with the velocity direction of the medium. The antenna is driven at its center by a voltage V_0 applied uniformly around an infinitesimally thin circumferential gap. This idealized generator is specified by

$$E_z(b,\Phi,z)=V_0\delta(z) \tag{14}$$

and is known as the delta-function generator. Such an idealized generator has been used extensively by King[11] in the study of linear antenna in free space. Seshadri[12] has discussed the case of such an antenna immersed in a warm plasma.

In this case, plugging (14) into (13), we obtain

$$\tfrac{1}{2}u(A)R(A,P)=\frac{1}{2}\int_0^{z_{A^-b}}u(z)\frac{\partial R}{\partial n}dz-\frac{iV_0\omega\epsilon a}{2b}R(P',P) \tag{15}$$

having chosen $z_{P'}=0^-$ (see Fig. 3).

6. RIEMANN-GREEN'S FUNCTIONS FOR $\hat{L}_z$ AND $\hat{L}_t$

In Sec. 4 an integral equation for H_Φ was obtained. Its solution implied knowing the R—G function of $\hat{L}_t$. On the other hand it can be easily shown that a similar integral equation can be written for E_z. In such a case the R—G function of $\hat{L}_z$ is needed.

In the present section some expressions of the R—G functions of $\hat{L}_z$ and $\hat{L}_t$ are derived. In this way, all the analytical tools for this integral approach are provided.

A. Riemann-Green's function of $\hat{L}_z$

1. Extension of the Henrici representation

The Riemann—Green's function $R_{(L_z)}$ of $\hat{L}_z$ has been obtained by Henrici (Ref. 9, Table 2) and reads

$$R_{(L_z)}(Q,P)=\left(\frac{r_Q}{r_P}\right)^{1/2}\Xi_2\left(\frac{1}{2},\frac{1}{2};1;\frac{w^2}{4r_Pr_Q},-\frac{1}{4}K^2w^2\right) \tag{16}$$

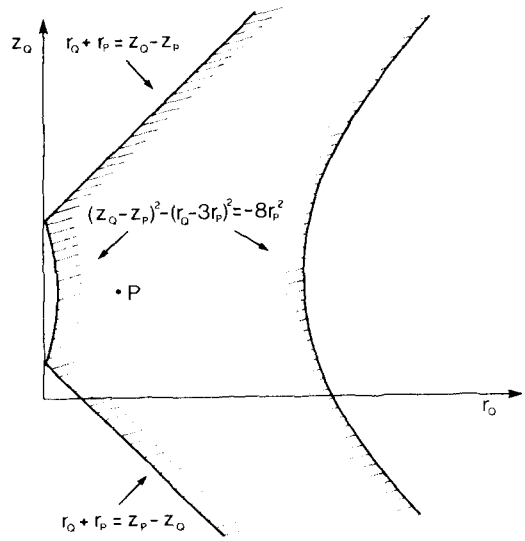


FIG. 4. Domain of existence of the Henrici Riemann—Green's function analytic expression.

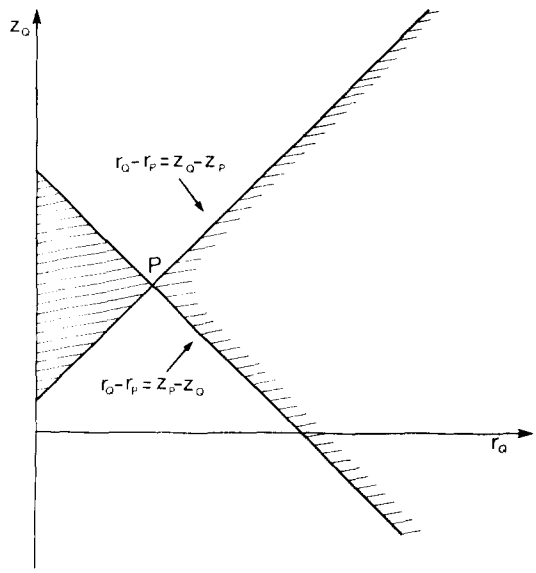


FIG. 5. Domain of existence of the Riemann—Green's function obtained in the present paper.

where Ξ_2 is the confluent hypergeometric function of two variables (Ref. 13, Vol. I, p. 226),

$$\Xi_2(\alpha,\beta;\gamma;x,y)=\sum_{m,n=0}^{\infty}\frac{(\alpha)_m(\beta)_m}{(\gamma)_{m+n}m!n!}x^m y^n, \quad |x|<1, \tag{17}$$

and $w^2=(r_P-r_Q)^2-(z_P-z_Q)^2$.

As the series (17) converges for $|x|<1$, (14) can be used for Q contained in the dashed region of Fig. 4.

Before we attempt to find a representation of $R_{(L_z)}$ valid in a different region, it is worth making a few general remarks. The operator $\hat{L}_z$ is strictly related to

$$\hat{L}_{\alpha,\beta}\equiv\frac{\partial^2}{\partial r^2}-\frac{\partial^2}{\partial t^2}+\frac{2\alpha}{r}\frac{\partial}{\partial r}-\frac{2\beta}{t}\frac{\partial}{\partial t}, \tag{18}$$

which has the important property of being more symmetric than $\hat{L}_z$. In fact, if we put $t=(\beta+iKz)/iK$ and make $\beta\to+\infty$, (18) becomes

$$\hat{L}_{\alpha,\beta}\underset{\beta\to\infty}{\longrightarrow}\hat{L}_\alpha=\frac{\partial^2}{\partial r^2}-\frac{\partial^2}{\partial z^2}+\frac{2\alpha}{r}\frac{\partial}{\partial r}-2iK\frac{\partial}{\partial z}. \tag{19}$$

Next, for $\alpha=\frac{1}{2}$, it follows

$$\hat{L}_z=\exp(ikz)\hat{L}_\alpha\exp(-iKz)\equiv\hat{L}'_\alpha. \tag{20}$$

Lastly, to get $R_{(L'_\alpha)}$ we transform $R_{(L_{\alpha\beta})}$ in agreement with the above variable transformation $t\to z$ and (20). $R_{(L_{\alpha\beta})}$ has been given by Copson as a Stieltjes integral,

$$R_{(L_{\alpha\beta})}(Q,P)=\left(\frac{r_Q}{r_P}\right)^\alpha\left(\frac{t_Q}{t_P}\right)^\beta\left(P_{-\alpha}(1+\eta)+\int_0^{\pi/2}P_{-\alpha}(1+\eta\cos^2\theta)\,dP_{-\beta}(1+\xi\sin^2\theta)\right), \tag{21}$$

where

$$\xi=\frac{(t_P-t_Q)^2-(r_P-r_Q)^2}{2t_Pt_Q}, \tag{22}$$

$$\eta=\frac{(r_P-r_Q)^2-(t_Q-t_P)^2}{2r_Pr_Q}>0. \tag{23}$$

Now, since

$$\left(\frac{t_Q}{t_P}\right)\underset{\beta\to\infty}{\longrightarrow}\exp iK(z_Q-z_P), \tag{24}$$

$$\xi\underset{\beta\to\infty}{\longrightarrow}-\frac{K^2}{2\beta^2}[(z_P-z_Q)^2-(r_P-r_Q)^2]=\frac{K^2w^2}{2\beta^2}, \tag{25}$$

$$\eta\underset{\beta\to\infty}{\longrightarrow}\frac{(r_P-r_Q)^2-(z_P-z_Q)^2}{2r_Pr_Q}\equiv\eta', \tag{26}$$

and (cf. Ref. 13, Vol. I, p. 148 and Vol. II, p. 55)

$$\begin{aligned}P_{-\beta}(1+\xi\sin^2\theta)&={}_2F_1(1-\beta,\beta,1;-\tfrac{1}{2}\xi\sin\theta^2)\\ &\underset{\beta\to\infty}{\longrightarrow}{}_2F_1(1-\beta,\beta,1,-(\tfrac{1}{4}\beta^2)K^2w^2\sin^2\theta)=I_0(Kw\sin\theta)\\ &={}_0F_1(1,\tfrac{1}{4}K^2w^2\sin^2\theta)\end{aligned} \tag{27}$$

we have

$$R_{(L_{\alpha\beta})}\underset{\beta\to\infty}{\longrightarrow}\exp[iK(z_Q-z_P)]\left(\frac{r_Q}{r_P}\right)^\alpha\left(P_{-\alpha}(1+\eta')+\int_0^{\pi/2}P_{-\alpha}(1+\eta'\cos^2\theta)\,dI_0(KW\sin\theta)\right). \tag{28}$$

Lastly, upon substitution of $\hat{L}_{\alpha\beta}$ with $\hat{L}'_\alpha$, (28) yields

$$R_{(L_\alpha)}=\left(\frac{r_Q}{r_P}\right)^\alpha\left(\int_0^{\pi/2}P_{\alpha-1}(1+\eta'\cos^2\theta)\,dI_0(KW\sin\theta)+P_{\alpha-1}(1+\eta')\right). \tag{29}$$

This last representation is more suited for our needs than (16). In fact, it holds for $|r_P-r_Q|\geq|z_P-z_Q|$ (see Fig. 5) and, consequently, it lends itself to give the R—G function values entering the field integral equation (13).

2. Series representation

The R—G function as expressed in (29) entails the calculation of a Stieltjes integral which is quite tedious to carry out. To simplify this representation we shall attempt to transform it into a series.

By developing the Stieltjes integral

$$\begin{aligned}&\int_0^{\pi/2}P_{\alpha-1}(1+\eta'\cos^2\theta)\,dI_0(Kw\sin\theta)\\ &\quad=\int_0^{\pi/2}P_{\alpha-1}(1+\eta'\cos^2\theta)I_1(Kw\sin\theta)\,d(Kz\sin\theta)\end{aligned} \tag{30}$$

and substituting the series expansion

$$I_1(Kw\sin\theta)=\sum_{n=0}^{\infty}\frac{(Kw\sin\theta)^{2n+1}}{2}\frac{1}{n!(n+1)!} \tag{31}$$

into (30), we obtain

$$\begin{aligned}\int_0^{\pi/2}P_{\alpha-1}(1+\eta'\cos^2\theta)\,dI_0(Kw\sin\theta)&=\sum_{n=0}^{\infty}\left(\frac{Kw}{2}\right)^{2n+1}\frac{1}{n!(n+1)!}\\ &\times\int_0^{\pi/2}P_{\alpha-1}(1+\eta'\cos^2\theta)(\sin\theta)^{2n+1}\,d(Kw\sin\theta)\\ &=\sum_{n=0}^{\infty}\left(\frac{Kw}{2}\right)^{2n+2}\frac{1}{n!(n+1)!}\int_0^1P_{\alpha-1}(1+\eta'x)(1-x)^n\,dx\end{aligned} \tag{32}$$

where $x=\cos^2\theta$. As (Ref. 15, p. 802, No. 9)

$$\int_0^1P_{\alpha-1}(1+\eta'x)(1-x)^n\,dx=\left(\frac{2+\eta'}{\eta'}\right)^{n+1/2}P_{\alpha-1}^{-n-1}(1+\eta')n!, \tag{33}$$

then

$$\int_0^{\pi/2} P_{\alpha-1}(1+\eta'\cos^2\theta)\,dI_0(Kw\sin\theta) = \sum_{n=0}^{\infty}\left(\frac{Kw}{2}\right)^{2n+2} \times\left(\frac{2+\eta'}{\eta'}\right)^{n+1/2} P_{-1}^{-n-1}(1+\eta')n!^{-1}. \tag{34}$$

Lastly, upon substitution of (34) into (29) we obtain

$$R_{(L'_\alpha)}(Q,P) = \left(\frac{r_Q}{r_P}\right)^{1/2}\sum_{n+0}^{\infty}\left(\frac{Kw}{2}\right)^{2n}\left(\frac{2+\eta'}{\eta'}\right)^{n/2} P_{\alpha-1}^{-n}(\cosh\gamma)n!^{-1} \tag{35}$$

where $\cosh\gamma = 1+\eta'$.

This last representation lends itself to the calculation $R_{(L_z)}$ step by step, starting with (cf. Ref. 14, p. 337)

$$P_{-1/2}(\cosh\gamma) = \frac{2}{\pi}\operatorname{sech}\frac{\gamma}{2}K\left(\tanh\frac{\gamma}{2}\right), \tag{36}$$

$$P_{1/2}(\cosh\gamma) = \frac{2}{\pi}\exp(\gamma/2)\,E[1-\exp(-2\gamma)]^{1/2} \tag{37}$$

where K and E are the complete elliptic integrals of the first and second kinds, and then making use of the relations between contiguous Legendre functions (cf. Ref. 13, Vol. I, p. 160)

$$P_\nu^{\mu+2}(\cosh\gamma) + 2(\mu+1)\coth\gamma\, P_\nu^{\mu+1}(\cosh\gamma) = (\nu-\mu)(\nu+\mu+1)\,P_\nu^{\mu}(\cosh\gamma), \tag{38}$$

$$\cosh\gamma\, P_{-1/2}(\cosh\gamma) - P_{1/2}(\cosh\gamma) = \tfrac{1}{2}\sinh\gamma\, P_{-1/2}^{1}(\cosh\gamma). \tag{39}$$

3. Integral representation

Naturally, we now try to recast (35) as an integral again. The resulting representation will turn out to be more suitable for studying the behavior of $R_{(L_z)}$. We start with the following integral representation (Ref. 13, Vol. I, p. 156):

$$P_{\alpha-1}^{-n}(\cosh\gamma) = \frac{(\sinh\gamma)^n}{\sqrt{\pi}\,2^n\Gamma(n+\frac{1}{2})}\int_0^{\pi}(\sin t)^{2n} \times(\cosh\gamma+\sinh\gamma\cos t)^{-n+\alpha-1}\,dt. \tag{40}$$

Substitution of (40) into (35) leads to the new integral representation

$$R_{(L_\alpha)}(Q,P) = \left(\frac{r_Q}{r_P}\right)^{\alpha}\frac{1}{\sqrt{\pi}}\int_0^{\pi} dt\,x^{\alpha-1}\sum_{n=0}^{\infty}\frac{y^n}{n!\,\Gamma(n+\frac{1}{2})}, \tag{41}$$

where

$$x = \cosh\gamma+\sinh\gamma\cos t, \tag{42a}$$

$$y = \frac{1}{2}\left(\frac{Kw}{2}\right)^2\left(\frac{2+\eta'}{\eta'}\right)^{1/2}\sinh\gamma(\sin t)^2x^{-1}. \tag{42b}$$

Since

$$\sum_{n=0}^{\infty}\frac{y^n}{n!\,\Gamma(n+\frac{1}{2})} = \frac{1}{\sqrt{\pi}}\cosh 2\sqrt{y}, \tag{43}$$

then

$$R_{(L_\alpha)}(Q,P) = \frac{1}{\pi}\left(\frac{r_Q}{r_P}\right)^{\alpha}\int_0^{\pi} dt\,x^{\alpha-1}\cosh(cx^{-1/2}\sin t), \tag{44}$$

c standing for $Kw[(2+\eta'/\eta')]^{1/4}(\sinh\gamma/2)^{1/2}$.

In particular, when Q is on the z axis, η' tends to infinity and

$$\left(\frac{r_Q}{r_P}\right)^{\alpha}(\cosh\gamma+\sinh\gamma\cos t)^{\alpha-1}\underset{r_Q\to 0}{\longrightarrow}0, \tag{45a}$$

$$c(\cosh\gamma+\sinh\gamma\cos t)^{1/2}\sin t \underset{r_Q\to 0}{\longrightarrow} Kw\sin t/2 \tag{45b}$$

Accordingly,

$$R_{(L_z)}(Q,P)\underset{r_Q\to 0}{\longrightarrow}0. \tag{46}$$

B. Riemann-Green's function for the operator $\hat{L}_t$

In view of a very important property[6] of the Riemann–Green's functions, $R_{(L_t)}(Q,P)$ coincides with $R_{(L_t^*)}(P,Q)$, where $\hat{L}_t^*$ is the formal adjoint of $\hat{L}_t$,

$$\hat{L}_t^* = \frac{\partial^2}{\partial r^2} - \frac{\partial}{\partial r}\frac{1}{r} - \frac{1}{r^2} - \frac{\partial^2}{\partial z^2} - K^2 = \frac{\partial^2}{\partial r^2} - \frac{1}{r}\frac{\partial}{\partial r} - \frac{\partial^2}{\partial z^2} - K^2. \tag{47}$$

The new operator $\hat{L}_t^*$ differs from $\hat{L}_z$ only in the minus sign in front of $r^{-1}\partial/\partial r$. This implies that many of the considerations of Sec. 6.A apply as well to $\hat{L}_t^*$. The only difference occurs when we put α [cf. (18)] equal to $-1/2$ instead of $+1/2$. In addition, when we exchange P with Q, η' and w remain unchanged. Accordingly, (29) and (35) transform as

$$R_{(L_t)}(Q,P) = R_{(L_t^*)}(P,Q) = \left(\frac{r_Q}{r_P}\right)^{1/2}\left(P_{1/2}(1+\eta') + \int_0^{\pi/2} P_{1/2}(1+\eta'\cos^2\theta)\,dI_0(Kw\sin\theta)\right) = \left(\frac{r_Q}{r_P}\right)^{1/2}\sum_{n=0}^{\infty}\left(\frac{Kw}{2}\right)^{2n}\left(\frac{2+\eta'}{\eta'}\right)^{n/2} P_{1/2}^{-n}(\cosh\gamma)n!^{-1}. \tag{49}$$

On the other hand, (Ref. 13, Vol I, p. 156)

$$P_{1/2}^{-n}(\cosh\gamma) = \frac{(\sinh\gamma)^n}{\sqrt{\pi}\,2^n\Gamma(n+\frac{1}{2})} \times\int_0^{\pi}(\sin t)^{2n}(\cosh\gamma+\sinh\gamma\cos t)^{-n+1/2}\,dt. \tag{50}$$

Then, by retracing the discussion leading to (44), we obtain

$$R_{(L_t)}(Q,P) = \frac{1}{\pi}\left(\frac{r_Q}{r_P}\right)^{1/2}\int_0^{\pi} dt\,x^{1/2}\cosh(cx^{-1/2}\sin t). \tag{51}$$

In addition, as

$$\left(\frac{r_Q}{r_P}\right)^{1/2}(\cosh\gamma+\sinh\gamma\cos t)^{1/2}\underset{r_Q\to 0}{\longrightarrow}\frac{w}{r_P}\cos t/2, \tag{52}$$

then

$$R_{(L_t)}(Q,P)\underset{r_Q\to 0}{\longrightarrow}\frac{w}{\pi r_P}\int_0^{\pi} dt\cos t/2\cosh(Kw\sin t/2) = \frac{2}{\pi K r_P}\sinh Kw. \tag{53}$$

Comparison of (53) with (46) shows that the two functions differ considerably in their behavior on the z axis: Whereas $R_{(L_t)}$ is a monotone function of z, $R_{(L_z)}$ vanishes identically.

CONCLUSIONS

It has been shown that the search for the electromagnetic field distribution in a medium moving with superluminal velocity rests on the solution of a Volterra integral equation. Even though this paper rests on the assumption that the magnetic field has only one magnetic component ($\underline{H} = u\hat{\phi}$), and the metallic boundaries are represented by a free curve, it is possible to extend the present approach. In fact, by dividing the EM field region into subregions bounded by free curves and characteristics, the integral approach can be applied separately to each of them. Some difficulties arise when we do not know the boundary conditions on some of these boundaries. In this case, some physical intuition has to be invoked.

Before concluding, it is worth noting that a similar approach could be proposed for radiation problems in uniaxially anisotropic plasmas. Unfortunately, in this case, we cannot require the field to vanish along a straight line [condition leading to Eq. (13)]. But a condition can be imposed on the far-field behavior. To proceed along this line would require knowing some asymptotic expressions for the R—G functions. This line of research is presently in progress.

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Professor C. H. Papas for the warm hospitality offered during his stay at the California Institute of Technology. Particular thanks are due to Professor G. Franceschetti for stimulating discussions.

*Work partially sponsored by the Italian National Council of Research.

[1]K. S. H. Lee and C. Papas, J. Math. Phys. **5**, 1668 (1964); C. T. Tai, Proc. IEEE **52**, 685 (1964); R. T. Compton, J. Math. Phys. **7**, 2145 (1966); C. Johannsen, J. Math. Phys. **13**, 78 (1972).
[2]H. Minkowski, Göttincher Nachrichten **53** (1908).
[3]S. Solimeno, Alta Frequenza (in print, Nov. 74 issue), Papers I and II.
[4]P. C. Clemmow, Proc. IEEE **110**, 101 (1963).
[5]G. F. B. Riemann, Abh. d. Kön. Ges. derWin. zu Göttingen **8** (1860).
[6]H. Bateman, *Partial Differential Equations of Mathematical Physics* (Cambridge U. P., Cambridge, 1964); R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1966), Vol. II.
[7]T. W. Chaundy, Quart. J. Math., Oxford Ser. **6**, 288 (1935); **7**, 306 (1936); **8**, 280 (1937); **9**, 234 (1938); **10**, 219 (1939); **11**, 101 (1940).
[8]E. C. Titchmarsh, *Theory of Fourier Integrals* (Oxford U. P., Oxford, 1937).
[9]P. Henrici, Z. Angew, Math. Phys. **8**, 169 (1957).
[10]E. T. Copson, J. Rat. Mech. Anal. **1**, 324 (1958).
[11]R. W. P. King, *Theory of Lincar Antennas* (Harvard U. P., Cambridge, 1956).
[12]S. R. Seshadri, IEEE Trans. AP-**13**, 189 (1965).
[13]A. Erdélyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953).
[14]M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1968).
[15]I. S. Gradshteyn and I. M. Ryzhiik, *Tables of Integrals, Series and Products* (Academic, New York, 1965).

Closed gravitational-wave universes: Analytic solutions with two-parameter symmetry*

Robert H. Gowdy

Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742
(Received 9 June 1974)

Einstein's vacuum field equations are solved for spacetimes with two-parameter spacelike symmetry, a space-reflection symmetry, and space sections homeomorphic to either $S^1\times S^2$ or S^3. All integrals are evaluated, and the spacetime metrics are presented in analytic form.

I. INTRODUCTION

The spacetimes considered here may be thought of as closed empty universes which are inhomogeneous due to the presence of gravitational waves. Locally, these waves are indistinguishable from Einstein–Rosen cylindrical waves. The boundary conditions which make such waves compatible with closed spacelike hypersurfaces have been discussed in a previous paper.[1] Just as in the Einstein–Rosen case, two of Einstein's equations can be reduced to quadratures, one equation has a canonical solution that fixes the cylindrical radius coordinate, and the remaining equation is an easily solved linear wave equation.[2] For the closed-universe boundary conditions considered here, the solutions to this wave equation can be expressed entirely in terms of polynomials. The integral which solves the remaining Einstein equations can then be evaluated analytically. This paper presents the result of evaluating this integral and thus obtains this class of spacetime metrics in a form which is suitable for exploring such properties as the behavior of geodesics and test fields.

II. METRIC AND FIELD EQUATIONS

Any spacetime with two spacelike Killing vectors $\partial/\partial\sigma$ and $\partial/\partial\delta$, a reflection symmetry $\sigma\rightsquigarrow-\sigma$, and at least one degenerate isometry group trajectory can be expressed in the form[1]

$$ds^2=L^2\{e^{2a}(d\theta^2-dt^2)+R[e^{2\psi}\,d\sigma^2+e^{-2\psi}\,d\delta^2]\}, \tag{1}$$

where L is a constant length and the functions a, ψ, and R depend only on the coordinates θ and t. The group coordinates σ and δ are angles which range from 0 to 2π. The coordinates θ and t range from 0 to π with no identifications. In the notation of the previous paper, this metric corresponds to $\Phi=0$, $W+b=\psi$.

Denote derivatives with respect to the advanced and retarded time coordinates $v=t+\theta$ and $u=t-\theta$ by $f_+=\partial f/\partial v$ and $f_-=\partial f/\partial u$ and write the independent Einstein equations in the form[1,2]

$$R_+a_+=R\psi_+^2+\tfrac{1}{2}R_{++}-\tfrac{1}{4}R(R_+/R)^2, \tag{2}$$

$$R_-a_-=R\psi_-^2+\tfrac{1}{2}R_{--}-\tfrac{1}{4}R(R_-/R)^2, \tag{3}$$

$$R_{+-}=0, \tag{4}$$

$$(\partial/\partial\theta)(R\partial\psi/\partial\theta)-(\partial/\partial t)(R\partial\psi/\partial t)=0. \tag{5}$$

III. SOLUTIONS TO THE WAVE EQUATIONS

Equations 4 and 5 are simply wave equations and their solution is straightforward. For spacetimes homeomorphic to $(0,1)\times S^1\times S^2$ or $(0,1)\times S^3$, the coordinates θ and t may always be chosen so that

$$R=\sin\theta\sin t \tag{6}$$

which solves Eq. (4).[1] Equation 5 has the solution

$$\psi=W+b, \tag{7}$$

where

$$W=\sum_j Z_j(\cos t)P_j(\cos\theta), \tag{8}$$

$$Z_j(w)=A_jP_j(w)+C_jQ_j(w), \tag{9}$$

P_j and Q_j are Legendre functions of the first and second kind, respectively, A_j and C_j are adjustable constants, and

$$b=-\tfrac{1}{2}\ln R \tag{10}$$

for the $S^1\times S^2$ topology and

$$b=-\tfrac{1}{2}\ln\tan(\theta/2) \tag{11}$$

for the S^3 topology.[1,2]

IV. MATCHING CONSTRAINTS

Now consider Eqs. (2) and (3). From Eq. (6), R_- vanishes everywhere on the null surface $u=0$ and R_+ vanishes everywhere on the null surface $v=\pi$. On these null surfaces, Eqs. (2) and (3) are constraints on the function ψ and thus on the constants A_j and C_j. One way to understand the origin of these constraints is to cast the spacetime metric into the canonical Einstein–Rosen form by using R and $T=\cos\theta\cos t$ as coordinates instead of θ and t. One finds that four distinct coordinate patches of this sort are needed to cover the spacetime. These patches are joined across the null surfaces $u=0$ and $v=\pi$ and the constraints which arise from Eqs. (2) and (3) are the conditions for a smooth match of metrics.

Because the matching constraints are preserved in time when Eqs. (4) and (5) are satisfied, it is sufficient to satisfy them on the 2-surface where the matching null surfaces intersect.[2] The resulting conditions on the function W are

$$(\partial W/\partial\theta)\big|_{\theta=t=\pi/2}=q+\tfrac{1}{2}\alpha,\quad (\partial W/\partial t)\big|_{\theta=t=\pi/2}=p, \tag{12}$$

where the pair (q,p) is either $(0,\pm1)$ or $(\pm1,0)$ and

$$\alpha=\begin{cases}0 & \text{for the } S^1\times S^2 \text{ topology}\\ 1 & \text{for the } S^3 \text{ topology.}\end{cases} \tag{13}$$

The corresponding constraints on the constants A_j and

C_j are

$$\sum_j L_j C_{2j+1} = q + \tfrac{1}{2}\alpha, \quad \sum_j L_j C_{2j} = p, \tag{14}$$

where

$$L_j = (2j+1)P_{2j}(0)Q_{2j+1}(0).$$

V. FORMAL INTEGRATION OF THE REMAINING EINSTEIN EQUATIONS

Once the matching constraints are satisfied, the right-hand side of Eq. (2) vanishes whenever R_+ vanishes and the right-hand side of Eq. (3) vanishes whenever R_- vanishes. Multiply Eq. (2) by R_- and Eq. (3) by R_+, take the difference and find

$$(\partial a/\partial\theta)R_+R_- = -\tfrac{1}{4}F\sin\theta, \tag{15}$$

where the function

$$F = -4(\sin\theta)^{-1}[R(R_-\psi_+^2 - R_+\psi_-^2) + \tfrac{1}{2}R_-R_{++} - \tfrac{1}{2}R_+R_{--} - \tfrac{1}{4}R^{-1}R_+R_-(\partial R/\partial\theta)] \tag{16}$$

vanishes whenever R_+R_- vanishes provided that the matching constraints hold. In terms of the coordinates $\tau = \cos t$ and $x = \cos\theta$, Eq. (15) becomes

$$\partial a/\partial x = (\tau^2 - x^2)^{-1}F$$

which can be integrated to

$$a(\theta, t) = a(0, t) - \int_{\cos\theta}^{1} dx F/(\tau^2 - x^2). \tag{17}$$

To express F in terms of x and τ, define the derivatives $f_x = \partial f/\partial x$, $f_T = (1-\tau^2)\partial f/\partial\tau$, and the quantity

$$Z = xW_T^2 + x(1-\tau^2)(1-x^2)W_x^2 - 2\tau(1-x^2)W_xW_T. \tag{18}$$

For the $S^1\times S^2$ topology,

$$F = Z - x + (x^2 - \tau^2)W_x. \tag{19}$$

For the S^3 topology,

$$F = Z - (3/4)x + x(1-\tau^2)W_x - \tau W_T. \tag{20}$$

If only a finite set of the constants A_j and C_j that appear in Eqs. (8) and (9) for W are nonzero, then W is a polynomial in x and so is F. Since $F=0$ whenever $R_+R_- = \tfrac{1}{4}(\tau^2 - x^2) = 0$, F must contain $\tau^2 - x^2$ as a factor. Thus, the integral in Eq. (17) can be evaluated by performing a polynomial long division of F by $\tau^2 - x^2$ and integrating the resulting polynomial in x.

VI. EXPLICIT GENERAL SOLUTION

In order to perform the integral in Eq. (17), let P_m^n denote $P_m^n(0)$, the associated Legendre polynomial at zero and use the expansion

$$P_j(x) = \sum_n (1/n!)P_j^n x^n,$$

where the conventions $P_j^n = 0$ for $n > j$, $1/n! = 0$ for $n < 0$, and $0! = 1$ are used. The result of substituting this expansion into Eqs. (8), (9), (17), (19), and (20), carrying out the polynomial long division, and integrating when $A_j = C_j = 0$ for all $j > J$ is

$$a(\theta, t) = a(0, t) + (1-\alpha)[W(0, t) - W(\theta, t)] - \sum_{k=0}^{2j-1}[\Gamma_{2k}(1 - \cos^{2k+1}\theta) + \Gamma_{2k+1}(1 - \cos^{2k+2}\theta)], \tag{21}$$

where

$$\Gamma_{2k} = (2k+1)^{-1}\sum_{j=0}^{k} \tau^{-2(k-j+1)}F_{2j}, \tag{22}$$

$$\Gamma_{2k+1} = (2k+2)^{-1}\sum_{j=0}^{k} \tau^{-2(k-j+1)}F_{2j+1}, \tag{23}$$

$$F_k = (\tfrac{1}{4}\alpha - 1)\delta_{1k} + \sum_{m=0}^{J} P_m^k\left(\frac{1-\tau^2}{(k-1)!}Z_m - \frac{\tau}{k!}Z_{mT}\right) + \sum_{m=0}^{J}\sum_{n=0}^{J}[K_{mnk}Z_{mT}Z_{nT} - 2\tau D_{mnk}Z_{mT}Z_n + (1-\tau^2)V_{mnk}Z_mZ_n], \tag{24}$$

$$K_{mnk} = \sum_j \frac{P_m^{k-j-1}P_m^j}{(k-j-1)!\,j!}, \tag{25}$$

$$D_{mnk} = \frac{n(n+1)}{2n+1}\sum_j \frac{(P_{n-1}^j - P_{n+1}^j)P_m^{k-j}}{(k-j)!\,j!}, \tag{26}$$

$$V_{mnk} = \frac{n(n+1)}{2n+1}\sum_j \frac{(P_{n-1}^j - P_{n+1}^j)P_m^{k-j}}{j!\,(k-j-1)!}. \tag{27}$$

VII. NONCONICALITY CONSTRAINTS

The function $a(0, t)$ in Eq. (21) is determined by requiring the spacetime metric to be nonconical at $\theta = 0$. For the $S^1\times S^2$ topology,

$$a(0, t) = -\ln\sin t. \tag{28}$$

For the S^3 topology,

$$a(0, t) = -\tfrac{1}{2}\ln(\tfrac{1}{2}\sin t). \tag{29}$$

It is also necessary to require nonconicality at the antipodal symmetry axis. This requirement takes the form

$$a(\pi, t) - a(0, t) = W(0, t) - (1-2\alpha)W(\pi, t). \tag{30}$$

This condition is preserved in time when Einstein's equations are satisfied.[1,2] Thus, one can choose to impose it at $t = \pi/2$ ($\tau = 0$) where it takes the form

$$\int_{-1}^{1} dx x^{-1}[W_T^2 + (1-x^2)W_x^2 + \alpha W_x - 1 + \tfrac{1}{4}\alpha] = 2\alpha W(\pi, \pi/2). \tag{31}$$

In terms of the constants A_j and C_j, the condition is

$$4\sum_s\sum_j (v_{js}A_{2j}C_{2s+1} - k_{js}A_{2s+1}C_{2j}) = 2\alpha\sum_j[(P_{2j}^0 - p_j)A_{2j} - Q_{2j+1}^0 C_{2j+1}], \tag{32}$$

where

$$v_{js} = P_{2j}^0 Q_{2s+1}^0 \sum_k (2k-1)^{-1}V_{2j,2s+1,2k}, \tag{33}$$

$$k_{js} = (2j+1)(2s+1)P_{2s}^0 Q_{2j+1}^0 \sum_k (2k-1)^{-1}K_{2j,2s+1,2k}, \tag{34}$$

and

$$p_j = P_{2j}^0 \sum_k \frac{P_{2j}^{2k}}{(2k-1)!\,(2k-1)}. \tag{35}$$

VIII. A LONG-WAVELENGTH EXAMPLE

Because the general expressions for $a(\theta,t)$ and for the constraints are lengthy, it is useful to display them for the case in which only the amplitudes A_0, A_1, A_2, C_0, C_1, and C_2 are nonzero. This case can be thought of as a universe which is dominated by the longest wavelength gravitational waves that can break homogeneity. The gravitational wave modes displayed in this example are among the ones which compete most directly with the homogeneous modes of the mixmaster universe (Bianchi type IX) studied by Misner.[3] Thus, they may have a bearing on the evolution of chaotic universes toward homogeneity and isotropy.

For this case, the matching constraints are just

$$C_1 = q + \tfrac{1}{2}\alpha, \tag{36}$$

$$C_0 + C_2 = p. \tag{37}$$

The functions Γ_k which appear in the expression for a (Equation 21) are

$$\begin{aligned}\Gamma_0 &= 2(C_1 - \tfrac{1}{2}\alpha)(C_0 + C_2)\tau^{-1} - 2A_1(C_0 + C_2) - 3(C_1 - \tfrac{1}{2}\alpha)A_2 \\ &\quad + [3A_1A_2 - 3C_2(C_1 - \tfrac{1}{2}\alpha)]\tau + [3A_1C_2 + 3A_2(C_1 - \tfrac{1}{2}\alpha)]\tau^2 \\ &\quad - 3A_1A_2\tau^3 + Q_0\{-2C_1(C_0 + C_2) - 3C_2(C_1 - \tfrac{1}{2}\alpha) \\ &\quad + 3(A_1C_2 + A_2C_1)\tau + [3C_1C_2 + 3C_2(C_1 - \tfrac{1}{2}\alpha)]\tau^2 \\ &\quad - 3(A_1C_2 + A_2C_1)\tau^3\} + 3C_1C_2Q_0^2\tau(1 - \tau^2),\end{aligned} \tag{38}$$

$$\begin{aligned}\Gamma_1 &= [(C_0 + C_2)^2 + (C_1 - \tfrac{1}{2}\alpha)^2 - 1]\tau^{-2} - (9/4)A_2^2 + 6C_2(C_0 + C_2) \\ &\quad + C_1^2 - A_1^2 - [2A_1C_1 + 6A_2(C_0 + C_2) + (27/2)A_2C_2]\tau \\ &\quad + (A_1^2 + (27/2)A_2^2 - (45/4)C_2^2)\tau^2 + (45/2)A_2C_2\tau^3 \\ &\quad - (45/4)A_2^2\tau^4 + Q_0\{-(2A_1C_1 + (9/2)A_2C_2) \\ &\quad - [6C_2(C_0 + C_2) + 2C_1^2 + (27/2)C_2^2]\tau + (2A_1C_1 + 27A_2C_2)\tau^2 \\ &\quad + (45/2)C_2^2\tau^3 - (45/4)A_2C_2)\tau^4\} + Q_0^2\{-C_1^2 - (9/4)C_2^2 \\ &\quad + (C_1^2 + (27/2)C_2^2)\tau^2 - (45/4)C_2^2\tau^4\},\end{aligned} \tag{39}$$

$$\begin{aligned}\Gamma_2 &= 2(C_1 - \tfrac{1}{2}\alpha)(C_0 + C_2)\tau^{-3} + 6A_1C_2 + 3A_2C_1 \\ &\quad + 9(C_1C_2 - A_1A_2)\tau - 9(A_1C_2 + A_2C_1)\tau^2 + 9A_1A_2\tau^3 \\ &\quad + 9Q_0[C_1C_2 - (A_1C_2 + A_2C_1)\tau - 2C_1C_2\tau^2 \\ &\quad + (A_1C_2 + A_2C_1)\tau^3] - 9C_1C_2Q_0^2\tau(1 - \tau^2),\end{aligned} \tag{40}$$

$$\begin{aligned}\Gamma_3 &= [(C_0 + C_2)^2 + (C_1 - \tfrac{1}{2}\alpha)^2 - 1]\tau^{-4} + (9/4)A_2^2 \\ &\quad - 9C_2^2 + (63/2)A_2C_2 + ((81/2)C_2^2 - (45/2)A_2^2)\tau^2 \\ &\quad - (81/2)A_2C_2\tau^3 + (81/4)A_2^2\tau^4 + Q_0[(9/2)A_2C_2 \\ &\quad + (63/2)C_2^2\tau - 45A_2C_2\tau^2 - (81/2)C_2^2\tau^3 \\ &\quad + (81/2)A_2C_2\tau^4] + 9C_2^2Q_0^2(\tfrac{1}{4} - (5/2)\tau^2 \\ &\quad + (9/4)\tau^4),\end{aligned} \tag{40}$$

$$\Gamma_4 = 2(C_1 - \tfrac{1}{2}\alpha)(C_0 + C_2)\tau^{-5}, \tag{41}$$

$$\Gamma_5 = [(C_0 + C_2)^2 + (C_1 - \tfrac{1}{2}\alpha)^2 - 1]\tau^{-6}. \tag{42}$$

Notice that when the matching constraints are satisfied, all of the negative powers of τ vanish and the series terminates.

The nonconicality constraints may be found either from Eq. (32) or by using Eqs. (38)—(42) in Eq. (21) and substituting the result in Eq. (30). In either case, one obtains

$$A_2(C_1 - \tfrac{1}{2}\alpha) + A_1(C_0 + C_2) - A_1C_2 = \tfrac{1}{2}\alpha(A_0 + A_2). \tag{43}$$

*Research supported by NSF Grant No. GP-34022X.

[1]R.H. Gowdy, Ann. Phys. (N.Y.) 83, 203 (1974).
[2]R.H. Gowdy, Phys. Rev. Lett. 27, 826 (1971).
[3]C.W. Misner, Phys.Rev. Lett. 22, 1071 (1969).

Clebsch-Gordan coefficients for crystal space groups*

Rhoda Berenson† and Joseph L. Birman‡

Physics Department, New York University, New York, New York 10003
(Received 5 August 1974)

A practical method for calculating Clebsch–Gordan coefficients for crystal space groups is presented. It is based on properties of the space group irreducible representations as induced from ray representations of subgroups. Using this method, we obtain all Clebsch–Gordan coefficients for a family of representations in a single calculation: For space groups, for a given triangle of stars $*k$, $*k'$, $*k''$, where $*k \otimes *k' \in *k''$, the coefficients for all allowable little group representations l, l', l'' are obtained. In the following paper this is applied to rocksalt O_h^5-Fm3m and diamond O_h^7-Fd3m space groups.

I. INTRODUCTION

In this article we shall report on our investigation of crystal Clebsch—Gordan coefficients. Our major objective is to develop a useful and practical method by which Clebsch—Gordan coefficients can be calculated for crystal space groups; in a following paper we shall give the results of such calculations for some physically important cases in rocksalt O_h^5 and diamond O_h^7 space groups which are typical symmorphic and nonsymmorphic space groups, respectively. In the course of developing a practical method of calculation we have had to reexamine the structure of the crystal Clebsch—Gordan coefficients with particular attention to the use of induced ray representation theory.

The organization of this paper is as follows: in Sec. 2 we give a brief precis of the definition and some properties of the crystal space group Clebsch—Gordan coefficients including especially the important matter of the transformation of the coefficients when each of the participating representations undergoes similarity transformation. The key equation is (2.18). In Sec. 3 we discuss the structure of induced ray representations: this is relevant since the irreducible representations of space groups are induced from ray representations of various underlying point groups. In the present context it should be recalled that the process of induction produces block matrices. The problem of finding the Clebsch—Gordan coefficients then partitions into sub-problems each one of which relates to a partial Clebsch—Gordan matrix. From a canonical matrix block [called the (1,1,1) block] by appropriate transformations we can find the other matrices [$(\sigma,\sigma',\sigma'')$ blocks] ultimately using equation (2.18). In Sec. 4 we develop the theory in case all the induced representations are irreducible. This applies, as is shown in Sec. 5 to the conventional formulation of the space group problem, in which all irreducible representations of the space group G are written as $D^{(*\mathbf{k})(l)}$ and each is induced from a $D^{(\mathbf{k})(l)}$ of $G(\mathbf{k})$. In Sec. 6 we develop the theory in case the induced representations are reducible. Evidently a reducible representation can be reduced by similarity transformation. Thus the relationship between the work of Secs. 4 and 6 is by means of a similarity transformation and this again reminds us of Eq. (2.18). In Sec. 7 we develop space group theory by inducing reducible representations. In the most extreme case, we induce from the translation group T [equivalently from the trivial point group $P(\epsilon)=\epsilon$]. This produces a "ray regular" representation which contains a family of representations (e.g., *all* representations at one star). It is then shown that the calculation of all the Clebsch—Gordan coefficients for that family can be carried out efficiently in one single formulation. Thus we can find the (1,1,1) block of the Clebsch—Gordan coefficients for $D^{(*\mathbf{k})(l)} \otimes D^{(*\mathbf{k}')(l')} \to D^{(*\mathbf{k}'')(l'')}$ for all (l,l',l'') for a fixed triplet of stars $(*\mathbf{k}, *\mathbf{k}', *\mathbf{k}'')$. The ease of doing this calculation is illustrated in an accompanying paper. A final Sec. 8 briefly compares our formulation with others recently given.

2. GLEBSCH-GORDAN COEFFICIENTS

In this section all representations and elements refer to a given finite group G. The definition of Clebsch—Gordan coefficients for the irreducible representations of a finite group G is well known.[1] Let the set of l functions $\{\psi_\mu^{(l)}\}$, $\mu = 1 \cdots l$ span the irreducible linear vector space $\sum^{(l)}$ which is a basis for the irreducible representation $D^{(l)}$ of G. Likewise for $\{\psi_{\mu'}^{(l')}\}$ and $\{\psi_{\mu''}^{(l'')}\}$, $\sum^{(l')}$ and $\sum^{(l'')}$, $D^{(l')}$ and $D^{(l'')}$.

If, in the reduction of the direct product

$$D^{(l)} \otimes D^{(l')} \equiv D^{(l \otimes l')} = \sum_{l''} (ll' \mid l'') D^{(l'')} \tag{2.1}$$

the integer reduction coefficient is

$$(ll' \mid l'') \geq 1 \tag{2.2}$$

then there exist linear combinations

$$\psi_{\mu''}^{(l'')\gamma} = \sum_{\mu\mu'} \begin{pmatrix} l & l' \\ \mu & \mu' \end{pmatrix} \begin{matrix} l'' \\ \mu'' \end{matrix} \begin{matrix} \gamma \\ \end{matrix} \Bigg) \psi_\mu^{(l)} \psi_{\mu'}^{(l')} \tag{2.3}$$

which are bases for $D^{(l'')}$. The coefficients $\left(\begin{smallmatrix} l & l' \\ \mu & \mu' \end{smallmatrix} \middle| \begin{smallmatrix} l'' & \gamma \\ \mu'' & \end{smallmatrix}\right)$ are the Clebsch—Gordan coefficients or vector coupling coefficients. If $(ll' \mid l'') = 1$ these coefficients are unique *modulo* a phase. In general if $(ll' \mid l'') > 1$ there will be $(ll' \mid l'')$ distinct linear combinations of the binary products $(\psi_\mu^{(l)} \cdot \psi_{\mu'}^{(l')})$ for each row μ'' of $D^{(l'')}$. To account for this multiplicity we introduce the index γ and take

$$\gamma = 1, \ldots, (ll' \mid l''). \tag{2.4}$$

For $(ll' \mid l'') > 1$ the Clebsch—Gordan coefficients are not unique, but each set (fixed γ) can be chosen to have correct orthonormality properties (*vide infra* and Appendix A).

Alternately, the Clebsch—Gordan coefficients are the elements of the ll' dimensional unitary matrix that reduces the direct product $D^{(l)} \otimes D^{(l')}$. If we write

$$U^{l \otimes l'}_{\mu\mu', l''\gamma\mu''} \equiv \begin{pmatrix} l & l' \\ \mu & \mu' \end{pmatrix} \begin{matrix} l'' \\ \mu'' \end{matrix} \begin{matrix} \gamma \\ \end{matrix} \Bigg) \tag{2.5}$$

then for an arbitrary symmetry element ϕ in G

$$D^{(l)}(\phi)\otimes D^{(l')}(\phi)=U\Delta(\phi)U^{-1}, \tag{2.6}$$

where Δ is the fully reduced, or decomposed matrix having elements

$$\Delta_{l''\gamma\nu'',\,\bar{l}''\bar{\gamma}\bar{\nu}''}=\delta_{l''\bar{l}''}\delta_{\gamma\bar{\gamma}}D^{(l'')}_{\nu''\bar{\nu}''}. \tag{2.7}$$

Using Eq. (2.7) in (2.6) we have

$$\sum_{l''_\mu\gamma}\sum_{\bar{\mu}''}U^{l\otimes l'}_{\nu\nu',\,l''\gamma\bar{\mu}''}D^{(l'')}(\phi)_{\bar{\mu}''\mu''}U^{l\otimes l'}_{\bar{\nu}\bar{\nu}',\,l''\gamma\mu''}$$

$$=D^{(l)}(\phi)_{\nu\bar{\nu}}D^{(l')}(\phi)_{\nu'\bar{\nu}'}. \tag{2.8}$$

Multiplying both sides of (2.8) by $D^{(l'')}(\phi)^*_{\nu''\bar{\nu}''}$, summing on ϕ and using the orthogonality theorem for irreducible representations, we obtain

$$\sum_\gamma U^{l\otimes l'}_{\nu\nu',\,l''\gamma\nu''}U^{l\otimes l'*}_{\bar{\nu}\bar{\nu}',\,l''\gamma\bar{\nu}''}$$

$$=\frac{l''}{g}\sum_\phi D^{(l)}(\phi)_{\nu\bar{\nu}}D^{(l')}(\phi)_{\nu'\bar{\nu}'}D^{(l'')}(\phi)^*_{\nu''\bar{\nu}''}. \tag{2.9}$$

This equation has been used by Koster, Dimmock, Wheeler, and Statz,[2] to obtain Clebsch–Gordan coefficients for the thirty two point groups and by Litvin and Zak[3] and Itzkan[4] for space groups. The usual procedure for calculating the Clebsch–Gordan coefficients from Eq. (2.9) is as follows.

First consider the case $(ll'\,|\,l'')=1$. Then

$$U^{l\otimes l'}_{\nu\nu',\,l''\nu''}U^{l\otimes l'*}_{\bar{\nu}\bar{\nu}',\,l''\bar{\nu}''}$$

$$=\frac{l''}{g}\sum_\phi D^{(l)}(\phi)_{\nu\bar{\nu}}D^{(l')}(\phi)_{\nu'\bar{\nu}'}D^{(l'')}(\phi)^*_{\nu''\bar{\nu}''}. \tag{2.10}$$

If in Eq. (2.10) we set $\nu=\nu$, $\nu'=\bar{\nu}'$, $\nu''=\bar{\nu}''$, we can determine which coefficients are nonzero. Having found a nonzero coefficient, say for $\bar{\nu}=\nu_0$, $\bar{\nu}'=\nu'_0$, and $\bar{\nu}''=\nu''_0$, we can choose its phase as real, fix $\bar{\nu}$, $\bar{\nu}'$, and $\bar{\nu}''$ and then let ν, ν', and ν'' range over all allowable values. This procedure yields the entire Clebsch–Gordan matrix for $(ll'\,|\,l'')=1$.

For $(ll'\,|\,l'')>1$ a systematic procedure to obtain all $(ll'\,|\,l'')$ sets of coefficients has also been given by Koster.[5] Notice that what appears on the right-hand side of (2.10) is $\sum_\gamma U_{\nu\nu',\,l''\gamma\nu''}U^*_{\bar{\nu}\bar{\nu}'l''\gamma\bar{\nu}}$ which for fixed $\bar{\nu}\bar{\nu}'\bar{\nu}''$ is a linear combination of Clebsch–Gordan coefficients. But since any linear combination of coefficients is equally good, we can separately calculate $[U_{\nu\nu',\,l''\gamma\nu''},\ldots,U_{\nu\nu',\,l''\gamma_{(ll'\,|\,l'')}\nu''}]$ as long as we assure that the matrices obtained for γ and γ' are orthogonal. Hence, we proceed as in the case for $(ll'\,|\,l'')=1$ by finding a non-zero coefficient, fixing $\bar{\nu}=\nu_0$, $\bar{\nu}'=\nu'_0$, $\bar{\nu}''=\nu''_0$, and letting ν, ν', and ν'' range over all values. This yields the Clebsch–Gordan matrix for $\gamma=1$. Now select ν_0, ν'_0, ν''_0 which are different from the first set, yielding a different nonzero coefficient. Set $\bar{\nu}=\nu_0$, $\bar{\nu}'=\nu'_0$, and $\bar{\nu}''=\nu''_0$ and let ν, ν', ν'' range over all allowed values, yielding a second set of coefficients, which are orthogonal to the first set. See Appendex A. We continue in this manner until all $(ll'\,|\,l'')$ sets of coefficients are obtained.

A. Transformation of Clebsch-Gordan coefficients

It will be of later interest to demonstrate how the matrix of Clebsch–Gordan coefficients transforms when the representations $D^{(l)}$, $D^{(l')}$, and $D^{(l'')}$ undergo similarity transformations. Let A, A', A'' be unitary matrices of appropriate dimensions, and let

$$\bar{D}^{(l)}=AD^{(l)}A^{-1};\quad D^{(l')}=A'D^{(l')}A'^{-1};$$
$$\bar{D}^{(l'')}=A''D^{(l'')}A''^{-1} \tag{2.11}$$

when each $D^{(l'')}$ in the fully reduced matrix Δ is subjected to transformation by the appropriate matrix A'';

$$\Delta\to\bar{\Delta}=\alpha\Delta\alpha^{-1}, \tag{2.12}$$

where

$$\bar{\Delta}_{l''\gamma\nu'',\,\bar{l}''\bar{\gamma}\bar{\nu}''}=(A''D^{(l'')}A''^{-1})_{\nu''\bar{\nu}''}\delta_{l''\bar{l}''}\delta_{\gamma\bar{\gamma}} \tag{2.13}$$

and α is the $(ll'\times ll')$ block diagonal matrix whose blocks are the matrices A'', each in its appropriate location; likewise α^{-1} is composed from blocks A''^{-1}.

The transformed basis functions corresponding to transformation by the A matrices are

$$\bar{\psi}^{(\bar{l})}_\mu=\sum_\lambda A^{-1}_{\lambda\mu}\psi^{(l)}_\lambda,$$
$$\bar{\psi}^{(\bar{l}')}_\mu=\sum_{\lambda'}A'^{-1}_{\lambda'\mu}\psi^{(l')}_{\lambda'},$$
$$\bar{\psi}^{(\bar{l}'')}_\mu=\sum_{\lambda''}A''^{-1}_{\lambda''\mu}\psi^{(l'')}_{\lambda''}. \tag{2.14}$$

For simplicity we neglect multiplicity, and also write the Clebsch–Gordan matrix for the original basis functions $\{\psi\}$ and $U^{(l\otimes l')}$, and for the transformed basis functions $\{\bar{\psi}\}$ as $\nu^{(\bar{l}\otimes\bar{l}')}$. Then we have for the products belonging to one resultant representation $\bar{D}^{(l'')}$:

$$\bar{\psi}^{(\bar{l}'')}_\mu=\sum_{\lambda''}A''^{-1}_{\lambda''\mu''}\psi^{(l'')}_{\lambda''}$$
$$=\sum_{\lambda''}A''^{-1}_{\lambda''\mu''}\sum U^{l\otimes l'}_{\lambda\lambda',\,l''\lambda''}\psi^{(l)}_\lambda\psi^{(l')}_{\lambda'},$$
$$=\sum_{\lambda\lambda'\lambda''}A''^{-1}_{\lambda''\mu''}U^{l\otimes l'}_{\lambda\lambda',\,l''\lambda''}\sum_{\mu\mu'}A_{\mu\lambda}A'_{\mu'\lambda'}\psi^{(\bar{l})}_\mu\psi^{(\bar{l}')}_{\mu'}. \tag{2.15}$$

But

$$\bar{\psi}^{(\bar{l}'')}_\mu=\sum_{\mu\mu'}V^{\bar{l}\otimes\bar{l}'}_{\mu\mu',\,\bar{l}''\mu''}\bar{\psi}^{(\bar{l})}_\mu\bar{\psi}^{(\bar{l}')}_{\mu'} \tag{2.16}$$

so that

$$V^{\bar{l}\otimes\bar{l}'}_{\mu\mu',\,l''\mu''}=\sum_{\lambda\lambda'\lambda''}(A\otimes A')_{\mu\mu',\,\lambda\lambda'}U^{l\otimes l'}_{\lambda\lambda',\,l''\lambda''}A''^{-1}_{\lambda''\mu''}. \tag{2.17}$$

Evidently Eq. (2.17) applies to one of the particular $\bar{D}^{(l'')}$. Expressing this as a matrix we have

$$V=(A\otimes A')U\alpha^{-1} \tag{2.18}$$

so that

$$\bar{D}^{(l)}\otimes\bar{D}^{(l')}=V\bar{\Delta}V^{-1} \tag{2.19}$$

and (2.19) should be compared with (2.6), which reads

$$D^{(l)}\otimes D^{(l')}=U\Delta U^{-1}. \tag{2.20}$$

The relation (2.18) between V and U matrices is quite important and we use it often. Note also when a *particular* resultant block $D^{(l'')}$ (or $\bar{D}^{(l'')}$) is kept fixed, then the relevant portion of the transformed V is given by Eq. (2.17), with elements of A''^{-1} appearing.

B. Reducible representations and Clebsch-Gordan coefficients

In our work on crystal space–group Clebsch–Gordan coefficients we will make use of reducible representations. Thus we now consider the reduction of the direct

product of two *reducible* representations $D^{(a)}$ and $D^{(b)}$. Assume the matrix $\bar{Z}$ brings their product to decomposed form

$$D^{(a)} \otimes D^{(b)} = \bar{Z} \Delta \bar{Z}^{-1}, \tag{2.21}$$

where, as before, the elements of Δ are given in (2.7). We require the relationship between $\bar{Z}$ and the Clebsch–Gordan matrix U which reduces the product of irreducible constituents.

Let $M^{(a)}$ be the matrix which brings $D^{(a)}$ into fully reduced form

$$M^{(a)} D^{(a)} M^{(a)-1} = \Delta^{(a)}, \tag{2.22}$$

where

$$\Delta^{(a)}_{l\alpha\bar{\mu}, l\alpha\mu} = D^{(l)}_{\bar{\mu}\mu} \delta_{\bar{l}l} \delta_{\bar{\alpha}\alpha} \tag{2.23}$$

and similarly for $M^{(b)}$. Let Z be the matrix which reduces the direct product of $\Delta^{(a)}$ and $\Delta^{(b)}$:

$$\Delta^{(a)} \otimes \Delta^{(b)} = Z \Delta Z^{-1}. \tag{2.24}$$

Now note: since $\Delta^{(a)}$ and $\Delta^{(b)}$ are already in fully reduced form, they are block diagonal matrices, with irreducible blocks down the diagonal. Therefore the matrix on the left hand side of (2.24) is a direct sum of matrices, each one of which is a direct product of irreducible constituents like (2.6) and each of these products is reduced by its Clebsch–Gordan matrix as in (2.20). It then follows by inspection that modulo a phase, the elements of the matrix Z are identical to elements of U. Actually Z is a direct sum of Clebsch–Gordan matrices U, i.e., a sort of super-Clebsch–Gordan matrix.

Furthermore by an argument entirely analogous to that which established (2.9) we can find the relationship between Z and $\bar{Z}$. Noting that the matrix Δ in both (2.21) and (2.24) is already in fully reduced form we have

$$Z = (M^{(a)} \otimes M^{(b)}) \bar{Z}. \tag{2.25}$$

When comparing with (2.18) note that α is the identity matrix in this case. Finally note that the elements of Z are Clebsch–Gordan coefficients

$$Z_{l\alpha\mu l'\alpha'\mu', l''\beta\mu''} \sim U^{l \otimes l'}_{\mu\mu', l''\gamma\mu''}, \tag{2.26}$$

where β is a multiplicity index for l'' occurring in $\Delta^a \otimes \Delta^b$; and γ is the usual multiplicity index for l'' in $D^{(l)} \otimes D^{(l')}$.

3. INDUCED RAY REPRESENTATIONS

In this section we recall some facts about induced ray representations and establish some results needed later.

Consider a group G with a subgroup H such that

$$G = \phi_1 H + \cdots + \phi_\sigma H + \cdots + \phi_s H. \tag{3.1}$$

Let $\mathcal{H}^m$ and $\mathcal{G}^{\bar{m}}$ be ray representations of H and G with ray factors ω^H and ω^G, respectively. Then

$$\mathcal{H}^m(\phi_\alpha)_{\mu\nu} \mathcal{H}^m(\phi_\beta)_{\nu\lambda} = \omega^H_{\alpha,\beta} \mathcal{H}^m(\phi_{\alpha\beta})_{\mu\lambda} \tag{3.2}$$

and

$$\mathcal{G}^{\bar{m}}(\phi_\alpha)_{\mu\nu} \mathcal{G}^{\bar{m}}(\phi_\beta)_{\nu\lambda} = \omega^G_{\alpha,\beta} \mathcal{G}^{\bar{m}}(\phi_{\alpha\beta})_{\mu\lambda}. \tag{3.3}$$

The operators P_{ϕ_α}, when applied to functions, multiply in the same manner as in equation (3.2)

$$P_{\phi_\alpha} P_{\phi_\beta} \psi = \omega_{\alpha,\beta} P_{\phi_{\alpha\beta}} \psi. \tag{3.4}$$

Suppose the l_m vectors $\{\psi^m_{11}, \ldots, \psi^m_{1l_m}\}$ form a basis for $\mathcal{H}^m$ so that for all $\phi \in H$ we have:

$$P_\phi \psi^m_{1\nu} = \sum_{\mu=1}^{l_m} \mathcal{H}^m(\phi)_{\mu\nu} \psi^m_{1\mu}. \tag{3.5}$$

Let l_m new vectors $\{\psi^m_{\sigma 1}, \ldots, \psi^m_{\sigma l_m}\}$ be defined for each coset representative in (3.1) where

$$P_{\phi_\sigma} \psi^m_{1\nu} = \psi^m_{\sigma\nu}. \tag{3.6}$$

If ϕ_g is any element in G we have

$$\begin{aligned} P_{\phi_g} \psi^m_{\tau\nu} &= P_{\phi_g} P_{\phi_\tau} \psi^m_{1\nu} \\ &= \frac{P_{\phi_\sigma} P_{\phi_{\bar{\sigma}}}}{\omega^G_{\sigma,\sigma^{-1}}} P_{\phi_g} P_{\phi_\tau} \psi^m_{1\nu}. \end{aligned} \tag{3.7}$$

If $\phi_\sigma^{-1} \phi_g \phi_\tau \in H$, then

$$\begin{aligned} P_{\phi_\sigma} \psi^m_{\tau\nu} &= \frac{P_{\phi_\sigma}}{\omega_{\sigma,\sigma^{-1}}} \sum_{\mu=1}^{l_m} \mathcal{H}^m(\phi_\sigma^{-1} \phi_g \phi_\tau)_{\mu\nu} \psi^m_{1\mu} \\ &= \frac{1}{\omega_{\sigma,\sigma^{-1}}} \sum_{\mu=1}^{l_m} \mathcal{H}^m(\phi_\sigma^{-1} \phi_g \phi_\tau)_{\mu\nu} \psi^m_{\sigma\mu}. \end{aligned} \tag{3.8}$$

Hence the vectors $\{\psi^m_{11} \cdots \psi^m_{sl_m}\}$ form a basis for a representation of G induced from $\mathcal{H}^m$.

The induced representation is usually written as $(\mathcal{H}^m \uparrow G)$ and has elements

$$\begin{aligned} &(\mathcal{H}^m \uparrow G)(\phi_g)_{\sigma\mu\tau\nu} \\ &= \frac{1}{\omega_{\sigma,\sigma^{-1}}} \dot{\mathcal{H}}^m(\phi_\sigma^{-1} \phi_g \phi_\tau)_{\mu\nu}, \end{aligned} \tag{3.9}$$

where

$$\begin{aligned} \dot{\mathcal{H}}^m(\phi)_{\mu\nu} &= \mathcal{H}^m(\phi)_{\mu\nu} \quad \text{for } \phi \in H \\ &= 0 \qquad\qquad \text{otherwise.} \end{aligned} \tag{3.10}$$

Note that for given ϕ_g, for each τ there is only one value of σ such that $\phi_\sigma^{-1} \phi_g \phi_\tau \in H$. Hence, the representation matrices of $(\mathcal{H}^m \uparrow G)$ are subdivided into S^2 blocks and in each row and column there is only one non-vanishing block. In general the induced representation is reducible.

In the work which follows we shall treat all the representations as induced from one or another subgroup. The block structure of each of these representations should then be kept in mind. Particularly important will be the corresponding factorization of the Clebsch–Gordan matrix.

4. IRREDUCIBLE INDUCED REPRESENTATIONS

In this section we shall consider determination of Clebsch–Gordan coefficients for a group G when each of the "factors" $D^{(m)}$ and $D^{(m')}$ is irreducible *and* induced from subgroups H and H' of G, respectively. Also we take the product $D^{(m'')}$ to be irreducible *and* induced from subgroup H''. This applies for a space group G for induction from subgroups $G(\mathbf{k})$ to G. We write:

$$D^m \equiv \mathcal{H}^m \uparrow G; \quad D^{m'} \equiv \mathcal{H}'^{m'} \uparrow G; \quad D^{m''} \equiv \mathcal{H}''^{m''} \uparrow G. \tag{4.1}$$

Consider the direct product $D^m \otimes D^{m'}$. Since D^m and $D^{m'}$

are irreducible, the matrix U which reduces this direct product is the matrix of Clebsch–Gordan coefficients and can be found from the usual equation (2.9) which we now write out using double index notation to take account of the induced structure.

$$\sum_\gamma U_{\sigma\mu\sigma'\mu',\, m''\gamma\sigma''\mu''} U^*_{\bar\sigma\bar\mu\bar\sigma'\bar\mu',\, m''\gamma\bar\sigma''\bar\mu''} = \frac{|l_{m''}|}{g}\sum_\phi D^m(\phi)_{\sigma\mu\bar\sigma\bar\mu} D^{m'}(\phi)_{\sigma'\mu'\bar\sigma'\bar\mu'} D^{m''}(\phi)^*_{\sigma''\mu''\bar\sigma''\bar\mu''}. \quad (4.2)$$

$|l_{m''}|$ is the dimension of $D^{m''}$ and g is the order of G. γ indicates the multiplicity of $D^{m''}$ in the direct product; the $\sum_\gamma$ will be handled in the usual way. D^m, $D^{m'}$, and $D^{m''}$ are all induced representations. Indices $(\sigma\bar\sigma)$ refer to block matrices, $\mu\bar\mu$ to matrix elements, etc. Equation (4.2) can be rewritten as

$$\sum_\gamma U_{\sigma\mu\sigma'\mu',\, m''\gamma\sigma''\mu''} U^*_{\bar\sigma\bar\mu\bar\sigma'\bar\mu',\, m''\gamma\bar\sigma''\bar\mu''} = \frac{|l_{m''}|}{g}\sum_\phi \dot H^m(\phi_\sigma^{-1}\phi\phi_{\bar\sigma})_{\mu\bar\mu}\dot H'^{m'}(\phi_{\sigma'}^{-1}\phi\phi_{\bar\sigma'})_{\mu'\bar\mu'}\dot H''^{m''}(\phi_{\sigma''}^{-1}\phi\phi_{\bar\sigma''})^*_{\mu''\bar\mu''}. \quad (4.3)$$

Hence we see that the Clebsch–Gordan matrix for the irreducible representations of G can be found from the irreducible representations of the subgroups H, H', and H''.

This calculation can be further simplified by using a method due to Itzkan.[4] First consider $\sigma = \sigma' = \sigma'' = \bar\sigma = \bar\sigma' = \bar\sigma'' = 1$. Then, for this (1,1,1) block

$$U_{1\mu 1\mu',\, m''1\mu''} U^*_{1\bar\mu 1\bar\mu',\, m''1\bar\mu''} = \frac{|l_{m''}|}{g}\sum_\phi D^m(\phi)_{1\mu 1\bar\mu} D^{m'}(\phi)_{1\mu' 1\bar\mu'} D^{m''}(\phi)^*_{1\mu''1\bar\mu''} = \frac{|l_{m''}|}{g}\sum_{\phi\in\mathcal{N}} \dot H^m(\phi)_{\mu\bar\mu}\dot H'^{m'}(\phi)_{\mu'\bar\mu'}\dot H''^{m''}(\phi)^*_{\mu''\bar\mu''}, \quad (4.4)$$

where the sum is over all ϕ belonging to $\mathcal{N}$, the intersection of H, H', and H''. We assume that we have solved Eq. (4.4) by the usual procedure to find the (1,1,1) block matrix. Next we need to determine the remaining blocks in the Clebsch–Gordan matrix.

We first adopt a procedure by which we generate another block, to be denoted $U(\sigma,\sigma',\sigma'')$. Consider those elements which belong to $\mathcal{N}$, the intersection of H, H', and H''. G can be decomposed into cosets with respect to $\mathcal{N}$ such that

$$G = \mathcal{N} + \cdots + \phi_\Sigma\mathcal{N} + \cdots + \phi_N\mathcal{N}. \quad (4.5)$$

If $\psi^m_{1\nu}$, $\psi^{m'}_{1\nu}$, and $\psi^{m''}_{1\nu}$ are bases for $\dot H^m$, $\dot H'^{m'}$, and $\dot H''^{m''}$ then

$$P_{\phi_\Sigma}\psi^m_{1\nu} = \psi^m_{\sigma\nu}; \quad \text{and} \quad P_{\phi_\Sigma}\psi^{m'}_{1\nu} = \psi^{m'}_{\sigma'\nu};$$

and

$$P_{\phi_\Sigma}\psi^{m''}_{1\nu} = \psi^{m''}_{\sigma''\nu}. \quad (4.6)$$

In effect this equation (4.6) defines cosets labelled by σ, σ', σ'' since the given element ϕ_Σ transforms the 1 (identity) element of H into coset $\phi_\sigma H$, likewise $1 \to \phi_{\sigma'}H'$, $1 \to \phi_{\sigma''}H''$. Thus all members of the coset $\phi_\Sigma\mathcal{N}$ take the (1,1,1) block into the $(\sigma,\sigma',\sigma'')$ block. The equations for coefficients in the $(\sigma,\sigma',\sigma'')$ block are

$$U_{\sigma\mu\sigma'\mu',\, m''\sigma''\mu''} U^*_{\bar\sigma\bar\mu\bar\sigma'\bar\mu',\, m''\bar\sigma''\bar\mu''} = \frac{|l_{m''}|}{g}\sum_\phi D^m(\phi)_{\sigma\mu\bar\sigma\bar\mu} D^{m'}(\phi)_{\sigma'\mu'\bar\sigma'\bar\mu'} D^{m''}(\phi)^*_{\sigma''\mu''\bar\sigma''\bar\mu''}. \quad (4.7)$$

From equation (3.10) for induced representations we see that this sum is nonzero only if

$$\phi_\sigma^{-1}\phi\phi_\sigma \in H; \quad \text{and} \quad \phi_{\sigma'}^{-1}\phi\phi_{\sigma'} \in H';$$

and

$$\phi_{\sigma''}^{-1}\phi\phi_{\sigma''} \in H''. \quad (4.8)$$

If ϕ is an element that belongs to $\mathcal{N}$, we can write

$$\phi_\Sigma^{-1}\phi\phi_\Sigma = \phi_n, \quad (4.9)$$

where $\phi_n \in \mathcal{N}$, or $\phi = \phi_\Sigma\phi_n\phi_\Sigma^{-1}$. Then

$$D^m(\phi)_{\sigma\mu\bar\sigma\bar\mu} = D^m(\phi_\Sigma\phi_n\phi_\Sigma^{-1})_{\sigma\mu\bar\sigma\bar\mu}. \quad (4.10)$$

But since $\phi_\Sigma\psi^m_{1\nu} = \psi^m_{\sigma\nu}$, ϕ_Σ must be in the σth coset of G with respect to H so that $\phi_\Sigma = \phi_\sigma\phi_h$ and

$$D^m(\phi)_{\sigma\mu\bar\sigma\bar\mu} = D^m(\phi_\sigma\phi_h\phi_n\phi_h^{-1}\phi_\sigma^{-1})_{\sigma\mu\bar\sigma\bar\mu} = \dot H^m(\phi_h\phi_n\phi_h^{-1})_{\mu\bar\mu} = \sum_{\lambda\bar\lambda}\dot H^m(\phi_h)_{\mu\lambda}\dot H^m(\phi_n)_{\lambda\bar\lambda}\dot H^m(\phi_h^{-1})_{\bar\lambda\bar\mu}. \quad (4.11)$$

For ray representations we have

$$\dot H^m(\phi_h^{-1}) = (\omega_{h,h^{-1}})^{-1}\dot H^m(\phi_h)^{-1} \quad (4.12)$$

and recalling $\phi_h = \phi_\sigma^{-1}\phi_\Sigma$, we can write

$$D^m(\phi)_{\sigma\mu\bar\sigma\bar\mu} = (\omega_{h,h^{-1}})^{-1}[\dot H^m(\phi_\sigma^{-1}\phi_\Sigma)\dot H^m(\phi_n)\dot H^m(\phi_\sigma^{-1}\phi_\Sigma)^{-1}]_{\sigma\mu\bar\sigma\bar\mu} \equiv (\omega_{h,h^{-1}})^{-1}\bar{\dot H}^m(\phi_n)_{\mu\bar\mu} \quad (4.13)$$

which defines the barred (conjugated) matrix. $D^{m'}(\phi)$ and $D^{m''}(\phi)$ can be rewritten similarly so that

$$U_{\sigma\mu\sigma'\mu',\, m''\sigma''\mu''} U^*_{\bar\sigma\bar\mu\bar\sigma'\bar\mu',\, m''\bar\sigma''\bar\mu''} = \frac{|l_{m''}|}{g}\frac{1}{\omega_{h,h^{-1}}}\frac{1}{\omega_{h',h'^{-1}}}\frac{1}{\omega_{h'',h''^{-1}}} \times \sum_{\phi\in\mathcal{N}}\bar{\dot H}^m(\phi)_{\mu\bar\mu}\bar{\dot H}'^{m'}(\phi)_{\mu'\bar\mu'}\bar{\dot H}''^{m''}(\phi)^*_{\mu''\bar\mu''}. \quad (4.14)$$

Now compare Eqs. (4.4) and (4.14), with (2.9) and (2.19). The barred matrices $\bar{\dot H}$ are similarity transforms of the corresponding unbarred matrices, just as in equation (2.19) compared to (2.9). Therefore we can determine the Clebsch–Gordan matrix $U(\sigma,\sigma',\sigma'')$ of Eq. (4.14) from the matrix $U(1,1,1)$ by using the result (2.18) which immediately yields

$$U(\sigma,\sigma',\sigma'') = \frac{1}{\omega_{h,h^{-1}}}\frac{1}{\omega_{h',h'^{-1}}}\frac{1}{\omega_{h'',h''^{-1}}} \times \dot H^m(\phi_h)\otimes\dot H'^{m'}(\phi_{h'})U(1,1,1)\dot H''^{m''}(\phi_{h''})^{-1}, \quad (4.15)$$

with

$$\phi_h = \phi_\sigma^{-1}\phi_\Sigma; \quad \phi_{h'} = \phi_{\sigma'}^{-1}\phi_\Sigma; \quad \text{and} \quad \phi_{h''} = \phi_{\sigma''}^{-1}\phi_\Sigma. \quad (4.16)$$

Hence the entire Clebsch–Gordan matrix for an irreducible representation can be obtained from the representations $\dot H^m$, $\dot H'^{m'}$, and $\dot H''^{m''}$ of the subgroups H, H', and H''.

5. CLEBSCH-GORDAN COEFFICIENTS FOR A SPACE GROUP G FROM $G(k)$

The results of Sec. 4 can now be immediately taken

over to determine Clebsch–Gordan coefficients for a space group G from representations of its subgroups $G(\mathbf{k})$. First we recall some well known facts.

A space group G consists of all symmetry operators $\{\phi\,|\,\mathbf{t}\}$ which leave a lattice invariant where

$$\mathbf{r}' = \{\phi\,|\,\mathbf{t}\}\mathbf{r} = \phi\mathbf{r} + \mathbf{t}, \tag{5.1}$$

$$\{\phi_\alpha\,|\,\mathbf{t}_\alpha\}\{\phi_\beta\,|\,\mathbf{t}_\beta\} = \{\phi_\alpha\phi_\beta\,|\,\phi_\alpha\mathbf{t}_\beta + \mathbf{t}_\alpha\}, \tag{5.2}$$

and $\mathbf{t} = \mathbf{R}_L + \boldsymbol{\tau}$.

$\mathbf{R}_L$ is a lattice vector and $\boldsymbol{\tau}$ is a fractional translation. The group T consisting of all operations $\{\epsilon\,|\,\mathbf{R}_L\}$ is a subgroup of G, so we can write

$$G = T + \{\phi_2\,|\,\boldsymbol{\tau}_2\}T + \cdots + \{\phi_k\,|\,\boldsymbol{\tau}_k\}T. \tag{5.3}$$

Irreducible representations of G are written as $D^{(*\mathbf{k})(l)}(\{\phi\,|\,\mathbf{t}\})$, and the representations of T have the form

$$D^{(*\mathbf{k})(l)}(\{\epsilon\,|\,\mathbf{R}_L\}) = \exp(i\mathbf{k}\cdot\mathbf{R}_L). \tag{5.4}$$

In addition we can write

$$G = G(\mathbf{k}) + \cdots + \{\phi_\lambda\,|\,\boldsymbol{\tau}_\lambda\}G(\mathbf{k}) + \cdots + \{\phi_s\,|\,\boldsymbol{\tau}_s\}G(\mathbf{k}), \tag{5.5}$$

where $G(\mathbf{k})$, the group of $\mathbf{k}$ consists of all those elements that leave wave vector $\mathbf{k}$ invariant modulo a reciprocal lattice vector.

The irreducible representations $D^{(\mathbf{k})(l)}$ of $G(\mathbf{k})$, are ray representations. This can be seen as follows: If we decompose $G(\mathbf{k})$ with respect to T then

$$G(\mathbf{k}) = T + \{\phi_2\,|\,\boldsymbol{\tau}_2\}T + \cdots + \{\phi_k\,|\,\boldsymbol{\tau}_k\}T. \tag{5.6}$$

Consider the product of 2 coset representatives in Eq. (5.6).

$$\begin{aligned}\{\phi_\alpha\,|\,\boldsymbol{\tau}_\alpha\}\{\phi_\beta\,|\,\boldsymbol{\tau}_\beta\} &= \{\phi_\alpha\phi_\beta\,|\,\phi_\alpha\boldsymbol{\tau}_\beta + \boldsymbol{\tau}_\alpha\}\\ &= \{\phi_\alpha\phi_\beta\,|\,\phi_\alpha\boldsymbol{\tau}_\beta + \boldsymbol{\tau}_\alpha - \boldsymbol{\tau}_{\alpha\beta} + \boldsymbol{\tau}_{\alpha\beta}\}\\ &= \{\epsilon\,|\,\mathbf{R}_{L_{\alpha\beta}}\}\{\phi_{\alpha\beta}\,|\,\boldsymbol{\tau}_{\alpha\beta}\},\end{aligned} \tag{5.7}$$

where

$$\mathbf{R}_{L_{\alpha\beta}} = \phi_\alpha\boldsymbol{\tau}_\beta + \boldsymbol{\tau}_\alpha - \boldsymbol{\tau}_{\alpha\beta}. \tag{5.8}$$

Since the representations of $G(\mathbf{k})$ must multiply in the same way as the elements of $G(\mathbf{k})$ we have

$$\begin{aligned}&D^{(\mathbf{k})(l)}(\{\phi_\alpha\,|\,\boldsymbol{\tau}_\alpha\})D^{(\mathbf{k})(l)}(\{\phi_\beta\,|\,\boldsymbol{\tau}_\beta\})\\ &\quad = \exp(-i\mathbf{k}\cdot\mathbf{R}_{L_{\alpha\beta}})D^{(\mathbf{k})(l)}(\{\phi_{\alpha\beta}\,|\,\boldsymbol{\tau}_{\alpha\beta}\}).\end{aligned} \tag{5.9}$$

Now consider the point group $P(\mathbf{k})$ which is isomorphic to $G(\mathbf{k})/T$ and which has elements $\{\epsilon, \phi_1, \ldots, \phi_k\}$ where

$$\phi_\alpha\phi_\beta = \phi_{\alpha\beta} \tag{5.10}$$

and for the representations of $P(\mathbf{k})$

$$D^{(j)}(\phi_\alpha)D^{(j)}(\phi_\beta) = D^{(j)}(\phi_{\alpha\beta}). \tag{5.11}$$

If we define a ray representation $\Gamma^{(j)}$ of $P(\mathbf{k})$ with factor set $\omega^k_{\alpha,\beta}$, then $\Gamma^{(j)}$ will multiply as

$$\Gamma^{(j)}(\phi_\alpha)\Gamma^{(j)}(\phi_\beta) = \omega^k_{\alpha,\beta}\Gamma^{(j)}(\phi_{\alpha\beta}), \tag{5.12}$$

where $|\omega^k_{\alpha,\beta}| = 1$ and

$$\omega^k_{\alpha,\beta}\omega^k_{\alpha\beta,\gamma} = \omega^k_{\alpha,\beta\gamma}\omega^k_{\beta,\gamma}. \tag{5.13}$$

Comparing Eqs. (5.9), (5.11), and (5.12) we see that the irreducible representations $D^{(\mathbf{k})(l)}$ of $G(\mathbf{k})$ are ray representations of $P(\mathbf{k})$ with factor set

$$\omega^k_{\alpha,\beta} = \exp(-i\mathbf{k}\cdot\mathbf{R}_{L_{\alpha\beta}}) \tag{5.14}$$

and in particular

$$\omega^k_{\alpha,\alpha^{-1}} = 1. \tag{5.15}$$

All the irreducible representations $D^{(*\mathbf{k})(l)}$ of a space group G can be induced from the irreducible representations $D^{(\mathbf{k})(l)}$ of $G(\mathbf{k})$. In this case the induced representations are irreducible and

$$D^{(*\mathbf{k})(l)}(\{\phi_\varepsilon\,|\,\mathbf{t}_\varepsilon\})_{\sigma\mu,\tau\nu} = \dot{D}^{(\mathbf{k})(l)}(\{\phi_\sigma\,|\,\boldsymbol{\tau}_\sigma\}^{-1}\{\phi_\varepsilon\,|\,\mathbf{t}_\varepsilon\}\{\phi_\tau\,|\,\boldsymbol{\tau}_\tau\})_{\mu\nu}. \tag{5.16}$$

The Clebsch–Gordan coefficients for the reduction of the direct product of 2 space group representations, $D^{(*\mathbf{k})(l)} \otimes D^{(*\mathbf{k}')(l')}$ into a direct sum of representations $D^{(*\mathbf{k}'')(l'')}$ is given by

$$\begin{aligned}&\sum_\gamma U_{\sigma\mu\sigma'\mu',\,l''\gamma\sigma''\mu''}U^*_{\bar\sigma\bar\mu\bar\sigma'\bar\mu',\,l''\gamma\bar\sigma''\bar\mu''}\\ &\quad = \frac{|l''|}{gt}\sum_x D^{(*\mathbf{k})(l)}(\{\phi_x\,|\,\mathbf{t}_x\})_{\sigma\mu\bar\sigma\bar\mu}D^{(*\mathbf{k}')(l')}(\{\phi_x\,|\,\mathbf{t}_x\})_{\sigma'\mu'\bar\sigma'\bar\mu'}\\ &\qquad\times D^{(*\mathbf{k}'')(l'')}(\{\phi_x\,|\,\mathbf{t}_x\})^*_{\sigma''\mu''\bar\sigma''\bar\mu''},\end{aligned} \tag{5.17}$$

where $|l''|$ is the dimension of $D^{(*\mathbf{k}'')(l'')}$ and gt is the order of the space group. The sum is on all elements in G. Any element $\{\phi_\sigma\,|\,\mathbf{t}_\sigma\}$ can be written as a product of a coset representative $\{\phi_\sigma\,|\,\boldsymbol{\tau}_\sigma\}$ and a pure translation $\{\epsilon\,|\,\mathbf{R}_L\}$ and then

$$\begin{aligned}&D^{(*\mathbf{k})(l)}(\{\phi_x\,|\,\mathbf{t}_x\})_{\sigma\mu\bar\sigma\bar\mu}\\ &\quad = \sum_{\beta b} D^{(*\mathbf{k})(l)}(\{\phi_x\,|\,\boldsymbol{\tau}_x\})_{\sigma\mu\beta b}D^{(*\mathbf{k})(l)}\{\epsilon\,|\,\mathbf{R}_L\})_{\beta b\bar\sigma\bar\mu},\\ &\quad = \sum_{\beta b} D^{(*\mathbf{k})(l)}(\{\phi_x\,|\,\boldsymbol{\tau}_x\})_{\sigma\mu\beta b}\delta_{\beta\bar\sigma}\delta_{b\bar\mu}\exp(-i\mathbf{k}_{\bar\sigma}\cdot\mathbf{R}_L),\\ &\quad = \exp(-i\mathbf{k}_{\bar\sigma}\cdot\mathbf{R}_L)D^{(*\mathbf{k})(l)}(\{\phi_x\,|\,\boldsymbol{\tau}_x\})_{\sigma\mu\bar\sigma\bar\mu}.\end{aligned} \tag{5.18}$$

Equation (5.17) can, therefore, be rewritten as

$$\begin{aligned}&\sum_\gamma U_{\sigma\mu\sigma'\mu',\,l''\gamma\sigma''\mu''}U^*_{\bar\sigma\bar\mu\bar\sigma'\bar\mu',\,l''\gamma\bar\sigma''\bar\mu''}\\ &\quad = \frac{|l''|}{gt}\sum_{\mathbf{R}_L}\exp[-i(\mathbf{k}_\sigma + \mathbf{k}'_{\sigma'} - \mathbf{k}''_{\sigma''})\cdot\mathbf{R}_L]\\ &\qquad\times\sum_x D^{(*\mathbf{k})(l)}(\{\phi_x\,|\,\boldsymbol{\tau}_x\})_{\sigma\mu\bar\sigma\bar\mu}D^{(*\mathbf{k}')(l')}(\{\phi_x\,|\,\boldsymbol{\tau}_x\})_{\sigma'\mu'\bar\sigma'\bar\mu'}\\ &\qquad\times D^{(*\mathbf{k}'')(l'')}(\{\phi_x\,|\,\boldsymbol{\tau}_x\})^*_{\sigma''\mu''\bar\sigma''\bar\mu''}.\end{aligned} \tag{5.19}$$

The sum on x is now only over coset representatives.

But $\sum_{\mathbf{R}_L}\exp(-i\mathbf{k}\cdot\mathbf{R}_L) = 0$ if $\mathbf{k} \neq 2\pi\mathbf{B}_H$ where $\mathbf{B}_H$ is a reciprocal lattice vector and $\sum_{\mathbf{R}_L}\exp(-i\mathbf{k}\cdot\mathbf{R}_L) = gt/g$ if $\mathbf{k} = 2\pi\mathbf{B}_H$, where g is the order of G/T. Therefore, the Clebsch–Gordan coefficients are zero unless $\mathbf{k}_{\bar\sigma} + \mathbf{k}'_{\bar\sigma'} - \mathbf{k}''_{\bar\sigma''} = 2\pi\mathbf{B}_H$. For these nonzero coefficients

$$\begin{aligned}&\sum_\gamma U_{\sigma\mu\sigma'\mu',\,l''\gamma\sigma''\mu''}U^*_{\bar\sigma\bar\mu\bar\sigma'\bar\mu',\,l''\gamma\bar\sigma''\bar\mu''}\\ &\quad = \frac{|l''|}{g}\sum_x D^{(*\mathbf{k})(l)}(\{\phi_x\,|\,\boldsymbol{\tau}_x\})_{\sigma\mu\bar\sigma\bar\mu}D^{(*\mathbf{k}')(l')}(\{\phi_x\,|\,\boldsymbol{\tau}_x\})_{\sigma'\mu'\bar\sigma'\bar\mu'}\\ &\qquad\times D^{(*\mathbf{k}'')(l'')}(\{\phi_x\,|\,\boldsymbol{\tau}_x\})^*_{\sigma''\mu''\bar\sigma''\bar\mu''}.\end{aligned} \tag{5.20}$$

However, the irreducible representations $D^{(*\mathbf{k})(l)}$, $D^{(*\mathbf{k}')(l')}$, and $D^{(*\mathbf{k}'')(l'')}$ are induced from $D^{(\mathbf{k})(l)}$, $D^{(\mathbf{k}')(l')}$, and $D^{(\mathbf{k}'')(l'')}$. Therefore we may immediately take over the discussion in Sec. 4. The (1, 1, 1) block of coefficients is given by solution of the equation

$$U_{1\mu 1\mu', l''\gamma 1\mu''}U^{*}_{1\bar{\mu}1\bar{\mu}', l''\gamma 1\bar{\mu}''}$$
$$= \frac{|l''|}{g}\sum_{x} D^{(\mathbf{k})(l)}(\{\phi_x|\tau_x\})_{\mu\bar{\mu}} D^{(\mathbf{k}')(l')}(\{\phi_x\ \tau_x\})_{\mu'\bar{\mu}'}$$
$$\times D^{(\mathbf{k}'')(l'')}(\{\phi_x|\tau_x\})^{*}_{\mu''\bar{\mu}''}. \tag{5.21}$$

Assuming we have solved (5.21) for the (1,1,1) block, we immediately obtain the $(\sigma,\sigma',\sigma'')$ block as

$$U_{\sigma\nu\sigma'\nu', l''\gamma\sigma''\nu''}$$
$$= [D^{(\mathbf{k})(l)}(\{\phi_k|t_k\})\otimes D^{(\mathbf{k}')(l')}(\{\phi_{k'}|t_{k'}\})]_{\nu\mu\nu'\mu'}U_{1\mu 1\mu', l''\gamma 1\mu''}$$
$$\times D^{(\mathbf{k}'')(l'')}(\{\phi_{k''}|t_{k''}\})^{-1}_{\mu''\nu''}, \tag{5.22}$$

where

$$\{\phi_k|t_k\} = \{\phi_\sigma|\tau_\sigma\}^{-1}\{\phi_\Sigma|\tau_\Sigma\},$$
$$\{\phi_{k'}|t_{k'}\} = \{\phi_{\sigma'}|\tau_{\sigma'}\}^{-1}\{\phi_\Sigma|\tau_\Sigma\},$$
$$\{\phi_{k''}|t_{k''}\} = \{\phi_{\sigma''}|\tau_{\sigma''}\}^{-1}\{\phi_\Sigma|\tau_\Sigma\}, \tag{5.23}$$

and $\phi_\Sigma\mathbf{k} = \mathbf{k}_\sigma$; $\phi_\Sigma\mathbf{k}' = \mathbf{k}'_{\sigma'}$; $\phi_\Sigma\mathbf{k}'' = \mathbf{k}''_{\sigma''}$. Hence the Clebsch–Gordan matrix for space group G can be found from the representations for the subgroups $G(\mathbf{k})$, $G(\mathbf{k}')$, and $G(\mathbf{k}'')$.

We can obtain some useful simplifications by explicitly using the ray representation nature of irreducible representations. We recall that the representations of $G(\mathbf{k})$, $G(\mathbf{k}')$, and $G(\mathbf{k}'')$ are ray representations of $P(\mathbf{k})$, $P(\mathbf{k}')$, and $P(\mathbf{k}'')$, respectively. In particular, it is convenient to use the factor system used by Kovalev[6] so that we write

$$D^{(\mathbf{k})(l)}(\{\phi_x|\tau_x\})_{\mu\bar{\mu}} = \exp(-i\mathbf{k}\cdot\tau_x)\Gamma^{(\mathbf{k})(l)}(\phi_x)_{\mu\bar{\mu}}, \tag{5.24}$$

where

$$\Gamma^{(\mathbf{k})(l)}(\phi_x)\Gamma^{(\mathbf{k})(l)}(\phi_y) = \bar{\omega}^{k}_{x,y}\Gamma^{(\mathbf{k})(l)}(\phi_{xy}) \tag{5.25}$$

and

$$\bar{\omega}^{k}_{x,y} = \exp[-i\mathbf{k}\cdot(\phi_x-\epsilon)\tau_y]. \tag{5.26}$$

Hence, if we use equation (5.24) in (5.21) we can write the (1,1,1) block of coefficients in terms of ray representations. Then we have

$$U_{1\mu 1\mu', 1\mu''}U^{*}_{1\bar{\mu}1\bar{\mu}', 1\bar{\mu}''}$$
$$= \frac{|l''|}{g}\sum_{x}\exp(-2\pi i\mathbf{B}_1\cdot\tau_x)\Gamma^{(\mathbf{k})(l)}(\phi_x)_{\mu\bar{\mu}}\Gamma^{(\mathbf{k}')(l')}(\phi_x)_{\mu'\bar{\mu}'}$$
$$\times\Gamma^{(\mathbf{k}'')(l'')}(\phi_x)^{*}_{\mu''\bar{\mu}''}, \tag{5.27}$$

where $2\pi\mathbf{B}_1 = \mathbf{k}+\mathbf{k}'-\mathbf{k}''$. If we have a symmorphic group, or if $\mathbf{B}_1 = 0$, then

$$U_{1\mu 1\mu', l''\gamma 1\mu''}U^{*}_{1\mu 1\mu', l''\gamma 1\mu''}$$
$$= \frac{|l''|}{g}\sum_{x}\Gamma^{(\mathbf{k})(l)}(\phi_x)_{\mu\bar{\mu}}\Gamma^{(\mathbf{k}')(l')}(\phi_x)_{\mu'\bar{\mu}'}\Gamma^{(\mathbf{k}'')(l'')}(\phi_x)_{\mu''\bar{\mu}''}. \tag{5.28}$$

In this case, for particular $\mathbf{k}$, $\mathbf{k}'$, and $\mathbf{k}''$, if two space groups have the same point groups $P(\mathbf{k})$, $P(\mathbf{k}')$, and $P(\mathbf{k}'')$ and consequently the same $\Gamma^{(\mathbf{k})}$, $\Gamma^{(\mathbf{k}')}$, and $\Gamma^{(\mathbf{k}'')}$, their Clebsch–Gordan coefficients for the (1,1,1) block are identical. For example, this block of coefficients for $D^{(*\Delta)(l)}\otimes D^{(*\Delta)(l')} = D^{(*\Delta)(l'')}$ in diamond (O^7_h—a nonsymmorphic group) are the same as those for rocksalt (O^5_h—a symmorphic group) when $\mathbf{k}+\mathbf{k}'-\mathbf{k}'' = 0$. We must consider the phase factor $\exp(-i2\pi\mathbf{B}_1\cdot\tau_x)$ in Eq. (5.27) explicitly only for Umklapp processes in nonsymmorphic crystals.

If we now express the $(\sigma,\sigma',\sigma'')$ block using the ray representations Γ, we have

$$U(\sigma,\sigma',\sigma'')$$
$$= \exp[-i(\mathbf{k}\cdot\mathbf{t}_k+\mathbf{k}'\cdot\mathbf{t}_{k'}-\mathbf{k}''\cdot\mathbf{t}_{k''})]\Gamma^{(\mathbf{k})(l)}(\phi_k)$$
$$\otimes\Gamma^{(\mathbf{k}')(l')}(\phi_{k'})U(1,1,1)\Gamma^{(\mathbf{k}'')(l'')}(\phi_{k''})^{-1}. \tag{5.29}$$

Since

$$\{\phi_\kappa|\mathbf{t}_\kappa\} = \{\phi_\sigma|\tau_\sigma\}^{-1}\{\phi_\Sigma|\tau_\Sigma\} \tag{5.30}$$

then

$$\mathbf{t}_\kappa = \phi_\sigma^{-1}(\tau_\Sigma-\tau_\sigma) \tag{5.31}$$

and similarly

$$\mathbf{t}_{\kappa'} = \phi_{\sigma'}^{-1}(\tau_\Sigma-\tau_{\sigma'}); \quad \mathbf{t}_{\kappa''} = \phi_{\sigma''}^{-1}(\tau_\Sigma-\tau_{\sigma''}). \tag{5.32}$$

It is convenient, therefore, when calculating Clebsch–Gordan coefficients to choose ϕ_σ, $\phi_{\sigma'}$, $\phi_{\sigma''}$, and ϕ_Σ which have zero fractionals whenever possible.

6. REDUCIBLE INDUCED REPRESENTATIONS

Now we consider the situation in which $H^{(m)}\uparrow G$, and $H'^{m'}\uparrow G$ are both reducible. The analysis parallels that given in Eqs. (2.20)–(2.25). Assume a unitary matrix M^m exists which reduces $H^m\uparrow G$

$$(M^m)^{-1}\Delta^m M^m = (H^m\uparrow G), \tag{6.1}$$

where

$$\Delta^m_{l\alpha\bar{\mu},\,l\alpha\mu} = D^{(l)}_{\bar{\mu}\mu}\delta_{\bar{\gamma}l}\delta_{\bar{\alpha}\alpha} \tag{6.2}$$

and $D^{(l)}$ is irreducible in G; $\alpha\equiv 1,\ldots,(\Delta^m|l)$ is the multiplicity of D^l in Δ^m. Similarly we define $\Delta^{m'}$.

From (6.1) we have

$$(H\uparrow G)(\phi)_{\sigma a\bar{\sigma}\bar{a}} = \sum_{l\alpha\mu\bar{\mu}} M^{*}_{l\alpha\mu,\sigma a}D^{(l)}(\phi)_{\mu\bar{\mu}}M_{l\alpha\bar{\mu},\bar{\sigma}\bar{a}}. \tag{6.3}$$

If we multiply both sides of (6.3) by $D^{\bar{l}}_{\bar{\mu}'\mu'}(\phi)^{*}$ sum on ϕ and use the orthogonality theorem, then

$$\sum_{\alpha} M^{*}_{l\alpha\mu,\sigma a}M_{l\alpha\bar{\mu},\bar{\sigma}\bar{a}} = \frac{|l|}{g}\sum_{\phi}(H\uparrow G)(\phi)_{\sigma a\bar{\sigma}\bar{a}}D^{(l)}(\phi)^{*}_{\mu\bar{\mu}}. \tag{6.4}$$

Hence, the elements of M can be found in the same way as one finds the elements of Clebsch–Gordan matrices; that is, setting $\mu=\bar{\mu}$, and $\sigma a = \bar{\sigma}\bar{a}$, we find a nonzero element, set its phase as real, fix μ,σ,a and let $\bar{\mu},\bar{\sigma},\bar{a}$ vary. In this manner we obtain M for $\alpha=1$. If D^l occurs more than once in Δ, we repeat this procedure for a different initial nonzero element to obtain another M matrix orthogonal to the first.

Consider the direct product of 2 induced representations, $[H^m\uparrow G\otimes H'^{m'}\uparrow G]$. There exists a unitary matrix $\bar{Z}$ which reduces this direct product into a direct sum of irreducible representations $D^{l''}$.

$$\bar{Z}^{-1}[H^m\uparrow G\otimes H'^{m'}\uparrow G]\bar{Z} = \sum_{l''\beta}\oplus D^{(l''_\beta)}. \tag{6.5}$$

There is also a matrix Z which reduces the direct product of reduced representations $\Delta^m\otimes\Delta^{m'}$

$$Z^{-1}[\Delta^m\otimes\Delta^{m'}]Z = \sum_{l''\beta}\oplus D^{(l''_\beta)} \tag{6.6}$$

or

$$\sum_\beta Z_{l\alpha\mu l'\alpha'\mu', l''\beta\mu''} Z^*_{\bar{l}\bar{\alpha}\bar{\mu}\bar{l}'\bar{\alpha}'\bar{\mu}', l''\beta\bar{\mu}''}$$

$$= \frac{|l''|}{g} \sum_\phi \Delta^m(\phi)_{l\alpha\mu,\bar{l}\bar{\alpha}\bar{\mu}} \Delta^{m'}(\phi)_{l'\alpha'\mu\bar{l}'\bar{\alpha}'\bar{\mu}} D^{(l'')}(\phi)^*_{\mu''\bar{\mu}''},$$

$$= \frac{|l''|}{g} \sum_\phi D^{(l)}(\phi)_{\mu\bar{\mu}} D^{(l')}(\phi)_{\mu'\bar{\mu}'} D^{(l'')}(\phi)^*_{\mu''\bar{\mu}''} \delta_{\alpha\bar{\alpha}} \delta_{l\bar{l}} \delta_{\alpha'\bar{\alpha}'} \delta_{l'\bar{l}'},$$

$$= \sum_\gamma U^{l\otimes l'}_{\mu\mu', l''\gamma\mu''} U^{l\otimes l'*}_{\bar{\mu}\bar{\mu}', l''\gamma\bar{\mu}''} \delta_{\alpha\bar{\alpha}} \delta_{l\bar{l}} \delta_{\alpha'\bar{\alpha}'} \delta_{l'\bar{l}'}. \tag{6.7}$$

Hence the elements of the matrix U which reduces $D^l \otimes D^{l'}$ can be read off from the elements of Z. But Δ^m and $\Delta^{m'}$ are related to $H^m \uparrow G$ and $H'^{m'} \uparrow G$ by unitary transformations by M^m and $M^{m'}$, respectively. Hence as in Sec. 2, the matrix Z is a transform of the matrix $\bar{Z}$ and

$$U \sim Z = (M^m \otimes M^{m'}) \bar{Z}. \tag{6.8}$$

Now consider $\bar{Z}$ in more detail. $\bar{Z}$ brings the direct product $(H^m \uparrow G) \otimes (H'^{m'} \uparrow G)$ into a fully reduced matrix. Let the first $l_{m''}$ rows and columns of this fully reduced matrix contain those irreducible representations which occur in the reduction of some induced representation $H''^{m''} \uparrow G$. Then this induced representation has a reduced form $\Delta^{m''}$ given by

$$\Delta^{m''} = M^{m''}[H''^{m''} \uparrow G] M^{m''-1} \tag{6.9}$$

and

$$\Delta^{m''}_{l''\alpha''\mu'', \bar{l}''\bar{\alpha}''\bar{\mu}''} = D^{(l'')}_{\mu\bar{\mu}} \delta_{l''\bar{l}''} \delta_{\alpha''\bar{\alpha}''}. \tag{6.10}$$

Hence for those rows and columns $\bar{Z}$ must satisfy

$$\sum_\beta \bar{Z}_{ij, l''\beta k} \bar{Z}^*_{\bar{i}\bar{j}, l''\beta\bar{k}}$$

$$= \sum_{d\bar{d}} M^{m''}_{l''\alpha''k, d} M^{m''-1}_{\bar{d}, l''\alpha''\bar{k}} \frac{|l''|}{g} \sum_\phi (H^m \uparrow G)(\phi)_{i\bar{i}}$$

$$\times (H'^{m'} \uparrow G)(\phi)_{j\bar{j}} (H''^{m''} \uparrow G)(\phi)^*_{d\bar{d}}. \tag{6.11}$$

Or, rewriting the induced representations in block form

$$\sum_\beta \bar{Z}_{\sigma a\sigma' b, l''\beta\sigma'' c} \bar{Z}^*_{\bar{\sigma}\bar{a}\bar{\sigma}'\bar{b}, l''\beta\bar{\sigma}''\bar{c}}$$

$$= \sum_{\lambda d\bar{\lambda}\bar{d}} M^{m''}_{l''\alpha''\sigma c, \lambda d} M^{m''-1}_{\bar{\lambda}\bar{d}, l''\alpha''\bar{\sigma}\bar{c}}$$

$$\times \frac{|l''|}{g} \sum_\phi \dot{H}^m(\phi^{-1}_\sigma \phi \phi_{\bar{\sigma}})_{a\bar{a}} \dot{H}'^{m'}(\phi^{-1}_{\sigma'} \phi \phi_{\bar{\sigma}'})_{b\bar{b}} \dot{H}''^{m''}(\phi^{-1}_\lambda \phi \phi_{\bar{\lambda}})^*_{d\bar{d}}. \tag{6.12}$$

Thus $\bar{Z}$ is calculated from the matrix $M^{m''}$ and the irreducible representations for the subgroups H, H', and H''. The sum on β is handled in the usual way. For any β, if we define

$$W_{\sigma a\sigma' b, m''\lambda d} W^*_{\bar{\sigma}\bar{a}\bar{\sigma}'\bar{b}, m''\bar{\lambda}\bar{d}}$$

$$= \frac{|l''|}{g} \sum_\phi \dot{H}^m(\phi^{-1}_\sigma \phi \phi_{\bar{\sigma}})_{a\bar{a}} \dot{H}'^{m'}(\phi^{-1}_{\sigma'} \phi \phi_{\bar{\sigma}'})_{b\bar{b}} \dot{H}''^{m''}(\phi^{-1}_\lambda \phi \phi_{\bar{\lambda}})^*_{d\bar{d}}, \tag{6.13}$$

then

$$\bar{Z} = W M^{m''-1} \tag{6.14}$$

and

$$Z = (M^m \otimes M^{m'}) W M^{m''-1}. \tag{6.15}$$

The elements of W are most easily calculated in blocks as in (Sec. 4) so that

$$W_{1a1b, m''1d} W^*_{1\bar{a}1\bar{b}, m''1\bar{d}}$$

$$= \frac{|l''|}{g} \sum_\phi \dot{H}^m(\phi)_{a\bar{a}} \dot{H}'^{m'}(\phi)_{b\bar{b}} \dot{H}''^{m''}(\phi)^*_{d\bar{d}} \tag{6.16}$$

and

$$W(\sigma, \sigma', \sigma'') = \frac{1}{\omega_{h,h^{-1}}} \frac{1}{\omega_{h',h'^{-1}}} \frac{1}{\omega_{h'',h''^{-1}}} H^m(\phi_h)$$

$$\otimes H'^{m'}(\phi_{h'}) W(1,1,1) H''^{m''}(\phi_{h''})^{-1}, \tag{6.17}$$

where ϕ_h, $\phi_{h'}$, and $\phi_{h''}$ are given in Eq. (4.16).

7. INDUCING FROM A GENERAL STAR TO A SPECIAL STAR

In this section we develop the theory by demonstrating how to obtain the Clebsch–Gordan coefficients by inducing from the subgroup of lowest symmetry T, or what is equivalent, by working with induced ray representations of the trivial point group $P(\epsilon) = \epsilon$, consisting of the identity. At first sight this may appear excessively cumbersome. However the advantages of this procedure are that the induced representations are ray "regular" representations and contain only a single nonzero element in each row and column and also that we can obtain in one calculational step an entire set of Clebsch–Gordan coefficients. In case of space groups this means that for a given triangle of stars *k, *k', and *k'', we obtain the coefficients for all allowable irreducible representations at each of those stars.

The irreducible representations $D^{(\mathbf{k})(l)}$ of $G(\mathbf{k})$ are irreducible ray representations of the point group $P(\mathbf{k})$ which is isomorphic to $G(\mathbf{k})/T$, and so the representations satisfy,

$$D^{(\mathbf{k})(l)}(\{\phi_\alpha | \tau_\alpha\}) D^{(\mathbf{k})(l)}(\{\phi_\beta | \tau_\beta\})$$

$$= \omega^k_{\alpha,\beta} D^{(\mathbf{k})(l)}(\{\phi_{\alpha\beta} | \tau_{\alpha\beta}\}), \tag{7.1}$$

where

$$\omega^k_{\alpha,\beta} = \exp[-i\mathbf{k} \cdot (\phi_\alpha \tau_\beta + \tau_\alpha - \tau_{\alpha\beta})] \tag{7.2}$$

and

$$\omega^k_{\alpha,\alpha^{-1}} = 1.$$

If $\mathbf{k}$ is a wave vector of a special star then $P(\mathbf{k})$ has a subgroup $P(\epsilon)$, the point group of a general wave vector $\mathbf{k}_\epsilon$. $P(\epsilon)$ contains only one element, the identity, and $G(\mathbf{k}_\epsilon)$ has only one representation, $D^{(\epsilon)}(\{\epsilon | 0\}) = 1$. Since $P(\epsilon)$ is a subgroup of $P(\mathbf{k})$

$$P(\mathbf{k}) = P(\epsilon) + \cdots + \phi_\sigma P(\epsilon) + \cdots + \phi_g P(\epsilon) \tag{7.3}$$

and the induced ray representation $D^{(\epsilon)} \uparrow G(\mathbf{k})$ has elements given by

$$[D^{(\epsilon)} \uparrow G(\mathbf{k})](\{\phi_\rho | \tau_\rho\})_{\sigma\tau}$$

$$= (\omega^k_{\sigma,\sigma^{-1}})^{-1} D^{(\epsilon)}(\{\phi_\sigma | \tau_\sigma\}^{-1} \{\phi_\rho | \tau_\rho\} \{\phi_\tau | \tau_\tau\})$$

$$= \begin{cases} \omega^k_{\sigma^{-1},\rho\tau} \omega^k_{\rho,\tau} & \text{for } \phi_{\sigma^{-1}} \phi_\rho \phi_\tau = \epsilon \\ 0 & \text{otherwise.} \end{cases} \tag{7.4}$$

Equation (7.4) defines a "ray-regular" representation of $G(\mathbf{k})$. (This differs from the usual definition of a regular representation by the factor ω.) This representation contains all the irreducible representations of $G(\mathbf{k})$ and each representation $D^{(\mathbf{k})(l)}$ occurs $|l|$ times.

There exists a unitary matrix M^k which brings the induced representation $[D^{(\epsilon)} \uparrow G(\mathbf{k})]$ into a fully reduced form Δ^k so that

$$\Delta^k = M^k[D^{(\epsilon)} \uparrow G(\mathbf{k})]M^{k-1} \tag{7.5}$$

and

$$\Delta^k_{\bar{l}\bar{\alpha}\bar{\mu},\, l\alpha\mu} = D^{(\mathbf{k})(l)}_{\bar{\mu}\mu}\delta_{\bar{l}l}\delta_{\bar{\alpha}\alpha}. \tag{7.6}$$

Now consider the Clebsch–Gordan matrix for $D^{(\mathbf{k})(l)} \otimes D^{(\mathbf{k}')(l')} \to D^{(\mathbf{k}'')(l'')}$. $D^{(\mathbf{k})(l)}$, $D^{(\mathbf{k}')(l')}$, and $D^{(\mathbf{k}'')(l'')}$ can all be induced from $D^{(\epsilon)}$. We have seen in the previous section that this Clebsch–Gordan matrix can be found from the representation $D^{(\epsilon)}$ and the matrices M^k, $M^{k'}$, and $M^{k''}$ which reduce $(D^{\epsilon} \uparrow G(\mathbf{k}))$, $(D^{\epsilon} \uparrow G(\mathbf{k}'))$, and $(D^{\epsilon} \uparrow G(k''))$, respectively. Specifically, from Eq. (6.15) we see that we need to calculate

$$Z = (M^k \otimes M^{k'})WM^{k''}, \tag{7.7}$$

where

$$\begin{aligned} &W_{1a1b,\, k''l''1d}W^*_{1\bar{a}1\bar{b},\, k''l''1\bar{d}} \\ &\quad = \frac{|l''|}{g}\sum_{\phi} D^{(\epsilon)}(\phi)_{a\bar{a}}D^{(\epsilon)}(\phi)_{b\bar{b}}D^{(\epsilon)}(\phi)^*_{d\bar{d}} \\ &\quad = \frac{|l''|}{g}\,\delta_{a\bar{a}}\delta_{b\bar{b}}\delta_{d\bar{d}} \end{aligned} \tag{7.8}$$

and

$$\begin{aligned} W(\sigma,\sigma',\sigma'') &= D^{(\epsilon)}(\{\phi_k|\mathbf{t}_k\}) \otimes D^{(\epsilon)}(\{\phi_{k'}|\mathbf{t}_{k'}\})W(1,1,1) \\ &\quad \times D^{(\epsilon)}(\{\phi_{k''}|\mathbf{t}_{k''}\})^{-1}. \end{aligned} \tag{7.9}$$

Since $G(\mathbf{k})$, $G(\mathbf{k}')$, and $G(\mathbf{k}'')$ are all induced from $P(\epsilon)$, the $(\sigma,\sigma',\sigma'')$ block is nonzero only for $\sigma = \sigma' = \sigma''$ and then $\{\phi_k|\mathbf{t}_k\} = \{\phi_{k'}|\mathbf{t}_{k'}\} = \{\phi_{k''}|\mathbf{t}_{k''}\} = \{\epsilon|0\}$, and

$$\begin{aligned} W(\sigma,\sigma',\sigma'') &= \left(\frac{|l''|}{g}\right)^{1/2} \quad \text{for } \sigma = \sigma' = \sigma'' \\ &= 0 \qquad\qquad \text{otherwise.} \end{aligned} \tag{7.10}$$

Then

$$\begin{aligned} &Z_{l\alpha\mu\, l'\alpha'\mu',\, l''\alpha''\mu''} \\ &\quad = \sum_{\sigma\sigma'\sigma''} M^k_{l\alpha\mu,\sigma}M^{k'}_{l'\alpha'\mu',\sigma'}W(\sigma,\sigma',\sigma'')M^{k''*}_{l''\alpha''\mu'',\sigma''} \\ &\quad = \left(\frac{|l''|}{g}\right)^{1/2}\sum_{\sigma} M^k_{l\alpha\mu,\sigma}M^{k'}_{l'\alpha'\mu',\sigma}M^{k''*}_{l''\alpha''\mu'',\sigma}, \end{aligned} \tag{7.11}$$

where we recall from Eq. (6.4) that the matrix M is most easily obtained from

$$\begin{aligned} &\sum_{\alpha} M^{k*}_{l\alpha\mu,\sigma}M^k_{l\alpha\bar{\mu},\bar{\sigma}} \\ &\quad = \frac{|l|}{g}\sum_{\rho} D^{(\mathbf{k})(l)}(\{\phi_\rho|\boldsymbol{\tau}_\rho\})^*_{\mu\bar{\mu}}[D^{(\epsilon)} \uparrow G(\mathbf{k})](\{\phi_\rho|\boldsymbol{\tau}_\rho\})_{\sigma\bar{\sigma}} \\ &\quad = \frac{|l|}{g}\sum_{\rho} D^{(\mathbf{k})(l)}(\{\phi_\rho|\boldsymbol{\tau}_\rho\})^*_{\mu\bar{\mu}}\omega^k_{\sigma^{-1},\rho\bar{\sigma}}\omega^k_{\rho,\bar{\sigma}} \quad \text{for } \phi_{\sigma^{-1}\rho\bar{\sigma}} = \epsilon. \end{aligned} \tag{7.12}$$

For $\alpha = \alpha_1$ choose $\sigma = \bar{\sigma} = 1$, $\mu = \bar{\mu} = \mu_1$ so that

$$\begin{aligned} |M^k_{l\alpha_1\mu_1,1}|^2 &= (|l|/g)D^{(\mathbf{k})(l)}(\{\epsilon|0\})^*_{11} \\ &= |l|/g \end{aligned} \tag{7.13}$$

and

$$M_{l\alpha_1\mu,\sigma} = \left(\frac{|l|}{g}\right)^{1/2}\sum_{\rho} D^{(\mathbf{k})(l)}(\{\phi_\rho|\boldsymbol{\tau}_\rho\})^*_{\mu\mu_1}\omega_{\sigma^{-1},\rho}, \quad \text{for } \phi_{\sigma^{-1}\rho} = \epsilon$$

$$= \left(\frac{|l|}{g}\right)^{1/2} D^{(\mathbf{k})(l)}(\{\phi_\sigma|\boldsymbol{\tau}_\sigma\})^*_{\mu\mu_1}. \tag{7.14}$$

For $\alpha = \alpha_2$ we choose $\sigma = \bar{\sigma} = 1$, $\mu = \bar{\mu} = \mu_2$ and repeat the above procedure to obtain $M^k_{l\alpha_2\mu,\sigma}$, etc.

Once the matrix Z is known we have the Clebsch–Gordan matrices $U^{l\otimes l'}_{\mu\mu',\, l''\gamma\mu''}$ for all (l,l',l'') associated with $(*\mathbf{k}, *\mathbf{k}', *\mathbf{k}'')$. In particular choosing $\alpha = \alpha_0$, $\alpha' = \alpha'_0$, and $\alpha'' = \alpha''_0$ so that Z is nonzero we have

$$\begin{aligned} &Z_{l\alpha_0\mu\, l'\alpha'_0\mu',\, l''\alpha''_0\mu''} \\ &\quad = \left(\frac{|l''|}{g}\right)^{1/2}\sum_{\sigma} M^k_{l\alpha_0\mu,\sigma}M^{k'}_{l'\alpha'_0\mu',\sigma}M^{k''*}_{l''\alpha''_0\mu'',\sigma}. \end{aligned} \tag{7.15}$$

But

$$\begin{aligned} M^k_{l\alpha_0\mu,\sigma} &= \left(\frac{|l|}{g}\right)^{1/2} D^{(\mathbf{k})(l)}(\phi_\sigma)_{\mu\mu_0}, \\ M^{k'}_{l'\alpha'_0\mu',\sigma} &= \left(\frac{|l'|}{g}\right)^{1/2} D^{(\mathbf{k}')(l')}(\phi_\sigma)_{\mu'\mu'_0}, \\ M^{k''}_{l''\alpha''_0\mu'',\sigma} &= \frac{|l''|}{g}^{1/2} D^{(\mathbf{k}'')(l'')}(\phi_\sigma)_{\mu''\mu''_0}. \end{aligned} \tag{7.16}$$

Therefore,

$$\begin{aligned} &Z_{l\alpha_0\mu\, l'\alpha'_0\mu',\, l''\alpha''_0\mu''} \\ &\quad = \left(\frac{|l|\,|l'|}{g^2}\right)^{1/2}\frac{|l''|}{g}\sum_{\sigma} D^{(\mathbf{k})(l)}(\phi_\sigma)_{\mu\mu_0}D^{(\mathbf{k}')(l')}(\phi_\sigma)_{\mu'\mu'_0} \\ &\qquad \times D^{(\mathbf{k}'')(l'')}(\phi_\sigma)^*_{\mu''\mu''_0} \\ &\quad = \left(\frac{|l|\,|l'|}{g^2}\right)U^{l\otimes l'}_{\mu\mu',\, l''\gamma_0\mu''}U^{l\otimes l'*}_{\mu_0\mu'_0,\, l''\gamma_0\mu''_0}. \end{aligned} \tag{7.17}$$

In this fashion all coefficients have been obtained for the space groups we studied. Some details of the calculation and results will be given in the following paper.

8. CONCLUSION

In the preceding sections we have demonstrated a method for obtaining Clebsch–Gordan coefficients for a group G by inducing from subgroup representations and we have shown how this method can be used to obtain space group Clebsch–Gordan coefficients. In particular, since the full space group G is induced from the ray representations for the point groups of k, the Clebsch–Gordan coefficients for G can be calculated from the representations of these smaller groups. Since the induced representations have a block form the Clebsch–Gordan matrix can also be calculated in blocks.

This block form has already been used by Litvin and Zak.[3] The procedure given by Litvin and Zak, however, requires a separate calculation for each of the $(1,1,1), \cdots (\sigma,\sigma',1) \cdots$ blocks followed by an appropriate transformation to obtain the remaining nonzero blocks. In the method presented here, only the $(1,1,1)$ block need be calculated explicitly and the remaining coefficients may all be obtained by a transformation.

In a recent paper Sakata[7] has also presented a method for calculating space group coefficients. His procedure, however, requires first the calculation of the $(1,1,1)$ block of coefficients and then the construction of the correct linear combination of product functions before

transforming to obtain the remaining blocks.

Neither Litvin and Zak[3] nor Sakata[7] make explicit use of ray representations.

In addition we have shown that the method of induced representations can be used to obtain the Clebsch–Gordan coefficients for the (1, 1, 1) block by inducing from the group of a general wave vector. This has the advantage of producing in one formulation the (1, 1, 1) block of coefficients for all representations associated with a fixed triplet of stars (*k, *k′, *k″).

APPENDIX: THE UNITARITY OF THE CLEBSCH-GORDAN MATRIX

In Sec. 2 we have indicated a systematic procedure for obtaining the Clebsch–Gordan matrix in the case of multiplicity, i. e., $(ll'|l'') > 1$. We will now demonstrate that this procedure produces a unitary matrix. Specifically we will show that

$$\sum_{\mu\mu'} U_{\mu\mu',\, l''\gamma_1\mu''} \bar{U}^*_{\mu\mu',\, \bar{l}''\gamma_2\nu''} = \delta_{l''\bar{l}''}\delta_{\mu''\nu''}\delta_{\gamma_1\gamma_2}, \tag{A1}$$

where $U_{\mu\mu',\, l''\gamma_1\mu''}$ and $\bar{U}_{\mu\mu',\, l''\gamma_2\mu''}$ are specified as follows:

From Eq. (2. 9) we have

$$\sum_{\gamma} U_{\mu\mu',\, l''\gamma\mu''} U^*_{\bar{\mu}\bar{\mu}',\, l''\gamma\bar{\mu}''} = \frac{|l''|}{g} \sum_{\phi} D^{(l)}(\phi)_{\mu\bar{\mu}} D^{(l')}(\phi)_{\mu'\bar{\mu}'} D^{(l'')}(\phi)^*_{\mu''\bar{\mu}''}. \tag{A2}$$

Using the procedure following Eq. (2. 9) we set $\gamma = \gamma_1$ and choose $\bar{\mu} = \mu_0$, $\bar{\mu}' = \mu_0'$, $\bar{\mu}'' = \mu_0''$.

Then

$$|U_0|^2 = U_{\mu_0\mu_0',\, l''\gamma_1\mu_0''} U^*_{\mu_0\mu_0',\, l''\gamma_1\mu_0''} = \frac{|l''|}{g} \sum_{\phi} D^{(l)}(\phi)_{\mu_0\mu_0} D^{(l')}(\phi)_{\mu_0'\mu_0'} D^{(l'')}(\phi)^*_{\mu_0''\mu_0''} \tag{A3}$$

and

$$U_{\mu\mu',\, l''\gamma_1\mu} U_0^* = \frac{|l''|}{g} \sum_{\phi} D^{(l)}(\phi)_{\mu\mu_0} D^{(l')}(\phi)_{\mu'\mu_0} D^{(l'')}(\phi)^*_{\mu''\mu_0''}. \tag{A4}$$

Similarly for $\gamma = \gamma_2$ we choose $\bar{\mu} = \nu_0$, $\bar{\mu}' = \nu_0'$, and $\bar{\mu}'' = \nu_0''$ so that

$$U_{\nu\nu',\, l''\gamma_2\nu''} \bar{U}_0^* = \frac{|l''|}{g} \sum_{\phi'} D^{(l)}(\phi')_{\nu\nu_0} D^{(l')}(\phi')_{\nu'\nu_0'} D^{(\bar{l}'')}(\phi')^*_{\nu''\nu_0''}. \tag{A5}$$

From (A4) and (A5) we have

$$\sum U_{\mu\mu',\, l''\gamma_1\mu''} \bar{U}^*_{\mu\mu',\, l''\gamma_2\nu''} = \frac{1}{U_0^* \bar{U}_0} \frac{|l''|\,|\bar{l}''|}{g^2} \sum_{\phi\phi'} \sum_{\mu\mu'} [D^{(l)}(\phi)_{\mu\mu_0} D^{(l)}(\phi')^*_{\mu\nu_0} \times D^{(l')}(\phi)_{\mu'\mu_0'} D^{(l')}(\phi')^*_{\mu'\nu_0'} D^{(l'')}(\phi)^*_{\mu''\mu_0''} D^{(\bar{l}'')}(\phi')_{\nu''\nu_0''}]. \tag{A6}$$

But

$$\sum_{\mu} D^{(l)}(\phi)_{\mu\mu_0} D^{(l)}(\phi')^*_{\mu\nu_0} = \sum_{\mu} D^{(l)}(\phi)_{\mu\mu_0} D^{(l)}(\phi'^{-1})_{\nu_0\mu} = D^{(l)}(\phi'^{-1}\phi)_{\nu_0\mu_0}, \tag{A7}$$

and similarly

$$\sum_{\mu'} D^{(l')}(\phi)_{\mu'\mu_0'} D^{(l')}(\phi')^*_{\mu'\nu_0'} = D^{(l')}(\phi'^{-1}\phi)_{\nu_0'\mu_0'}. \tag{A8}$$

Hence

$$\sum_{\mu\mu'} U_{\mu\mu',\, l''\gamma_1\mu''} \bar{U}^*_{\mu\mu',\, \bar{l}''\gamma_2\nu''} = \frac{1}{U_0^* \bar{U}_0} \frac{|l''|\,|\bar{l}''|}{g^2} \sum_{\phi\phi'} [D^{(l)}(\phi'^{-1}\phi)_{\nu_0\mu_0} D^{(l')}(\phi'^{-1}\phi)_{\nu_0'\mu_0'} \times D^{(l'')}(\phi)^*_{\mu''\mu_0''} D^{(\bar{l}'')}(\phi')_{\nu''\nu_0''}]. \tag{A9}$$

Let $\phi'' = \phi'^{-1}\phi$ so that

$$\sum_{\phi\phi'} D^{(l)}(\phi'^{-1}\phi)_{\nu_0\mu_0} D^{(l')}(\phi'^{-1}\phi)_{\nu_0'\mu_0'} D^{(l'')}(\phi)^*_{\mu''\mu_0''} D^{(\bar{l}'')}(\phi')_{\nu''\nu_0''} = \sum_{\phi'\phi''} D^{(l)}(\phi'')_{\nu_0\mu_0} D^{(l')}(\phi'')_{\nu_0'\mu_0'} D^{(l'')}(\phi')^*_{\mu''\lambda} D^{(l'')}(\phi'')^*_{\lambda\mu_0''} \times D^{(\bar{l}'')}(\phi')_{\nu''\nu_0''}. \tag{A10}$$

But

$$\sum_{\phi'} D^{(l'')}(\phi')^*_{\mu''\lambda} D^{(\bar{l}'')}(\phi')_{\nu''\nu_0''} = \frac{g}{|l''|} \delta_{\mu''\nu''}\delta_{\lambda\nu_0''}\delta_{l''\bar{l}''}. \tag{A11}$$

Therefore

$$\sum_{\mu\mu'} U_{\mu\mu',\, l''\gamma_1\mu''} U^*_{\mu\mu',\, \bar{l}''\gamma_2\nu''} = \frac{1}{U_0^* \bar{U}_0} \frac{|l''|}{g} \sum_{\phi''} D^{(l)}(\phi'')_{\nu_0\mu_0} D^{(l')}(\phi'')_{\nu_0'\mu_0'} D^{(l'')}(\phi'')^*_{\nu_0''\mu_0''} \times \delta_{\mu''\nu''}\delta_{l''\bar{l}''}. \tag{A12}$$

Hence for Eq. (A1) to hold, for given μ_0, μ_0', μ_0'', we must choose ν_0, ν_0', and ν_0'' so that

$$\sum_{\phi} D^{(l)}(\phi)_{\nu_0\mu_0} D^{(l')}(\phi)_{\nu_0'\mu_0'} D^{(l'')}(\phi)^*_{\nu_0''\mu_0''} = 0 \quad \text{for } \gamma_2 \neq \gamma_1. \tag{A13}$$

This procedure will then produce a unitary Clebsch–Gordan matrix.

Note added in proof: It has recently come to our attention that a different method for calculating Clebsch–Gordan coefficients, also utilizing ray representations, has recently been given by Patricia Gard, in J. Phys. A 6, 1837 (1973).

*Supported in Part by NSF-GH-31742 grant and AROD-DAHCO4-74-G-0109 grant.

†This paper is based in part upon a Ph.D. thesis submitted by Rhoda Berenson in partial fulfillment of the requirements for the Ph.D. degree in the Physics Department, New York University, February 1974.

‡Address after 1 September 1974: Department of Physics, City College of the CUNY, Convent Ave. and 138 St., New York, N.Y. 10031.

[1]M. Hammermesh, *Group Theory and its Application to Physical Problems* (Addison-Wesley, Reading, Mass., 1962), pp. 148–151.
[2]G.F. Koster, J.O. Dimmock, R.G. Wheeler, and H. Statz, *Properties of the Thirty-two Point Groups* (Technology Press, Cambridge, Mass., 1963).
[3]D.B. Litvin and J. Zak, J. Math. Phys. 9, 212 (1968).
[4]I. Itzkan, "Clebsch-Gordan Coefficients for Space Groups", unpublished Ph.D. thesis, New York University, 1969.
[5]G.F. Koster, Phys. Rev. 109, 227 (1958).
[6]O.V. Kovalev, *Irreducible Representations of the Space Groups* (Gordon and Breach, New York, 1965).
[7]I. Sakata, J. Math. Phys. 15, 1702, 1710 (1974).

Clebsch-Gordan coefficients for $^*X \otimes {}^*X$ in diamond O_h^7-Fd3m and rocksalt O_h^5-Fm3m*

Rhoda Berenson[a†], Irving Itzkan[b‡], and Joseph L. Birman[a§]

[a] *Physics Department, New York University, New York, New York 10003*

[b] *Avco Everett Research Laboratory, Inc., Everett, Massachusetts 02149*

(Received 5 August 1974)

By using the method described in the previous paper, based on properties of space group irreducible representations as induced from ray representations of subgroups, Clebsch–Gordan coefficients are calculated for $^*X \otimes {}^*X$ in diamond O_h^7-Fd3m and rocksalt O_h^5-Fm3m structures. Tables of coefficients for these stars are presented. An example of explicit calculation of the coefficients is given for these symmorphic and nonsymmorphic groups with multiplicity included in the former.

1. INTRODUCTION

In a previous paper[1] we have given a method for calculating crystal space group Clebsch–Gordan coefficients based on the theory of induced ray representations. In particular we have shown how the coefficients for the full space group G can be induced from representations of $G(\mathbf{k})$ which in turn can be induced from a group of lower symmetry such as the translation group T.

In this paper we will give the results of calculations of Clebsch–Gordan coefficients for products $^*X \otimes {}^*X$ in rocksalt O_h^5 and diamond O_h^7. The cases are intended to illustrate the practicality and ease of calculation of the method given. These cases typify the complexities encountered in calculating these coefficients for symmorphic and nonsymmorphic groups, respectively.

2. STRUCTURE OF INDUCED REPRESENTATIONS AND INDUCED CLEBSCH-GORDAN MATRICES

We will now briefly recapitulate those results of the previous paper[1] which will be used in the calculations that follow.

Consider a group G with subgroup H where we can write

$$G = H + \phi_2 H + \cdots \phi_\sigma H + \cdots \phi_s H. \tag{2.1}$$

If $\mathcal{H}^m$ is an irreducible representation of H, then the induced ray representation $\mathcal{H}^m \uparrow G$ has elements given by

$$[\mathcal{H}^m \uparrow G](\phi_g)_{\sigma\mu,\tau\nu} = (\omega_{\sigma,\sigma^{-1}})^{-1}\, \mathring{\mathcal{H}}^m(\phi_\sigma^{-1}\phi_g\phi_\tau)_{\mu\nu}, \tag{2.2}$$

where

$$\mathring{\mathcal{H}}^m(\phi)_{\mu\nu} = \mathcal{H}^m(\phi)_{\mu\nu}, \text{ for } \phi \in H$$
$$= 0 \text{ otherwise.} \tag{2.3}$$

$(\mathcal{H}^m \uparrow G)$ is in general a reducible ray representation of G and ω is the ray factor associated with this representation. In Eq. (2.2) for a given ϕ_g, for each τ there is only one value of σ such that $\phi_\sigma^{-1}\phi_g\phi_\tau \in H$. Hence, the induced representations are subdivided into s^2 blocks and in each row and column there is only one nonvanishing block.

A space group G has a subgroup $G(\mathbf{k})$ consisting of all those elements that leave $\mathbf{k}$ invariant modulo a reciprocal lattice vector so that we can write

$$G = G(\mathbf{k}) + \cdots + \{\phi_\sigma|\boldsymbol{\tau}_\sigma\}G(\mathbf{k}) + \cdots + \{\phi_s|\boldsymbol{\tau}_s\}G(\mathbf{k}). \tag{2.4}$$

When inducing from $G(\mathbf{k})$ to G, the induced representation $D^{(*\mathbf{k})(l)} \equiv D^{(\mathbf{k})(l)} \uparrow G$ is irreducible and

$$D^{(*\mathbf{k})(l)}(\{\phi_g|\mathbf{t}_g\})_{\sigma\mu\tau\nu} = \mathring{D}^{(\mathbf{k})(l)}(\{\phi_\sigma|\boldsymbol{\tau}_\sigma\}^{-1}\{\phi_g|\mathbf{t}_g\}\{\phi_\tau|\boldsymbol{\tau}_\tau\})_{\mu\nu}. \tag{2.5}$$

$D^{(*\mathbf{k})(l)}$ therefore has a block structure and the Clebsch–Gordan matrix U which reduces the products $D^{(*\mathbf{k})(l)} \otimes D^{(*\mathbf{k}')(l')} \to D^{(*\mathbf{k}'')(l'')}$ also has a block structure and is found from

$$\sum_\gamma U^{l\otimes l'}_{\sigma a\bar{\sigma}\bar{a}',l''\gamma\sigma''a''} U^{l\otimes l'*}_{\bar{\sigma}\bar{a}\bar{\sigma}'\bar{a}',l''\gamma\bar{\sigma}''\bar{a}''} = \frac{|l''|}{g}\sum_{\{\phi|\mathbf{t}\}} D^{(*\mathbf{k})(l)}(\{\phi|\mathbf{t}\})_{\sigma a\bar{\sigma}\bar{a}} D^{(*\mathbf{k}')(l')}(\{\phi|\mathbf{t}\})_{\sigma' a'\bar{\sigma}'\bar{a}'} \times D^{(*\mathbf{k}'')(l'')}(\{\phi|\mathbf{t}\})^*_{\sigma''a''\bar{\sigma}''\bar{a}''}, \tag{2.6}$$

where $|l''|$ is the dimension of $D^{(*\mathbf{k}'')(l'')}$, g is the order of the space group G, and γ indicates the multiplicity of $D^{(*\mathbf{k}'')(l'')}$ in this product matrix.

The calculation of this Clebsch–Gordan matrix can be simplified by first finding the (1, 1, 1) block of coefficients defined by

$$\sum_\gamma U^{l\otimes l'}_{1a1a',l''\gamma 1a''} U^{l\otimes l'*}_{1\bar{a}1\bar{a}',l''\gamma 1\bar{a}''} = \frac{|l''|}{g}\sum_x D^{(\mathbf{k})(l)}(\{\phi_x|\boldsymbol{\tau}_x\}_{a\bar{a}} D^{(\mathbf{k}')(l')}(\{\phi_x|\boldsymbol{\tau}_x\})_{a'\bar{a}'} \times D^{(\mathbf{k}'')(l'')}(\{\phi_x|\boldsymbol{\tau}_x\})^*_{a''\bar{a}''}, \tag{2.7}$$

for $\mathbf{k}+\mathbf{k}'-\mathbf{k}'' = 2\pi\mathbf{B}_H$. x is summed over all members of $\mathcal{N}$, the intersection group of $G(\mathbf{k})$, $G(\mathbf{k}')$, and $G(\mathbf{k}'')$. Then the $(\sigma, \sigma', \sigma'')$ block, $U^{l\otimes l'}_{ab\sigma''b',l''\gamma\sigma''b''}$ is obtained from

$$U_{\sigma b\sigma''b',l''\gamma\sigma''b''} = \sum_{aa'a''}[D^{(\mathbf{k})(l)}(\{\phi_k|\mathbf{t}_k\}) \otimes D^{(\mathbf{k}')(l')}(\{\phi_{k'}|\mathbf{t}_{k'}\})]_{bb'aa'} \times U^{l\otimes l'}_{1a1a',l''\gamma 1a''} D^{(\mathbf{k}'')(l'')}(\{\phi_{k''}|\mathbf{t}_{k''}\})^{-1}_{a''b''}, \tag{2.8}$$

where

$$\{\phi_k|\mathbf{t}_k\} = \{\phi_\sigma|\boldsymbol{\tau}_\sigma\}^{-1}\{\phi_\Sigma|\boldsymbol{\tau}_\Sigma\},$$
$$\{\phi_{k'}|\mathbf{t}_{k'}\} = \{\phi_{\sigma'}|\boldsymbol{\tau}_{\sigma'}\}\{\phi_\Sigma|\boldsymbol{\tau}_\Sigma\},$$
$$\{\phi_{k''}|\mathbf{t}_{k''}\} = \{\phi_{\sigma''}|\boldsymbol{\tau}_{\sigma''}\}\{\phi_\Sigma|\boldsymbol{\tau}_\Sigma\}, \tag{2.9}$$

and

$$\phi_\Sigma\mathbf{k} = \mathbf{k}_\sigma;\ \phi_\Sigma\mathbf{k}' = \mathbf{k}'_{\sigma'};\ \phi_\Sigma\mathbf{k}'' = \mathbf{k}''_{\sigma''}.$$

The (1, 1, 1) block of coefficients is readily calculated by inducing from the trivial point group $P(\epsilon) = \epsilon$ having irreducible representation $D^{(\epsilon)}$. This method has the

TABLE I. Ray representations for $D^{(\mathbf{k}_z)(X(l))}$ in O_h^7.

	$X(1)$	$X(2)$	$X(3)$	$X(4)$
Type $\{\varphi \mid 0\}$				
$h_1=\epsilon$	1 0 0 1	1 0 0 1	1 0 0 1	1 0 0 1
$h_2=\delta_{2x}$	0 1 1 0	0 1 1 0	−1 0 0 1	1 0 0 −1
$h_3=\delta_{2y}$	0 1 1 0	0 1 1 0	1 0 0 −1	−1 0 0 1
$h_4=\delta_{2z}$	1 0 0 1	1 0 0 1	−1 0 0 −1	−1 0 0 −1
$h_{38}=\sigma_{4z}$	0 1 1 0	0 −1 −1 0	0 −i −i 0	0 −i −i 0
$h_{39}=\sigma_{4z}^{-1}$	0 1 1 0	0 −1 −1 0	0 i i 0	0 i i 0
$h_{40}=\rho_{xy}$	1 0 0 1	−1 0 0 −1	0 −i i 0	0 i −i 0
$h_{37}=\rho_{x\bar{y}}$	1 0 0 1	−1 0 0 −1	0 i −i 0	0 −i i 0
Type $\{\varphi \mid \mathcal{L}\}$				
$\mathcal{L}=(\frac{1}{4},\frac{1}{4},\frac{1}{4},)a$				
$h_{14}=\delta_{4z}$	−i 0 0 i	−i 0 0 i	−1 0 0 1	−1 0 0 1
$h_{15}=\delta_{4z}^{-1}$	−i 0 0 i	−i 0 0 i	1 0 0 −1	1 0 0 −1
$h_{16}=\delta_{2xy}$	0 −i i 0	0 −i i 0	1 0 0 1	−1 0 0 −1
$h_{13}=\delta_{2x\bar{y}}$	0 −i i 0	0 −i i 0	−1 0 0 −1	1 0 0 1
$h_{25}=i$	0 −i i 0	0 i −i 0	0 −i i 0	0 −i i 0
$h_{26}=\rho_x$	−i 0 0 i	i 0 0 −i	0 −i −i 0	0 i i 0
$h_{27}=\rho_y$	−i 0 0 i	i 0 0 −i	0 i i 0	0 −i −i 0
$h_{28}=\rho_z$	0 −i i 0	0 i −i 0	0 i −i 0	0 i −i 0

advantage of obtaining in one calculational step, an entire set of coefficients for *all* representations associated with a given triplet of stars $(*k, *k', *k'')$. The Clebsch–Gordan matrix $U^{l\otimes l'}_{l\mu l'\mu', l''\gamma l''\mu''}$ is then proportional to a "super Clebsch–Gordan matrix"

$$Z_{l\alpha\mu l'\alpha'\mu', l''\alpha''\mu''} = \left(\frac{|l''|}{g}\right)^{1/2} \sum_\sigma M^k_{l\alpha\mu,\sigma} M^{k'}_{l'\alpha'\mu',\sigma} M^{k''*}_{l''\alpha''\mu'',\sigma}, \tag{2.10}$$

where the unitary matrices M^k, $M^{k'}$, and $M^{k''}$ reduce $D^{(\epsilon)}\uparrow G(\mathbf{k})$, $D^{(\epsilon)}\uparrow G(\mathbf{k}')$, and $D^{(\epsilon)}\uparrow G(\mathbf{k}'')$ into direct sums of irreducible representations $D^{(\mathbf{k})(l)}$. α indicates the multiplicity of $D^{(\mathbf{k})(l)}$ in $D^{(\epsilon)}\uparrow G^{(\mathbf{k})}$ and similarly for α' and α''. M^k is then most easily calculated from

$$\sum_\alpha M^k_{l\alpha\mu,\sigma} M^k_{l\alpha\mu_0,1} = \frac{|l|}{g} D^{(\mathbf{k})(l)}(\{\phi_\sigma \mid \boldsymbol{\tau}_\sigma\})^*_{\mu\mu_0}. \tag{2.11}$$

As demonstrated in the previous paper, the Clebsch–Gordan $U^{l\otimes l'}_{\mu\mu', l''\gamma\mu''}$ is then related to the matrix Z by

$$\left(\frac{|l||l'|}{g^2}\right)^{1/2} U^{l\otimes l'}_{\mu\mu', l''\gamma_0\mu''} U^{l\otimes l'*}_{\mu_0\mu_0', l''\gamma_0\mu_0''} = Z_{l\alpha_0\mu l'\alpha_0'\mu', l''\alpha_0''\mu''}. \tag{2.12}$$

In the case of multiplicity when $D^{(*\mathbf{k}'')(l'')}$ appears say twice in $D^{(*\mathbf{k})(l)}\otimes D^{(*\mathbf{k}')(l')}$ there is another choice of $\alpha_0, \alpha_0', \alpha_0''$ corresponding to a different choice of $\mu_0\mu_0'$ and μ_0'' producing a second Clebsch–Gordan matrix which is orthogonal to the first. This will be demonstrated explicitly in Sec. 4 in a calculation for rocksalt, O_h^5.

3. $*X \otimes *X$ IN DIAMOND, O_h^7

The diamond structure is a face-centered cubic lattice[2] having a nonsymmorphic space group O_h^7. The point X is on the zone boundary in the (100) direction. The three arms of $*\mathbf{X}$ are $\mathbf{k}_x = (2\pi/a)(100)$, $\mathbf{k}_y = (2\pi/a) \times (010)$, $\mathbf{k}_z = (2\pi/a)(001)$. The point Γ is at the zone center and $\mathbf{k}_\Gamma = (000)$.

Table I lists the elements of $G(\mathbf{k}_z)/T$ along with their representations. These representations are of the form $D^{(\mathbf{k}_z)(l)}(\{\phi \mid \boldsymbol{\tau}\}) = \exp(-i\mathbf{k}_z\cdot\boldsymbol{\tau})\hat{\tau}^{(l)}(\phi)$ where $\hat{\tau}^{(l)}$ is a representation tabulated by Kovalev.[3] Representations $X(l)$ in Herring's notation[2] correspond to those of Kovalev if $X(1)=\hat{\tau}(3)$, $X(2)=\hat{\tau}(4)$, $X(3)=\hat{\tau}(2)$, and $X(4)=\hat{\tau}(1)$.

Since X is at the zone boundary of a nonsymmorphic group the $D^{(\mathbf{k})(l)}$ are ray representations with a nontrivial factor system.

The diamond space group G can be decomposed into cosets with respect to $G(\mathbf{X})$ as

$$G = \{\phi_1 \mid \boldsymbol{\tau}_1\}G(\mathbf{X}) + \{\phi_2 \mid \boldsymbol{\tau}_2\}G(\mathbf{X}) + \{\phi_3 \mid \boldsymbol{\tau}_3\}G(\mathbf{X}), \tag{3.1}$$

where

$$\{\phi_1 \mid \boldsymbol{\tau}_1\} = \{\epsilon \mid 0\},$$
$$\{\phi_2 \mid \boldsymbol{\tau}_2\} = \{\delta_{3xyz} \mid 0\},$$
$$\{\phi_3 \mid \boldsymbol{\tau}_3\} = \{\delta^{-1}_{3xyz} \mid 0\}. \tag{3.2}$$

Equation (3.1) holds for $G(\mathbf{X})$ with $\mathbf{X}=\mathbf{k}_x$, or $\mathbf{k}_y$, or $\mathbf{k}_z$.

The selection rules for products at X have been given by Birman.[4] In particular, the wave vector selection rule at X is

$$*\mathbf{X}\otimes *\mathbf{X} = 2*\mathbf{X} + 3\Gamma. \tag{3.3}$$

When calculating the Clebsch–Gordan matrix for $*\mathbf{X}\otimes *\mathbf{X} = *\mathbf{X}$ we see that $\mathbf{k}_\sigma$, $\mathbf{k}'_{\sigma'}$, and $\mathbf{k}''_{\sigma''}$ belong to the same star but the only nonzero coefficients are those for which $\mathbf{k}_\sigma$, $\mathbf{k}'_{\sigma'}$, and $\mathbf{k}''_{\sigma''}$ are different arms of the star. On the other hand, for $*\mathbf{X}\otimes *\mathbf{X} = \Gamma$ the only zero contributions will be for $\mathbf{k}_\sigma = \mathbf{k}'_{\sigma'}$. The general procedure for calculating the Clebsch–Gordan matrix is as follows: First calculate those coefficients for the principle (1,1,1) block. These can be obtained by inducing from a general $\mathbf{k}$ vector to $G(\mathbf{X})$. Once the coefficients for this block are known the remaining nonzero blocks can be found by inducing to the full space group.

A. The Clebsch-Gordan matrix for $*X \otimes *X \to *X$

For $D^{(\mathbf{k}_x)(l)}\otimes D^{(\mathbf{k}_y)(l')} \to D^{(\mathbf{k}_z)(l'')}$ the Clebsch–Gordan coefficients can be obtained by inducing from $D^{(\epsilon)}$, the representation for a general $\mathbf{k}$ vector. The induced coefficients are then given by

$$Z_{l\alpha\mu l'\alpha'\mu', l''\alpha''\mu''} = \left(\frac{|l''|}{g}\right)^{1/2} \sum_\sigma M^{k_x}_{l\alpha\mu,\sigma} M^{k_y}_{l'\alpha'\mu',\sigma} M^{k_z*}_{l''\alpha''\mu'',\sigma}, \tag{3.4}$$

where M^{k_x} is the matrix that reduces $D^{(\epsilon)}\uparrow G(\mathbf{k}_x)$ so that

$$M^{k_x}[D^{(\epsilon)} \uparrow G(\mathbf{k}_x)]M^{k_x-1}$$

$$= \begin{bmatrix} D^{(\mathbf{k}_x)(x(1))} & & & & & & & \\ & D^{(\mathbf{k}_x)(x(2))} & & & & & & \\ & & D^{(\mathbf{k}_x)(x(3))} & & & & & \\ & & & D^{(\mathbf{k}_x)(x(4)} & & & & \\ & & & & D^{(\mathbf{k}_x)(x(1))} & & & \\ & & & & & D^{(\mathbf{k}_x)(x(2))} & & \\ & & & & & & D^{(\mathbf{k}_x)(x(3))} & \\ & & & & & & & D^{(\mathbf{k}_x)(x(4))} \end{bmatrix}.$$

Similarly M^{k_y} and M^{k_z} reduce $D^{(\epsilon)} \uparrow G(\mathbf{k}_y)$ and $D^{(\epsilon)} \uparrow G(\mathbf{k}_z)$. Note that in Eq. (3.4) the sum on σ is over all elements belonging to N the intersection group of $G(\mathbf{k}_x)/T$, $G(\mathbf{k}_y)/T$, and $G(\mathbf{k}_z)/T$. Since all $D^{(\mathbf{k})(l)}$ are two dimensional, each appears two times in the induced regular representation. In what follows we take $\mathbf{k}=\mathbf{k}_x$, $\mathbf{k}_y$, or $\mathbf{k}_z$. Then

$$M^{k}_{l\alpha_1\mu,\sigma} = \left(\frac{|l|}{g}\right)^{1/2} D^{(\mathbf{k})(l)}(\{\phi_\sigma|\tau_\sigma\})^*_{\mu 1}$$

and (3.5)

$$M^{k}_{l\alpha_2\mu,\sigma} = \left(\frac{|l|}{g}\right)^{1/2} D^{(\mathbf{k})(l)}(\{\phi_\sigma|\tau_\sigma\})^*_{\mu 2}.$$

The intersection group, N, for $G(\mathbf{k}_x)/T$, $G(\mathbf{k}_y)/T$, and $G(\mathbf{k}_z)/T$ can be written as

$$N = S(\mathbf{k}) + \{i|\tau\}S(\mathbf{k}), \tag{3.6}$$

where

$$S(\mathbf{k}) = \{\epsilon|0\} + \{\delta_{2x}|0\} + \{\delta_{2y}|0\} + \{\delta_{2z}|0\}. \tag{3.7}$$

Then we can write

$$M^{k}_{l\alpha_1\mu,\sigma} = \left(\frac{|l|}{g}\right)^{1/2} D^{(\mathbf{k})(l)}(\{\phi_\sigma|\tau_\sigma\})^*_{\mu 1}, \text{ for } \sigma=1,2,3,4 \tag{3.8}$$

and

$$\begin{aligned} M^{k}_{l\alpha_1\mu,\bar{\sigma}} &= \left(\frac{|l|}{g}\right)^{1/2} D^{(\mathbf{k})(l)}(\{\phi_{\bar{\sigma}}|\tau_{\bar{\sigma}}\})^*_{\mu 1}, \quad \bar{\sigma}=5,6,7,8 \\ &= \left(\frac{|l|}{g}\right)^{1/2} D^{(\mathbf{k})(l)}(\{i|\tau\})^*_{\mu\lambda} D^{(\mathbf{k})(l)}(\{\phi_\sigma|\tau_\sigma\})^*_{\lambda 1}, \\ &= \sum_\lambda D^{(\mathbf{k})(l)}(\{i|\tau\})^*_{\mu\lambda} M^{k}_{l\alpha_1\lambda,\sigma}. \end{aligned} \tag{3.9}$$

$M_{l\alpha_2\mu,\sigma}$ can be written similarly. Hence we can write

$$\begin{aligned} Z_{l\alpha_\mu l''\alpha''\mu'',l'''\alpha'''\mu'''} &= Z'_{l\alpha_\mu l''\alpha''\mu'',l'''\alpha'''\mu'''} + \sum_{\mu\bar{\mu}'\bar{\mu}} [D^{(\mathbf{k}_x)(l)}(\{i|\tau\})^* \\ &\otimes D^{(\mathbf{k}_y)(l'')}(\{i|\tau\})^*]_{\mu\mu'',\bar{\mu}\bar{\mu}'} Z_{l\alpha\bar{\mu}l''\alpha''\bar{\mu}',l'''\alpha'''\bar{\mu}'''} \\ &\times D^{(\mathbf{k}_z)(l)}(\{i|\tau\})^{-1}_{\bar{\mu}'''\mu'''}, \end{aligned} \tag{3.10}$$

where

$$\begin{aligned} Z'_{l\alpha_\mu l''\alpha''\mu'',l'''\alpha'''\mu'''} &= \left(\frac{|l''|}{g}\right)^{1/2} \sum_{\sigma=1}^{4} M^{k_x}_{l\alpha_\mu,\sigma} M^{k_y}_{l''\alpha''\mu'',\sigma} M^{k_z}_{l'''\alpha'''\mu''',\sigma}. \end{aligned} \tag{3.11}$$

The M matrices for $\sigma=1,2,3,4$ are given below.

$$\sqrt{8}M^{k_x}_{l\alpha_1\mu,\sigma} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sqrt{8}\, M^{k_y}_{l\alpha_1\mu,\sigma} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\sqrt{8}\, M^{k_z}_{l\alpha_1\mu,\sigma} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sqrt{8}\, M^{k_x}_{l\alpha_2\mu,\sigma} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$$\sqrt{8}\, M^{k_y}_{l\alpha_2\mu,\sigma} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & -1 \end{bmatrix} \sqrt{8}\, M^{k_z}_{l\alpha_2\mu,\sigma} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 \end{bmatrix} \tag{3.12}$$

For example consider $D^{(\mathbf{k}_x)(X_1)} \otimes D^{(\mathbf{k}_y)(X_2)} \to D^{(\mathbf{k}_z)(X_2)}$. From Eqs. (3.10), (3.11), and (3.12) we fix $l=1$, $l'=2$, $l''=2$ so that

$$Z'_{X_1\alpha_1\mu X_2\alpha_1\mu',X_2\alpha_1\mu''} = \frac{1}{64} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{3.13}$$

where we have written Z' as a 4×2 matrix having elements

$$[\mu\mu',\mu''] \begin{bmatrix} 111 & 112 \\ 121 & 122 \\ 211 & 212 \\ 221 & 222 \end{bmatrix}. \tag{3.14}$$

Then

$$\begin{aligned} &64 Z_{X_1\alpha_1\mu X_2\alpha_1\mu',X_2\alpha_1\mu''} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} = \begin{bmatrix} 1 & i \\ i & 1 \\ i & 1 \\ 1 & i \end{bmatrix}. \end{aligned} \tag{3.15}$$

Normalizing we find

$$U^{(\mathbf{k}_x)(X_1)\otimes(\mathbf{k}_y)(X_2)}_{1\mu 1\mu',(\mathbf{k}_z)(X_2)1\mu''} = \frac{1}{\sqrt{8}} \begin{pmatrix} 1 & i \\ i & 1 \\ i & 1 \\ 1 & i \end{pmatrix}. \tag{3.16}$$

The Clebsch–Gordan coefficients for the principle (1,1,1) block for all $X(l)\otimes X(l')=X(l'')$ for canonical wave vectors $\mathbf{k}=\mathbf{k}_x$, $\mathbf{k}'=\mathbf{k}_y$, $\mathbf{k}''=\mathbf{k}_z$ are given in Table II. All entries are written as a 4×2 matrix with $(\mu\mu',\mu'')$ give as in Eq. (3.14).

The other $(\sigma,\sigma',\sigma'')$ blocks are found from $U(\sigma,\sigma',\sigma'')$

TABLE II. Principle block of coefficients for $*X\otimes *X\to *X$ in O_h^7.

	$X(1)$	$X(2)$	$X(3)$	$X(4)$
$X(1)\otimes X(1)$	$a\begin{pmatrix}1 & i\\ i & 1\\ i & 1\\ 1 & i\end{pmatrix}$	$a\begin{pmatrix}1-i\\ -i\ \ 1\\ -i\ \ 1\\ 1-i\end{pmatrix}$	$a\begin{pmatrix}1-i\\ -1\ \ -i\\ 1\ \ i\\ -1\ \ i\end{pmatrix}$	$a\begin{pmatrix}1-i\\ 1\ \ i\\ -1-i\\ -1\ \ i\end{pmatrix}$
$X(1)\otimes X(2)$	$a\begin{pmatrix}1-i\\ -i\ \ 1\\ -i\ \ 1\\ 1-i\end{pmatrix}$	$a\begin{pmatrix}1 & i\\ i & 1\\ i & 1\\ 1 & i\end{pmatrix}$	$a\begin{pmatrix}1 & i\\ -1 & i\\ 1-i\\ -1-i\end{pmatrix}$	$a\begin{pmatrix}1 & i\\ 1-i\\ -1 & i\\ -1-i\end{pmatrix}$
$X(1)\otimes X(3)$	$a\begin{pmatrix}1 & 1\\ -i & i\\ -1-1\\ -i & i\end{pmatrix}$	$a\begin{pmatrix}1 & 1\\ i-i\\ -1-1\\ i-i\end{pmatrix}$	$b\begin{pmatrix}0 & 1\\ i & 0\\ 0 & 1\\ -i & 0\end{pmatrix}$	$b\begin{pmatrix}1 & 0\\ 0-i\\ 1 & 0\\ 0 & i\end{pmatrix}$
$X(1)\otimes X(4)$	$a\begin{pmatrix}1-1\\ -i-i\\ 1-1\\ i & i\end{pmatrix}$	$a\begin{pmatrix}1-1\\ i & i\\ 1-1\\ -i-i\end{pmatrix}$	$b\begin{pmatrix}1 & 0\\ 0 & i\\ -1 & 0\\ 0 & i\end{pmatrix}$	$b\begin{pmatrix}0 & 1\\ -i & 0\\ 0-1\\ -i & 0\end{pmatrix}$
$X(2)\otimes X(3)$	$a\begin{pmatrix}1 & 1\\ i-i\\ -1-1\\ i-i\end{pmatrix}$	$a\begin{pmatrix}1 & 1\\ -i & i\\ -1-1\\ -i & i\end{pmatrix}$	$b\begin{pmatrix}0 & 1\\ -i & 0\\ 0 & 1\\ i & 0\end{pmatrix}$	$b\begin{pmatrix}1 & 0\\ 0 & i\\ 1 & 0\\ 0-i\end{pmatrix}$
$X(2)\otimes X(4)$	$a\begin{pmatrix}1-1\\ i & i\\ 1-1\\ -i-i\end{pmatrix}$	$a\begin{pmatrix}1-1\\ -i-i\\ 1-1\\ i & i\end{pmatrix}$	$b\begin{pmatrix}1 & 0\\ 0-i\\ -1 & 0\\ 0-i\end{pmatrix}$	$b\begin{pmatrix}0 & 1\\ i & 0\\ 0-1\\ i & 0\end{pmatrix}$
$X(3)\otimes X(3)$	$b\begin{pmatrix}0 & 0\\ 1 & 1\\ i-i\\ 0 & 0\end{pmatrix}$	$b\begin{pmatrix}0 & 0\\ 1 & 1\\ -i & i\\ 0 & 0\end{pmatrix}$	$c\begin{pmatrix}1 & 0\\ 0 & 0\\ 0 & 0\\ 0 & i\end{pmatrix}$	$c\begin{pmatrix}0 & 1\\ 0 & 0\\ 0 & 0\\ -i & 0\end{pmatrix}$
$X(3)\otimes X(4)$	$b\begin{pmatrix}1 & 1\\ 0 & 0\\ 0 & 0\\ -i & i\end{pmatrix}$	$b\begin{pmatrix}1 & 1\\ 0 & 0\\ 0 & 0\\ i-i\end{pmatrix}$	$c\begin{pmatrix}0 & 0\\ 1 & 0\\ 0-i\\ 0 & 0\end{pmatrix}$	$c\begin{pmatrix}0 & 0\\ 0-i\\ 1 & 0\\ 0 & 0\end{pmatrix}$
$X(4)\otimes X(4)$	$b\begin{pmatrix}0 & 0\\ i-i\\ 1 & 1\\ 0 & 0\end{pmatrix}$	$b\begin{pmatrix}0 & 0\\ -i & i\\ 1 & 1\\ 0 & 0\end{pmatrix}$	$c\begin{pmatrix}0 & 1\\ 0 & 0\\ 0 & 0\\ -i & 0\end{pmatrix}$	$c\begin{pmatrix}1 & 0\\ 0 & 0\\ 0 & 0\\ 0 & i\end{pmatrix}$

$a=1/\sqrt{8}$ $\quad b=1/2$ $\quad c=1/\sqrt{2}$

$X(2)\otimes X(2)=X(1)\otimes X(1)$

TABLE III. Calculating the $(\sigma,\sigma',\sigma'')$ block for $*X\otimes *X=*X$ in O_h^7.

$\sigma\ \sigma'\ \sigma''$	$\{\phi_\Sigma \mid \tau_\Sigma\}$	$\{\phi_k \mid t_k\}$	$\{\phi_{-k'} \mid t_{k'}\}$	$\{\phi_{k''} \mid t_{k''}\}$
1 1 1	$\{\epsilon \mid 0\}$	$\{\epsilon \mid 0\}$	$\{\epsilon \mid 0\}$	$\{\epsilon \mid 0\}$
2 2 2	$\{\delta_{3\bar{x}y\bar{z}} \mid 0\}$	$\{\epsilon \mid 0\}$	$\{\epsilon \mid 0\}$	$\{\epsilon \mid 0\}$
3 3 3	$\{\delta_{3\bar{x}\bar{y}z} \mid 0\}$	$\{\epsilon \mid 0\}$	$\{\epsilon \mid 0\}$	$\{\epsilon \mid 0\}$
2 3 1	$\{\rho_{x\bar{y}} \mid 0\}$	$\{\rho_{y\bar{z}} \mid 0\}$	$\{\rho_{x\bar{z}} \mid 0\}$	$\{\rho_{x\bar{y}} \mid 0\}$
3 1 2	$\{\rho_{x\bar{z}} \mid 0\}$	$\{\rho_{y\bar{z}} \mid 0\}$	$\{\rho_{x\bar{z}} \mid 0\}$	$\{\rho_{x\bar{y}} \mid 0\}$
1 2 3	$\{\rho_{y\bar{z}} \mid 0\}$	$\{\rho_{y\bar{z}} \mid 0\}$	$\{\rho_{x\bar{z}} \mid 0\}$	$\{\rho_{x\bar{y}} \mid 0\}$

	$D^{(\mathbf{k}_x)(l)}(\{\rho_{y\bar{z}} \mid 0\})$	$D^{(\mathbf{k}_y)(l')}(\{\rho_{x\bar{z}} \mid 0\})$	$D^{(\mathbf{k}_z)(l)}(\{\rho_{x\bar{y}} \mid 0\})$
$l=X(1)$	$\begin{matrix}1 & 0\\ 0 & 1\end{matrix}$	$\begin{matrix}1 & 0\\ 0 & 1\end{matrix}$	$\begin{matrix}1 & 0\\ 0 & 1\end{matrix}$
$l=X(2)$	$\begin{matrix}-1 & 0\\ 0-1\end{matrix}$	$\begin{matrix}-1 & 0\\ 0-1\end{matrix}$	$\begin{matrix}-1 & 0\\ 0-1\end{matrix}$
$l=X(3)$	$\begin{matrix}0 & i\\ -i & 0\end{matrix}$	$\begin{matrix}0 & i\\ -i & 0\end{matrix}$	$\begin{matrix}0 & i\\ -i & 0\end{matrix}$
$l=X(4)$	$\begin{matrix}0-i\\ i & 0\end{matrix}$	$\begin{matrix}0-i\\ i & 0\end{matrix}$	$\begin{matrix}0-i\\ i & 0\end{matrix}$

$$=D^{(\mathbf{k}_x)(l)}(\{\phi_k \mid t_k\})\otimes D^{(\mathbf{k}_y)(l')}(\{\phi_{k'} \mid t_{k'}\})U(1,1,1)\times D^{(\mathbf{k}_z)(l'')}(\{\phi_{k''} \mid t_{k''}\})^{-1}, \tag{3.17}$$

where from Eq. (2.9) we obtain the elements $\{\phi_k \mid t_k\}$, $\{\phi_{k'} \mid t_{k'}\}$, $\{\phi_{k''} \mid t_{k''}\}$, and $\phi_\Sigma \mathbf{k}_x \doteq \mathbf{k}_\sigma$; $\phi_\Sigma \mathbf{k}_y \doteq \mathbf{k}_{\sigma'}$; $\phi_\Sigma \mathbf{k}_z \doteq \mathbf{k}_{\sigma''}$.

Table III lists all the necessary information for calculating the nonzero $(\sigma'\sigma',\sigma'')$ blocks. As the table indicates, the (1,1,1), (2,2,2), and (3,3,3) blocks have identical coefficients as do the (1,2,3), (2,3,1), and (3,1,2) blocks where

$$U(1,2,3,)=[D^{(\mathbf{k}_x)(l)}(\{\rho_{y\bar{z}} \mid 0\})\otimes D^{(\mathbf{k}_y)(l')}(\{\rho_{x\bar{z}} \mid 0\})]U(1,1,1)\times D^{(\mathbf{k}_z)(l'')}(\{\rho_{x\bar{y}} \mid 0\})^{-1}. \tag{3.18}$$

The coefficients for the (1,2,3) block are given in Table IV.

TABLE IV. (1, 2, 3) Block of coefficients for $*X\otimes *X=*X$ in O_h^7.

	$X(1)$	$X(2)$	$X(3)$	$X(4)$
$X(1)\otimes X(1)$	$a\begin{bmatrix}1 & i\\ i & 1\\ i & 1\\ 1 & i\end{bmatrix}$	$a\begin{bmatrix}-1 & i\\ i-1\\ i-1\\ -1 & i\end{bmatrix}$	$a\begin{bmatrix}-1 & i\\ -1-i\\ 1 & i\\ 1-i\end{bmatrix}$	$a\begin{bmatrix}1-i\\ -1-i\\ 1 & i\\ -1 & i\end{bmatrix}$
$X(1)\otimes X(2)$	$a\begin{bmatrix}-1 & i\\ i-1\\ i-1\\ 1 & i\end{bmatrix}$	$a\begin{bmatrix}1 & i\\ i & 1\\ i & 1\\ 1 & i\end{bmatrix}$	$a\begin{bmatrix}-1-i\\ -1 & i\\ 1-i\\ 1 & i\end{bmatrix}$	$a\begin{bmatrix}1 & i\\ -1 & i\\ 1-i\\ -1-i\end{bmatrix}$
$X(1)\otimes X(3)$	$a\begin{bmatrix}1-1\\ i-i\\ 1-1\\ i & i\end{bmatrix}$	$a\begin{bmatrix}1-1\\ i & i\\ 0 & i\\ i-i\end{bmatrix}$	$b\begin{bmatrix}0-i\\ -1 & 0\\ 0 & i\\ 1 & 0\end{bmatrix}$	$b\begin{bmatrix}i & 0\\ 0-1\\ -i & 0\\ 0-1\end{bmatrix}$
$X(1)\otimes X(4)$	$a\begin{bmatrix}1-1\\ i-i\\ 1 & 1\\ i-i\end{bmatrix}$	$a\begin{bmatrix}1-1\\ -i & i\\ 1 & 1\\ i & i\end{bmatrix}$	$b\begin{bmatrix}i & 0\\ 0-1\\ -i & 0\\ 0 & 1\end{bmatrix}$	$b\begin{bmatrix}0 & i\\ -1 & 0\\ 0 & i\\ 1 & 0\end{bmatrix}$
$X(2)\otimes X(3)$	$a\begin{bmatrix}1-1\\ i & i\\ 1-1\\ i-i\end{bmatrix}$	$a\begin{bmatrix}-1 & 1\\ i & i\\ -1 & 1\\ i-i\end{bmatrix}$	$b\begin{bmatrix}0-i\\ 1 & 0\\ 0 & i\\ 1 & 0\end{bmatrix}$	$b\begin{bmatrix}i & 0\\ 0 & 1\\ -i & 0\\ 0 & 1\end{bmatrix}$
$X(2)\otimes X(4)$	$a\begin{bmatrix}-1-1\\ -i & i\\ 1 & 1\\ i & i\end{bmatrix}$	$a\begin{bmatrix}1-1\\ i-i\\ 1 & 1\\ i-i\end{bmatrix}$	$b\begin{bmatrix}i & 0\\ 0 & 1\\ -i & 0\\ 0-1\end{bmatrix}$	$b\begin{bmatrix}0 & i\\ 1 & 0\\ 0 & i\\ 1 & 0\end{bmatrix}$
$X(3)\otimes X(3)$	$b\begin{bmatrix}0 & 0\\ i-i\\ 1 & 1\\ 0 & 0\end{bmatrix}$	$b\begin{bmatrix}0 & 0\\ i-i\\ -1-1\\ 0 & 0\end{bmatrix}$	$c\begin{bmatrix}-1 & 0\\ 0 & 0\\ 0 & 0\\ 0-i\end{bmatrix}$	$c\begin{bmatrix}0 & 1\\ 0 & 0\\ 0 & 0\\ -i & 0\end{bmatrix}$
$X(3)\otimes X(4)$	$b\begin{bmatrix}-i & i\\ 0 & 0\\ 0 & 0\\ 1 & 1\end{bmatrix}$	$b\begin{bmatrix}-i & i\\ 0 & 0\\ 0 & 0\\ -1-1\end{bmatrix}$	$c\begin{bmatrix}0 & 0\\ 1 & 0\\ 0-i\\ 0 & 0\end{bmatrix}$	$c\begin{bmatrix}0 & 0\\ 0 & i\\ -1 & 0\\ 0 & 0\end{bmatrix}$
$X(4)\otimes X(4)$	$b\begin{bmatrix}0 & 0\\ 1 & 1\\ i-i\\ 0 & 0\end{bmatrix}$	$b\begin{bmatrix}0 & 0\\ -1-1\\ i-i\\ 0 & 0\end{bmatrix}$	$c\begin{bmatrix}0-1\\ 0 & 0\\ 0 & 0\\ i & 0\end{bmatrix}$	$c\begin{bmatrix}1 & 0\\ 0 & 0\\ 0 & 0\\ 0 & i\end{bmatrix}$

$a=1/\sqrt{8}$ $\quad b=1/2$ $\quad c=1/\sqrt{2}$

$X(2)\otimes X(2)=X(1)\otimes X(1)$

TABLE V. Principle block of coefficients for $*X \otimes *X = \Gamma$ in O_h^7.

	Γ^{1+}	Γ^{1-}	Γ^{2+}	Γ^{2-}	Γ^{12+}	Γ^{12-}	Γ^{15+}	Γ^{15-}	Γ^{25+}	Γ^{25-}
$X(1)\otimes X(1)$	$a\begin{bmatrix}0\\1\\1\\0\end{bmatrix}$			$a\begin{bmatrix}1\\0\\0\\1\end{bmatrix}$	$a\begin{bmatrix}0&0\\1&1\\1&1\\0&0\end{bmatrix}$	$a\begin{bmatrix}1&-1\\0&0\\0&0\\1&-1\end{bmatrix}$		$b\begin{bmatrix}0&0&0\\0&0&1\\0&0&-1\\0&0&0\end{bmatrix}$	$b\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\\0&0&-1\end{bmatrix}$	
$X(1)\otimes X(2)$		$a\begin{bmatrix}0\\1\\1\\0\end{bmatrix}a\begin{bmatrix}1\\0\\0\\1\end{bmatrix}$		$a\begin{bmatrix}1&-1\\0&0\\0&0\\1&-1\end{bmatrix}$	$a\begin{bmatrix}0&0\\1&1\\1&1\\0&0\end{bmatrix}$	$b\begin{bmatrix}0&0&0\\0&0&1\\0&0&-1\\0&0&0\end{bmatrix}$				$b\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\\0&0&-1\end{bmatrix}$
$X(1)\otimes X(3)$							$c\begin{bmatrix}-i&1&0\\i&1&0\\i&1&0\\i-1&0\end{bmatrix}$	$c\begin{bmatrix}-i&1&0\\-i-1&0\\i&1&0\\-i&1&0\end{bmatrix}$	$c\begin{bmatrix}i&1&0\\-i&1&0\\-i&1&0\\-i-1&0\end{bmatrix}$	$c\begin{bmatrix}i&1&0\\i-1&0\\-i&1&0\\i&1&0\end{bmatrix}$
$X(1)\otimes X(4)$							$c\begin{bmatrix}-i&1&0\\-i-1&0\\-i-1&0\\i-1&0\end{bmatrix}$	$c\begin{bmatrix}-i&1&0\\i&1&0\\-i-1&0\\-i&1&0\end{bmatrix}$	$c\begin{bmatrix}i&1&0\\i-1&0\\i-1&0\\-i-1&0\end{bmatrix}$	$c\begin{bmatrix}i&1&0\\-i&1&0\\i-1&0\\i&1&0\end{bmatrix}$
$X(3)\otimes X(3)$	$a\begin{bmatrix}1\\0\\0\\1\end{bmatrix}a\begin{bmatrix}1\\0\\0\\1\end{bmatrix}$			$a\begin{bmatrix}1&1\\0&0\\0&0\\1&-1\end{bmatrix}$	$a\begin{bmatrix}1&1\\0&0\\0&0\\1&1\end{bmatrix}$			$b\begin{bmatrix}0&0&0\\0&0&1\\0&0&1\\0&0&0\end{bmatrix}$	$b\begin{bmatrix}0&0&0\\0&0&1\\0&0&-1\\0&0&0\end{bmatrix}$	
$X(3)\otimes X(4)$		$a\begin{bmatrix}0\\1\\1\\0\end{bmatrix}a\begin{bmatrix}0\\1\\-1\\0\end{bmatrix}$		$a\begin{bmatrix}0&0\\1&-1\\1&-1\\0&0\end{bmatrix}$	$a\begin{bmatrix}0&0\\1&-1\\-1&1\\0&0\end{bmatrix}$	$b\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\\0&0&-1\end{bmatrix}$	$b\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\\0&0&1\end{bmatrix}$			

$a = 1/\sqrt{6}$ $\quad b = 1/\sqrt{2}$ $\quad c = 1/\sqrt{8}$

$U(X(2)\otimes X(2) = \Gamma(j)) = U(X(1)\otimes X(1) = \Gamma(j))$
$U(X(2)\otimes X(3) = \Gamma(j^{\pm})) = U(X(1)\otimes X(3) = \Gamma(j^{\mp}))$
$U(X(2)\otimes X(4) = \Gamma(j^{\pm})) = U(X(1)\otimes X(4) = \Gamma(j^{\mp}))$
$U(X(4)\otimes X(4) = \Gamma(j)) = U(X(3)\otimes X(3) = \Gamma(j))$

B. The Clebsch-Gordan coefficients for $*X \otimes *X \to \Gamma$

The coefficients for $D^{(*X)(l)} \otimes D^{(*X)(l')} = \otimes D^{(\Gamma)(l'')}$ can be obtained in a similar manner. In this case we choose $\mathbf{k} = \mathbf{k}_z$, $\mathbf{k}' = \mathbf{k}_z$, $\mathbf{k}'' = 0$ so that the (1,1,1) block is nonzero.

Table V lists the coefficients for the principle (1,1,1) block and Table VI lists the information necessary to obtain the remaining nonzero blocks. The Clebsch–Gordan coefficients for $*\mathbf{X} \otimes *\mathbf{X}$ in O_h^7 have already been obtained by Saulevich, Sviridov, and Smirnov.[5] Our results disagree with theirs in many cases, and although we have not located the source of disagreement, we believe our coefficients are correct.

4. $*X \otimes *X$ IN ROCKSALT O_h^5

The rocksalt structure is also a face centered cubic lattice having symmorphic space group O_h^5. The point X and the wave vector rules are the same as in diamond. The irreducible representations at X are the ten representations of point group D_{4h}. These representations are found in Table VII. The selection rules for these representations have been given by Chen, Berenson, and Birman.[6]

Using the methods of the previous sections the nonzero blocks of coefficients for $*X \otimes *X \to *X$ were calculated and are given in Table VIII.

We will work out a calculation for the (1,1,1) block for $D^{(*\mathbf{k}_x)(X_5)} \otimes D^{(*\mathbf{k}_y)(X_5)} \to D^{(*\mathbf{k}_z)(X_5)}$. This case is interesting since it includes multiplicity. As in the calculation for diamond we can write the intersection group N as

$$N = S(\mathbf{k}) + \{i\,|\,0\}S(\mathbf{k}), \tag{4.1}$$

where

$$S(\mathbf{k}) = \{\epsilon\,|\,0\} + \{\delta_{2x}\,|\,0\} + \{\delta_{2y}\,|\,0\} + \{\delta_{2z}\,|\,0\} \tag{4.2}$$

TABLE VI. Calculating the $(\sigma, \sigma', \sigma'')$ block for $*X \otimes *X = \Gamma$ in O_h^7.

$\sigma\ \sigma'\ \sigma''$	$\{\phi_\Sigma\,\vert\,\mathbf{t}_k\}$	$\{\phi_k\,\vert\,\mathbf{t}_k\}$	$\{\phi_{k'}\,\vert\,\mathbf{t}_{k'}\}$	$\{\phi_{k''}\,\vert\,\mathbf{t}_{k''}\}$
1 1 1	$\{\epsilon\,\vert\,0\}$	$\{\epsilon\,\vert\,0\}$	$\{\epsilon\,\vert\,0\}$	$\{\epsilon\,\vert\,0\}$
2 2 1	$\{\delta_{3xyz}^{-1}\,\vert\,0\}$	$\{\epsilon\,\vert\,0\}$	$\{\epsilon\,\vert\,0\}$	$\{\delta_{3xyz}^{-1}\,\vert\,0\}$
3 3 1	$\{\delta_{3xyz}\,\vert\,0\}$	$\{\epsilon\,\vert\,0\}$	$\{\epsilon\,\vert\,0\}$	$\{\delta_{3xyz}\,\vert\,0\}$

	$\{\delta_{3xyz}\,\vert\,0\}$	$\{\delta_{3xyz}^{-1}\,\vert\,0\}$
$\Gamma^{(1^{\pm})}$	1	1
$\Gamma^{(2^{\pm})}$	1	1
$\Gamma^{(12^{\pm})}$	$\begin{pmatrix}\epsilon^2&0\\0&\epsilon\end{pmatrix}$	$\begin{pmatrix}\epsilon&0\\0&\epsilon^2\end{pmatrix}$
$\Gamma^{(15^{\pm})}$	$\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$	$\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}$
$\Gamma^{(25^{\pm})}$	$\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$	$\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}$

$\epsilon = e^{2\pi i/3}$

TABLE VII. Representations for $D^{(k_\alpha)(X(l))}$ in O_h^5.

	$X_1^\pm$	$X_2^\pm$	$X_3^\pm$	$X_4^\pm$	$X_5^\pm$
ϵ	1	1	1	1	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$
δ_{4z}	1	−1	−1	1	$\begin{pmatrix}0&-1\\1&0\end{pmatrix}$
δ_{4z}^{-1}	1	−1	−1	1	$\begin{pmatrix}0&1\\-1&0\end{pmatrix}$
δ_{2z}	1	1	1	1	$\begin{pmatrix}-1&0\\0&-1\end{pmatrix}$
δ_{2x}	1	1	−1	−1	$\begin{pmatrix}1&0\\0&-1\end{pmatrix}$
δ_{2y}	1	1	−1	−1	$\begin{pmatrix}-1&0\\0&1\end{pmatrix}$
δ_{2xy}	1	−1	1	−1	$\begin{pmatrix}0&1\\1&0\end{pmatrix}$
$\delta_{2x\bar{y}}$	1	−1	1	−1	$\begin{pmatrix}0&-1\\-1&0\end{pmatrix}$
i	±1	±1	±1	±1	$\begin{pmatrix}\pm1&0\\0&\mp1\end{pmatrix}$
σ_{4z}	±1	∓1	∓1	±1	$\begin{pmatrix}0&\mp1\\\pm1&0\end{pmatrix}$
σ_{4z}^{-1}	±1	∓1	∓1	±1	$\begin{pmatrix}0&\pm1\\\mp1&0\end{pmatrix}$
ρ_z	±1	±1	±1	±1	$\begin{pmatrix}\mp1&0\\0&\mp1\end{pmatrix}$
ρ_x	±1	±1	∓1	∓1	$\begin{pmatrix}\pm1&0\\0&\mp1\end{pmatrix}$
ρ_y	±1	±1	∓1	∓1	$\begin{pmatrix}\mp1&0\\0&\pm1\end{pmatrix}$
ρ_{xy}	±1	∓1	±1	∓1	$\begin{pmatrix}0&\pm1\\\pm1&0\end{pmatrix}$
$\rho_{x\bar{y}}$	±1	∓1	±1	∓1	$\begin{pmatrix}0&\mp1\\\mp1&0\end{pmatrix}$

and $M^k_{l\alpha\mu,\sigma}$ is then given as in (3.8) and (3.9) with $\tau=0$. For all representations $D^{(k)(l)}(\{i|0\})=\pm1$ so that from Eq. (3.10) $Z=2Z'$ for (l,l',l'') containing an even number of odd representations and $Z=0$ otherwise.

Writing those portions of M^k that correspond specifically to X_5 we have

$$\sqrt{8}M^{kx}_{X_5\alpha_1\mu,\sigma}=\begin{bmatrix}1&-1&-1&1\\0&0&0&0\end{bmatrix};$$
$$\sqrt{8}M^{ky}_{X_5\alpha_1\mu,\sigma}=\begin{bmatrix}1&-1&1&-1\\0&0&0&0\end{bmatrix};$$
$$\sqrt{8}M^{kz}_{X_5\alpha,\mu,\sigma}=\begin{bmatrix}1&1&-1&-1\\0&0&0&0\end{bmatrix};$$
$$\sqrt{8}M^{kx}_{X_5\alpha_2\mu,\sigma}=\begin{bmatrix}0&0&0&0\\1&1&-1&-1\end{bmatrix};$$
$$\sqrt{8}M^{ky}_{X_5\alpha_2\mu,\sigma}=\begin{bmatrix}0&0&0&0\\1&-1&-1&1\end{bmatrix};$$
$$\sqrt{8}M^{kz}_{X_5\alpha_2\mu,\sigma}=\begin{bmatrix}0&0&0&0\\1&-1&1&-1\end{bmatrix};\tag{4.3}$$

Hence

$$Z_{X_5\alpha_1\mu X_5\alpha_1\mu',X_5\alpha_1\mu''}=\frac{1}{8}\begin{bmatrix}1&0\\0&0\\0&0\\0&0\end{bmatrix}\tag{4.4}$$

and

TABLE VIII. Clebsch–Gordan coefficients for $*X\otimes *X\to *X$ in rocksalt, O_h^5.

	$X(1)$		$X(2)$		$X(3)$		$X(4)$		$X(5)$	
	(111)	(123)	(111)	(123)	(111)	(123)	(111)	(123)	(111)	(123)
$X(1)\otimes X(1)$	a	a	a	a						
$X(1)\otimes X(2)$	a	$-a$	a	a						
$X(2)\otimes X(2)$	a	a	a	$-a$						
$X(3)\otimes X(3)$					a	a	a	$-a$		
$X(3)\otimes X(4)$					a	$-a$	a	a		
$X(4)\otimes X(4)$					a	a	a	$-a$		
$X(1)\otimes X(3)$									[0 1]	[−1 0]
$X(1)\otimes X(4)$									[0 1]	[1 0]
$X(2)\otimes X(3)$									[0 1]	[1 0]
$X(2)\otimes X(4)$									[0 1]	[−1 0]
$X(1)\otimes X(5)$					$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}-1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0&a\\0&0\end{bmatrix}$	$\begin{bmatrix}0&0\\a&0\end{bmatrix}$
$X(2)\otimes X(5)$					$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}-1\\0\end{bmatrix}$	$\begin{bmatrix}0&a\\0&0\end{bmatrix}$	$\begin{bmatrix}0&0\\-a&0\end{bmatrix}$
$X(3)\otimes X(5)$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}-1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$					$\begin{bmatrix}0&0\\a&0\end{bmatrix}$	$\begin{bmatrix}0&a\\0&0\end{bmatrix}$
$X(4)\otimes X(5)$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}-1\\0\end{bmatrix}$					$\begin{bmatrix}0&0\\a&0\end{bmatrix}$	$\begin{bmatrix}0&-a\\0&0\end{bmatrix}$
$X(5)\otimes X(5)$	$\begin{bmatrix}0\\0\\a\\0\end{bmatrix}$	$\begin{bmatrix}0\\a\\0\\0\end{bmatrix}$	$\begin{bmatrix}0\\0\\a\\0\end{bmatrix}$	$\begin{bmatrix}0\\-a\\0\\0\end{bmatrix}$	$\begin{bmatrix}0\\a\\0\\0\end{bmatrix}$	$\begin{bmatrix}0\\0\\a\\0\end{bmatrix}$	$\begin{bmatrix}0\\a\\0\\0\end{bmatrix}$	$\begin{bmatrix}0\\0\\-a\\0\end{bmatrix}$	$\begin{bmatrix}1&0\\0&0\\0&0\\0&0\end{bmatrix}$	$\begin{bmatrix}-1&0\\0&0\\0&0\\0&0\end{bmatrix}$
									$\begin{bmatrix}0&0\\0&0\\0&0\\0&1\end{bmatrix}$	$\begin{bmatrix}0&0\\0&0\\0&0\\0-1\end{bmatrix}$

$a=1/\sqrt{2}$

$X_i^+\otimes X_j^+\to X_k^+=X_i^+\otimes X_j^-\to X_k^-=X_i^-\otimes X_j^+\to X_k^-=X_i^-\otimes X_j^-\to X_k^+$

$$Z_{X_5\alpha_2\mu X_5\alpha_2\mu',X_5\alpha_2\mu''}=\frac{1}{8}\begin{bmatrix}0&0\\0&0\\0&0\\0&1\end{bmatrix}. \tag{4.5}$$

Normalizing we find

$$U^{(\mathbf{k}_x)(X_5)\otimes(\mathbf{k}_y)(X_5)}_{1\mu 1\mu',(\mathbf{k}_z)(X_5)\gamma_1 1\mu''}=\begin{bmatrix}1&0\\0&0\\0&0\\0&0\end{bmatrix} \tag{4.6}$$

and

$$U^{(\mathbf{k}_x)(X_5)\otimes(\mathbf{k}_y)(X_5)}_{1\mu 1\mu',(\mathbf{k}_z)(X_5)\gamma_2 1\mu''}=\begin{bmatrix}0&0\\0&0\\0&0\\0&1\end{bmatrix}. \tag{4.7}$$

5. USE OF TABLES

We will now demonstrate how the tables of Clebsch–Gordan coefficients can be used to obtain the correct linear combinations of product functions. That is, if the l functions $\{\psi^l_\mu\}$ are bases for $D^{(l)}$ and the l' functions $\{\psi^{l'}_{\mu'}\}$ are bases for $D^{(l')}$ then the l'' functions $\psi^{l''\gamma}_{\mu''}$ given by

$$\psi^{l''\gamma}_{\mu''}=\sum_{\mu\mu'}U^{l\otimes l'}_{\mu\mu',l''\gamma\mu''}\psi^l_\mu\psi^{l'}_{\mu'} \tag{5.1}$$

are bases for $D^{(l'')}$ which occurs γ times in the reduction of the product $D^{(l)}\otimes D^{(l')}$. In particular, for space group representations (5.1) can be rewritten using block notation as follows:

$$\psi^{(\mathbf{k}'')(l'')\gamma}_{\sigma''\mu''}=\sum_{\substack{\sigma\mu\\ \sigma'\mu'}}U^{(\mathbf{k})(l)\otimes(\mathbf{k}')(l')}_{\sigma\mu\sigma'\mu',(\mathbf{k}'')(l'')\gamma\sigma''\mu''}\psi^{(\mathbf{k})(l)}_{\sigma\mu}\psi^{(\mathbf{k}')(l')}_{\sigma'\mu'}, \tag{5.2}$$

where $\psi^{(\mathbf{k}'')(l'')}_{\sigma''\mu''}$ is a basis functions for $D^{(*\mathbf{k}'')(l'')}$ which appears γ times in the reductions of the product $D^{(*\mathbf{k})(l)}\otimes D^{(*\mathbf{k}')(l')}$. As particular examples we will consider the correct linear combinations which arise in $X(3)\otimes X(4)\to X(1)$ and $X(3)\otimes X(4)\to\Gamma(12+)$ in diamond.

For $X(3)\otimes X(4)\to X(1)$ we see from Tables II and IV that

$$U^{(\mathbf{k}_x)(X_3)\otimes(\mathbf{k}_y)(X_4)}_{\sigma\mu\sigma\mu',(\mathbf{k}_z)(X_1)\sigma\mu''}=\frac{1}{2}\begin{bmatrix}1&1\\0&0\\0&0\\-i&i\end{bmatrix},\quad\text{for }\sigma=1,2,3 \tag{5.3}$$

and

$$U^{(\mathbf{k}_x)(X_3)\otimes(\mathbf{k}_y)(X_4)}_{\sigma\mu\sigma'\mu',(\mathbf{k}_z)(X_1)\sigma''\mu''}=\frac{1}{2}\begin{bmatrix}-i&i\\0&0\\0&0\\1&1\end{bmatrix},$$

$$\text{for }(\sigma\sigma'\sigma'')=(1\,2\,3),\ (2\,3\,1),\ (3\,1\,2), \tag{5.4}$$

where the elements $(\mu\mu',\mu'')$ are given as in Eq. (3.14). All other coefficients are zero.

We can use the coefficients in (5.3) and (5.4) to find the correct linear combinations of products $\{\psi^{(\mathbf{k}_x)(X_1)}_{\sigma\mu}\times\psi^{(\mathbf{k}_y)(X_4)}_{\sigma'\mu'}\}$ that transform as a basis for $D^{(*\mathbf{k}_z)(X_1)}$, say $\psi^{(\mathbf{k}_z)(X_1)}_{11}$. Then

$$\psi^{(\mathbf{k}_z)(X_1)}_{11}=\sum_{\substack{\sigma\mu\\ \sigma'\mu'}}U^{(\mathbf{k}_x)(X_3)\otimes(\mathbf{k}_y)(X_4)}_{\sigma\mu\sigma'\mu',(\mathbf{k}_z)(X_1)11}\psi^{(\mathbf{k}_x)(X_3)}_{\sigma\mu}\psi^{(\mathbf{k}_y)(X_4)}_{\sigma'\mu'}$$

$$=\frac{1}{2}\left[\psi^{(\mathbf{k}_x)(X_3)}_{11}\psi^{(\mathbf{k}_y)(X_4)}_{11}-i\psi^{(\mathbf{k}_x)(X_3)}_{12}\psi^{(\mathbf{k}_y)(X_4)}_{12}\right.$$

$$\left.-i\psi^{(\mathbf{k}_x)(X_3)}_{21}\psi^{(\mathbf{k}_y)(X_4)}_{31}+\psi^{(\mathbf{k}_x)(X_3)}_{22}\psi^{(\mathbf{k}_y)(X_4)}_{32}\right]. \tag{5.5}$$

Similarly for $X(3)\otimes X(4)\to\Gamma(12+)$ we have from Table V.

$$U^{(\mathbf{k}_z)(X_3)\otimes(\mathbf{k}_z)(X_4)}_{1\mu 1\mu',\Gamma(12+)1\mu''}=\frac{1}{\sqrt{6}}\begin{bmatrix}0&0\\1&-1\\1&-1\\0&0\end{bmatrix} \tag{5.6}$$

and using Table VI

$$U_{2\mu 2\mu',\Gamma(12+)1\mu''}=\frac{1}{\sqrt{6}}\begin{bmatrix}0&0\\\epsilon^2&-\epsilon\\\epsilon^2&-\epsilon\\0&0\end{bmatrix} \tag{5.7}$$

and

$$U_{3\mu 3\mu',\Gamma(12+)1\mu''}=\frac{1}{\sqrt{6}}\begin{bmatrix}0&0\\\epsilon&-\epsilon^2\\\epsilon&-\epsilon^2\\0&0\end{bmatrix},$$

where $\epsilon=\exp(2\pi i/3)$. Therefore

$$\psi^{\Gamma(12+)}_1=\frac{1}{\sqrt{6}}\left[\psi^{(\mathbf{k}_z)(X_3)}_{11}\psi^{(\mathbf{k}_z)(X_4)}_{12}+\psi^{(\mathbf{k}_z)(X_3)}_{12}\psi^{(\mathbf{k}_z)(X_4)}_{11}\right.$$

$$+\epsilon^2\left(\psi^{(\mathbf{k}_z)(X_3)}_{21}\psi^{(\mathbf{k}_z)(X_4)}_{22}+\psi^{(\mathbf{k}_z)(X_3)}_{22}\psi^{(\mathbf{k}_z)(X_4)}_{21}\right)$$

$$\left.+\epsilon\left(\psi^{(\mathbf{k}_z)(X_3)}_{31}\psi^{(\mathbf{k}_z)(X_4)}_{32}+\psi^{(\mathbf{k}_z)(X_3)}_{32}\psi^{(\mathbf{k}_z)(X_4)}_{31}\right)\right]. \tag{5.8}$$

$X(3)\otimes X(4)\to\Gamma(12+)$ is a particular example of results which differ from those of Saulevich, Sviridov and Smirnov.[5] Our coefficients have been checked by applying projection operators to the wave function in (5.8) and for a similar wave function $\psi^{\Gamma(12+)}_2$ and we thereby demonstrated that we do indeed have the correct linear combinations.

*Supported in part by NSF-GH-31742 grant and AROD-DAHCO4-74-G-0109 grant.

†This paper is based in part upon a Ph.D. thesis submitted by Rhoda Berenson in partial fulfillment of the requirements for the Ph.D. degree in the Physics Department, New York University, February 1974.

‡This paper is based in part upon a Ph.D. thesis submitted by Irving Itzkan in partial fulfillment of the requirements for the Ph.D. degree in the Physics Department, New York University, October 1969.

§Address after 1 September 1974: City College of the CUNY, Physics Department, Convent Ave. and 138 St., New York, N.Y. 10031.

[1]R. Berenson and J. L. Birman, J. Math. Phys. 16, 227 (1975), preceding paper.
[2]C. Herring, J. Franklin Inst. 233, 525 (1942).
[3]O. V. Kovalev, *Irreducible Representations of the Space Groups* (Gordan and Breach, New York, 1965).
[4]J. L. Birman, Phys. Rev. 127, 1093 (1962).
[5]L. K. Saulevich, D. T. Sviridov, and Yu. F. Smirnov, Sov. Phys.—Crystall. 15, 355 (1970).
[6]L. C. Chen, R. Berenson, and J. L. Birman, Phys. Rev. 170, 639 (1968).

On the spinning axis representation

J. A. de Azcárraga

Física Matemática and GIFT, Facultad de Ciencias, Salamanca, Spain

L. Navarro

Departamento de Física Teórica and GIFT, Facaltad de Ciencias, Barcelona, Spain
(Received 10 April 1974)

A new spinning axis representation is introduced. It allows us to calculate the evolution operator of a system with slowly varying time-dependent Hamiltonian with the desired degree of approximation in the parameter used for describing its dynamical evolution. The procedure is compared with a previously existing one and applied to a simple example.

1. INTRODUCTION

The application of perturbative methods to systems described by slowly varying time-dependent Hamiltonians allows one to obtain asymptotic rather than convergent expansions for the solution of the Schrödinger equation. This is usually done making use of the "spinning axis representation" (SAR) which is defined by the fact that, in it, the Hamiltonian eigenvectors are independent of time. The application of the SAR to the Schrödinger equation satisfied by the evolution operator leads to an equation which makes easy the use of time-dependent perturbation theory. In this way the ordinary adiabatic theorem is obtained, which can be established in the following form: "If the state of a system is represented at $t=t_0$ by the ket $|\epsilon_j(t_0)\rangle$ which is a solution of the eigenvalue equation $H(t)|\epsilon_j(t)\rangle=\epsilon_j(t)|\epsilon_j(t)\rangle$ for $t=t_0$, at $t=t_1$ the state of the system is represented by the ket $|\epsilon_j(t_1)\rangle$ which is solution of the same equation for $t=t_1$." This result is more accurate the stronger is the inequality

$$|\max(\alpha)/\min(\omega)|^2 \ll 1, \tag{1.1}$$

where $\max(\alpha)$ is the maximum value of the "angular velocity" of the eigenaxes of the Hamiltonian, and $\min(\omega)$ is the minimum value of the Bohr frequency for transitions between eigenstates.[1]

In the usual treatment of the adiabatic theorem it is supposed that the eigenvalues of the Hamiltonian remain separated during the transition period $T=t_1-t_0$. However, the adiabatic theorem remains valid if this assumption is violated[2]; in the same way, the hypothesis of discrete spectrum is not necessary.[3] These restrictions are introduced for the sake of simplicity, and we shall maintain them here.

The ordinary adiabatic theorem, valid in the first approximation of order $1/T$, may be generalized to any order without requiring special properties for the Hamiltonian. To this aim, instead of applying perturbation theory to the Schrödinger equation in the SAR as is done in the usual adiabatic theorem, a new change of representation is made[4] which fixes the eigenstates of the operator which plays the role of the Hamiltonian in the SAR of the Schrodinger equation. This transformation defines a new SAR, and the procedure can be repeated to get the desired order of approximation. In this way, to obtain the evolution operator U of a system up to terms of order $(1/T)^n$, *n succesive transformations* (n SAR) are required. This makes the method rather cumbersome in practice, since to obtain the form of the nth transformation it is necessary to know the explicit form of the previous $(n-1)$ SAR.

In this paper we propose a method which allows one to obtain the evolution operator U to the desired degree of approximation by means of only *one* change of representation which defines a generalized SAR (GSAR). The action of the operator defining the GSAR is obviously equivalent to the combined action of the n operators R^i defining the succesive SAR of Ref. 4.

In order to show this, we describe briefly in Sec. 2 the ordinary SAR and introduce in Sec. 3 the GSAR. Finally, in Sec. 4 we consider a simple example which illustrates the method.

2. THE ORDINARY SPINNING AXIS REPRESENTATION

Let us consider a quantum system whose dynamical evolution is determined by its Hamiltonian $H(t)$. We shall assume a good behavior for it and its derivatives in the interval $t_0 \leq t \leq t_1$ in which we consider the variation of $H(t)$. We shall measure the time by means of the variable $\tau=t-t_0/T$, where $T=t_1-t_0$. We shall suppose T large (adiabatic evolution) in the sense that the inequality (1.1) holds.

Let $U(\tau)$ be the evolution operator of the system described by $H(\tau)$. $U(\tau)$ is then the solution of

$$i\hbar\frac{d}{d\tau}U(\tau)=TH(\tau)U(\tau) \tag{2.1}$$

with the initial condition

$$U(0)=I. \tag{2.2}$$

After applying the unitary transformation $A(\tau)$, (2.1) takes the form

$$i\hbar\frac{d}{d\tau}U^{(A)}(\tau)=T[H^{(A)}(\tau)-(1/T)K^{(A)}(\tau)]U^{(A)}(\tau) \tag{2.3}$$

where $U^{(A)}(\tau)\equiv A^\dagger(\tau)U(\tau)$ is the new evolution operator, $H^{(A)}(\tau)\equiv A^\dagger(\tau)H(\tau)A(\tau)$ is the transformed Hamiltonian, and $K^{(A)}(\tau)\equiv A^\dagger(\tau)K(\tau)A(\tau)$, with $K(\tau)$ being the generator of the transformation $A(\tau)$, i. e., the solution of the differential equation

$$i\hbar\frac{d}{d\tau}A(\tau)=K(\tau)A(\tau) \tag{2.4}$$

subjected to the initial condition $K(0)=0$. Note that in (2.3) the operator which plays the role of the Hamilton-

ian of the new representation differs from the transformed Hamiltonian in the small term $-(1/T)K^{(A)}(\tau)$.

It is clear that the equation satisfied by the evolution operator is easily integrated if the projectors of the eigenvectors of the Hamiltonian do not depend on time. Thus, to calculate U in the adiabatic approximation, A (SAR) is defined in such a way that $A|\epsilon_j(0)\rangle = |\epsilon_j(\tau)\rangle$, $j=1,2,\cdots$, and then the transformed Hamiltonian $H^{(A)}=\sum_j \epsilon_j(\tau)|\epsilon_j(0)\rangle\langle\epsilon_j(0)|$ is treated by perturbation theory. The procedure can be improved by finding a new SAR for the operator $[H^{(A)}(\tau)-(1/T)K^{(A)}(\tau)]$ of (2.3). In this way one is naturally led to the method of Ref. 4 in which, to get an approximation of nth order, n succesive SAR are needed. Instead of doing this, however, we look in the next section for a unique transformation which will allow us to calculate U with the same degree of approximation.

3. THE NEW TRANSFORMATION

Let us write (2.3) in the form

$$\begin{aligned} i\hbar\frac{d}{d\tau}U^{(S)}(\tau) &= TS^{\dagger}(\tau)\left(H(\tau)-\frac{1}{T}K(\tau)\right)S(\tau)U^{(S)}(\tau) \\ &\equiv TS^{\dagger}(\tau)H'(\tau)S(\tau)U^{(S)}(\tau) \\ &= TH'^{(S)}(\tau)U^{(S)}(\tau). \end{aligned} \tag{3.1}$$

If it were possible to find a transformation S such that $H'^{(S)}(\tau)$ [not $H^{(S)}(\tau)$] had time-independent eigenvectors, i.e., such that

$$S^{\dagger}(\tau)|\epsilon_j(\tau)\rangle = |\epsilon_j(0)\rangle, \quad j=1,2,\cdots, \tag{3.2}$$

where

$$H'(\tau)|\epsilon_j(\tau)\rangle = \epsilon_j(\tau)|\epsilon_j(\tau)\rangle, \tag{3.3}$$

(3.1) could be exactly solved giving for U,

$$U(\tau)=S(\tau)\sum_j \exp[-iT\varphi_j(\tau)/\hbar]|\epsilon_j(0)\rangle\langle\epsilon_j(0)| \tag{3.4}$$

with

$$\varphi_j(\tau)=\int_0^{\tau}\epsilon_j(\tau')\,d\tau'. \tag{3.5}$$

It is obvious that S, which defines the GSAR, cannot be found exactly: A simple inspection of (3.1)–(3.3) shows that, to find $S(\tau)$, one needs to solve the eigenvalue problem for $H'(\tau)$ and that, with this aim, it is necessary to know the generator K of the transformation S which enters in the definition of $H'(\tau)$. We show next, however, that it is possible to determine S (and thus U) with the desired degree of approximation taking into account that T is large in the already specified sense.

A. First approximation

Let $|\epsilon_j^{(1)}(\tau)\rangle$, $S^{(1)}(\tau)$, $K^{(1)}(\tau)$, and $\epsilon_j^{(1)}(\tau)$ be the results of this approximation (the order corresponds to that in which the practical calculation is to be made). Since

$$H'(\tau)=H(\tau)-(1/T)K(\tau) \tag{3.6}$$

in the first approximation, we find that $|\epsilon_j^{(1)}(\tau)\rangle$ are defined through

$$H(\tau)|\epsilon_j^{(1)}(\tau)\rangle=\epsilon_j^{(1)}(\tau)|\epsilon_j^{(1)}(\tau)\rangle, \quad j=1,2,\cdots, \tag{3.7}$$

and

$$S^{(1)}(\tau)|\epsilon_j^{(1)}(0)\rangle = |\epsilon_j^{(1)}(\tau)\rangle. \tag{3.8}$$

$S^{(1)}(\tau)$ clearly corresponds to the transformation $A(\tau)$ mentioned in Sec. 2 and to $R^{(1)}(\tau)$ of Ref. 4. The generator $K^{(1)}(\tau)$ can also be calculated, and is given by

$$K^{(1)}(\tau)=i\hbar\sum_j\left(\frac{d}{d\tau}|\epsilon_j^{(1)}(\tau)\rangle\right)\langle\epsilon_j^{(1)}(\tau)|. \tag{3.9}$$

The choice in (3.9) implies that the arbitrariness of the phase of $|\epsilon_j^{(1)}(\tau)\rangle$ has been eliminated in such a way that

$$\langle\epsilon_j^{(1)}(\tau)|\left(\frac{d}{d\tau}|\epsilon_j^{(1)}(\tau)\rangle\right)=0. \tag{3.10}$$

Finally,

$$\varphi_j^{(1)}(\tau)=\int_0^{\tau}d\tau'\,\epsilon_j^{(1)}(\tau'), \tag{3.11}$$

where $\epsilon_j^{(1)}(\tau)$ is the perturbation value for $\epsilon_j(\tau)$, which, since

$$\langle\epsilon_j^{(1)}(\tau)|(1/T)K^{(1)}(\tau)|\epsilon_j^{(1)}(\tau)\rangle=0 \tag{3.12}$$

is precisely the eigenvalue of $H(\tau)$ [(3.7)].

In this approximation, the evolution operator is

$$U^{(1)}(\tau)=S^{(1)}(\tau)\sum_j \exp[-iT\varphi_j^{(1)}(\tau)/\hbar]|\epsilon_j^{(1)}(0)\rangle\langle\epsilon_j^{(1)}(0)| \tag{3.13}$$

and, as one would expect, coincides with the result which is obtained in the ordinary adiabatic theorem,[1] which corresponds to an approximation of order $1/T$.

B. Second approximation

We proceed now to calculate $|\epsilon_j^{(2)}(\tau)\rangle$, $S^{(2)}(\tau)$, $K^{(2)}(\tau)$, and $\epsilon_j^{(2)}(\tau)$. Straightforward application of ordinary perturbation theory to (3.6) gives

$$\begin{aligned} |\epsilon_j^{(2)}(\tau)\rangle &= |\epsilon_j^{(1)}(\tau)\rangle+\sum_{i\neq j}\frac{\langle\epsilon_i^{(1)}(\tau)|-(1/T)K^{(1)}(\tau)|\epsilon_j^{(1)}(\tau)\rangle}{\epsilon_j^{(1)}(\tau)-\epsilon_i^{(1)}(\tau)} \\ &\quad\times|\epsilon_i^{(1)}(\tau)\rangle, \end{aligned} \tag{3.14}$$

which, with (3.9), is

$$|\epsilon_j^{(2)}(\tau)\rangle = |\epsilon_j^{(1)}(\tau)\rangle-(i/T)\sum_{i\neq j}c_{ij}(\tau)|\epsilon_i^{(1)}(\tau)\rangle, \tag{3.15}$$

where

$$c_{ij}^{(1)}(\tau)=\alpha_{ij}^{(1)}(\tau)/\omega_{ji}^{(1)}(\tau), \quad \alpha_{ij}^{(1)}(\tau)=\langle\epsilon_i^{(1)}(\tau)|\left(\frac{d}{d\tau}|\epsilon_j^{(1)}(\tau)\rangle\right),$$

$$\omega_{ij}^{(1)}(\tau)=\frac{\epsilon_i^{(1)}(\tau)-\epsilon_j^{(1)}(\tau)}{\hbar}. \tag{3.16}$$

In this order, $S^{(2)}(\tau)$ is given by

$$S^{(2)\dagger}(\tau)|\epsilon_j^{(2)}(\tau)\rangle = |\epsilon_j^{(2)}(0)\rangle \tag{3.17}$$

and $K^{(2)}(\tau)$ is

$$\begin{aligned} K^{(2)}(\tau) &= i\hbar\sum_j\left(\frac{d}{d\tau}|\epsilon_j^{(2)}(\tau)\rangle\right)\langle\epsilon_j^{(2)}(\tau)| \\ &\approx K^{(1)}(\tau)+\frac{\hbar}{T}\sum_j\sum_{i\neq j}\left\{\left(\frac{d}{d\tau}c_{ij}^{(1)}(\tau)\right)|\epsilon_i^{(1)}(\tau)\rangle\langle\epsilon_j^{(1)}(\tau)|\right. \\ &\quad -c_{ij}^{(1)*}(\tau)\left(\frac{d}{d\tau}|\epsilon_j^{(1)}(\tau)\rangle\right)\langle\epsilon_i^{(1)}(\tau)| \\ &\quad \left.+c_{ij}^{(1)}(\tau)\left(\frac{d}{d\tau}|\epsilon_i^{(1)}(\tau)\rangle\right)\langle\epsilon_j^{(1)}(\tau)|\right\} \\ &\equiv K^{(1)}(\tau)+\frac{\hbar}{T}\sum_j\sum_{i\neq j}F_{ij}^{(1)}(\tau). \end{aligned} \tag{3.18}$$

Putting in (3.6) $(1/T)K^{(2)}$ instead of $(1/T)K$ and applying perturbation theory, one finds that to order $1/T^2$ the contributions to $\epsilon_j^{(2)}(\tau)$ will come from the diagonal and nondiagonal terms of $K^{(1)}(\tau)$ and from the diagonal elements of $F_{ij}^{(1)}(\tau)$. Of all these, only the nondiagonal elements of $K^{(1)}$ contribute, giving the final result

$$\epsilon_j^{(2)}(\tau)=\epsilon_j^{(1)}(\tau)+\frac{\hbar}{T^2}\sum_{i\neq j}\frac{|\alpha_{ij}(\tau)|^2}{\omega_{ji}(\tau)}\,. \tag{3.19}$$

The expression for the evolution operator in this approximation is thus

$$U^{(2)}(\tau)=S^{(2)}(\tau)\times\sum_j\exp\left[-iT\int_0^\tau\left(\epsilon_j^{(1)}(\tau')+\frac{\hbar}{T^2}\sum_{i\neq j}\frac{|\alpha_{ij}(\tau')|^2}{\omega_{ji}}\right)d\tau'\right]\times|\epsilon_j^{(2)}(0)\rangle\langle\epsilon_j^{(2)}(0)|, \tag{3.20}$$

where everything has been previously calculated.

C. nth approximation

For nth order, $|\epsilon_j^{(n)}\rangle$ is obtained from $|\epsilon_j^{(n-1)}\rangle$ by taking $(1/T)K^{(n-1)}(\tau)$ as the perturbation and calculating to order $1/T^n$. $S^{(n)}$ is then determined by the relation

$$S^{(n)}|\epsilon_j^{(n)}(0)\rangle=|\epsilon_j^{(n)}(\tau)\rangle \tag{3.21}$$

and $K^{(n)}(\tau)$ given by

$$K^{(n)}(\tau)=i\hbar\sum_j\left(\frac{d}{d\tau}|\epsilon_j^{(n)}(\tau)\rangle\right)\langle\epsilon_j^{(n)}(\tau)|.$$

This allows one to calculate $\epsilon_j^{(n)}$ through the perturbing term $-(1/T)K^{(n)}(\tau)$, and finally one would obtain

$$U^{(n)}(\tau)=S^{(n)}(\tau)\sum_j\exp[-iT\textstyle\int_0^\tau\epsilon_j^{(n)}(\tau')\,d\tau']\times|\epsilon_j^{(n)}(0)\rangle\langle\epsilon_j^{(n)}(0)|, \tag{3.22}$$

which is the explicit expression for the evolution operator to order $(1/T)^n$.

4. AN EXAMPLE

As an application of the described method we calculate now, in second order, the evolution operator of a system consisting of an atom in a magnetic field whose direction is reversed adiabatically.[1]

The Hamiltonian is given by

$$H(t)=H^{(0)}+A(\mathbf{L}\cdot\mathbf{S})-\frac{e}{2\mu c}(L_z+2S_z)\mathcal{H}(t), \tag{4.1}$$

where the magnetic field $\mathcal{H}$ is always parallel to the z axis and changes adiabatically from $\mathcal{H}_0$ to $-\mathcal{H}_0$ according to the law

$$\mathcal{H}(t)=\mathcal{H}_0(2\tau-1), \tag{4.2}$$

τ being $t-t_0/T$, T large.

Let us suppose that the system is initially in a state ^2P with $J_z=\frac{1}{2}$. There are two eigenvectors of $H(t)$ corresponding to this value that will be linear combinations of the eigenstates $(|L_z,S_z\rangle)|0,\frac{1}{2}\rangle$ and $|1,-\frac{1}{2}\rangle$ of $H^{(0)}$. Taking the corresponding $H^{(0)}$ eigenvalue as the zero of energy, we get, in this basis,

$$H(\tau)=\tfrac{1}{4}A\hbar^2\{[-1-2\rho(\tau)]+b(\boldsymbol{\sigma}\mathbf{u})\}, \tag{4.3}$$

where

$$\rho(\tau)\equiv\frac{e\hbar}{2\mu c}\frac{\mathcal{H}(\tau)}{A\hbar^2},\quad \mathbf{u}\equiv\left(\frac{2\sqrt{2}}{b},0,\frac{1-2\rho}{b}\right),\quad b\equiv[8+(1-2\rho)^2]^{1/2},\quad |\mathbf{u}|=1, \tag{4.4}$$

and the $\boldsymbol{\sigma}$ are the Pauli matrices.

The starting point of our procedure is the eigenvalue problem for $H(\tau)$ given by (4.3). The two solutions, labelled $\pm$, are

$$\epsilon_\pm^{(1)}(\tau)=\tfrac{1}{4}A\hbar^2[-1-2\rho(\tau)\pm b(\tau)], \tag{4.5}$$

$$|\epsilon_+^{(1)}(\tau)\rangle=\frac{1}{\sqrt{2(1+u_z)}}\begin{pmatrix}1+u_z\\u_x+iu_y\end{pmatrix}=\frac{1}{\sqrt{2b(1-2\rho+b)}}\begin{pmatrix}1-2\rho+b\\2\sqrt{2}\end{pmatrix} \tag{4.6a}$$

$$|\epsilon_-^{(1)}(\tau)\rangle=\frac{1}{\sqrt{2(1+u_z)}}\begin{pmatrix}-u_x+iu_y\\1+u_z\end{pmatrix}=\frac{1}{\sqrt{2b(1-2\rho+b)}}\begin{pmatrix}-2\sqrt{2}\\1-2\rho+b\end{pmatrix}. \tag{4.6b}$$

Now, $S^{(1)}(\tau)$ is determined by (3.8), which in this case is

$$S^{(1)}(\tau)|\epsilon_\pm^{(1)}(0)\rangle=|\epsilon_\pm^{(1)}(\tau)\rangle, \tag{4.7}$$

and gives

$$S^{(1)}(\tau)=\frac{1}{\sqrt{3b(1-2\rho+b)}}\times\begin{pmatrix}3+b-2\rho & -(1/\sqrt{2})(3-b+2\rho)\\(1/\sqrt{2})(3-b+2\rho) & 3+b-2\rho\end{pmatrix}. \tag{4.8}$$

The evolution operator is now determined by (3.13), a result which corresponds to the ordinary adiabatic theorem, valid in order $1/T$.

To obtain U in second order, we need $\alpha_{ij}^{(1)}$ and $\omega_{ij}^{(1)}$. From (3.16), (4.5), and (4.6) we get

$$\alpha_{+-}^{(1)}=-2\sqrt{2}\,\frac{[(d/d\tau)\rho]}{b^2}=-\alpha_{-+}^{(1)},\qquad \alpha_{++}^{(1)}=\alpha_{--}^{(1)}=0, \tag{4.9}$$

$$\omega_{+-}^{(1)}=\tfrac{1}{2}A\hbar b=-\omega_{-+}^{(1)}, \tag{4.10}$$

from which

$$C_{+-}^{(1)}=\frac{-4\sqrt{2}[(d/d\tau)\rho]}{b^3A\hbar}=C_{-+}^{(1)}\equiv C. \tag{4.11}$$

With (4.11), Eq. (3.15) gives

$$|\epsilon_+^{(2)}(\tau)\rangle=\frac{1}{\sqrt{2b(1-2\rho+b)}}\begin{pmatrix}(1-2\rho+b)+(i/T)2\sqrt{2}c\\2\sqrt{2}-(i/T)(1-2\rho+b)c\end{pmatrix}\equiv\begin{pmatrix}a_1(\tau)\\a_2(\tau)\end{pmatrix}, \tag{4.12a}$$

$$|\epsilon_-^{(2)}(\tau)\rangle=\frac{1}{\sqrt{2b(1-2\rho+b)}}\begin{pmatrix}-2\sqrt{2}-(i/T)(1-2\rho+b)c\\(1-2\rho+b)-(i/T)2\sqrt{2}c\end{pmatrix}\equiv\begin{pmatrix}b_1(\tau)\\b_2(\tau)\end{pmatrix}. \tag{4.12b}$$

Finally, (3.17) defines $S^{(2)}(\tau)$, which in terms of the variables introduced in (4.12) is given by

$$S^{(2)}(\tau)=\frac{1}{a_1(0)b_2(0)-a_2(0)b_1(0)}\times\begin{pmatrix}a_1(\tau)b_2(0)-a_2(0)b_1(\tau) & a_1(0)b_1(\tau)-a_1(\tau)b_1(0)\\a_2(\tau)b_2(0)-a_2(0)b_2(\tau) & a_1(0)b_2(\tau)-a_2(\tau)b_1(0)\end{pmatrix}, \tag{4.13}$$

and the eigenvalues $\epsilon_{\pm}^{(2)}(\tau)$ are obtained from (3.19), (4.9), and (4.10) with the result

$$\epsilon_{\pm}^{(2)}(\tau) = \frac{A\hbar^2}{4}(-1-2\rho \pm b) \pm \frac{\hbar}{T^2}\frac{16[(d/d\tau)\rho]}{A\hbar b^3}. \tag{4.14}$$

Equations (4.13) and (4.14) determine completely the evolution operator $U^{(2)}$ by (3.20),

$$U^{(2)}(\tau) = S^{(2)}(\tau)\{\exp[-iT\varphi_{+}^{(2)}(\tau)/\hbar]\,|\epsilon_{+}^{(2)}(0)\rangle\langle\epsilon_{+}^{(2)}(0)| + \exp[-iT\varphi_{-}^{(2)}(\tau)/\hbar]\,|\epsilon_{-}^{(2)}(0)\rangle\langle\epsilon_{-}^{(2)}(0)|\},$$
$$\varphi_{\pm}^{(2)}(\tau) = \int_0^{\tau} d\tau'\,\epsilon_{\pm}^{(2)}(\tau'). \tag{4.15}$$

The procedure can be continued to higher orders with increasing computational difficulties.

5. CONCLUSIONS

As stated previously, the advantage of the spinning axis representation proposed here lies in the fact that it allows one to calculate the evolution operator in the approximation $1/T^n$ performing only one transformation $S^{(n)}$ on the original system. In the second approximation only $S^{(2)}$ is necessary; it is simple to see that the combined action of the $R^{(2)}$, $R^{(1)}$ of Ref. 4 is equivalent to the action of $S^{(2)}$, etc.

It should be remarked that although in the previous example we have calculated the explicit form of $S^{(1)}$ for the sake of completeness, it is not necessary to know it to obtain $S^{(2)}$ and that, in general, $S^{(n)}$ can be obtained directly without going through the succesive changes of representation of Ref. 4. This is the advantage of the method, since to calculate $S^{(n)}$ it is only necessary to apply straightforward perturbation theory, the perturbation being "finer" at each succesive step.

We finally mention that, taking into account that the role of the SAR is played in classical mechanics by a canonical transformation, the method of this paper could be extended to classical systems in a form similar to the one used in a previous work.[5]

[1] A. Messiah, *Mécanique quantique* (Dunod, Paris, 1964), Vol. II, Chap. XVII.
[2] K. O. Friedrichs, Reports IMM-NYU 281, New York University (1955). IMM-NYU 230, New York University (1956).
[3] T. Kato, J. Phys. Soc. Jap. **5**, 435 (1950).
[4] L. Garrido, J. Math. Phys. **5**, 355 (1964).
[5] L. Navarro and L. Garrido, J. Phys. A (Proc. Phys. Soc.) **1**, 326 (1968).

An exact solution of the spin–spin autocorrelation function for a one-dimensional system of hard rods

G. Subramanian*, D. Levitt*, and H. T. Davis†

Departments of Physiology and Chemical Engineering,† University of Minnesota, Minneapolis, Minnesota 55455*
(Received 31 May 1974)

It is shown exactly that, in a one-dimensional system of hard rods with spins, the autocorrelation function of any function of spin $F(w)$ decays as t^{-1} at long times provided that $\langle F\rangle_{eq}$ exists and that $g(0) \neq 0$, where $g(v)$ is the linear velocity distribution function. As a consequence of this, when $F(w) = w$, the spin diffusion coefficient defined by the Kubo relation $D_s = \int_0^\infty \langle w(0)w(t)\rangle \times dt$ does not exist. The results are true for arbitrary initial equilibrium velocity and spin distributions, the only restriction being that they be symmetric.

The framework for the exact dynamic analysis of a one-dimensional system of hard rods was laid by Jepsen,[1] who examined Poincaré cycles, pseudostochastic behavior and nonequilibrium properties of the system. Levitt[2] and Percus[3] have generalized the system to include independent stochastic forces acting on each particle and have shown that under certain circumstances non-Fickian diffusion behavior is possible.

Theoretical attempts[4–6] to explain the long time behavior as t^{-1} of velocity autocorrelation functions in two dimensions observed in the computer experiments of Alder[7] have so far used either hydrodynamical models or kinetic equations with approximations on the time scales involved in the loss of correlations. In this paper, by examining a one-dimensional hard rod system with spins using Jepsen's technique, we show exactly that the spin–spin correlation $\langle w(0)\, w(t)\rangle$ goes as t^{-1} at long times and consequently the spin–spin diffusion coefficient defined by the Kubo relation,

$$D_s = \int_0^\infty \langle w(0)\, w(t)\rangle\, dt, \tag{1}$$

does not exist. Previous exact derivations[2,3] on one-dimensional hard rod systems with random background show a different type of anomalous behavior of the diffusion coefficient. Here the ensemble average $\langle x^2(t)\rangle$, where x is the displacement, was shown to be proportional to $\sqrt{t}$ rather than t as in canonical diffusive behavior and the diffusion coefficient would be zero in such a system. In the following, references to Jepsen's paper will be denoted by (J. _).

We consider N point particles numbered $0, 1, 2, \ldots N-1$, all of the same mass, impenetrable so that they are constrained to move along a line, length L, but with periodic boundary conditions. Though the analysis is done for point particles, the finite size of the particles is taken into account easily by substituting for the point density the expression $\rho/(1-\rho a)$, where ρ is the rod density and a is the rod length.

When a pair of particles collide, their kinetic energy, rotational kinetic energy, and linear and angular momenta are conserved so that they merely exchange linear and angular velocities (spins). There is no interchange of energy from one mode to the other.

The dynamics of the system is clear from Fig. 1, where the position of the particles is plotted against time. If no collision occurs, the motion of one of the particles is represented by a trajectory starting on the X-axis at the initial position of the particle with a slope equal to the reciprocal of the velocity. When a collision occurs, two of these lines cross, and since the collisions are perfectly elastic, the two particles merely exchange trajectories (i.e., velocities) and spins. The trajectories are labeled by the number of the particle occupying it at the initial time. With each trajectory one can also associate the initial value of the spin of the particle occupying it at the initial time. Therefore, at any time t, if we know which particle is on a given trajectory, both the velocity and spin of that particle are determined.

Define a characteristic function $A_{jk}(t)$ such that

$A_{jk}(t) = 1$ if particle j is on trajectory k at time t

$= 0$ otherwise.

The knowledge of A_{jk} at all times determines the complete dynamics. The average over the initial conditions, $\langle A_{jk}(t)\rangle$, gives the probability of finding particle j on trajectory k, at time t, when an ensemble of systems is considered. We assume that the initial distributions of velocity and spin, $g(v)$ and $G(w)$ that each particle can have in an ensemble are independent and satisfy the symmetry requirements

$$\int_{-\infty}^{\infty} v\, g(v)\, dv = 0, \tag{2}$$

$$\int_{-\infty}^{\infty} w\, G(w)\, dw = 0. \tag{3}$$

Since there is no interconversion of translational and rotational kinetic energy, the spins can be considered

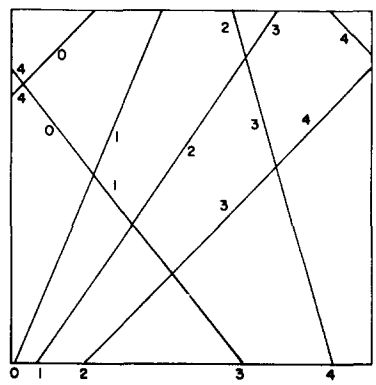


FIG. 1. Diagram showing the nature of the dynamics of the system.

to ride piggyback fashion on the particle and they do not influence the course of the trajectories. The function $A_{jk}(t)$ is therefore exactly the same as that given by Jepsen, i.e.,

$$A_{jk}(t)=\frac{1}{N}\sum_{l=0}^{N-1}\exp(-2\pi ijl/N)\prod_{h=0}^{N-1}S[u,w_{kh}], \tag{4}$$

$$S[u,w_{kh}]=\exp(inu) \text{ when } (n-1)L<w_{kh}\leq nL \text{ for each } n,$$

where

$$w_{kh}=x_k-x_h+(v_k-v_h)t.$$

x_k, x_h are initial positions of particles k and h, and

$$u\equiv 2\pi l/N.$$

ENSEMBLE CALCULATIONS

We consider properties which can be written in the form

$$E=\langle f[v_0(0),w_0(0);\ x_j(t),v_j(t),w_j(t)]\rangle. \tag{5}$$

The particle zero starts at the origin at zero time. The ensemble average (5) can be written as

$$E=\sum_k\langle f(v_0,w_0;\ x_k+v_kt,v_k,w_k)A_{jk}\rangle. \tag{6}$$

The form of f in Eq. (6) is useful in evaluating properties that are dependent on the time correlations of single particles like the diffusion coefficient. As particle j keeps changing trajectories, it "carries" with it the two properties of the different trajectories, the velocity v_k and spin w_k, and the summation in Eq. (6) over k takes this into account. The ensemble average in Eq. (6) is taken over positions, velocities, and spins of all the particles.

Since the spins and velocities are independent, we can make use of the same approximations for terms independent of w as made by Jepsen (J. 50, J. 51) when N and L are large and also the transformations and identities (J. 47—J. 49). Introducing Eq. (4) for A_{jk} in (6), exexpanding the ensemble average in Eq. (6), and proceeding to the limit $N\to\infty$, $L\to\infty$ with $N/L=\rho$, we get after converting summations to integrals,

$$\begin{aligned}E=\frac{\rho}{2\pi}\int_0^{2\pi}du\exp(-iju)\left(\int_0^\infty dx_k+\exp(-iu)\int_{-\infty}^0 dx_k\right)\\ \times\int_{-\infty}^\infty g(v_k)\,dv_k\int_{-\infty}^\infty G(w_k)dw_kQ(u,x_k+v_kt,v_k,w_k)\\ \times\exp\{\rho(1-e^{iu})T[u,x_k+v_kt]\}+\frac{1}{2\pi}\int_0^{2\pi}du\exp(-iju)\\ \times\int_{-\infty}^\infty g(v_0)\,dv_0\int_{-\infty}^\infty G(w_0)dw_0f(v_0,w_0,v_0t,w_0)\\ \times\exp\{\rho(1-e^{-iu})T(u,v_0t)\}.\end{aligned} \tag{7}$$

where

$$T(u,z)\equiv\int_{-\infty}^\infty(z-v_ht)S[u,z-v_ht]g(v_h)dv_h,$$

$$\begin{aligned}Q(u,x_k+v_kt,v_k,w_k)\\ =\int_{-\infty}^\infty g(v_0)dv_0\int_{-\infty}^\infty G(w_0)dw_0F(v_0,w_0,x_k+v_kt,v_k,w_k)\\ \times S[u,x_k+v_kt-v_0t]dv_0.\end{aligned} \tag{8}$$

For the detailed algebra, the reader is referred to Ref. 8. Equation (7) reduces to Jepsen's result (J.52) in the absence of spins and provides the general expression for the ensemble average Eq. (5) of f for the system of one-dimensional rods in which translation and spin do not interact.

Consider a function $F(w)$, only of spin w, such that the equilibrium thermal average $\langle F(w)\rangle$ vanishes. Any arbitrary function of spin can be cast in this form by subtracting from it its equilibrium thermal average. Then we let

$$f[v_0,w_0,x_j(t),v_j(t),w_j(t)]=F[w_j(0)]F[w_j(t)]. \tag{9}$$

Equation (6) becomes

$$E=\sum_k\langle F(w_k)F(w_k)A_{jk}\rangle \tag{10}$$

and

$$\begin{aligned}Q(u,x_k+v_kt,v_k,w_k)\\ =\int_{-\infty}^\infty dv_0\,g(v_0)\int_{-\infty}^\infty dw_0\,F(w_0)F(w_k)G(w_0)S[u,x_k\\ +v_kt-v_0t]=0\end{aligned} \tag{11}$$

by virtue of $\langle F(w_0)\rangle=\int_{-\infty}^\infty dw_0\,F(w_0)\,G(w_0)=0$. The first term of E in (7) then vanishes. Since $f(v_0,w_0,v_0t,w_0)=F(w_0)^2$,

$$\begin{aligned}E=\frac{1}{2\pi}\int_0^{2\pi}du\exp(-iju)\int_{-\infty}^\infty dv_0\,g(v_0)\int_{-\infty}^\infty dw_0\,G(w_0)F(w_0)^2\\ \times\exp\{\rho(1-e^{-iu})T(u,v_0t)\}.\end{aligned} \tag{12}$$

Using the simplified expression for T of Jepsen (J. 56), we have

$$T(u,z)=-A(z)+(z+A)\,e^{iu}. \tag{13}$$

By defining $A(z)$ by

$$A(z)=\int_{z/t}^\infty(-z+v_ht)g(v_h)dv_h \tag{14}$$

and using the identity (J. 57) with $m=j$ and setting $j=0$, the autocorrelation function of F becomes

$$\begin{aligned}E=\langle F[w_0(0)]F[w_0(t)]\rangle\\ =C\int_{-\infty}^\infty dv_0\,g(v_0)I_0[2\rho\{A(v_0t)[v_0t+A(v_0t)]\}^{1/2}]\\ \times\exp[-\rho v_0t-2A(v_0t)\rho],\end{aligned} \tag{15}$$

where

$$C\equiv\int_{-\infty}^\infty F(w_0)^2G(w_0)dw_0\equiv\langle F(w_0)^2\rangle. \tag{16}$$

I_0 is the zeroth order Bessel function of imaginary argument. To examine the asymptotic behavior of (15) in the limit of large time, we use the fact that for large arguments

$$I_0(x)\to e^x/\sqrt{2\pi x}$$

i.e.,

$$\begin{aligned}I_0[2\rho\{AB\}^{1/2}]\exp[-A+B)\rho]\\ \approx\exp[-(\sqrt{A}-\sqrt{B})^2\rho]/[4\pi\rho\{AB\}^{1/2}]^{1/2},\end{aligned} \tag{17}$$

where $B\equiv v_0t+A(v_0t)$. A is of the form $A't$, B of the form $B't$, where A', B' are independent of t:

$$\exp[-(\sqrt{A}-\sqrt{B})^2\rho]\approx\exp[-(\sqrt{A'}-\sqrt{B'})^2\rho t]. \tag{18}$$

For large times Eq. (17) is zero except for $\sqrt{A'}\approx\sqrt{B'}$,

i.e., for $v_0 \approx 0$ when

$$A \sim t \int_0^\infty v_h g(v_h) dv_h. \tag{19}$$

Defining

$$D \equiv \frac{1}{\rho} \int_0^\infty v_h g(v_h) dv_h, \tag{20}$$

$$A' \sim \rho D, \quad B' \sim v_0 + A', \tag{21}$$

and

$$\exp[-(\sqrt{A'} - \sqrt{B'})^2 t] \sim \exp(-V_0^2 t/4D). \tag{22}$$

With relation (22), Eq. (15) becomes at long times

$$E \sim \frac{C}{\rho} \int_{-\infty}^{\infty} dv_0 g(v_0) \frac{\exp(-v_0^2 t/4D)}{\sqrt{4\pi D t}} . \tag{23}$$

At long times, the integrand in Eq. (23) is negligible except for $v_0 \approx 0$, and Eq. (23) reduces to

$$E \sim [C g(0)/\rho] t^{-1} \tag{24}$$

as long as $g(0) \neq 0$. Letting $F(w) = w$, the spin–spin autocorrelation $\langle w_0(0) w_0(t) \rangle$ decays at long times as t^{-1} and the spin diffusion coefficient defined by

$$D_s = \int_0^\infty \langle w(0) w_0(t) \rangle \, dt$$

will not exist, as long as $g(0) \neq 0$.

Equation (24) will now be verified by the following argument which uses directly the long time behavior of the function $A_{00}(t)$, the probability that a particle returns to the trajectory it started out on. The function $A_{00}(t)$ is given by Levitt, [9] and at long times it reduces to

$$A_{00}(t) = \frac{\exp(-v_0^2 t/4D)}{\rho \sqrt{4\pi D t}} . \tag{25}$$

Let us focus attention on a particular particle, say particle 0. The probability that particle 0 has spin w at time t given that it started out with spin w' at time 0 is given by

$$p(w, t; w', 0) = \sum_i \langle \delta(w_i(t) - w) \delta(w_0(0) - w') \rangle \tag{26}$$

where the summation is over all trajectories i,

$$p(w, t; w', 0) = \sum_i \langle \delta(w_i - w) \delta(w_0 - w') A_{0i}(t) \rangle \tag{27}$$

with w_i and w_0 being the initial values of the spins associated with the particles occupying the trajectories i and 0 at initial time and $A_{0i}(t)$ the probability that particle 0 is on trajectory i at time t given that it started out on trajectory 0. In Eq. (27) after averaging over w_i and w_0 the summation is split as

$$\begin{aligned} p(w, t; w', 0) &= \delta(w - w') G(w) \langle A_{00}(t) \rangle_v + G(w) G(w') \sum_{i \neq 0} \langle A_{i0}(t) \rangle_v \\ &= \delta(w - w') G(w) \langle A_{00}(t) \rangle_v + G(w) G(w') \langle 1 - A_{00}(t) \rangle_v , \end{aligned}$$

where $\langle A_{00}(t) \rangle_v = \int A_{00}(v, t) g(v) \, dv$.

The autocorrelation function of any function of spin, $F(w)$ can be written as

$$\langle F(w(t)) F(w(0)) \rangle = \langle F(w') p(w, t; w', 0) F(w') \rangle_{w, w'} , \tag{29}$$

and by using Eq. (28) this reduces to

$$\langle F(w(t)) F(w(0)) \rangle = [\langle F(w)^2 \rangle - \langle F(w) \rangle^2] \langle A_{00}(t) \rangle_v + \langle F(w) \rangle^2 . \tag{30}$$

As explained earlier, $F(w)$ can always be reduced to a form such that $\langle F(w) \rangle = 0$, and Eq. (30) becomes

$$\langle F(w(t)) F(w(0)) \rangle = \langle F(w) \rangle^2 \langle A_{00}(t) \rangle_v = \langle F(w)^2 A_{00}(t) \rangle . \tag{31}$$

The result (31) can be explained as follows. Since the spin distributions of the individual particles are independent of each other as well as the velocity distributions, the spin of the particle at the initial time or any function of spin $F(w)$ is totally uncorrelated at all subsequent times, other than times at which the particle returns to the trajectory it started out on, when there is total correlation. By using result (25), Eq. (31) becomes

$$\langle F(0) F(t) \rangle = \iint F(w_0)^2 \frac{\exp(-v_0^2 t/4D)}{\sqrt{4\pi D t}} \frac{1}{\rho} G(w_0) g(v_0) \, dw_0 \, dv_0 . \tag{32}$$

We note that the long time behavior of $\langle F(0) F(t) \rangle$ as $1/t$ is the same as that of $\langle A_{00}(t) \rangle_v$. The ensemble average in (32) is taken over all possible initial velocities and spins. This is seen to be identical with Eq. (23). Integration over w_0 yields the constant C and the relation (24) is valid as before. By letting $F(w_0) = w_0$ the spin diffusion coefficient defined in (1) will not exist as long as $g(0) \neq 0$.

However, if $g(0) = 0$, $\langle A_{00}(t) \rangle$ will not decay as $1/t$ at long times, and consequently a spin diffusion coefficient can exist. In that case,

$$D_s = \int_0^\infty \langle F(w(t) F(w(0)) \rangle \, dt = \langle F(w)^2 \rangle \int_0^\infty \langle A_{00}(t) \rangle_v \, dt, \tag{33}$$

by using Eq. (31). By using the expression for $A_{00}(t)$ from Levitt[9]

$$D_s = \langle F(w)^2 \rangle \langle 1/\rho \, | v | \rangle, \tag{34}$$

it is seen that if $g(0) \neq 0$, then $\langle 1/\rho \, | v | \rangle$ and D_s will not exist.

The probability given by (25) is the same as (J. 62),

$$p(y) = (4\pi D t)^{-1/2} \exp(-y^2/4Dt), \tag{35}$$

with $y = v_0 t$, which is the probability of finding a particle at position y at time t.

The physical argument given above cannot be used to derive the long time behavior of the linear velocity correlation function or correlation of any function of the linear velocity because the velocity of a particle is correlated at all times after it leaves its initial trajectory, whether it returns to the initial trajectory or not. This is because, unlike in the spin case, the subsequent velocities depend on the orientation of the initial trajectory.

Thus it has been demonstrated exactly for this one-dimensional system that the rotational diffusion constant D_s does not exist. We have also shown more generally that at long times the autocorrelation of any function of spin alone will decay as t^{-1}. This means that for example the autocorrelation function of the rotational kinetic energy Iw^2 will decay as t^{-1} at long times.

[1]D.W. Jepsen, J. Math. Phys. **6**, 405 (1965).
[2]D.G. Levitt, Phys. Rev. A **8**, 3050 (1973).
[3]J.K. Percus, Phys. Rev. A **9**, 557 (1974).
[4]M.H. Ernst, E.H. Hauge, and J.M.J. van Leeuwen, Phys. Lett. **25**, 1254 (1970).
[5]J.R. Dorfman and E.G.D. Cohen, Phys. Rev. Lett. **25**, 1257 (1970).
[6]R. Zwanzig, "Nonlinear Dynamics of Collective Modes," presented at *IUPAP Conf. Stat. Mech.* (University of Chicago Press, Chicago, 1971).
[7]B.J. Alder and T.E. Wainwright, Phys. Rev. A **1**, 18 (1970).
[8]G. Subramanian, Ph.D. dissertation (unpublished), University of Minnesota (1973).
[9]D.G. Levitt, J. Stat. Phys. **7**, 329 (1973).

Ray and explosive solutions of nonlinear evolutional equations in Hilbert spaces

P. K. C. Wang

School of Engineering and Applied Science, University of California, Los Angeles, California
(Received 1 August 1974)

Many nonlinear phenomena occurring in physical systems are describable by a set of ordinary, partial or functional differential equations which can be regarded as an evolutional equation in a suitable abstract vector space. In this paper, we consider nonlinear evolutional equations defined on Hilbert spaces. Attention is focused on developing conditions for the existence of solutions which lie along half-rays emanating from the origin of the space. The results are used to establish sufficient conditions for the existence or nonexistence of explosive solutions or solutions having finite escape time. The paper concludes with a discussion of the application of some of the results to specific classes of evolutional equations arising from physical situations.

I. INTRODUCTION

Many nonlinear phenomena in continuum systems are describable by a set of nonlinear differential equations of the form

$$\frac{du_j}{dt} = a_j(u_1, \ldots, u_n) + f_j(u_1, \ldots, u_n), \quad j = 1, \ldots, n, \tag{1}$$

where a_j and f_j are linear and nonlinear operators, respectively. For example, the macroscopic plasma models[1] consisting of the Maxwell's equations and the hydromagnetic equations can be cast into the form of (1). Also, the hydrodynamic equations describing the nonlinear interaction of gravity waves in an ideal fluid of finite depth have a similar form.[2] In these cases, (1) corresponds to a set of partial differential equations which can be regarded as an evolutional equation in a suitable function space. In the case of a dynamical system having finite degrees of freedom, (1) corresponds to a set of ordinary differential equations or an evolutional equation in a finite-dimensional vector space. It is known that even the solutions to the linearized equations are well behaved, the presence of the nonlinear terms could induce explosive instabilities or the finite-escape-time phenomenon in which some of the solutions become unbounded over a finite time interval.[3,4] Recently, this phenomenon occurring in the nonlinear interaction of waves in plasmas and dielectric media has been investigated both theoretically and experimentally.[5-10] In this paper, we formulate (1) as an evolutional equation in a Hilbert space. Attention is focused on developing conditions for the existence of ray solutions for a certain class of equations. The results are used to establish sufficient conditions for the existence or nonexistence of explosive solutions. This paper concludes with a discussion of specific classes of evolutional equations arising from physical situations.

II. PRELIMINARIES

Let H^n denote the n-fold Cartesian product of a Hilbert space H with inner product $(\cdot, \cdot)_H$. For any $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{u}' = (u_1', \ldots, u_n')$ in H^n, their inner product is defined by

$$(\mathbf{u}, \mathbf{u}') = \sum_{j=1}^{n} (u_j, u_j')_H.$$

In particular, H may be taken to be the Hilbert space $L_2(\Omega)$ [resp. $L_2(\Omega)$] of complex-valued [resp. real-valued] functions defined on an open subset Ω of the Euclidean m-space R^m. For $H = L_2(\Omega)$, the inner product on H^n is defined by

$$(\mathbf{u}, \mathbf{u}') = \sum_{j=1}^{n} \int_\Omega u_j(x)\, \overline{u_j'(x)}\, dx,$$

where $\mathbf{x} = (x_1, \ldots, x_m)$ denotes a point in Ω; $\overline{u_j'}$ is the complex conjugate of u_j' and $d\mathbf{x}$ is the element of Lebesgue m-measure on Ω. Also, we shall consider the case where H^n is an n-dimensional real or complex Euclidean space R^n or C^n, respectively, with the usual inner product. The open [resp. closed] ball with radius r centered about the origin $\boldsymbol{\theta}$ in H^n will be denoted by Σ_r [resp. $\overline{\Sigma}_r$], and the sphere $\{u \in H^n: \|\mathbf{u}\| = r\}$ by $\partial\Sigma_r$.

Let $I = [0, T[$, $T \leq \infty$ and $C_1(I; H^n)$ denote the space of all continuously differentiable H^n-valued functions on I. We formulate (1) as a nonlinear evolutional equation in H^n:

$$\frac{d\mathbf{u}}{dt} = \mathbf{A}\mathbf{u} + \mathbf{f}(\mathbf{u}) = \mathbf{h}(\mathbf{u}) \tag{2}$$

with initial data $\mathbf{u}(0) = \mathbf{u}_0 \in D_h$—the domain of $\mathbf{h}$. $\mathbf{A}$ is a linear operator with domain $D_\mathbf{A} = \{\mathbf{u} \in H^n: \mathbf{A}\mathbf{u} \in H^n\}$ and $\mathbf{f}$ is a nonlinear operator with domain $D_\mathbf{f} = \{\mathbf{u} \in H^n: \mathbf{f}(\mathbf{u}) \in H^n\}$. In cases where (2) corresponds to a set of partial differential equations, we shall incorporate any boundary conditions in the definition of $D_\mathbf{A}$ and $D_\mathbf{f}$. By a solution of the initial-value problem (2), we mean a function $\mathbf{u} \in C_1(I; H^n)$ such that $\mathbf{u}(t)$ satisfies (2) for each fixed $t \in]0, T[$ and initial data $\mathbf{u}_0$ at $t = 0$. We shall call a nonequilibrium solution $\mathbf{u}(t)$, defined for $t \in I$, a *ray solution*, if $\mathbf{u}(t)$ lies on a half-ray $\{\mathbf{u} \in H^n: \mathbf{u} = \lambda\mathbf{v}, \mathbf{v} \neq \theta, \lambda \geq 0\}$ for each $t \in I$. Evidently, such a solution can be written in the form $\mathbf{u}(t) = \eta(t)\,\mathbf{v}$, where $\mathbf{v}$ is a nonzero vector in H^n and η is a real-valued C_1 function of t such that

$$0 \leq \inf\{\eta(t); t \in I\} < \sup\{\eta(t); t \in I\} \leq \infty. \tag{3}$$

For the subsequent development, we introduce the following assumptions:

(A-1) $D_\mathbf{h} = D_\mathbf{f} \cap D_\mathbf{A}$ is a linear subspace of H^n.

(A-2) The operator $\mathbf{f}$ has the homogeneity property: $\mathbf{f}(c\mathbf{u}) = c^k\mathbf{f}(\mathbf{u})$ for some real number $k > 0$, all real numbers c and all $\mathbf{u} \in D_\mathbf{f}$.

These assumptions imply that $\mathbf{f}(\boldsymbol{\theta}) = \boldsymbol{\theta}$ and $\mathbf{f}$ is an even

[resp. odd] operator {i.e, $\mathbf{f}(-\mathbf{u})=\mathbf{f}(\mathbf{u})$ [resp. $\mathbf{f}(-\mathbf{u})=-\mathbf{f}(\mathbf{u})$] for all $\mathbf{u}\in D_{\mathbf{f}}$} if k is an even [resp. odd] positive integer.

A few simple examples of $\mathbf{f}$ satisfying (A-2) are given below.

Example 1: f is a mapping from $D_f\subset L_2(\Omega)$ into $L_2(\Omega)$ defined by

$$f(u)=\Delta(g(u)), \tag{4}$$

where Δ is the Laplacian operator and g is a homogeneous polynomial of u with degree $k>0$.

Example 2: Ω is a bounded open subset of R^m and f is a Hammerstein integral operator on $D_f\subseteq \mathcal{L}_2(\Omega)$ into $\mathcal{L}_2(\Omega)$ given by

$$(f(u))(\mathbf{x})=\int_\Omega K(\mathbf{x},\mathbf{x}')\,g(u(\mathbf{x}'))\,d\mathbf{x}', \tag{5}$$

where g is as in Example 1 and the kernel $K\in \mathcal{L}_2(\Omega\times\Omega)$.

III. RAY SOLUTIONS

Now, we shall establish conditions for the existence of ray solutions of (2) defined in a complex Hilbert space H^n. Let $r(t)=\|\mathbf{u}(t)\|=(\mathbf{u}(t),\mathbf{u}(t))^{1/2}$. We introduce the transformation $\mathbf{v}=\mathbf{u}/r$ which projects the solutions of (2) onto the unit sphere $\partial\Sigma_1$.

Proposition: Let $\mathbf{u}(t)$ be a nontrivial solution of (2). Then, $r(t)$ and $\mathbf{v}(t)$ satisfy

$$\frac{dr}{dt}=r\{\mathrm{Re}(\mathbf{Av},\mathbf{v})+r^{k-1}\,\mathrm{Re}(\mathbf{f(v)},\mathbf{v})\}, \tag{6}$$

$$\frac{d\mathbf{v}}{dt}=\{\mathbf{Av}-\mathbf{v}\,\mathrm{Re}(\mathbf{Av},\mathbf{v})\}+r^{k-1}\{\mathbf{f(v)}-\mathbf{v}\,\mathrm{Re}(\mathbf{f(v)},\mathbf{v})\}. \tag{7}$$

Proof: By direct computation

$$\begin{aligned} r\frac{dr}{dt}&=\frac{1}{2}\frac{d}{dt}(\mathbf{u},\mathbf{u})=\mathrm{Re}(\mathbf{Au},\mathbf{u})+\mathrm{Re}(\mathbf{f(u)},\mathbf{u})\\ &=r^2\,\mathrm{Re}(\mathbf{Av},\mathbf{v})+\mathrm{Re}(\mathbf{f}(r\mathbf{v}),r\mathbf{v}). \end{aligned} \tag{8}$$

Equation (6) follows from (8) and assumption (A-2). Using (2) and (6), we have

$$\begin{aligned} \frac{d\mathbf{v}}{dt}&=r^{-1}\frac{d\mathbf{u}}{dt}-\mathbf{u}r^{-2}\frac{dr}{dt}\\ &=r^{-1}(\mathbf{Au}+\mathbf{f(u)})-\mathbf{u}r^{-2}\{\mathrm{Re}(\mathbf{Au},\mathbf{u})+\mathrm{Re}(\mathbf{f(u)},\mathbf{u})\}\\ &=\mathbf{Av}+r^{-1}\mathbf{f}(r\mathbf{v})-r^{-1}\mathbf{v}\{r^2\,\mathrm{Re}(\mathbf{Av},\mathbf{v})+\mathrm{Re}(\mathbf{f}(r\mathbf{v}),r\mathbf{v})\}. \end{aligned} \tag{9}$$

Equation (7) follows directly from (9) and assumption (A-2). ||

Theorem 1: Let $\alpha(\mathbf{v})=(\mathbf{Av},\mathbf{v})$, $\beta(\mathbf{v})=(\mathbf{f(v)},\mathbf{v})$ and $r(0)=\|\mathbf{u}_0\|>0$. If $\mathbf{A}$ and $\mathbf{f}$ have a common invariant direction $\mathbf{v}_*\in D_{\mathbf{h}}\cap\partial\Sigma_1$, then (2) has a ray solution $\mathbf{u}(t)=r(t)\,\mathbf{v}_*$ defined on some interval $I\subseteq[0,\infty]$ with positive measure, where, for $k\neq 1$, $r(t)$ is given by

$$\begin{aligned} r(t)=r(0)&\times\exp(\alpha(\mathbf{v}_*)\,t)\{1+\alpha(\mathbf{v}_*)^{-1}\,\beta(\mathbf{v}_*)\,r(0)^{k-1}\\ &\times(1-\exp[\alpha(\mathbf{v}_*)(k-1)\,t])\}^{-1/(k-1)} \end{aligned} \tag{10}$$

if $\alpha(\mathbf{v}_*)\neq 0$, or

$$r(t)=r(0)\{1-(k-1)\,\beta(\mathbf{v}_*)\,r(0)^{k-1}\,t\}^{-1/(k-1)} \tag{11}$$

if $\alpha(\mathbf{v}_*)=0$, where we take $\{\cdots\}^{-1/(k-1)}$ to be the nonnegative real root when $k>1$. When $k=1$, $r(t)$ is given by

$$r(t)=r(0)\exp[(\alpha(\mathbf{v}_*)+\beta(\mathbf{v}_*)]\,t\},\quad t\geq 0. \tag{12}$$

Proof: Let $\mathbf{A}$ be a nonzero linear operator having in common with $\mathbf{f}$ an invariant direction $\mathbf{v}_*\in D_{\mathbf{h}}\cap\partial\Sigma_1$. By definition, there exist nonzero real numbers λ and μ such that $\mathbf{Av}_*=\lambda\mathbf{v}_*$ and $\mathbf{f}(\mathbf{v}_*)=\mu\mathbf{v}_*$. It follows that $(\mathbf{Av}_*,\mathbf{v}_*)=\lambda\|\mathbf{v}_*\|^2$ and $(\mathbf{f}(\mathbf{v}_*),\mathbf{v}_*)=\mu\|\mathbf{v}_*\|^2$. Since $\mathbf{v}_*$ lies on the unit sphere $\partial\Sigma_1$, $\alpha(\mathbf{v}_*)=(\mathbf{Av}_*,\mathbf{v}_*)$ and $\beta(\mathbf{v}_*)=(\mathbf{f}(\mathbf{v}_*),\mathbf{v}_*)$ are real and equal to λ and μ, respectively. From (7), if $\mathbf{v}(t_1)=\mathbf{v}_*$ for any $t_1\geq 0$, then $\mathbf{v}(t)=\mathbf{v}_*$ for all $t\geq t_1$. The corresponding solution $\mathbf{u}(t)=r(t)\,\mathbf{v}_*$, $t\geq t_1$, lies on the half-ray from the origin passing through the point $\mathbf{v}_*$, where $r(t)$ satisfies (6) or

$$\frac{dr}{dt}=\alpha(\mathbf{v}_*)\,r+\beta(\mathbf{v}_*)\,r^k. \tag{13}$$

Integrating (13) with initial condition $r(0)$ at $t=0$ leads directly to (10) or (12). For the case where A is the zero operator on H^n, every invariant direction of $\mathbf{f}$ on $D_{\mathbf{f}}\cap\partial\Sigma_1$ is an invariant direction of $\mathbf{A}$. Integrating (13) with $\alpha(\mathbf{v}_*)=0$ and initial condition $r(0)$ gives (11). The interval of definition I for (10) or (11) is $[0,\infty[$ if the $\{\cdots\}$ term does not vanish for all $t>0$, and $I\subset[0,\infty[$ if the $\{\cdots\}$ term vanishes for some finite $t=\tau>0$. ||

Corollary 1: If $\mathbf{u}_*\in D_{\mathbf{h}}$ is a nonzero fixed point of both $\mathbf{A}$ and $\mathbf{f}$, then $\mathbf{v}_*=\mathbf{u}_*/\|\mathbf{u}_*\|$ is a common invariant direction of $\mathbf{A}$ and $\mathbf{f}$ on $\partial\Sigma_1$, and $\mathbf{u}(t)=r(t)\,\mathbf{v}_*$ is a ray solution of (2), where $r(t)$ is given by (10) or (12) with $\alpha(\mathbf{v}_*)=1$ and $\beta(\mathbf{v}_*)=\|\mathbf{u}_*\|^{1-k}$.

Proof: Since $\|\mathbf{u}_*\|\neq 0$ and $\mathbf{Au}_*=\mathbf{u}_*=\mathbf{f}(\mathbf{u}_*)$, therefore $\mathbf{v}_*$ is defined and satisfies $\mathbf{Av}_*=\mathbf{v}_*$, $\mathbf{f}(\mathbf{v}_*)=\|\mathbf{u}_*\|^{-k}\mathbf{f}(\mathbf{u}_*)=\|\mathbf{u}_*\|^{1-k}\mathbf{v}_*$ and $\alpha(\mathbf{v}_*)=(\mathbf{Av}_*,\mathbf{v}_*)=1$, $\beta(\mathbf{v}_*)=(\mathbf{f}(\mathbf{v}_*),\mathbf{v}_*)=\|\mathbf{u}_*\|^{1-k}$. Thus $\mathbf{v}_*$ is an invariant direction of both $\mathbf{A}$ and $\mathbf{f}$ on the unit sphere $\partial\Sigma_1$. The desired result follows from Theorem 1. ||

It is evident from Theorem 1 that the existence of ray solutions of (2) can be established by showing the existence of common invariant directions of $\mathbf{A}$ and $\mathbf{f}$ on $D_{\mathbf{h}}\cap\partial\Sigma_1$. In general, a ray solution may not be associated with any common invariant direction of $\mathbf{A}$ and $\mathbf{f}$. This can be demonstrated by the following simple example of (2) defined on R^2 with $\mathbf{Au}=(u_2,0)$ and $\mathbf{f}(\mathbf{u})=(0,u_1)$ for all $\mathbf{u}=(u_1,u_2)\in R^2$. Its ray solutions are given by $\mathbf{u}(t)=c\mathbf{v}\exp(t)$ and $\mathbf{u}(t)=c'\mathbf{v}'\exp(-t)$, $t\geq 0$, where c, c' are positive numbers; $\mathbf{v}=\pm(1,1)$ and $\mathbf{v}'=\pm(1,-1)$ are eigenvectors of $(\mathbf{A}+\mathbf{f})$. But the invariant directions of $\mathbf{A}$ and $\mathbf{f}$ correspond to their eigenvectors which are nonzero scalar multiples of (0, 1) and (1, 0) respectively. Hence $\mathbf{A}$ and $\mathbf{f}$ have no common invariant directions.

When $\mathbf{A}$ is the zero or the identity map on H^n, then every ray solution of (2) is associated with an invariant direction of $\mathbf{f}$ provided that $\mathbf{f}(\mathbf{u})\neq\boldsymbol{\theta}$ for all $\mathbf{u}\neq\boldsymbol{\theta}$. This can be verified by considering a ray solution of (2) (with $\mathbf{A}$ being the identity map) of the form $\mathbf{u}(t)=\eta(t)\mathbf{v}$, $t\in I\subseteq[0,\infty[$, $\mathbf{v}\neq\boldsymbol{\theta}$ and η is a real C_1 function satisfying (3). Substituting $\mathbf{u}(t)$ into (2), we arrive at the identity

$$\left\{\left(\frac{d\eta(t)}{dt}-\eta(t)\right)\,\eta(t)^{-k}\right\}\mathbf{v}=\mathbf{f}(\mathbf{v})\quad\text{for all } t\in I.$$

Since $\mathbf{f}(\mathbf{v}) \neq \theta$ for all $\mathbf{v} \neq \theta$, hence the $\{\cdots\}$ term must equal to a nonzero real number for all $t \in I$ or $\mathbf{v}$ is an invariant direction of $\mathbf{f}$.

Finally, we note that if $\mathbf{u}(t) = r(t)\,\mathbf{v}_*$ is a ray solution of (2) such that $\mathbf{v}_* \in D_h \cap \partial\Sigma_1$ is an invariant direction of both $\mathbf{A}$ and $\mathbf{f}$ with $\alpha(\mathbf{v}_*) = 1$, then $\hat{\mathbf{v}} = \mathbf{v}_* \beta(\mathbf{v}_*)^{-1/(k-1)}$ is a fixed point of both $\mathbf{A}$ and $\mathbf{f}$, since

$$\mathbf{A}\hat{\mathbf{v}} = \beta(\mathbf{v}_*)^{-1/(k-1)}\,\mathbf{A}\mathbf{v}_* = \beta(\mathbf{v}_*)^{-1/(k-1)}\,(\mathbf{A}\mathbf{v}_*, \mathbf{v}_*)\,\mathbf{v}_* = \hat{\mathbf{v}},$$

$$\mathbf{f}(\hat{\mathbf{v}}) = \mathbf{f}(\beta(\mathbf{v}_*)^{-1/(k-1)}\,\mathbf{v}_*) = \beta(\mathbf{v}_*)^{-k/(k-1)}\,\mathbf{f}(\mathbf{v}_*)$$
$$= \beta(\mathbf{v}_*)^{-1/(k-1)}\,\mathbf{v}_* = \hat{\mathbf{v}},$$

where $\beta(\mathbf{v}_*) = (\mathbf{f}(\mathbf{v}_*), \mathbf{v}_*)$.

Now, we shall establish sufficient conditions for the existence of ray solutions of (2) in which f belongs to certain special classes of operators.

1. *Compact Operators*: We shall make use of the following result:

Lemma 1: Let $\mathbf{f}$ be a compact mapping of a Banach space V into itself such that $\|\mathbf{f}(\mathbf{v})\|_V \geq c > 0$ for all $\mathbf{v} \in \partial\Sigma_1 = \{\mathbf{v} \in V : \|\mathbf{v}\|_V = 1\}$. Then, $\mathbf{f}$ has an invariant direction on $\partial\Sigma_1$ or the equation $\mathbf{f}(\mathbf{v}) = \lambda\mathbf{v}$ has a solution $\mathbf{v}_* \in \partial\Sigma_1$ for some nonzero real number λ.

The above result is a generalization of that due to Birkhoff and Kellogg.[11] A detailed proof can be found in Ref. 12. We note that if $\|\mathbf{f}(\mathbf{v})\|_V \geq c > 0$ for all $\mathbf{v} \in \bar{\Sigma}_1 = \{\mathbf{v} \in V : \|\mathbf{v}\|_V \leq 1\}$, then the result follows trivially from Schauder's fixed point theorem. Let $\tilde{\mathbf{f}}(\mathbf{v}) = \mathbf{f}(\mathbf{v})/\|\mathbf{f}(\mathbf{v})\|_V$. Since $\mathbf{f}$ is compact and $\|\mathbf{f}(\mathbf{v})\|_V \geq c > 0$ for all $\mathbf{v} \in \bar{\Sigma}_1$, then $\tilde{\mathbf{f}}$ is a compact mapping of $\bar{\Sigma}_1$ into itself. By Schauder's fixed point theorem, there exists a $\mathbf{v}_* \in \bar{\Sigma}_1$ such that $\tilde{\mathbf{f}}(\mathbf{v}_*) = \mathbf{v}_*$. Since $\|\tilde{\mathbf{f}}(\mathbf{v}_*)\|_V = 1$, hence $\mathbf{v}_* \in \partial\Sigma_1$ and $\mathbf{f}(\mathbf{v}_*) = \|\mathbf{f}(\mathbf{v}_*)\|_V\,\mathbf{v}_*$ with $\|\mathbf{f}(\mathbf{v}_*)\|_V \geq c > 0$ implying that $\mathbf{v}_*$ is an invariant direction of $\mathbf{f}$. The foregoing proof is invalid for a homgeneous $\mathbf{f}$, since $\mathbf{f}(\theta) = \theta$. In Ref. 12, using the so-called Sweeping Theorem for Banach spaces, Lemma 1 is proved for a more general case where $\partial\Sigma_1$ is replaced by the boundary of any bounded subset of V containing the origin.

Theorem 2: Let $\mathbf{A}$ be the zero transformation on H^n and $\mathbf{f}$ is a compact mapping on H^n into itself satisfying assumption (A2) for all $\mathbf{u} \in H^n$. Moreover, $\|\mathbf{f}(\mathbf{u})\| \geq c > 0$ for all $\mathbf{u} \in \partial\Sigma_1$. Then (2) has at least one ray solution $\mathbf{u}(t) = r(t)\,\mathbf{v}_*$, $t \in I$, where $\mathbf{v}_* \in \partial\Sigma_1$ and $r(t)$ is given by (11) if $k \neq 1$, and by (12) if $k = 1$. Furthermore, if $\mathbf{f}$ is an odd operator, then (2) has at least one pair of ray solutions $\mathbf{u}(t) = \pm\, r(t)\,\mathbf{v}_*$, $t \in I$.

Proof: From Lemma 1, f has an invariant direction $\mathbf{v}_* \in \partial\Sigma_1$ such that $\mathbf{f}(\mathbf{v}_*) = \lambda\mathbf{v}_*$ with $|\lambda| = \|\mathbf{f}(\mathbf{v}_*)\| \geq c > 0$. Hence, $|\beta(\mathbf{v}_*)| = |(\mathbf{f}(\mathbf{v}_*), \mathbf{v}_*)| = \|\mathbf{f}(\mathbf{v}_*)\| > 0$. The existence of a ray solution follows directly from Theorem 1. When $\mathbf{f}$ is an odd operator, $-\mathbf{v}_*$ is also an invariant direction of $\mathbf{f}$. Consequently, a pair of ray solutions $\mathbf{u}(t) = \pm\, r(t)\,\mathbf{v}_*$ exist. ||

2. *Gradient Operators*: Here, we consider the case in which the rhs (right-hand side) of (2) consists of the term $\mathbf{f}(\mathbf{u})$ only, where $\mathbf{f}$ is the strong gradient of a functional F defined on a real Hilbert space H^n, i.e., the Fréchet differential $(\mathbf{f}(\mathbf{u}), \mathbf{u}') = (\mathrm{grad}\,F(\mathbf{u}), \mathbf{u}')$ exists for all $\mathbf{u}, \mathbf{u}' \in H^n$. Such an evolutional equation describes the smooth trajectories of the steepest ascent method for maximizing F over H^n.[13] We note that for $F(\theta) = 0$, $F(\mathbf{u})$ is expressible in the form[14]:

$$F(\mathbf{u}) = \int_0^1 (\mathbf{u}, \mathbf{f}(s\mathbf{u}))\,ds, \tag{14}$$

and if $\mathbf{f}$ satisfies assumption (A-2), (14) reduces to

$$F(\mathbf{u}) = (\mathbf{u}, \mathbf{f}(\mathbf{u}))/(k+1), \quad k > 0. \tag{15}$$

In what follows, we shall establish a sufficient condition for the existence of ray solutions.

First, consider the problem of extremizing $F(\mathbf{u})$ on the unit sphere $\partial\Sigma_1$ in H^n or subject to the constraint $\phi(\mathbf{u}) = (\mathbf{u}, \mathbf{u}) = 1$. Applying the abstract Lagrange multiplier rule (Ref. 14, p. 96), if $\mathbf{u}_*$ is an extremum point of F with respect to $\partial\Sigma_1$ and F is Fréchet differentiable at $\mathbf{u}_*$, then there exists a real number λ such that

$$\mathrm{grad}F(\mathbf{u}_*) = \lambda\,\mathrm{grad}\phi(\mathbf{u}_*) = 2\lambda\mathbf{u}_*, \tag{16}$$

or $\mathbf{u}_*$ is a critical point of $(F + \lambda\phi)$. Also, $\mathbf{u}_*$ is an invariant direction of f on $\partial\Sigma_1$ if $\mathrm{grad}F(\mathbf{u}_*) \neq \theta$. Since the closed unit ball $\bar{\Sigma}_1$ in H^n is weakly compact and weakly closed, by the generalized Weierstrass theorem (Ref. 14, p. 100), $F|_{\bar{\Sigma}_1}$, the restriction of F to $\bar{\Sigma}_1$, attains its infimum and supremum on $\bar{\Sigma}_1$ provided that F is weakly continuous on $\bar{\Sigma}_1$. Moreover, if F is Fréchet differentiable on Σ_1 and attains its infimum or supremum at an interior point $\tilde{\mathbf{u}}$ of $\bar{\Sigma}_1$, then $\mathrm{grad}F(\tilde{\mathbf{u}}) = \theta$. It follows that if $\mathrm{grad}F(\theta) = \theta$ and $\|\mathrm{grad}F(\mathbf{u})\| > 0$ for all nonzero $\mathbf{u} \in \bar{\Sigma}_1$, then $F|_{\bar{\Sigma}_1}$ attains its infimum or supremum at one or more points on $\partial\Sigma_1$. Such points are invariant directions of $\mathbf{f} = \mathrm{grad}F$. This result is summarized below:

Lemma 2: Let $\mathbf{f}$ be the strong gradient of a weakly continuous functional F on a real Hilbert space H^n. If $\mathbf{f}(\theta) = \theta$ and $\|\mathbf{f}(\mathbf{u})\| > 0$ for all nonzero $\mathbf{u} \in \bar{\Sigma}_1$, then $\mathbf{f}$ has at least one invariant direction on $\partial\Sigma_1$

Direct application of Lemma 2 and Theorem 1 leads to the following result:

Theorem 3: Let $\mathbf{A}$ be the zero transformation on a real Hilbert space H^n and $\mathbf{f}$ be a gradient operator satisfying the hypotheses of Lemma 2 and assumption (A-2) for all $\mathbf{u} \in H^n$. Then, (2) has at least one ray solution $\mathbf{u}(t) = r(t)v_*$, $t \in I$, where $r(t)$ is given by (11) if $k \neq 1$ or by (12) if $k = 1$, and $\mathbf{v}_*$ is a minimum or maximum point of $F|_{\bar{\Sigma}_1}$ or an invariant direction of $\mathbf{f} = \mathrm{grad}F$ on $\partial\Sigma_1$.

Theorem 4: Let H^n be a finite-dimensional real Hilbert space and $\mathbf{f}$ be the gradient of a real C_1 function F on H^n. If $\mathbf{f}$ satisfies assumption (A-2) for all $\mathbf{u} \in H^n$ and $\mathbf{f}(\mathbf{u}) = \mathrm{grad}F(\mathbf{u}) \neq \theta$ for all $\mathbf{u} \in \partial\Sigma_1$, then (2) with $\mathbf{f}(\mathbf{u})$ as its rhs has at least two distinct ray solutions.

Proof: Since $F|_{\partial\Sigma_1}$ is continuous on the compact set $\partial\Sigma_1$, by Weierstrass theorem, $F|_{\partial\Sigma_1}$ attains its minimum and maximum on $\partial\Sigma_1$ at some points $\check{v}$ and $\hat{v}$ respectively. Since F is C_1 and $\mathrm{grad}F(\mathbf{u}) \neq \theta$ on $\partial\Sigma_1$, then, in view of (16), there exist nonzero real numbers $\check{\lambda}$ and $\hat{\lambda}$ such that $\mathrm{grad}F(\check{\mathbf{v}}) = \check{\lambda}\check{\mathbf{v}}$ and $\mathrm{grad}F(\hat{\mathbf{v}}) = \hat{\lambda}\hat{\mathbf{v}}$. Thus, $\hat{\mathbf{v}}$ and $\check{\mathbf{v}}$ are invariant directions of $\mathbf{f} = \mathrm{grad}F$ on $\partial\Sigma_1$. The desired result follows from Theorem 1. ||

Now, if $\mathbf{f}$ is a C_1 function on R^n into R^n, then a necessary and sufficient condition for $\mathbf{f}$ to be the gradient

of a real-valued function F is that the Jacobian matrix $[\partial f_i/\partial u_j]$ is symmetric on R^n and F is given by (14). If such a f is an odd gradient operator, then we have the following result:

Theorem 5: Let $\mathbf{f}$ be a C_1 gradient operator satisfying assumption (A-2) on R^n. If $\mathbf{f}$ is an odd operator such that $\mathbf{f}(\mathbf{u}) \neq \theta$ for all $\mathbf{u} \in \partial\Sigma_1$, then (2) with $f(u)$ as its rhs has at least $2n$ distinct ray solutions.

Proof: By hypothesis, $\mathbf{f}$ is the gradient of a real-valued C_2 function F on R^n given by (14). Since $\mathbf{f}$ is odd, then F is even. Thus, F can be regarded as a real-valued function on the $(n-1)$-dimensional projective space P^{n-1} obtained by identifying the antipodal points of $\partial\Sigma_1$. Since the category of P^{n-1} (in the sense of Lusternik and Schnirelmann[14-16]) is n, F has at least n distinct critical points $\mathbf{u}^j$, $j=1,\ldots,n$ on P^{n-1}, or $2n$ distinct critical points $\pm\mathbf{u}^j$, $j=1,\ldots,n$ on $\partial\Sigma_1$. For each pair $\pm\mathbf{u}^j$, we have, in view of (16), $\mathbf{f}(\pm\mathbf{u}^j) = \text{grad}F(\pm\mathbf{u}^j) = \pm\lambda_j\mathbf{u}^j$ for some real number λ_j. Since $\mathbf{f}(\mathbf{u}) \neq \theta$ on $\partial\Sigma_1$, $\lambda_j \neq 0$. Thus, $\pm\mathbf{u}^j$ are invariant directions of $\mathbf{f}$ on $\partial\Sigma_1$. The existence of $2n$ distinct ray solutions of (2) follows from Theorem 1. | |

Finally, we note that Theorem 3 can be generalized to the case where the rhs of (2) is the product of a linear operator and a gradient operator. Using Theorem 1 and a result due to Vainberg (Ref. 14, Thm. 15.1, p. 123), we have the following result on the existence of ray solutions of (2):

Theorem 6: Let $\mathbf{f} = \mathbf{K}\tilde{\mathbf{f}}$, the rhs of (2), be a mapping on a real Hilbert space H^n, where $\tilde{\mathbf{f}}$ is the strong gradient of a weakly continuous functional F on H^n, and $\mathbf{K}$ is a positive self-adjoint linear operator whose domain is H^n. If $\tilde{\mathbf{f}}$ satisfies assumption (A-2) on H^n and $(\tilde{\mathbf{f}}(\mathbf{u}), \mathbf{u}) > 0$ for all $\mathbf{u} \neq 0$, then (2) has at least one ray solution.

Remarks:

(R-1) Theorem 3 is a special case of Theorem 6 with K being the identity map on H^n.

(R-2) Under the hypotheses of Theorem 6, Vainberg's result (Ref. 14, Thm. 15.1, p. 123) ensures the existence of an invariant direction $\mathbf{u}_*$ of $\mathbf{K}\tilde{\mathbf{f}}$ in the form $\mathbf{u}_* = \mathbf{K}^{1/2}\mathbf{v}$, where $\mathbf{v}$ is a maximum point of the functional $\mathbf{F}\mathbf{K}^{1/2}$ over $\partial\Sigma_1$, and $\mathbf{K}^{1/2}$ is the positive square root of $\mathbf{K}$. Thus, the corresponding ray solution of (2) has the form $\mathbf{u}(t) = r(t)(\mathbf{K}^{1/2}\mathbf{v})/\|\mathbf{K}^{1/2}\mathbf{v}\|$.

(R-3) If, in addition to the conditions of Theorem 6, $\tilde{\mathbf{f}}$ is a strongly continuous odd operator on H^n, and $\mathbf{K}$ is a continuous compact operator on H^n having a countable set of positive eigenvalues, then it can be shown (Ref. 14, Thm. 15.6, p. 127) that $\tilde{\mathbf{f}} = \mathbf{K}\tilde{\mathbf{f}}$ has at least a countably infinite number of distinct invariant directions on $\partial\Sigma_1$. Consequently, (2) has at least a countably infinite number of distinct ray solutions.

(R-4) Theorems 2—6 are applicable when D_f, the domain of $\mathbf{f}$, is the whole space H^n. When D_f is a proper subspace of H^n, the problem of establishing conditions for the existence of invariant directions of $\mathbf{f}$ on $D_f \cap \partial\Sigma_1$ or the ray solutions of (2) is considerably more complex. However, results are obtainable for special classes of operators such as those having invariant cones in H^n.

(R-5) Theorems 1 and 2 remain valid for (2) defined on certain non-Hilbert spaces. In particular, consider the case where $D_h = D_f \cap D_A$ is the real Banach space $L_p(\Omega)$, $p > 2$, and Ω is a bounded open subset of R^m. Let

$$r(t) = \|u(t)\|_{L_p} = \{\int_\Omega |u(t, \mathbf{x})|^p d\mathbf{x}\}^{1/p}$$

and $v(t) = u(t)/r(t)$. Here, the equations for $r(t)$ and $v(t)$ are identical to (6) and (7) except that the terms $\text{Re}(Av, v)$ and $\text{Re}(f(v), v)$ are replaced respectively by

$$\tilde{\alpha}(v) = \int_\Omega (\text{sgn}v^{p-1})\,|v|^{p-1}Av\,d\mathbf{x}$$

and

$$\tilde{\beta}(v) = \int_\Omega (\text{sgn}v^{p-1})\,|v|^{p-1}f(v)\,d\mathbf{x}.$$

If A and f have a common invariant direction $v_* \in \partial\Sigma_1 \subset L_p(\Omega)$, then (2) has a ray solution as given in Theorem 1.

IV. EXPLOSIVE SOLUTIONS

By an explosive solution $u(t)$ of (2), we mean there exists a finite $t_1 > 0$ such that $\|\mathbf{u}(t)\| \to \infty$ as $t \to t_1^-$, or the solution has finite escape time. Sufficient conditions for the existence of such solutions can be readily deduced from the results of Sec. III. From Theorem 1, the ray solution (10) for $\alpha(\mathbf{v}_*) \neq 0$ and $k \neq 1$ has finite escape time if there exists a finite $\tau > 0$ such that

$$\alpha(\mathbf{v}_*) + \beta(\mathbf{v}_*)r(0)^{k-1}(1 - \exp[\alpha(\mathbf{v}_*)](k-1)\tau) = 0. \tag{17}$$

Evidently, if $r(0) > 0$, $k > 1$, and $\text{sgn}\alpha(\mathbf{v}_*) = \text{sgn}\beta(\mathbf{v}_*)$, such a τ exists and is given by

$$\tau = [(k-1)\alpha(\mathbf{v}_*)]^{-1}\ln[1 + \alpha(\mathbf{v}_*)\beta(\mathbf{v}_*)^{-1}r(0)^{1-k}]. \tag{18}$$

For $k > 1$, $\alpha(\mathbf{v}_*) = 0$ and $\beta(\mathbf{v}_*) > 0$, it is apparent that the ray solution (11) has finite escape time τ given by

$$\tau = [(k-1)\beta(\mathbf{v}_*)]^{-1}r(0)^{k-1}. \tag{19}$$

The foregoing observations are summarized in the following theorem:

Theorem 7: Let $r(0) = \|\mathbf{u}_0\| > 0$ and $\mathbf{f}$ satisfies assumptions (A-1) and (A-2) with $k > 1$. If $\mathbf{A}$ and $\mathbf{f}$ have a common invariant direction $\mathbf{v}_* \in D_h \cap \partial\Sigma_1$ such that (i) $\alpha(\mathbf{v}_*) \neq 0$ and $\text{sgn}\alpha(\mathbf{v}_*) = \text{sgn}\beta(\mathbf{v}_*)$ or (ii) $\alpha(\mathbf{v}_*) = 0$ and $\beta(\mathbf{v}_*) > 0$, then (2) has an explosive ray solution $\mathbf{u}(t) = r(t)\mathbf{v}_*$ defined on $[0, \tau]$, where $r(t)$ is given by (10) or (11), and the escape time τ is given by (18) or (19) respectively.

From Corollary 1, we have a special case of the above result with $\alpha(\mathbf{v}_*) = 1$ and $\beta(\mathbf{v}_*) = \|\mathbf{v}_*\|^{1-k}$.

Corollary 2: Let $r(0) = \|\mathbf{u}_0\| > 0$ and $\mathbf{f}$ satisfies (A-1) and (A-2) with $k = 1$. If $\mathbf{A}$ and $\mathbf{f}$ have a common nonzero fixed point $\mathbf{u}_* \in D_h$, then the ray solution $u(t) = r(t)\,\mathbf{u}_*/\|\mathbf{u}_*\|$ with $r(t)$ given by (10) is explosive and its escape time τ is given by

$$\tau = (k-1)^{-1}\ln[1 + \|\mathbf{u}_*\|^{k-1}r(0)^{1-k}]. \tag{20}$$

Similar results can be established for various special classes of operators discussed earlier. We note that it is possible to have explosive solutions which are not ray solutions. A simple example is provided by the solutions of the following equation defined on R^2:

$$\frac{du_1}{dt} = u_1^2, \quad u_1(0) = u_{10},$$

$$\frac{du_2}{dt} = u_1^2, \quad u_2(0) = u_{20}.$$

This equation has an explosive ray solution for $u_{10} = u_{20} > 0$. When $u_{10} > 0$ and $u_{10} \neq u_{20} \neq 0$, the corresponding solutions are explosive but they are not ray solutions.

For the case where $r(0) > 0$ and the degree of homogeneity of f satisfies $0 < k < 1$, $r(t)$ as given by (10) or (11) goes to zero in finite $t = \tau > 0$ if $\alpha(\mathbf{v}_*) \leq 0$ and $\beta(\mathbf{v}_*) < 0$, where τ is given by (18) or (19). Since the origin of H^n is an equilibrium solution of (2), such a ray solution remains at the origin for all $t > \tau$.

Now, we give a simple sufficient condition for the nonexistence of explosive solutions of (2). Let

$$\hat{\alpha} = \sup\{\mathrm{Re}(\mathbf{Av}, \mathbf{v}); \mathbf{v} \in \partial\Sigma_1\},$$
$$\hat{\beta} = \sup\{\mathrm{Re}(\mathbf{f}(\mathbf{v}), \mathbf{v}); \mathbf{v} \in \partial\Sigma_1\}. \tag{21}$$

From (6), we have

$$\frac{dr(t)}{dt} \leq [\hat{\alpha} + \hat{\beta} r(t)^{k-1}] r(t).$$

For $k > 1$, if $-\infty < \hat{\alpha} < 0$ and $0 < \hat{\beta} < \infty$, then for any initial point u_0 in the set $\Gamma = \{\mathbf{u} \in H^n : \|\mathbf{u}\| < \gamma\}$, where γ is the positive $(k-1)$th root of $|\hat{\alpha}|/\hat{\beta}$, the corresponding solution $\mathbf{u}(t) \in \Gamma$ for all $t \geq 0$, since $dr(t)/dt \leq 0$ for all $t > 0$ along such a solution. The boundedness of Γ implies that all the solutions initiating from Γ are nonexplosive. Similarly, for $0 < k < 1$, if $0 < \hat{\alpha} < \infty$ and $-\infty < \hat{\beta} < 0$, all the solutions initiating from $\tilde{\Gamma} = \{\mathbf{u} \in H^n : \|\mathbf{u}\| < \tilde{\gamma}\}$ are nonexplosive, where $\tilde{\gamma}$ is the positive $(1-k)$th root of $\hat{\alpha}/|\hat{\beta}|$. Finally, for $k = 1$, if $\hat{\alpha} + \hat{\beta} < 0$, all the solutions are nonexplosive.

V. APPLICATIONS

In this section, we apply some of the results in Secs. III and IV to specific evolutional equations arising from physical situations.

(1) Consider a system of ordinary differential equations of the form (2) defined on the complex n-dimensional space C^n, in which $\mathbf{A}$ is represented by a complex $n \times n$ matrix and $\mathbf{f}$ is a quadratic function of $\mathbf{u}$ given by

$$\mathbf{f}(\mathbf{u}) = (f_1(\mathbf{u}), \ldots, f_n(\mathbf{u})), \quad f_j(\mathbf{u}) = (\mathbf{u}, \mathbf{Q}_j\mathbf{u})_{C^n}, \quad j = 1, \ldots, n, \tag{22}$$

where the $\mathbf{Q}_j$'s are $n \times n$ complex matrices. Many physical phenomena such as the nonlinear interaction of a finite number of waves in plasmas and in other media are describable by this type of equations.[2,5-10] It is of interest to establish conditions for the existence and nonexistence of explosive solutions.

First, we write $\mathbf{Q}_j = \mathbf{Q}_{j1} + i\mathbf{Q}_{j2}$, where $\mathbf{Q}_{j1}$ and $\mathbf{Q}_{j2}$ are Hermitian matrices given by

$$\mathbf{Q}_{j1} = (\mathbf{Q}_j + \mathbf{Q}_j^*)/2, \quad \mathbf{Q}_{j2} = (\mathbf{Q}_j - \mathbf{Q}_j^*)/2i, \tag{23}$$

where $i = \sqrt{-1}$ and $(\cdot)^*$ denotes conjugate transposition. If $\mathbf{Q}_{j1}$ and $\mathbf{Q}_{j2}$ are positive semidefinite, then we have the estimate

$$\sum_{j=1}^{n} (\check{\lambda}_{j1}^2 + \check{\lambda}_{j2}^2) \|\mathbf{u}\|^4 \leq \|\mathbf{f}(\mathbf{u})\|^2 = \sum_{j=1}^{n} |(\mathbf{u}, \mathbf{Q}_j\mathbf{u})_{C^n}|^2 \leq \sum_{j=1}^{n} (\hat{\lambda}_{j1}^2 + \hat{\lambda}_{j2}^2) \|\mathbf{u}\|^4 \tag{24}$$

for all $\mathbf{u} \in C^n$, where $\hat{\lambda}_{jk}$ and $\check{\lambda}_{jk}$ are real numbers corresponding to the maximum and minimum eigenvalues of $\mathbf{Q}_{jk}$, $k = 1, 2$, respectively. Also, we have $\check{\lambda}_{j1}^2 + \check{\lambda}_{j2}^2 \leq |\lambda_s(\mathbf{Q}_j)|^2 \leq \hat{\lambda}_{j1}^2 + \hat{\lambda}_{j2}^2$ for $s = 1, \ldots, n$, where $\lambda_s(\mathbf{Q}_j)$ is the sth eigenvalue of $\mathbf{Q}_j$. Obviously, $\mathbf{f}$ is a compact mapping on C^n into C^n. From Theorem 2, a sufficient condition for the foregoing system (with $\mathbf{A}$ being the zero transformation on C^n) to have at least one ray solution $\mathbf{u}(t) = r(t)\mathbf{v}_*$, with $r(t)$ given by (11) and $\mathbf{v}_*$ an invariant direction of $\mathbf{f}$ on $\partial\Sigma_1$, is that $\check{\lambda}_{jk} \neq 0$ or $\mathbf{Q}_{jk}$ is nonsingular for some j. Such a solution is explosive if $\beta(\mathbf{v}_*) = (\mathbf{f}(\mathbf{v}_*), \mathbf{v}_*)_{C^n} > 0$. To establish a sufficient condition for the nonexistence of explosive solutions of (2) with $\mathbf{f}$ given by (22), we consider $\hat{\alpha}$ and $\hat{\beta}$ as defined in (21). From (24), we have

$$\hat{\beta} \leq \sup\{\|\mathbf{f}(\mathbf{v})\|; \mathbf{v} \in \partial\Sigma_1\} \leq \left\{\sum_{j=1}^{n} (\hat{\lambda}_{j1}^2 + \hat{\lambda}_{j2}^2)\right\}^{1/2}.$$

If $\mathbf{Q}_{jk}$ is nonsingular for some (j, k) and $\hat{\alpha} < 0$, then the solution corresponding to any initial point $\mathbf{u}_0 \in \Gamma = \{\mathbf{u} \in C^n : \|\mathbf{u}\| < |\hat{\alpha}|/\hat{\beta}\}$ is nonexplosive.

(2) Let $\Omega =]0, 1[$. Consider a single nonlinear partial differential equation of the form

$$\frac{\partial u(t, x)}{\partial t} = \gamma k^{-1} \frac{\partial u(t, x)^k}{\partial x}, \quad x \in \Omega, \tag{25}$$

where k is an integer > 1 and γ is a nonzero real number. For $k = 2$ and $\gamma = -1$, we have the equation describing an one-dimensional, pressureless, inviscous fluid. Let the domain of f [defined by the right-hand side of (25)] be $D_f = \{u \in L_2(\Omega) : f(u) \in L_2(\Omega)\}$. To determine the ray solutions of (25), we first seek the invariant directions of f on $D_f \cap \partial\Sigma_1$. This corresponds to finding a nonzero $v_* \in D_f$ satisfying

$$\gamma k^{-1} \frac{dv_*(x)^k}{dx} = \lambda v_*(x), \quad x \in \Omega, \tag{26}$$

for some nonzero real λ, and

$$\int_0^1 |v_*(x)|^2 dx = 1. \tag{27}$$

Let $\tilde{\lambda} = \lambda/\gamma$. By direct integration of (26), we have for $k > 1$

$$v_*(x) = \{\tilde{\lambda}(k-1)(x+c)\}^{1/(k-1)}, \quad x \in \Omega, \tag{28}$$

and for $k = 1$

$$v_*(x) = c \exp(\tilde{\lambda} x), \quad x \in \Omega. \tag{29}$$

The constant c in (28) or (29) is determined by (27), and $\lambda = (f(v_*), v_*)$. For $k > 1$, we have from (27) and (28)

$$\int_0^1 |v_*(x)|^2 dx = (k+1)^{-1}(k-1)^{(k+1)/(k-1)} \tilde{\lambda}^{2/(k-1)} \times \{(1+c)^{(k+1)/(k-1)} - c^{(k+1)/(k-1)}\}^{-1} = 1. \tag{30}$$

It is evident that the $\{\cdots\}$ term is > 1 for all real $c \geq 0$. Consequently, $|\tilde{\lambda}| \neq 0$ and there exist an uncountably infinite number of invariant directions of f on $D_f \cap \partial\Sigma_1$. By Theorem 1, for $r(0) > 0$ and each $c \geq 0$, there is a ray solution given by

$$u(t,x)=r(t)v_*(x)=r(0)[\tilde{\lambda}(k-1)(x+c)/$$
$$\{1-(k-1)\tilde{\lambda}r(0)^{k-1}t\}]^{1/(k-1)},\quad t\geq 0,\ x\in\Omega, \tag{31}$$

where $\tilde{\lambda}$ is a nonzero real number satisfying (30) for the given c. For $\tilde{\lambda}>0$ (resp. $\tilde{\lambda}<0$), the corresponding ray solution is explosive (resp. nonexplosive). For the case with $k=2$, (30) reduces to $\tilde{\lambda}^2=3/[(1+c)^3-c^3]$. For $c=0$, $\tilde{\lambda}=\pm\sqrt{3}$ and there are two invariant directions $v_*(x)=\pm\sqrt{3}x$, $x\in\Omega$. They correspond to an explosive and a nonexplosive ray solution given by

$$u(t,x)=\pm\sqrt{3}r(0)\,x/(1\mp\sqrt{3}r(0)t),\quad t\geq 0,\ x\in\Omega,$$

where $r(0)$ is a positive number.

For the case where $k=1$, (27) and (29) give: $c^2=2\tilde{\lambda}/[\exp(2\tilde{\lambda})-1]$. Evidently, $c^2>0$ for $\tilde{\lambda}\neq 0$. Hence (25) has an infinite number of nonexplosive ray solutions. Finally, we note that if we take $\Omega=R$, then v_* given by (28) or (29) no longer belongs to $L_2(\Omega)$. Consequently, there are no invariant directions of f on $D_f\cap\partial\Sigma_1$. Also, if we add a linear term Au to the rhs of (25) with $k=2$ and $\gamma=-1$, where $Au=\mu\partial^2u/\partial x^2$ or $-\mu\partial^3u/\partial x^3$ with μ being a nonzero real number, then we have the well-known Burgers or Korteweg–de Vries equation respectively. It can be readily verified that such A's have no common invariant directions with f on $D_f\cap\partial\Sigma_1$ in $L_2(\Omega)$.

(3) Let Ω be a given bounded open subset of R^m, whose closure is denoted by $\bar{\Omega}$. Consider the following partial differential–integral equation

$$\frac{\partial u(t,\mathbf{x})}{\partial t}=f(u(t,\cdot))(\mathbf{x}),\quad \mathbf{x}\in\Omega, \tag{32}$$

with initial data $u(0,\mathbf{x})=u_0(\mathbf{x})$, $\mathbf{x}\in\Omega$, at $t=0$, where $u_0\in L_{2p}(\Omega)$ and f is a Hammerstein integral operator on L_{2p} of the form

$$f(u(t,\cdot))(\mathbf{x})=\int_\Omega K(\mathbf{x},\mathbf{x}')\,u^{2p}(t,\mathbf{x}')\,d\mathbf{x}', \tag{33}$$

where p is an integer ≥ 1. Assume that the kernel K is real and continuous on $\bar{\Omega}\times\bar{\Omega}$ and there exist constants $\check{M}$ and $\hat{M}$ such that $0<\check{M}\leq K(\mathbf{x},\mathbf{x}')\leq\hat{M}<\infty$ for all $(\mathbf{x},\mathbf{x}')\in\bar{\Omega}\times\bar{\Omega}$. Since $\mu(\Omega)$, the total measure of Ω, is finite, we have

$$||f(u)||_{L_{2p}}=\{\int_\Omega(\int_\Omega K(\mathbf{x},\mathbf{x}')\,u^{2p}(\mathbf{x}')\,d\mathbf{x}')^{2p}\,d\mathbf{x}\}^{1/2p}$$
$$\geq\check{M}(\mu(\Omega))^{1/2p}\,||\mathbf{u}||_{2p}^{2p}=\hat{M}(\mu(\Omega))^{1/2p}>0$$

for all $u\in\partial\Sigma_1\subset L_{2p}(\Omega)$. It can be readily verified by using the Ascoli–Arzelà theorem that f is a continuous compact mapping of $L_{2p}(\Omega)$ into itself. In view of Theorems 2 and 7 and Remark (R-5), (32) has at least one explosive ray solution $u(t)=r(t)v_*$, $0\leq t<\tau$, with $r(t)$ being a positive root of

$$r(t)^{2p-1}=r(0)^{2p-1}\{1-(2p-1)\,\tilde{\beta}(v_*)\,r(0)^{2p-1}t\}^{-1}$$

and

$$\tau=[(2p-1)\,\tilde{\beta}(v_*)]^{-1}\,r(0)^{2p-1},$$

where

$$\tilde{\beta}(v_*)=\int_\Omega[(\mathrm{sgn}v_*(\mathbf{x}))\,|v_*(\mathbf{x})|^{2p-1}$$
$$\times\int_\Omega K(\mathbf{x},\mathbf{x}')v_*^{2p}(\mathbf{x}')\,d\mathbf{x}']\,d\mathbf{x}.$$

The point v_* is an invariant direction of f on $\partial\Sigma_1$. It can be found by first obtaining any nontrivial solution u_* to the Hammerstein integral equation $u=f(u)$ and then setting $v_*=u_*/||\mathbf{u}_*||_{L_{2p}}$. Analogous results can be obtained for certain systems of equation of a similar form, for example,

$$\frac{\partial u_j(t,\mathbf{x})}{\partial t}$$
$$=\sum_{k=1}^{n}\int_\Omega K_{jk}(\mathbf{x},\mathbf{x}')\,u_j(t,\mathbf{x}')\,u_k(t,\mathbf{x}')\,d\mathbf{x}',\quad j=1,\ldots,n,$$

defined on $L_2^n(\Omega)$, where $K_{jk}\in C_0(\bar{\Omega}\times\bar{\Omega})$ and its corresponding matrix is uniformly positive-definite on $\bar{\Omega}\times\bar{\Omega}$.

ACKNOWLEDGMENT

This work was supported by AFOSR Grant No. AFOSR 74-2662.

[1] R. Jancel, and Th. Kahan, *Electrodynamics of Plasmas* (Wiley, New York, 1966).
[2] K. Hasselmann, J. Fluid Mech. **12**, 481 (1962); **15**, 273 (1963); **15**, 385 (1963).
[3] L. Markus, Rend. Sem. Mat. Politecnico Torino **11**, 271 (1952).
[4] W. Hahn, *Stability of Motion* (Springer, New York, 1967).
[5] V.N. Tsytovich, *Nonlinear Effects in Plasma* (Plenum, New York, 1970).
[6] R.C. Davidson, *Methods in Nonlinear Plasma Theory* (Academic, New York, 1972).
[7] H. Wilhelmsson, J. Stenflo, and F. Englemann, J. Math. Phys. **11**, 1738 (1970).
[8] J. Fukai, S. Krishan, and E.G. Harris, Phys. Fluids **13**, 3031 (1970).
[9] P.K.C. Wang, J. Math. Phys. **13**, 943 (1972).
[10] J.A. Armstrong, N. Bloembergen, J. Ducuing, and P.S. Pershan, Phys. Rev. **127**, 1918 (1962).
[11] G.D. Birkhoff and O.D. Kellogg, Trans. Amer. Math. Soc. **23**, 96 (1922).
[12] A. Granas, "Introduction to Topology of Functional Spaces," University of Chicago, Mathematics Lecture Notes, Spring, 1961.
[13] P.C. Rosenbloom, Proc. Symp. Appl. Math. **VI**, 127 (1956).
[14] M.M. Vainberg, *Variational Methods for the Study of Nonlinear Operators* (Holden-Day, San Francisco, 1964).
[15] L.A. Lusternik and L.G. Schnirelmann, *Methodes topologiques dans les problemes variationels* (Hermann, Paris, 1934).
[16] M.S. Berger and M.S. Berger, *Perspectives in Nonlinearity, An Introduction to Nonlinear Analysis* (Benjamin, New York, 1968).

Construction of spin–orbit potentials from the phase shifts at fixed energy*

M. A. Hooshyar

Physics Department, Pahlavi University, Shiraz, Iran
(Received 30 July 1974)

The nonrelativistic scattering of spin-1/2 particles by central and spin–orbit potentials is considered. The form of central and spin–orbit potentials is deduced from a knowledge of the S matrix as a function of angular momentum at a fixed energy. Similar to the case of central potentials, the problem of constructing central and spin–orbit potentials from information on the phase shifts at a fixed energy has an infinity of solutions, depending on an infinite number of parameters.

1. INTRODUCTION

The problem of constructing interparticle forces from the scattering information is of obvious physical importance. Construction of nonrelativistic central potentials from the information on phase shifts at fixed energy was first carried out by Newton,[1] and later Sabatier[2] generalized the method to contain a larger set of central potentials.

An important tool in construction of central potentials from the phase shifts at fixed energy is the so called Regge–Newton equation. For the case of spin–orbit potential the analog of the mentioned equation already exists and was first found by Sabatier.[3] In this work we will make use of this important equation, the Sabatier equation, in order to deduce central and spin–orbit potentials from the knowledge of phase shifts at a fixed energy.

The procedure will be as follows. Since we need to depend heavily on the work of Sabatier,[3] for the ease of reader in Sec. 2 we review the relevant parts of that work. Section 3 is devoted to finding a set of input coefficients from the information on the phase shifts, and in Sec. 4 we use these input coefficients to construct the spin–orbit and the central potentials, which are associated with the desired phase shifts. In Sec. 4 we also show that when we have only one set of phase shifts, that is, "no spin–orbit interaction is present," then the method can reduce to what is already known about central potentials. In this case if we do not make the method to reduce to the central potential problem, then it will give us the transparent spin–orbit field, that is, spin–orbit interactions whose presence can not be detected by doing scattering experiment at a fixed energy. Since the construction procedure is more involved than for the case of only central potentials, in Sec. 5 we have summarized the construction procedure.

The conclusion of the paper is as follows. Provided that the phase shifts tend to zero rapidly with increasing angular momentum, and if they are related to spin–orbit potentials which go to zero faster than r^{-3} for large r, then we can find a set of central and spin–orbit potentials which will correspond to the same phase shifts. The corresponding set of central and spin-orbit potentials is not unique. In fact the method can generate transparent central and spin–orbit potentials at a fixed energy. The exact conditions that the phase shifts should satisfy so that the method can give us the potentials, in other words, the class of potentials which can be constructed by this method, are not considered in this paper and will be the subject of a forthcoming communication.

2. A SURVEY OF PREVIOUS RESULTS

When considering the scattering of spin-$\frac{1}{2}$ particles by central and spin–orbit potentials, then the Schrödinger equation can be written in the following form:

$$\Delta\Psi(\mathbf{r}) + [1 - U_c(r) - 2\mathbf{L}\cdot\mathbf{S}\,U_s(r)]\Psi(\mathbf{r}) = 0. \tag{2.1}$$

The differential cross section is then given by[4]

$$I(\theta) = |f(\theta)|^2 + |g(\theta)|^2, \tag{2.2}$$

where

$$f(\theta) = \frac{1}{2i}\sum_0^\infty \{(l+1)[\exp(2i\delta_l^+) - 1] + l[\exp(2i\delta_l^-) - 1]\}P_l(\cos\theta) \tag{2.3}$$

and

$$g(\theta) = \frac{1}{2i}\sum_0^\infty [\exp(2i\delta_l^+) - \exp(2i\delta_l^-)]P_l^1(\cos\theta). \tag{2.4}$$

In the direct scattering problem one usually knows the U_c and U_s and is asked to find the cross section. In that case, using Eq. (2.1), one finds the wavefunction, and from its asymptotic behavior one finds the phase shifts $\delta_l^\pm$, from which the cross section can be calculated. The problem of our interest is the inverse of this case; That is, given the phase shifts, find the interaction U_c and U_s. In other words, find the potentials in such a way that the asymptotic behavior of the regular solutions $\psi_\lambda^\pm(r)$, which satisfy the differential equations given below, will have the desired form:

$$r^2\frac{d^2}{dr^2}\psi_\lambda^+(r) + r^2[1 - U_c(r) + U_s(r) - \lambda U_s(r)]\psi_\lambda^+(r) = (\lambda^2 - \tfrac{1}{4})\psi_\lambda^+(r), \tag{2.5}$$

$$r^2\frac{d^2}{dr^2}\psi_\lambda^-(r) + r^2[1 - U_c(r) + U_s(r) + \lambda U_s(r)]\psi_\lambda^-(r) = (\lambda^2 - \tfrac{1}{4})\psi_\lambda^-(r). \tag{2.6}$$

To put things in a more standard notation, let us state that

$$V(r) = U_c(r) - U_s(r) \quad \text{and} \quad Q(r) = \tfrac{1}{2}U_s(r). \tag{2.7}$$

By using the above notation, Eqs. (2.5) and (2.6) take

the following form:

$$r^2\frac{d^2}{dr^2}\psi_\lambda^\pm(r)+r^2[1-V(r)\mp 2\lambda Q(r)]\psi_\lambda^\pm(r)$$
$$=(\lambda^2-\tfrac{1}{4})\psi_\lambda^\pm(r). \tag{2.8}$$

Having put the equations in the form considered by Sabatier,[3] we note that the wavefunctions can be represented as

$$\psi_\lambda^\pm(r)=F^\pm(r)U_\lambda(r)-\int_0^r K^\pm(r,s)U_\lambda(s)s^{-2}\,ds, \tag{2.9}$$

where

$$F^\pm(r)=\exp[\pm\int_0^r s\,Q(s)\,ds] \tag{2.10}$$

and $U_\lambda(r)$ is the regular solution of the following equation:

$$r^2\frac{d^2}{dr^2}U_\lambda(r)+r^2[1-V_k(r)]U_\lambda(r)=(\lambda^2-\tfrac{1}{4})U_\lambda(r). \tag{2.11}$$

$V_k(r)$, which Sabatier calls the comparison potential, can be any physical potential, but in order to simplify the mathematics in our case, we choose $V_k(r)$ to be zero. The work which follows can be easily generalized to all comparison potentials. If the comparison potential is zero, then

$$U_\lambda(r)=(\tfrac{1}{2}\pi r)^{1/2}J_\lambda(r). \tag{2.12}$$

The function $K^\pm(r,s)$ is defined through the following integral equations: If we denote the function $K^\pm(r,s)$ by

$$K^\pm(r,s)=I^\pm(r,s)-J^\pm(r,s), \tag{2.13}$$

then

$$J^\pm(r,r')=F^\pm(r)\,e(r,r')-\int_0^r K^\pm(r,s)\,e(s,r')s^{-2}\,ds \tag{2.14}$$

and

$$I^\pm(r,r')=F^\mp(r)g^\pm(r,r')-\int_0^r K^\mp(r,s)\,g^\pm(s,r')s^{-2}\,ds, \tag{2.15}$$

where

$$e(r,r')=\sum_{\lambda\in S}a_\lambda U_\lambda(r)U_\lambda(r'), \tag{2.16}$$

$$g^\pm(r,r')=\sum_{\lambda\in S}b_\lambda^\pm U_\lambda(r)U_\lambda(r') \tag{2.17}$$

with

$$S=\{1/2,\ 1,\ 3/2,\ 2,\ 5/2,\ 3,\ \cdots\}.$$

For our special case, where $V_k=0$, a_λ is nothing but[5]

$$a_\lambda=\begin{cases}2\lambda/\pi & \text{for positive integer } \lambda\\ 0 & \text{otherwise.}\end{cases} \tag{2.18}$$

Finally the set of coefficients $b_\lambda^\pm$, which we shall call potential coefficients, are some constants containing the information on V and Q in such a way that the following holds:

$$K^\pm(r,r)=\tfrac{1}{2}r\,F^\mp(r)\{\pm r^2Q(r)+\int_0^r[s^3Q^2(s)-sV(s)]\,ds\}. \tag{2.19}$$

To summarize, it is proved by Sabatier that, given a V and a Q, we can find a set of coefficients $b_\lambda^\pm$, in such a way that the functions $K^\mp(r,r')$ defined by Eqs. (2.13), (2.14), and (2.15) are such that they will satisfy Eq. (2.19) and the regular solutions of Eq. (2.8) can be represented by them through Eq. (2.9).

The way that the above results are presented, they seem more useful for the case of direct scattering problem than that of inverse scattering problem, because even if we could find the potential coefficients from the information on phase shifts, so that we could define the input function $g^\pm$, we would still not be able to use the definition of $K^\pm$ for finding them. Since in that definition $F^\pm$ which are related to Q are unknown, it is our task to find them. Knowing the input function $g^\pm$, we will show how to change the definition of $K^\pm$, so that the unknown functions do not appear in the analogs of (2.14) and (2.15). Then we find functions related to $K^\pm$ from information on $g^\pm$. Having found these functions, we will then use the analog of (2.19) to find the potentials V and Q. In any case the problem of constructing the input functions from the information on the phase shifts has to be dealt with first. To see the connection between the input functions, or the potential coefficients, with the phase shifts, let us substitute Eq. (2.16) and (2.17) in (2.14) and (2.15). Noting Eq. (2.9), one finds $K^\pm$ in terms of $\psi_\lambda^\pm$. Substitution of this new representation of $K^\pm$ in (2.9) will give us the following important relation:

$$\psi_\lambda^\pm(r)=F^\pm(r)U_\lambda(r)-\sum_{\mu\in S}[\psi_\mu^\mp(r)b_\mu^\pm-\psi_\mu^\pm(r)a_\mu]L_\mu^\lambda(r), \tag{2.20}$$

where

$$L_\mu^\lambda(r)=\int_0^r U_\lambda(s)U_\mu(s)s^{-2}\,ds. \tag{2.21}$$

Since the asymptotic form of $\psi_\lambda^\pm$ is assumed to be known, for large values of r, Eq. (2.20) is an statement about what the potential coefficients should be. In other words, find a set of coefficients $b_\lambda^\pm$ such that with a given set of phase shifts Eqs. (2.20) are satisfied for large values of r, and then the related potentials will give the desired scattering information. Thus the problem can be stated as such: Given a set of phase shifts, find the potential coefficients from Eq. (2.20). From the potential coefficients found, one has to find $K^\pm$ from which the interaction potentials will be defined. The potentials thus determined will give us the original phase shifts, because, by construction, the associated wavefunctions have the desired asympototic form.

3. DETERMINATION OF POTENTIAL COEFFICIENTS

Our aim in this section is to find $b_\lambda^\pm$ from Eqs. (2.20). But, as we will see, these equations can only give us $b_\lambda^\pm$ up to an unknown constant multiplicative factor. Before showing how we can get $b_\lambda^\pm$ from phase shift information, let us mention that, in order to be able to use asymptotic forms of $\psi_\lambda^\pm$ and U_λ in Eqs. (2.20), we need to make the following assumptions[6]:

$$\left|\int_0^r sQ(s)\,ds\right|<\infty \quad \text{for all } r,$$
$$Q(r)=o(r^{-3}) \quad \text{for large } r \tag{3.1}$$

As shown in the Appendix, using these assumptions, we can take the limit inside the summation in (2.20) if

$$b_\lambda^\pm=h_\pm^2(a_\lambda+c_\lambda^\pm)=h_\pm d_\lambda^\pm, \tag{3.2}$$

where

$$c_\lambda^\pm = o(\lambda^{1/3}) \quad \text{for large } \lambda \quad \text{and} \quad h_\pm = \exp[\pm \int_0^\infty sQ(s)\,ds].$$

With the above assumptions we take the limit as $r \to \infty$ of (2.20) and equate the coefficients of e^{ir} and e^{-ir} separately. Doing so, we get

$$A_\lambda^\pm \exp[i(\delta_\lambda^\pm - \pi\lambda/2)] = h_\pm e^{-i\pi\lambda/2} - \sum_{\mu\in S} [A_\mu^\mp h_\pm^2 d_\mu^\pm \exp(i\delta_\mu^\mp) - A_\mu^\pm a_\mu \exp(i\delta_\mu^\pm)] \exp(-i\mu\pi/2) L_\mu^\lambda \tag{3.3}$$

In deriving (3.3) we have used the following asymptotic forms for large r:

$$\psi_\lambda^\pm \sim A_\lambda^\pm \sin[r + \delta_\lambda^\pm - \pi(\lambda - \tfrac{1}{2})/2],$$
$$U_\lambda \sim \sin[r - \pi(\lambda - \tfrac{1}{2})/2], \tag{3.4}$$
$$L_\mu^\lambda(\infty) = L_\mu^\lambda = \frac{\sin\frac{1}{2}\pi(\lambda-\mu)}{\lambda^2 - \mu^2}(1 - \delta_\mu^\lambda) + (\pi/4\lambda)\delta_\mu^\lambda.$$

Let us call

$$B_\lambda^\pm = A_\lambda^\pm h_\mp. \tag{3.5}$$

Next multiply (3.3) by $h_\mp$ and since $h_- h_+ = 1$, Eqs. (3.3) reduces to

$$B_\lambda^\pm \exp[i(\delta_\lambda^\pm - \lambda\pi/2)] = \exp(-i\lambda\pi/2) - \sum_{\mu\in S} [B_\mu^\mp d_\mu^\pm \exp(i\delta_\mu^\mp) - B_\mu^\pm a_\mu \exp(i\delta_\mu^\pm)] \exp(-i\mu\pi/2) L_\mu^\lambda. \tag{3.6}$$

It is the aim of this section to show that if $\delta_\lambda^\pm$ for physical values of λ are given, $\lambda \in S_\rho = \{1/2,\ 3/2,\ 5/2,\ \cdots\}$, then we can find $d_\lambda^\pm$ for $\lambda \in S_\rho$. Note that for nonphysical values of $\lambda \in S_o = \{1, 2, 3, \cdots\}$, except for the conditions of (3.2), the values of $d_\lambda^\pm$ are at our disposal. In order to find $d_\lambda^\pm$ for physical values of λ, we follow a method similar to the one first used by Sabatier,[2] and that is to define two column vectors $\phi_o^\pm$ and $\phi_\rho^\pm$ in such a way that the elements of $\phi_o^\pm$ are $B_\lambda^\pm \exp[i(\delta_\lambda^\pm - \lambda\pi/2)]$ with $\lambda \in S_o$, and the elements of $\phi_\rho^\pm$ are similar except that $\lambda \in S_\rho$. Then we can write (3.6) in a vector form:

$$\phi_o^\pm = \xi_o - L_o^o(d_o^\pm \phi_o^\mp - a_o\phi_o^\pm) - L_\rho^o d_\rho^\pm \phi_\rho^\mp, \tag{3.7}$$
$$\phi_\rho^\pm = \xi_\rho - L_o^\rho(d_o^\pm \phi_o^\mp - a_o\phi_o^\pm) - L_\rho^\rho d_\rho^\pm \phi_\rho^\mp, \tag{3.8}$$

where ξ_o is a column vector with elements $\exp(-i\lambda\pi/2)$ with $\lambda \in S_o$, L_ρ^o is a matrix with elements L_μ^λ where $\mu \in S_\rho$ and $\lambda \in S_o$. $d_o^\pm$ and a_o are diagonal matrices with elements $d_\lambda^\pm$ and a_λ where $\lambda \in S_o$. Other elements are defined likewise. It should be noted that in deriving (3.7) and (3.8) we have made use of the fact that a_λ is zero for physical values of λ.

Using Eqs. (3.7), we find

$$\bar\Omega^\pm(d_o^\pm\phi_o^\mp - a_o\phi_o^\pm) = \bar\xi_o^\pm + \bar T^\pm d_\rho^\pm \phi_\rho^\mp + \bar U^\pm d_\rho^\mp \phi_\rho^\pm, \tag{3.9}$$

where

$$\bar\Omega^\pm = a_o\beta^\pm[1 - d_o^\pm(\beta^\pm a_o L_o^o + L_o^o a_o \beta^\pm + L_o^o d_o^\mp L_o^o - L_o^o a_o \beta^\pm a_o L_o^o)],$$
$$\bar\xi_o^\pm = [P^\pm(d_o^\pm - a_o) - a_o L_o^o(d_o^\mp - a_o)]\xi_o,$$
$$\bar T^\pm = [P^\pm a_o + a_o L_o^o d_o^\mp] L_\rho^o,$$
$$\bar U^\pm = -[P^\pm d_o^\pm + a_o L_o^o a_o] L_\rho^o$$

with $P^\pm = (1 - a_o L_o^o)a_o\beta^\pm$ and $\beta^\pm = (d_o^\pm)^{-1}$. Since $d_o^\pm$ are at our disposal, we can assume their inverses exist. Also, a_o is invertable; therefore, in general, $\bar\Omega^\pm$ can be inverted. Doing so, we get:

$$d_o^\pm\phi_o^\mp - a_o\phi_o^\pm = \xi_o^\pm + T^\pm d_\rho^\pm \phi_\rho^\mp + U^\pm d_\rho^\mp \phi_\rho^\pm, \tag{3.10}$$

where

$$\xi_o^\pm = \Omega^\pm \bar\xi_o^\pm, \quad T^\pm = \Omega^\pm \bar T^\pm,$$
$$U^\pm = \Omega^\pm \bar U^\pm, \quad \text{and} \quad \Omega^\pm = (\bar\Omega^\pm)^{-1}.$$

At this point we find it interesting to note that if we choose $d_o^\pm = a_o$, which is the spin–orbit analog of the class of central potentials considered by Newton,[1] then $\bar\Omega^\pm$ is a matrix very similar to the now famous matrix M which Newton needed to invert. A trivial modification of inversion method given by Sabatier[7] for matrix M can be used to invert matrix $\bar\Omega^\pm$, for the above mentioned case. In any case, let us continue the construction by substituting (3.10) in (3.8):

$$\phi_\rho^\pm = \bar\eta^\pm - (L_\rho^\rho + {}_\pm\bar S) d_\rho^\pm \phi_\rho^\mp - {}_\pm\bar R\, d_\rho^\mp \phi_\rho^\pm, \tag{3.11}$$

where

$$\bar\eta^\pm = \xi_\rho - L_o^\rho \xi_o^\pm, \quad {}_\pm\bar S = L_o^\rho T^\pm, \quad \text{and} \quad {}_\pm\bar R = L_o^\rho U^\pm.$$

Having eliminated the asymptotic form of the wavefunction for nonphysical values of λ, let us write (3.11) in terms of its elements:

$$B_\lambda^\pm = \eta_\lambda^\pm \exp[i(\Delta_\lambda^\pm - \delta_\lambda^\pm)] - \sum_{\mu\in S_\rho} B_\mu^\mp d_\mu^\pm \exp[i(\delta_\mu^\mp - \delta_\lambda^\pm)]\, i^{\lambda-\mu}(L_\mu^\lambda + {}_\pm\bar S_\mu^\lambda) + \sum_{\mu\in S_\rho} B_\mu^\pm d_\mu^\mp \exp[i(\delta_\mu^\pm - \delta_\lambda^\pm)]\, i^{\lambda-\mu}\, {}_\pm\bar R_\mu^\lambda, \tag{3.12}$$

where the element of $\bar\eta^\pm$ is represented by $\eta_\lambda^\pm \exp[i(\Delta_\lambda^\pm - \lambda\pi/2)]$.

Next let us define the following:

$$L_\mu^\lambda = i^{\mu-\lambda-1} M_\mu^\lambda + (\pi/4\lambda)\delta_\mu^\lambda$$
$${}_\pm N_\mu^\lambda = i^{\lambda-\mu}\, {}_\pm\bar S_\mu^\lambda + (\pi/4\lambda)\delta_\mu^\lambda \quad \text{if } \lambda-\mu \text{ is even, 0 otherwise,}$$
$${}_\pm S_\mu^\lambda = i^{\lambda-\mu-1}\, {}_\pm\bar S_\mu^\lambda \quad \text{if } \lambda-\mu \text{ is odd, 0 otherwise,}$$
$${}_\pm X_\mu^\lambda = i^{\lambda-\mu}\, {}_\pm\bar R_\mu^\lambda \quad \text{if } \lambda-\mu \text{ is even, 0 otherwise,}$$
$${}_\pm R_\mu^\lambda = i^{\lambda-\mu-1}\, {}_\pm\bar R_\mu^\lambda \quad \text{if } \lambda-\mu \text{ is odd, 0 otherwise.}$$

With this notation we separate the real and imaginary parts of (3.12). If we divide the imaginary part by $\cos\delta_\lambda^\pm$, then we get

$$\gamma_\lambda^\pm = \sum_{\mu\in S_\rho} D_\mu^\pm[(1 + \tan\delta_\mu^\pm \tan\delta_\lambda^\pm)\, {}_\pm R_\mu^\lambda - (\tan\delta_\mu^\pm - \tan\delta_\lambda^\pm)\, {}_\pm X_\mu^\lambda] + \sum_{\mu\in S_\rho} D_\mu^\mp[(1 + \tan\delta_\mu^\mp \tan\delta_\lambda^\pm)(M_\mu^\lambda + {}_\pm S_\mu^\lambda) - (\tan\delta_\mu^\mp - \tan\delta_\lambda^\pm)\, {}_\pm N_\mu^\lambda], \tag{3.13}$$

where

$$D^{\pm}_{\lambda} = B^{\pm}_{\lambda}\, d^{\mp}_{\lambda} \cos\delta^{\pm}_{\lambda}$$

and

$$\gamma^{\pm}_{\lambda} = \eta^{\pm}_{\lambda}(\cos\Delta^{\pm}_{\lambda} \tan\delta^{\pm}_{\lambda} - \sin\Delta^{\pm}_{\lambda}).$$

Writing the above equations in matrix form, we have

$$\gamma^{\pm} = {}_{\pm}Y\, D^{\pm} + M\, {}_{\pm}Z\, D^{\mp}, \tag{3.14}$$

where

$${}_{\pm}Y = {}_{\pm}R + \tan\delta^{\pm}\, {}_{\pm}R \tan\delta^{\pm} - {}_{\pm}X \tan\delta^{\pm} + \tan\delta^{\pm}\, {}_{\pm}X$$

and

$${}_{\pm}Z = 1 + M^{-1}[{}_{\pm}S + \tan\delta^{\pm}(M + {}_{\pm}S)\tan\delta^{\mp} - {}_{\pm}N \tan\delta^{\mp} + \tan\delta^{\pm}\, {}_{\pm}N]$$

From (3.14) we solve for $D^{\pm}$:

$$D^{\pm} = \theta^{\mp}(\gamma^{\mp} - {}_{\mp}Y\, {}_{\pm}Z^{-1} M^{-1} \gamma^{\pm}) \tag{3.15}$$

with

$$\theta^{\pm} = (1 - {}_{\pm}Z^{-1} M^{-1}\, {}_{\pm}Y\, {}_{\mp}Z^{-1} M^{-1}\, {}_{\mp}Y)^{-1}\, {}_{\pm}Z^{-1} M^{-1}.$$

Again since $d^{\pm}_{\lambda}$ for nonphysical values of λ are under our control, and if the values of $\delta^{\pm}_{\lambda}$ tend to zero sufficiently fast as λ tends to infinity, then the inverse of all the necessary matrices in general exist.

Having found $D^{\pm}$, we now turn to the real part of Eqs. (3.12) which give us $B^{\pm}$:

$$\begin{aligned} B^{\pm}_{\lambda} = {} & \eta^{\pm}_{\lambda} \cos(\Delta^{\pm}_{\lambda} - \delta^{\pm}_{\lambda}) \\ & - \sum_{\mu \in S_p} D^{\pm}_{\mu}[(\cos\delta^{\pm}_{\lambda} + \tan\delta^{\pm}_{\mu} \sin\delta^{\pm}_{\lambda})\, {}_{\pm}X^{\lambda}_{\mu} \\ & + (\tan\delta^{\pm}_{\mu} \cos\delta^{\pm}_{\lambda} - \sin\delta^{\pm}_{\lambda})\, {}_{\pm}R^{\lambda}_{\mu}] \\ & - \sum_{\mu \in S_p} D^{\mp}_{\mu}[(\cos\delta^{\pm}_{\lambda} + \tan\delta^{\mp}_{\mu} \sin\delta^{\pm}_{\lambda})\, {}_{\pm}N^{\lambda}_{\mu} \\ & + (\tan\delta^{\mp}_{\mu} \cos\delta^{\pm}_{\lambda} - \sin\delta^{\pm}_{\lambda})\,(M^{\lambda}_{\mu} + {}_{\pm}S^{\lambda}_{\mu})]. \end{aligned} \tag{3.16}$$

We are finally in a position to find $d^{\pm}_{\lambda}$ from the knowledge of $B^{\pm}_{\lambda}$ and $D^{\pm}_{\lambda}$. Needless to say that besides the usual indetermination that exists because of introducing M^{-1}, we also have further indetermination in finding $d^{\pm}_{\lambda}$ due to the fact that we can choose $d^{\pm}_{\lambda}$ for nonphysical values of λ, to a large extent, arbitrarily. In other words the information on phase shifts by no means will specify a unique set of $d^{\pm}_{\lambda}$ for physical values of λ.

4. CONSTRUCTION OF POTENTIALS

In Sec. 3 we showed that from information on phase shifts one cannot get the potential coefficients $b^{\pm}_{\lambda}$, but one can only find them up to unknown factor $h_{\pm}$. It is the purpose of this section to show that the information that we can get is enough to specify the potentials.

In order to see the problem clearly, let us write the equations for $K^{\pm}(r, r')$ in a more compact form. Using (2.13), (2.14), and (2.15), we write the integral equations for $K^{\pm}(r, r')$ in the following way:

$$\mathbf{K}(r, r') = \mathbf{F}(r)\, E(r, r') - \int_0^r \mathbf{K}(r, s) E(s, r') s^{-2}\, ds, \tag{4.1}$$

where

$$\mathbf{K}(r, r') = [K^{+}(r, r'), K^{-}(r, r')], \qquad \mathbf{F}(r) = [F^{+}(r), F^{-}(r)],$$

and

$$E(r, r') = \begin{bmatrix} -e(r, r') & g^{-}(r, r') \\ g^{+}(r, r') & -e(r, r') \end{bmatrix}.$$

Since we do not know what $F^{\pm}(r)$ and $g^{\pm}(r, r')$ are in Eq. (4.1), one cannot use this equation to find $\mathbf{K}(r, r')$ which would have enabled us to find the potentials. In order to use our partial information on $g^{\pm}$, let us introduce the auxiliary matrix $H(r, r')$ through the following integral equation:

$$H(r, r') = G(r, r') - \int_0^r H(r, s)\, G(s, r') s^{-2}\, ds, \tag{4.2}$$

where

$$G(r, r') = \begin{bmatrix} -e(r, r') & f^{-}(r, r') \\ f^{+}(r, r') & -e(r, r') \end{bmatrix}$$

and

$$f^{\pm}(r, r') = \sum_{\lambda \in S} U_{\lambda}(r) d^{\pm}_{\lambda} U_{\lambda}(r').$$

We should note that from information on phase shifts we can find $d^{\pm}_{\lambda}$; therefore, $G(r, r')$ is a known matrix, and, solving (4.2), we can find $H(r, r')$. Let us next note that the following is true:

$$\begin{aligned} \mathbf{F}(r) h^{-1} H(r, r') h = {} & \mathbf{F}(r)\, E(r, r') \\ & - \int_0^r \mathbf{F}(r) h^{-1} H(r, s) h E(s, r') s^{-2}\, ds, \end{aligned} \tag{4.3}$$

where

$$h = \begin{bmatrix} h_{+} & 0 \\ 0 & h_{-} \end{bmatrix}.$$

It should be noted that in deriving (4.3) from (4.2) the following fact was used:

$$h^{-1} G(r, r') h = E(r, r'). \tag{4.4}$$

Comparing (4.3) and (4.1), we come to the conclusion that

$$\mathbf{K}(r, r') = \mathbf{F}(r) h^{-1} H(r, r') h. \tag{4.5}$$

Let us mention that in this paper we assume that the homogeneous version of Eq. (4.1) does not have a nontrivial solution for real values of r. In other words the associated Fredholm determinant is nonzero for real values of r. As for the case of central potentials this assumption is justified on physical grounds. Otherwise, we would be dealing with potentials which have poles for finite values of r. Clearly this is not acceptable.

So up to now we have been able to find $K^{\pm}(r, r')$ up to some unknown function of r. We will show that this is all we need to specify the form of spin–orbit interaction. To see that this is indeed the case, let us subtract the equations in (2.14) from each other:

$$2r^{-2}[F^{-}(r) K^{+}(r, r) - F^{+}(r) K^{-}(r, r)] = 2rQ(r). \tag{4.6}$$

Substituting (4.5) in (4.6) yields the following:

$$\frac{d}{dr}\, t(r) + H_1(r)\, t(r) + H_2(r)\, t^2(r) = H_3(r),$$

$$\text{with } \lim_{r \to \infty} t(r) = 1, \tag{4.7}$$

where

$$H_1(r)=2r^{-2}[H_{22}(r,r)-H_{11}(r,r)],$$
$$H_2(r)=2r^{-2}H_{12}^{*}(r,r),$$
$$H_3(r)=2r^{-2}H_{21}(r,r),$$

and

$$t(r)=h_-^2\exp[2\int_0^r sQ(s)\,ds].$$

We find it remarkable that the differential equation (4.7) is nothing but the generalized Riccati equation,[8] which was first studied by d'Alembert. At this point it is important to note that in Eq. (4.7) all the coefficients can be found from the information on the phase shifts; therefore, we can solve the equation and find $t(r)$. Having found $t(r)$, we find the spin–orbit interaction to be

$$Q(r)=\frac{1}{2rt(r)}\frac{d}{dr}t(r). \tag{4.8}$$

Knowing $Q(r)$, one can find the functions $F^{\pm}(r)$ from (2.10), and from their asymptotic limits the constants $h_{\pm}$ can be found. Substituting the found values of $h_{\pm}$ in (4.5) will give us $K^{\pm}(r,r')$. Knowing $K^{\pm}(r,r)$, we are able to find the central potential from (2.19), if we first add the equations together and then take derivative respect to r. Doing so leads us to the following equation:

$$V(r)=r^2Q^2(r)-\frac{1}{r}\frac{d}{dr}[r^{-1}F^{-}(r)K^{+}(r,r)+r^{-1}F^{+}(r)K^{-}(r,r)] \tag{4.9}$$

Knowing $K^{\pm}(r,r)$ also enables us to find $\psi_\lambda^{\pm}(r)$ from (2.8). By construction, we are guaranteed that the wavefunctions $\psi_\lambda^{\pm}(r)$ associated with the calculated central and spin–orbit potentials will have the desired asymptotic form. In other words, the calculated central and spin–orbit potentials will give us the desired scattering information.

An interesting point becomes apparent if we ask the following question: In the case when the spin–orbit potential is not present, does this method reduce to Newton's method? In order to put this question in a more proper form, from the point of view of inverse scattering problem, let us rephrase the question in this manner. Given one set of phase shifts, $\delta_\lambda^{+}=\delta_\lambda^{-}$, does this method reduce to Newton's method? A short answer would be yes, if we wanted it to. In other words, even for the case of one set of phase shifts, this method is still a generalization of the Newton[1] and Sabatier[2] work on the inverse scattering problem for central potentials. To explain this in more detail, let us assume that $\delta_\lambda^{+}=\delta_\lambda^{-}$ for all physical values of λ. If this is the case and we want to solve the inverse problem, then we notice that there are two cases to consider. Either we choose $d_\lambda^{+}=d_\lambda^{-}$ for all the nonphysical values of λ, which are at our disposal, or we do not. If we make the above choice then a glance at Sec. 3 is enough for us to note that, in that case, all the quantities with plus and minus are identical. Therefore, $d_\lambda^{+}=d_\lambda^{-}$ for physical values of λ too. From this it follows that $f^{+}(r,r')=f^{-}(r,r')$ and from (4.2) we find that

$$H_{11}(r,r')=H_{22}(r,r') \text{ and } H_{12}(r,r')=H_{21}(r,r') \text{ if } \delta_\lambda^{+}=\delta_\lambda^{-} \tag{4.10}$$

Substitution of (4.10) in (4.7) leads us to

$$\frac{d}{dr}t(r)=H_2(r)[1-t^2(r)] \quad \text{with } t(\infty)=1, \quad \text{if } \delta_\lambda^{+}=\delta_\lambda^{-}. \tag{4.11}$$

$t(r)=1$ is clearly the solution to (4.11). From (4.8) it follows that $Q(r)=0$. From this information it then follows that $F^{\pm}(r)=1$. In this case, it is easy to see that this method is identical with the previous works[1,2] if $\delta_\lambda^{+}=\delta_\lambda^{-}$ for all physical values of λ, which is the case when we "only have central potentials."

The interesting result becomes apparent if we do not choose d_λ^{+} to be the same as d_λ^{-} for nonphysical values of λ, when $\delta_\lambda^{+}=\delta_\lambda^{-}$ for physical values of λ. In this case, in general, $f^{+}(r,r')$ and $f^{-}(r,r')$ are different from each other, and therefore $t(r)=1$ will no longer be the solution to (4.7). In other words, even when $\delta_\lambda^{+}=\delta_\lambda^{-}$ for all physical values of λ, it is still possible to have spin–orbit interaction. We call these types of potentials transparent spin–orbit potentials because one cannot detect their presence from scattering information at a fixed energy. In other words, the information that the differential cross section at a fixed energy has a form which is usually associated with central potential, and that there is no change in polarization, is not enough information to make us exclude the presence of spin–orbit interaction in that scattering experiment.

5. SUMMARY

Since the method presented for construction of central and spin–orbit potentials from information on phase shifts at a fixed energy is rather involved, we feel it is beneficial to summarize the method.

Given a set of phase shifts $\delta_\lambda^{\pm}$, one chooses a set of acceptable constants $d_\lambda^{\pm}$ for nonphysical values of angular momentum, $l=\lambda+\frac{1}{2}$. These $d_\lambda^{\pm}$ should satisfy condition (3.2) and be such that the relevant matrices could be inverted. Next using these $d_\lambda^{\pm}$ and $\delta_\lambda^{\pm}$ in (3.15) and (3.16) will lead us to another set of constants $d_\lambda^{\pm}$ for physical values of λ. Having found $d_\lambda^{\pm}$, then we are able to find the function $H(r,r')$ from integral equation (4.2). Substituting $H(r,r)$ in (4.7) enables us to find a function $t(r)$, from which the form of spin–orbit interaction, $Q(r)$, can be easily found using (4.8). Knowing $Q(r)$, then we are able to find $F^{\pm}(r)$ and their asymptotic forms, $h_{\pm}$, from (2.10) and (3.2). We are finally in a position to find the functions $K^{\pm}(r,r')$ and the form of the central potential from (4.5) and (4.9).

ACKNOWLEDGMENTS

The author would like to thank Professor R.G. Newton for many stimulating discussions. The author is also indebted to Professor P.C. Sabatier for encouraging him to work on this problem.

APPENDIX

In this appendix we would like to prove that we can use the asymptotic forms of $\psi_\lambda^{\pm}(r)$, $U_\lambda(r)$, and $L_\mu^\lambda(r)$ in (2.20), if the conditions given in (3.2) are satisfied. To prove this fact, let us substitute (2.9) and (3.2) in (2.20), and

write the summation in the following form:

$$\sum_{\mu\in S}[\psi_\mu^{\mp}(r)b_\mu^{\pm}-a_\mu\psi_\mu^{\pm}(r)]L_\mu^{\lambda}(r)$$
$$=\sum_{\mu\in S}[F^{\mp}(r)h_\pm^2-F^{\pm}(r)]a_\mu U_\mu(r)L_\mu^{\lambda}(r)$$
$$-\sum_{\mu\in S}\int_0^r ds\, s^{-2}[K^{\mp}(r,s)h_\pm^2-K^{\pm}(r,s)]$$
$$\times a_\mu U_\mu(s)L_\mu^{\lambda}(r)+\sum_{\mu\in S}\psi_\mu^{\mp}(r)h_\pm^2 c_\mu^{\pm}L_\mu^{\lambda}(r). \tag{A1}$$

Based on Sabatier's works,[7] the last summation is uniformly convergent, and therefore interchange of limit from the outside to the inside of summation is justified. We would like to show that the second term from the right could also be uniformly convergent if the spin–orbit interaction has the proper form. To see this, let us use the Schwarz's inequality to find that

$$\left|\int_0^r ds\, s^{-2}[K^{\mp}(r,s)h_\pm^2-K^{\pm}(r,s)]U_\lambda(s)\right|$$
$$\leq h_\pm\left[\int_0^r \bar{K}^2(r,s)\,s^{-7/3+\epsilon}\,ds\right]^{1/2}$$
$$\times\left[\int_0^r U_\lambda^2(s)\,s^{-5/3-\epsilon}\,ds\right]^{1/2} \tag{A2}$$

with

$$\bar{K}(r,s)=K^{-}(r,s)h_+-K^{+}(r,s)h_-.$$

If we use the integral formulas for Bessel's functions[9] and use Stirling's formula to find a bound for the answer, we get[10]

$$\int_0^r U_\lambda^2(s)\,s^{-5/3-\epsilon}\,ds\leq\int_0^\infty U_\lambda^2(s)\,s^{-5/3-\epsilon}\,ds\leq C\lambda^{-2/3-\epsilon}. \tag{A3}$$

From (A3) it follows that the second summation in (A1) is uniformly convergent if the first integral in right-hand side of (A2) is bounded. Clearly this integral can be bounded if the function $Q(r)$ has the proper form. Because, as we have already seen, $\bar{K}(r,s)$ depends very much on $Q(r)$, and if $Q(r)=0$ then $\bar{K}(r,s)$ is also zero. In order to see what condition on $Q(r)$ is needed to make the mentioned integral bounded, let us assume that, for large values of r and s, $\bar{K}(r,s)$ increases as $r^\alpha s^\sigma$ where α and σ are some constants. Since the function $\bar{K}(r,s)$ is finite for finite values of r and s, it follows that the mentioned integral is bounded if

$$(\alpha+\sigma)<2/3. \tag{A4}$$

A glance at (4.6), and noting that for large values of r the quantities at the left side of (4.6) are nothing but $\pm\bar{K}(r,r)$, leads us to the conclusion that (A4) can be satisfied if we can have:

$$Q(r)=o(r^{-7/3})\quad\text{for}\quad r\to\infty. \tag{A5}$$

In any case, if the phase shifts and $d_\lambda^\pm$ for nonphysical values of λ are such that the condition (A4) could be satisfied, then we can have

$$\int_0^r ds\, s^{-2}[K^{\mp}(r,s)h_\pm^2-K^{\pm}(r,s)]U_\lambda(s)\leq C\lambda^{-1/3-\epsilon}. \tag{A6}$$

From (A6) it follows that the second summation from the right in (A1) is uniformly bounded, and therefore the interchange of limit is possible. Again let us mention that finding the necessary conditions on the phase shifts and $d_\lambda^\pm$ for nonphysical values of λ, so that condition (A4) is satisfied, is beyond the scope of this work. It is best to deal with such questions in a separate communication, in which we will consider the class of central and spin–orbit potentials which can be dealt with using this method.

To show that interchange of limit in the remaining term in (A1) is also possible, we note that for large values of r, the following inequalities hold:

$$\sum_{\lambda\in S}U_\lambda(r)a_\lambda L_\lambda^{\mu}(r)\leq Cr \tag{A7}$$

and

$$\left|F^{\mp}(r)h_\pm^2-F^{\pm}(r)\right|\leq C\left|\int_r^\infty ds\, sQ(s)\right|.$$

From (A7) it follows that:

$$\lim_{r\to\infty}\{[F^{\mp}(r)h_\pm^2-F^{\pm}(r)]\sum_{\mu\in S}U_\mu(r)a_\mu L_\mu^{\lambda}(r)\}=0 \tag{A8}$$

if

$$Q(r)=o(r^{-3})\quad\text{for large } r.$$

On the other hand, since each term in the above summation is bounded, we have

$$\sum_{\mu\in S}\lim_{r\to\infty}\{[F^{\mp}(r)h_\pm^2-F^{\pm}(r)]U_\mu(r)a_\mu L_\mu^{\lambda}(r)\}=0. \tag{A9}$$

From (A8) and (A9) it follows that the interchange of limit in the above summation is justified. Since the interchange of limit in all the terms in the right-hand side of (A1) is possible, it follows that interchange of limit in the summation on the left-hand side of (A1) is also justified.

*Supported in part by the Pahlavi University Research Council, Shiraz, Iran.

[1]R.G. Newton, J. Math. Phys. 3, 75 (1962).
[2]P.C. Sabatier, J. Math. Phys. 8, 905 (1967).
[3]P.C. Sabatier, J. Math. Phys. 9, 1241 (1968).
[4]N.F. Mott and H.S.W. Massey, *The Theory of Atomic Collisions* (Clarendon, Oxford, 1965).
[5]P.C. Sabatier, J. Math. Phys. 8, 1957 (1967).
[6]By $f(x)=o(g(x))$ for large values of x, we mean that $\lim_{x\to\infty}[f(x)/g(x)]=0$.
[7]P.C. Sabatier, J. Math. Phys. 7, 1515 (1966).
[8]E.L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956).
[9]Bateman Manuscript Projects, *Higher Transcendental Functions*, edited by A. Ederlyi (McGraw-Hill, New York, 1953).
[10]C is some finite constant. It does not have the same value every time it is used.

A theorem on the adiabatic scaling of classical orbits*

T. A. Witten Jr.

Joseph Henry Laboratories of Physics, Princeton University, Princeton, New Jersey 08540
(Received 19 July 1974)

The theorem describes the effect on the motion when the Hamiltonian is scaled, that is, when dimensioned parameters in the Hamiltonian are varied without changing any dimensionless parameters in it. The motion changes in a way that is not in general simple or predictable via dimensional analysis. But under certain conditions the motion scales with the Hamiltonian: Classical orbits scale in coordinate space and constants of the motion scale according to their dimensions. The theorem is based on the observation that the observable D which generates dilations may be expressed as a time derivative of a simple quantity. We find it most direct to use quantum mechanics to describe dynamics in the adiabatic limit. In so doing we develop methods that may be useful for other classical adiabatic problems.

I. INTRODUCTION

Sometimes one is led to study a secular variation of the Hamiltonian which is equivalent to a change of scale. In cosmology, for example, one would like to know the effect of a secularly changing gravitational constant G on celestial motions in general and on Kepler orbits in particular. The answer for the Kepler problem is already known.[1,2] In the limit of slowly changing G an orbit simply scales in time along with the Hamiltonian. Its shape and orientation in space remain constant while its size grows or shrinks to give the proper total energy as a function of time. These facts follow from the adiabatic invariance of the action variables $J_\alpha \equiv \oint q_\alpha \, dp_\alpha$. For certain judicious choices of the coordinates q_α the corresponding action variables can be shown to be locally computable[3] and to have simple interpretations in terms of the properties of the orbit. For example, the action J_ϕ associated with the azimuthal angle is precisely the angular momentum. Likewise the total energy and the Lenz vector

$$A_0 \equiv \frac{\bar{p} \times (\bar{x} \times \bar{p})}{m^2} - \frac{GmM\bar{r}}{|r|}$$

(which points toward the apogee) can be simply stated in terms of J_ϕ and the canonical actions associated with the parabolic coordinates η and ξ.[4] From the invariance of the J's one verifies directly, e.g., that the eccentricity remains constant, and that the apogee does not precess.

A parallel line of reasoning applies to motion in an isotropic oscillator potential. There also relations have been discovered between the orbit parameters and the canonical action integrals. For this system the end result is the same as for the Kepler problem: The orbits scale when the strength of the interaction (the potential) is changed adiabatically.

These two cases suggest that any classical orbit scales if its Hamiltonian is adiabatically scaled in time. The purpose of this note is to examine this hypothesis. We find that it is partly correct. If the scaling transformation is of a restricted "canonical" form and if the interactions depend only on position then the orbits do scale with the Hamiltonian. Applied to the Kepler problem and the isotropic oscillator, this result only reiterates what is known. Its virtue is that it avoids the case-by-case approach of previous work, which relies on knowing the specific functional form of conserved quantities peculiar to each problem. Further, our result is in principle more general, since it applies to closed-orbit problems with an arbitrary number of particles in an arbitrary potential and does not rely on special symmetry conditions.[5] Our theorem is not a simple result of dimensional analysis. From dimensional analysis one can learn that a scaled orbit is a *possible* motion in a scaled Hamiltonian. The theorem shows that the *actual* motion scales during a slow scaling of the Hamiltonian in time under the proper conditions.

One necessary condition is that the change take place adiabatically. For in the Kepler problem if G is *suddenly* increased by a large factor the orbit certainly does *not* scale. Instead, the point along the orbit where the change was made becomes the apogee while the eccentricity increases nearly to one. Even with adiabatic changes of the Hamiltonian the orbit need not scale in general. To see this, consider a Hamiltonian of the form

$$H = \frac{p^2}{2m} - \frac{e^2}{(x^2+y^2)^{1/2}} + \tfrac{1}{2}m\omega^2 z^2. \tag{1}$$

The projection of the motion in the x-y plane is a Kepler orbit while the z motion is simple harmonic with angular frequency w. The initial conditions may be set to make the total motion periodic. From the three parameters specifying this Hamiltonian—m, e, and ω—it is not possible to form a dimensionless combination. Thus a change in any of these may be accomplished by a change of scale, i.e., by a change of units. Suppose then that the parameter ω is slowly changed. Then the period of the z motion changes while the period of the x-y motion remains unaffected. Certainly there is no scaling of the orbit; indeed, there is no longer an orbit. Thus not any type of scaling of the Hamiltonian produces scaling of the motion; we now go on to investigate a type of scaling that does produce scaling of the motion.

II. CANONICAL SCALING

We begin with a Hamiltonian H defined in $2d$-dimensional phase space which has at least one dynamical trajectory that is a closed orbit. Mathematically the Hamiltonian has some numerical value for every numerical value of the coordinates x_i and conjugate momenta p_i. We define a "scale transformation" as one which changes the numerical Hamiltonian function as though the units of physical quantities had been changed. That is, $H'(x_i p_i) \equiv c\,H(a x_i, b p_i)$. This transformation is equiva-

lent to changing the unit of length by a factor of a, the unit of momentum by a factor of b, and the unit of energy by a factor of c. The consequences of making a transformation like the above are not in general simple, as we saw above with the Hamiltonian of Eq. (1). We thus restrict ourselves to "canonical" scale changes—i.e., those which leave the units of action invariant. Expressed in the above form, a canonical scale change is one in which $b=a^{-1}$.

This "canonical" scale change consists of two types of transformation. One of these (the c factor) amounts simply to multiplying the Hamiltonian by a constant; the other is a canonical dilation in phase space. This latter can be generated by a function on phase space just as a rotation can be generated by the angular momentum function. Thus, an infinitesimal canonical scale change can be expressed in the form

$$H' = H + \gamma H + \alpha\{H, D\}, \tag{2}$$

where the small quantity γ is equal to $c-1$ and α is $a-1$. The function D appearing in the Poisson bracket is the generating function for dilations. If x_i denote the Cartesian position coordinates of the particles and p_i their conjugate momenta, then the function D which generates a dilation outward from the origin is given by

$$D = \tfrac{1}{2}\sum_i (x_i p_i + p_i x_i). \tag{3}$$

With this order of x and p the formula is valid both classically and quantum-mechanically. Indeed, it is readily verified that a wavefunction $\psi(x)$ is dilated by a factor of $1+\alpha$ under the action of $(1-i\alpha D)$:

$$(1-i\alpha D)\psi(x) = \psi(x) - \frac{\alpha}{2}\left(x\frac{\partial\psi}{\partial x} + \frac{\partial}{\partial x}(x\psi)\right)$$
$$= (1-\alpha)^{1/2}\psi((1-\alpha)x) + O(\alpha^2). \tag{4}$$

(The function D arises in classical mechanics in the proof of the virial theorem.[6])

What happens to the motion when the Hamiltonian is transformed continually in time by this scaling transformation? We can study the dynamics of this adiabatic change most conveniently from the point of view of quantum mechanics. Then we may readily recover the classical motion using the correspondence principle. Messiah's textbook[7] gives a concise formulation of the motion in the adiabatic limit. For this purpose the Hamiltonian is expressed in terms of projection operators $\underline{P}_n$

$$\underline{H} = \sum_n E_n(t)\underline{P}_n(t) \tag{5}$$

The $\underline{P}_n(t)$ project a state into the nth invariant subspace of the Hamiltonian, which has the eigenvalue $E_n(t)$. The $\underline{P}_n$ may be represented by $\sum_j |n,j\rangle\langle n,j|$, where the index j enumerates a basis of states in the nth invariant subspace. In this work we confine ourselves to the case of bounded motion, and hence the sum on n is a discrete sum. If the E_n and $\underline{P}_n$ vary sufficiently slowly,[8] an eigenstate $|n,j\rangle$ will develop in time according to

$$i\frac{\partial}{\partial t}|n,j\rangle = \frac{E_n}{\hbar}(t)|n,j\rangle + \frac{i\partial\underline{P}_n}{\partial t}|n,j\rangle. \tag{6}$$

The first term describes the ordinary time development that the state would undergo if the Hamiltonian were time dependent. The second term arises because the invariant subspaces are changing with time, and because a state in the nth subspace must stay in that subspace. This term produces a change in $|n,j\rangle$ just sufficient to move it into the current invariant subspace. The state changes as little as possible consistent with the requirement that it remain in the subspace $\underline{P}_n(t)$.

In our canonical scaling process the effect of the variation in the energy scale is simply to multiply all the original E_n by the time-dependent factor $c(t)$. The effect of the change in length and momentum scales is to transform the projection operators $\underline{P}_n$ by a dilation proportional to the length factor $a(t)$:

$$i\frac{\partial\underline{P}_n}{\partial t} = [\underline{P}_n(t), \underline{D}]\frac{1}{\hbar a}\frac{da}{dt} \tag{7}$$

where $\underline{D}$ is the dilation generator of Eq. (3). This operation dilates all the states in the $\underline{P}_n$ subspace, but it does not necessarily change the operator $\underline{P}_n$ itself. For example, if the various states spanning $\underline{P}_n$ simply rotated into one another under dilation, then the subspace as a whole would remain unchanged, $\underline{P}_n$ would commute with $\underline{D}$, and the operator $\underline{P}_n$ would remain unaffected by the dilation. In this case the state $|n,j\rangle$ would not be changed during an adiabatic dilation. In the general case $|n,j\rangle$ develops according to

$$i\frac{\partial|n,j\rangle}{\partial t} = c(t)\frac{E_n}{\hbar}|n,j\rangle - \frac{\dot{a}}{\hbar a}(\underline{D}|n,j\rangle - \underline{P}_n\underline{D}|n,j\rangle). \tag{8}$$

The first term in this expression changes the overall phase of the eigenstate; the second one causes it to dilate according to the change in length scale. The third term removes that part of the dilation that does not contribute to $\partial\underline{P}_n/\partial t$. We are interested here in the conditions necessary to cause the motion to scale along with the Hamiltonian. The third term spoils the scaling of $|n,j\rangle$; accordingly, we shall seek conditions under which this term vanishes.[9]

By forming a linear combination of the states $|n,j\rangle$ we may deduce the time development equation for a general state

$$i\frac{\partial|\psi\rangle}{\partial t} = \frac{H(t)}{\hbar}|\psi\rangle - \frac{\dot{a}}{\hbar a}\left(\underline{D} - \sum_n \underline{P}_n(t)\underline{D}\,\underline{P}_n(t)\right)|\psi\rangle. \tag{9}$$

The effect of the adiabatic change is evidently to replace the Hamiltonian by a new time development operator $\underline{H}_{\rm eff}$ given by

$$H_{\rm eff} = H - \frac{\dot{a}}{a}(D - \bar{D}), \tag{10}$$

where we have denoted the projected dilation operator $\sum\underline{P}_n\underline{D}\underline{P}_n$ by $\bar{D}$. This same $H_{\rm eff}$ must generate the time development of a Heisenberg operator $\underline{A}(t)$

$$\frac{dA}{dt} = -\frac{i}{\hbar}[A, H_{\rm eff}] + \frac{\partial A}{\partial t}, \tag{11}$$

where the last term takes account of explicit changes in the definition of A as a function on phase space. This term is important, for example, in computing the change in the Lenz vector. When G is varied in the Kepler problem the Lenz vector changes not only because the motion changes, but also because G appears in its definition. The time development equation must be valid classical-

ly also provided ($-i/\hbar$ times commutator) is replaced by a Poisson bracket everywhere. (In practice, one must also find the classical analog of the operator $\underline{D}$.)

III. THE VANISHING OF $\bar{D}$

We have argued that the motion scales with the Hamiltonian provided the term in $\bar{D}$ gives no contribution. In this section we find the classical interpretation of $\bar{D}$, and we find a broad class of Hamiltonians for which $\bar{D}$ vanishes. We may find the physical significance of $\bar{D}$ by taking its expectation value in an arbitrary state. (Questions of adiabatic change do not concern us at the moment, so that we may consider the Hamiltonian to be fixed.)

$$\langle\psi(t)|\bar{D}|\psi(t)\rangle=\sum_{\substack{nn'\\jj'\\l}} a^*_{n'j'}a_{nj}\exp(iE_{n'}t)\exp(-iE_nt)\times\langle n',j'|\underline{P}_l\underline{D}\underline{P}_l|nj\rangle. \tag{12}$$

Evidently the only terms that can contribute in this sum are those for which $n'=n=l$. This means that $E_{n'}=E_n$, and the time dependence drops out:

$$\langle\psi(t)|\bar{D}|\psi(t)\rangle=\sum_{njj'} a^*_{nj'}a_{nj}\langle n,j'|\underline{D}|nj\rangle. \tag{13}$$

The expectation value of the bare dilation operator $\bar{D}$ may be expressed as in Eq. (12), except that the $\underline{P}_l$'s are absent and thus all the time-dependent terms are present. Apparently one may extract $\langle\bar{D}\rangle$ from $\langle D\rangle$ by simply taking the time-independent terms. In other words, $\bar{D}$ is the time average of D.

Now we seek a condition that would make the time average of D vanish. In periodic motion the average would vanish if D were the time derivative of some quantity I, for then

$$\langle\bar{D}\rangle=\frac{1}{T}\int_0^T\langle D\rangle\,dt=\frac{1}{T}\int_0^T\frac{d}{dt}\langle I\rangle\,dt=\frac{I(T)-I(0)}{T}=0. \tag{14}$$

As it happens, D does have this form for interactions that depend only on coordinates. Whenever the interaction is of this form the canonical momentum p_i is proportional to the corresponding velocity $\dot{x}_i$. Then the quantity I we seek is simply the trace of the inertial tensor.

$$I=\tfrac{1}{2}\sum_i m_i x_i^2. \tag{15}$$

To verify rigorously that the operator $\bar{D}$ must vanish in this case, we express it in the form

$$\bar{\underline{D}}=\sum_n\underline{P}_n\left(-\frac{i}{\hbar}(\underline{I},\underline{H}]\right)\underline{P}_n=-\frac{i}{\hbar}\sum_n[\underline{P}_n\underline{I}\underline{P}_n,\underline{H}]=0. \tag{16}$$

The last step follows because the projection operators force $\underline{P}_n\underline{I}\underline{P}_n$ to be diagonal with $\underline{H}$.

We have suggested that it was the presence of $\bar{D}$ that prevented the motion from scaling when the Hamiltonian is scaled. We now substantiate this assertion by showing that if $\bar{D}=0$, orbits scale and conserved quantities scale. A conserved quantity A obeys the equation of motion

$$\frac{dA}{dt}=+\frac{i\dot{a}}{\hbar a}[A,D]+\frac{\partial A}{\partial t}, \tag{17}$$

where we have assumed that $\bar{D}$ vanishes and used $[\underline{H},\underline{A}]=0$. We wish to prove that A scales according to its dimensions. That is, if A has dimensions of (length)$^m\times$(momentum)$^n\times$(energy)^k, then it expands according to

$$A(t)=(a(t))^{n-m}(c(t))^kA(0). \tag{18}$$

If A is a vector, its direction must evidently remain unchanged. The above is just what will happen if all the quantities defining A scale in time according to *their* dimensions. This must be true since A must be a dimensionally consistent function of its arguments. These arguments are of two kinds: (1) phase space coordinates x_i and p_i, and (2) parameters s_j taken from the Hamiltonian. The quantity A will expand properly according to Eq. (18) if it obeys.

$$\frac{dA}{dt}=-\frac{\dot{a}}{a}\sum_i\frac{\partial A}{\partial x_i}x_i-\frac{1}{a^{-1}}\frac{d}{dt}a^{-1}\sum_i\frac{\partial A}{\partial p_i}p_i+\sum_j\frac{\partial A}{\partial s_j}\dot{s}_j, \tag{19}$$

where the s_j are presumed to expand analogously to Eq. (18), this being what we mean by a scaling transformation of the Hamiltonian.

The actual time development of A may be expressed as

$$\frac{dA}{dt}=\sum\frac{\partial A}{\partial x_i}\dot{x}_i+\sum\frac{\partial A}{\partial p_i}\dot{p}_i+\frac{\partial A}{\partial t}. \tag{20}$$

We may show this equivalent to the desired form, Eq. (19), by writing out $\dot{x}_i$ and $\dot{p}_i$ in terms of H_{eff} and writing $\partial A/\partial t$ in terms of the s_j:

$$\frac{dA}{dt}=\sum_i\frac{\partial A}{\partial x_i}\left(-\frac{i}{\hbar}\left[x_i,\left(H-\frac{\dot{a}}{a}D\right)\right]\right)+\sum_i\frac{\partial A}{\partial p_i}\times\left(-\frac{i}{\hbar}\left[p_i,\left(\underline{H}-\frac{\dot{a}}{a}\underline{D}\right)\right]\right)+\sum_j\frac{\partial A}{\partial s_j}\dot{s}_j. \tag{21}$$

Rearranging

$$\frac{dA}{dt}=\left\{\sum_i\frac{\partial A}{\partial x_i}\left(-\frac{i}{\hbar}[x_i,H]\right)+\frac{\partial A}{\partial p_i}\left(-\frac{i}{\hbar}[p_i,H]\right)\right\}-\frac{\dot{a}}{a}\left[\sum_i\frac{\partial A}{\partial x_i}\left(-\frac{i}{\hbar}[x_i,\underline{D}]\right)+\frac{\partial A}{\partial p_i}\left(-\frac{i}{\hbar}[p_i,\underline{D}]\right)\right]+\sum_j\frac{\partial A}{\partial s_j}\dot{s}_j. \tag{22}$$

The expression in braces vanishes, since by hypothesis A is a conserved quantity. The commutator-times-minus $i/\hbar$ (Poisson bracket) of $\underline{D}$ with $\underline{x}_i$ gives $\underline{x}_i$, and with $\underline{p}_i$ it gives $-\underline{p}_i$. Thus

$$\frac{dA}{dt}=-\frac{\dot{a}}{a}\sum_i\frac{\partial A}{\partial x_i}x_i+\frac{\dot{a}}{a}\sum_i\frac{\partial A}{\partial p_i}p_i+\sum_j\frac{\partial A}{\partial s_j}\dot{s}_j. \tag{23}$$

This expression is clearly equal to Eq. (19), thus proving our assertion that A scales according to its dimensions.

We now turn to the scaling of classical orbits. We assume that the initial motion was in a closed orbit with orbital period T. Then we wish to show that during the adiabatic scale change the orbit retains its orientation and shape but expands by a factor of a^{-1}. Similarly, the orbital period should expand by a factor of c^{-1}. One way to show this scaling might be to characterize the orbit

in terms of conserved quantities A and then to use the scaling of A that has just been established. We prefer instead to use a more physically transparent method based on the classical limit of quantum mechanics. This method relies on the fact that an "orbit eigenstate" $|n\rangle(t)$ may be constructed which (1) is an eigenstate of the quantum Hamiltonian, and (2) describes a particle confined near the locus of the classical orbit in d-dimensional configuration space. We may construct this state as a linear combination of wave packets $\psi(x,t)$ corresponding to particles moving along the orbit

$$\langle x|n\rangle(t)=\int_0^T f(s)\,\psi(x,t-s)\,ds. \tag{24}$$

The states $\psi(x,t)$ may be chosen to be tightly localized near the position of the classical particle at time t; this follows from the correspondence principle. Thus it is clear that the particle described by $|n\rangle$ must lie near the locus of the orbit. We show in the Appendix that the coefficients $f(s)$ may be chosen to make $|n\rangle$ an eigenstate corresponding to an eigenvalue E_n close to the classical energy of the orbit. We also show there that the energy spacing $E_{n+1}-E_n$ near E_n is equal to the orbital frequency times $2\pi\hbar$.

From Eq. (8) the state $|n\rangle$ evolves according to

$$i\frac{\partial|n\rangle}{\partial t}=\frac{c(t)}{\hbar}E_n|n\rangle-\frac{\dot a}{\hbar a}\underline{D}|n\rangle \tag{25}$$

where we have assumed that $\bar D$ vanishes. The first term gives $|n\rangle$ a time-dependent phase factor, which has no effect on the probability distribution in space. The second term causes the state to dilate by a factor of a^{-1} over time. Thus,

$$|\langle x|n\rangle(t)|^2=a^d|\langle ax|n\rangle(0)|^2. \tag{26}$$

Evidently after the scaling process the particle remains confined to an orbit, and this orbit is a dilation of the original one. Since the energies and energy difference scale by a factor of c, the orbital frequency must also. This is what we wished to prove.

IV. CONCLUSION

In the above development we have established the following Theorem. Given (1) a system of particles interacting via a potential that depends only on position, (2) that the initial motion is periodic with period T, and (3) that parameters in the Hamiltonian are changed adiabatically in a way equivalent to a continuous change of length and energy (but not action) units, then (1) a conserved quantity remains a conserved quantity and scales according to its physical dimensions, and (2) the motion remains periodic with a suitably scaled period and an orbit unchanged except for an overall expansion in size given by the expansion of the length scale. The theorem applies to the Kepler problem and the harmonic oscillator problem and thus it establishes their scaling properties within a general framework. The result shows directly the scaling of other known periodic motions such as that arising from the Hamiltonian of Eq. (1). The result may prove useful for more realistic N-particle motions in celestial mechanics[10] and in plasma physics.[11] Moreover, the techniques used to obtain this theorem may prove valuable in other problems involving adiabatic invariants. The use of quantum dynamics to describe adiabatic motions appears attractive, since it is simply stated and avoids the use of the delicate remappings and time averages often needed in the classical case.[3]

ACKNOWLEDGMENTS

It is a pleasure to thank Freeman Dyson for suggesting this problem, and Alan Glasser and Richard Palmer for discussions.

APPENDIX: CONSTRUCTION OF THE ORBIT WAVEFUNCTIONS

Let us suppose that the classical particle has an energy equal to the nth eigenvalue E_n of the quantum Hamiltonian. The wave packet describing this particle is composed of a superposition of eigenstates having energies near E_n,

$$\psi(\mathbf{r},t)=\sum_{k,j} b_{kj}\langle\mathbf{r}|k,j\rangle\exp(iE_k t), \tag{A1}$$

where k labels the energy level and j denotes the other quantum numbers necessary to specify the state. The classical motion is periodic with period T; thus the wave packet function ψ must also have this period[12] (up to a phase factor). This means that the spacing of the energy levels near E_n must be $2\pi/T$ or multiples thereof, as is evident from Eq. (A1). We now form the superposition of ψ states which is to yield the orbit wavefunction, denoted Φ,

$$\Phi(\mathbf{r},t)=\int_0^T\exp(iE_n\tau)\,\psi(r,t-\tau)\,d\tau. \tag{A2}$$

(The function Φ is not normalized, since the integral is not taken with respect to orthogonal states.) In terms of the $|k,j\rangle$ states Φ has the form

$$\Phi(\mathbf{r},t)=\sum_{k,j} b_{kj}\langle\mathbf{r}|k,j\rangle\exp(iE_k t)$$

$$\times\int_0^T\exp(iE_n\tau)\exp(-iE_k\tau)\,d\tau. \tag{A3}$$

Because of the spacing of the levels, the integral must vanish when $k\neq n$, and hence

$$\Phi(\mathbf{r},t)=T\sum_j b_{kj}\langle\mathbf{r}|n,j\rangle\exp(iE_n t). \tag{A4}$$

Evidently Φ represents a pure eigenstate. Moreover, it must vanish except near the classical orbit since all the functions ψ do. The orbit eigenstate thus has the properties claimed for it.

*Research supported in part by the Air Force Office of Scientific Research under Contract F44620-71-C-0108.

[1]Roman U. Sexl, Acta Phys. Austr. 22, 159 (1966).
[2]M. J. Englefield, *Group Theory and the Coulomb Problem* (Wiley, New York, 1972).
[3]M. Kruskal, J. Math. Phys. 3, 806 (1962).
[4]In detail, if ϕ is the angle measured around the z axis and if η and ξ are defined symmetrically about the z axis, then

$$E=\frac{G^2m^3M^3}{2(m+M)(J_\eta+J_\xi+|J_\phi|)^2}, \quad (A_\phi)_z=GmM\left(\frac{J_\eta-J_\xi}{J_\eta+J_\xi+|J_\phi|}\right).$$

See L.D. Landau, and E.M. Lifshitz *Quantum Mechanics* (Addison-Wesley, Reading, Mass., 1965), 2nd ed., p. 125ff. The eccentricity ϵ is given in terms of E and the angular momentum L: $\epsilon=(1+2EL^2/G^2M^2m^3)^{1/2}$.

[5]The existence of periodic motions in the first place may, however, require a special symmetry in the problem.

[6]H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, Mass., 1950), p. 69.

[7]Albert Messiah, *Quantum Mechanics*, 5th printing (Wiley, New York, 1966), p. 744.

[8]$\underline{P}_n$ must not vary appreciably during the time $\hbar(E_{n+1}-E_n)^{-1}$ for any populated state $|n,j\rangle$.

[9]Actually, it is possible for the state to scale even when the $\underline{P}_nD$ term does not vanish. It is sufficient for it to have the same form as the E_n term, as this would simply introduce an extra phase factor into the eigenstate.

[10]See, e.g., A. Ollongren, in *Galactic Structure*, edited by A. Blaauw and M. Schmidt (University of Chicago Press, Chicago, 1965).

[11]T.G. Northrop and E. Teller, Phys. Rev. 117, 215, (1960).

[12]For finite n the wave packet may undergo a slow spreading motion in addition to its periodic motion. This spreading motion may be made as small as desired by increasing n and reducing the range of E_k. We shall assume that this spreading motion has been made negligibly small on the time scale of the adiabatic variation.

Functional integration, Padé approximants, and the Schrödinger equation

A. Bove, G. Fano, and G. Turchetti

Istituto de Fisica dell'Università, Bologna, Italy

A. G. Teolis

Centro di Calcolo del CNEN, Bologna, Italy

(Received 15 July 1974)

A method for finding the eigenvalues and the eigenvectors of the Schrödinger equation is presented. If $H = T + V$ is an M-body Hamiltonian, we use Trotter's formula in the form $e^{-\beta T/n}\, e^{-\beta V/n} \sim e^{-\beta H/n}$ (for n large). This allows the computation of the matrix elements of $e^{-\beta H}$ in the configuration representation, and the moments $\mu_r = (\psi,(e^{-\beta H})^r\psi)$ $(r=1,2,\cdots)$ for any wavefunction ψ. From the moments μ_r we compute the $[N - 1/N]$ Padé approximant, whose poles are the approximate eigenvalues of $e^{-\beta H}$. The convergence of the method is proved and asymptotic formulas for the matrix elements of $e^{-\beta T}$ projected on states of given angular momentum are derived.

INTRODUCTION

The aim of this work is to present a method for finding the eigenvalues of the Schrödinger equation of a few-body system, combining the Trotter's formula and the method of moments.

Trotter's theorem not only provides a theoretical basis for proving the Feynman–Kac formula (see Ref. 1), but is also a powerful tool for the actual computation of the functional integral appearing in the Feynman–Kac formula.

Consider a quantum system of M spinless equal particles of mass m. Choosing $\hbar = m = 1$, the kinetic energy operator is

$$T = -\frac{1}{2}\sum_{i=1}^{3M}\frac{\partial^2}{\partial q_i^2}. \tag{1}$$

We denote by V the potential energy operator

$$(V\psi)(q) = V(q)\,\psi(q), \quad q \in R^{3M}, \quad \psi \in L^2(R^{3M})$$

and by $H = T + V$ the total Hamiltonian. $V(q)$ is a real measurable function bounded from below; throughout this paper we shall assume V such that H is self-adjoint and bounded from below. The scalar product in $L^2(R^{3M})$ will be denoted by $(\ ,\ )$.

Trotter's theorem can be formulated as follows (see Ref. 1):

Trotter's Theorem: Let T, V and $T + V$ be linear operators defined in a Hilbert space K, all self-adjoint and bounded from below. Then for all $\beta > 0$

$$\operatorname{s-lim}_{n\to\infty}[\exp(-\beta T/n)\exp(-\beta V/n)]^n = \exp[-\beta(T+V)]. \tag{2}$$

It follows that

$$\operatorname{s-lim}_{n\to\infty}[\exp(-\beta V/2n)\exp(-\beta T/n)\exp(-\beta V/2n)]^n = \exp[-\beta(T+V)]. \tag{2'}$$

We denote by $A^1(\beta)$ the operator $\exp(-\beta V/2)\exp(-\beta T)\times\exp(-\beta V/2)$, by $A(\beta)$ the operator $\exp(-\beta H)$, and by $|q|$ the distance $(\sum_{i=1}^{3M} q_i^2)^{1/2}$; then the kernel corresponding to the operator $A^1(\beta/n)$ is

$$A^1(\beta/n, q_1, q_2) = \left(\frac{2\pi\beta}{n}\right)^{-3M/2} \exp[-(\beta/2n)V(q_1)] \times\exp(-|q_1 - q_2|^2 n/2\beta)\exp[-(\beta/n)V(q_2)]. \tag{3}$$

Formulas (2), (3) have been widely used by Storer and collaborators[2–5] in order to compute the bound states of a two-or three-body system. They devoted their attention especially to the ground state problem.

Choosing a finite set of base points q_j $(j = 1, 2, \ldots, R)$ for the numerical integration, they replace the kernel (3) by an $R \times R$ matrix; taking n very large, they obtain an approximation scheme for computing $\langle\psi_i, e^{-\beta H}\psi_j\rangle$ in any orthonormal basis $\{\psi_i\}$ of $L^2(R^{3M})$. Furthermore, for large β and normalized ψ, the expectation value $(\psi, e^{-\beta H}\psi)$ tends to $e^{-\beta E_0}$ where E_0 is the ground state energy.

We propose here an improvement of the method which makes use of the good features of the Padé approximants, and is especially suited to compute the excited energy levels. We consider the set of moments

$$\mu_r = (\psi, A(\beta)^r\psi) = (\psi, A(r\beta)\psi), \quad r = 0, 1, 2, \ldots, 2N-1, \quad \psi \in L^2(R^{3M}), \ (\psi,\psi) = 1. \tag{4}$$

This set uniquely determines the eigenvalues of the operator $A_N = P_N A(\beta) P_N$, where P_N is the projection operator on the subspace H_N generated by ψ, $A(\beta)\psi$, $A(\beta)^2\psi, \ldots, A(\beta)^N\psi$. It is well known (see Ref. 6) that if ψ is such that $\cup_{N=1}^{\infty} H_N$ is dense in $L^2(R^{3M})$, A_N converges strongly to $A(\beta)$ as $N \to \infty$. Since the operator $A(\beta)$ is bounded, no domain questions are involved in the choice of ψ.

Formulas (2)–(4) allow us to compute numerically the moments μ_r $(r = 1, 2, \ldots, 2N-1)$ and the $[(N-1)/N]$ Padé approximant gives the expectation value

$$[(N-1)/N] = P_{N-1}(z)/Q_N(z) = (\psi, (I - zA_N)^{-1}\psi), \quad z \in C, \tag{5}$$

where P_{N-1} and Q_N are polynomials of degree $N-1$ and N, respectively.

As is well known, Q_N is given by

$$Q_N(z)=\det\begin{vmatrix}\mu_0 & \mu_1 & \cdots & \mu_N\\ \mu_1 & \mu_2 & \cdots & \mu_{N+1}\\ \cdot & \cdot & \cdot & \cdot\\ \mu_{N-1} & \mu_N & \cdots & \mu_{2N-1}\\ z^N & z^{N-1} & \cdots & 1\end{vmatrix} \tag{6}$$

2. CONVERGENCE OF THE METHOD AND ESTIMATE OF THE ERROR

We make the following additional assumption on the Hamiltonian $H=T+V$:

Assumption 1: In the interval $]-\infty,\ 0[$ there is only a discrete spectrum with a finite number of points $E_1 < E_2 \cdots < E_{n_D} < 0$.

Let $\beta>0$, and denote by $\sigma(A(\beta))$ the spectrum of $A(\beta)=e^{-\beta H}$. It follows from Assumption 1 that $A(\beta)$ has a finite number of eigenvalues $\lambda_1>\lambda_2\cdots>\lambda_{n_D}$ with $\lambda_{n_D}>1$. The remaining part of $\sigma(A(\beta))$ is contained in the interval $[0,1]$.

The following results are well known (see Refs. 6–8):
(a) If ψ is such that $\cup_{N=1}^{\infty} H_N$ is dense in the Hilbert space, $(I-zA_N)^{-1}$ strongly converges to $(I-zA)^{-1}$ for z^{-1} having positive distance from $\sigma(A(\beta))$. The convergence is locally uniform in z.

(b) Between two consecutive eigenvalues λ_{i-1}, λ_i $(i\leq n_D)$ there can be at most one eigenvalue $\lambda_i^{(N)}$ of A_N.

(c) For any $i\leq n_D$ there exists an $\bar{N}$ such that for $N>\bar{N}$ we have

$$\lambda_1^{(N)}\geq\lambda_1\geq\lambda_2^{(N)}\geq\lambda_2\cdots\geq\lambda_{n_D}^{(N)}\geq\lambda_{n_D}>1 \tag{7}$$

and

$$\lim_{N\to\infty}\lambda_i^{(N)}=\lambda_i\quad(i=1,\ 2,\ \ldots,\ n_D) \tag{8}$$

monotonically from above, unless ψ is orthogonal to the eigenvectors of H.

In order to estimate the "errors" $\Delta\lambda_i^{(N)}=\lambda_i^{(N)}-\lambda_i$ $(i=1,\ 2,\ \ldots,\ n_D)$, the following lemma is useful.

Lemma: For any integer $m>0$. there exists an $(m+n_D)\times(m+n_D)$ diagonal matrix $\tilde{A}$ and a mapping $\varphi:\psi\in L^2(R^{3M})\to v\in C^{m+n_D}$ such that

$$|\mu_r-\tilde{\mu}_r|\leq r/m\quad(r=1,\ 2,\ \ldots,\ 2N-1), \tag{9}$$

where μ_r is given by (4) and $\tilde{\mu}_r=\langle v,\tilde{A}^r v\rangle$. (By $\langle\ ,\ \rangle$ we denote the scalar product in C^{m+n_D}.)

Proof: Let

$$A=A(\beta)=\int_0^1\lambda dP(\lambda)+\sum_{i=1}^{n_D}\lambda_i P_i \tag{10}$$

be the spectral decomposition of $A(\beta)$ [with $P(\lambda)$ continuous from the right]. Let $\Delta_k^{(m)}$ denote the semiclosed interval $](k-1)/m,\ k/m]$, $k=1,\ 2,\ \ldots,\ m$, and $\xi(\Delta_k^{(m)})$ its characteristic function. Setting $P_k^{(m)}=P(k/m)-P((k-1)/m)=\xi(\Delta_k^{(m)})(A)$, the operator $B=\sum_{k=1}^{m}P_k^{(m)}k/m+\sum_{i=1}^{n_D}\lambda_i P_i$ uniformly approximates A:

$$\|B-A\|\leq 1/m. \tag{11}$$

On the other hand,

$$(\psi,B^r\psi)=\sum_{k=1}^{m}\left(\frac{k}{m}\right)^r\left|P_k^{(m)}\psi\right|^2+\sum_{i=1}^{n_D}\lambda_i^r\left|P_i\psi\right|^2 \tag{12}$$

$$(r=1,\ 2,\ \ldots,\ 2N-1).$$

Therefore, introducing the diagonal matrix $\tilde{A}$,

$$\tilde{A}_{11}=\frac{1}{m},\ \tilde{A}_{22}=\frac{2}{m}\cdots\tilde{A}_{m,m}=\frac{m}{m}=1,$$
$$\tilde{A}_{m+1,m+1}=\lambda_{n_D}\cdots\tilde{A}_{m+n_D,m+n_D}=\lambda_1, \tag{13}$$

and a mapping $\varphi:\psi\to v$ with

$$v_k=\left|P_k^{(m)}\psi\right|\quad(k=1,\ 2,\ \ldots,\ m),$$
$$v_{m+s}=\left|P_{n_D+1-s}\psi\right|\quad(s=1,\ 2,\ \ldots,\ n_D),$$

we can write

$$\tilde{\mu}_r=\langle v,\tilde{A}^r v\rangle=\sum_{k=1}^{m+n_D}(\tilde{A}_{kk})^r v_k^2=(\psi,B^r\psi). \tag{14}$$

It remains to prove that $|\mu_r-\tilde{\mu}_r|\leq r/m$ $(r=1,\ 2,\ \ldots,\ 2N-1)$. Let us put $P_c=I-\sum_{i=1}^{n_D}P_i$. Since $\|AP_c\|\leq 1$, $\|BP_c\|\leq 1$, and A and B commute, we have

$$|\mu_r-\tilde{\mu}_r|=|(\psi,(A^r-B^r)\psi)|=|(P_c\psi,(A^r-B^r)P_c\psi)|$$
$$=\left|\left(P_c\psi,(A-B)\left(\sum_{k=0}^{r-1}A^{r-k-1}B^k\right)P_c\psi\right)\right|$$
$$\leq\|A-B\|\sum_{k=0}^{r-1}\|A^{r-k-1}B^kP_c\|$$
$$\leq\frac{1}{m}\sum_{k=0}^{r-1}\|(AP_c)^{r-k-1}(BP_c)^k\|\leq\frac{r}{m}\cdot\qquad\text{QED}$$

In our case $\lambda_{n_D}=e^{-\beta E_{n_D}}$; therefore, it is possible to choose β in such a way that $\lambda_{n_D}>2$. Assuming that we are already in such a situation, we can prove the following.

Theorem: We define, for a normalized $\psi\in L^2(R^{3M})$,

$$a_i=\left|P_i\psi\right|,\quad i=1,\ 2,\ \ldots,\ n_D,$$

$$\gamma=\sup_{\substack{i,j\\ i\neq j}}\left|\frac{\lambda_i}{\lambda_i-\lambda_j}\right|,\quad\delta=\lambda_{n_D}-2>0.$$

The following inequality holds:

$$\left|\frac{\Delta\lambda_i^{(N)}}{\lambda_i}\right|\leq\frac{\gamma^{2(n_D-1)}}{a_i^2}\left(\frac{1}{1+\delta}\right)^{2(N-n_D)}. \tag{15}$$

Proof: We denote by $\tilde{\lambda}_i^{(N)}(m)$ the zeros of $\tilde{Q}_N(z)$ where $\tilde{Q}_N$ is obtained by Q_N [see Eq. (6)] replacing the moments μ_r by the moments $\tilde{\mu}_r$ considered in the Lemma. We have

$$|\Delta\lambda_i^{(N)}|=|\lambda_i-\lambda_i^{(N)}|\leq|\lambda_i-\tilde{\lambda}_i^{(N)}(m)|+|\tilde{\lambda}_i^{(N)}(m)-\lambda_i^{(N)}|,$$
$$i=1,\ 2,\ \ldots,\ n_D. \tag{16}$$

To the first term we can apply the following estimate (see Ref. 6):

$$|\lambda_i-\lambda_i^{(N)}(m)|\leq\frac{\lambda_i}{a_i^2}\frac{1}{(\lambda_i-1)^{2(N-n_D)}}\prod_{\substack{j=1\\ j\neq i}}^{n_D}\frac{\lambda_j^2}{(\lambda_j-\lambda_i)^2}$$
$$\leq\frac{\lambda_i}{a_i^2}\gamma^{2(n_D-1)}\left(\frac{1}{1+\delta}\right)^{2(N-n_D)}. \tag{17}$$

We notice that the rhs of (17) does not depend on m. Therefore, we can perform the limit for $m\to\infty$ in (16). Since for fixed N the coefficients of $\tilde{Q}_N$ are continuous functions of the moments, and the zeros $\tilde{\lambda}_i^{(N)}(m)$ are continuous functions of these coefficients we have, using the preceeding Lemma, that $\lim_{m\to\infty}|\tilde{\lambda}_i^{(N)}(m)-\lambda_i^{(N)}|=0$.

The theorem follows. It shows that we have at least geometric convergence to the true eigenvalues unless $a_i = 0$.

3. PROJECTION ON THE STATES WITH GIVEN ANGULAR MOMENTUM

Assume now that V is a two-body potential depending only on the interparticle distance. Let P_j be the projection operator on the subspace of $L^2(R^{3M})$ corresponding to states of total angular momentum j. We restrict the problem to such states. For simplicity we consider the case of Boltzmann statistics.

We need the kernel corresponding to the operator $A_j^1(\beta) = A^1(\beta)P_j = P_j A^1(\beta)P_j$. We denote by G the rotation group in R^3, by g an element of G, and by $R(g)$ the usual vector representation of G by 3×3 orthogonal matrices. Then the above kernel can be obtained applying $R(g)$ to $q_2 = (\mathbf{q}_{21}, \mathbf{q}_{22}, \ldots, \mathbf{q}_{2M})$ (with $\mathbf{q}_{2i} \in R^3$) in formula (3), multiplying by the character $\chi_j(g)$ of the irreducible representation of G with spin j, and averaging over the group. We obtain

$$A_j^1(\beta/n, q_1, q_2) = (2j+1)(2\pi\beta/n)^{-3M/2} \times \exp\left(-\frac{\beta}{2n}\sum_{i<j} V(|\mathbf{q}_{1i}-\mathbf{q}_{1j}|)\right) \times \left[\int d\mu_g\, \chi_j(g) \exp\left(-\sum_{i=1}^{M} |\mathbf{q}_{1i} - R(g)\mathbf{q}_{2i}|^2 \frac{n}{2\beta}\right)\right] \times \exp\left(-\frac{\beta}{2n}\sum_{i<j} V(|\mathbf{q}_{2i}-\mathbf{q}_{2j}|)\right), \tag{18}$$

where μ_g denotes the invariant measure over G.

In actual computations it is very useful to have approximate expressions of

$$I_j(\beta) = \int d\mu_g\, \chi_j(g) \exp\left(-\sum_{i=1}^{M} |\mathbf{q}_{1i} - R(g)\mathbf{q}_{2i}|^2 \frac{1}{2\beta}\right). \tag{19}$$

An asymptotic expansion of (19) for small β can be obtained by the saddle-point method. The expression

$$U(q_1, q_2, g) = \sum_{i=1}^{M} |\mathbf{q}_{1i} - R(g)\mathbf{q}_{2i}|^2 = \sum_{i=1}^{M} (|\mathbf{q}_{1i}|^2 + |\mathbf{q}_{2i}|^2 - 2\mathbf{q}_{1i}\cdot R(g)\mathbf{q}_{2i}) \tag{20}$$

can be interpreted as the harmonic potential energy of two rigid configurations q_1 and q_2 where the points $\mathbf{q}_{1i}$ and $\mathbf{q}_{2i}$ are linked by elastic strings. The configuration q_2 is free to rotate (for instance around the common center of mass at the origin of the coordinates). Hence there exists at least a minimum of U with respect to a set α_1, α_2, α_3 of parameters labeling an element g of the rotation group. Expanding U around the minimum point $g(\alpha^0)$, we have $(\alpha^0 = \alpha_1^0, \alpha_2^0, \alpha_3^0)$

$$U(q_1, q_2, g(\alpha)) = U(q_1, q_2, g(\alpha^0)) - \sum_{i,j=1}^{3}\sum_{l=1}^{M} \mathbf{q}_{1l}\cdot\frac{\partial^2 R}{\partial\alpha_i\partial\alpha_j}\bigg|_{\alpha^0}\mathbf{q}_{2l}(\alpha_i-\alpha_i^0)(\alpha_j-\alpha_j^0). \tag{21}$$

Denoting by $\rho(\alpha)$ the "density of elements" of the group G, the saddle-point method gives

$$I_j(\beta) \sim (2\pi\beta)^{3/2}\chi_j(\alpha^0)\rho(\alpha^0)|\det K|^{-1/2} \times \exp\left(-\sum_{i=1}^{M} |\mathbf{q}_{1i} - R(g_0)\mathbf{q}_{2i}|^2 \frac{1}{2\beta}\right) \tag{22}$$

where K is the 3×3 matrix

$$K_{im} = \sum_{l=1}^{M} \mathbf{q}_{1l}\cdot\frac{\partial^2 R}{\partial\alpha_i\partial\alpha_m}\bigg|_{\alpha^0}\mathbf{q}_{2l}.$$

Of course formula (22) holds only if α^0 gives a stable equilibrium position, i.e., $K<0$.

We notice that, due to the invariance of the measure μ_g, $A_{j=0}^1(\beta, q_1, q_2)$ depends only on the spatial invariants that can be formed from q_1 and q_2 separately (angles, volumes, etc.). $A_{j=0}^1(\beta, q_1, q_2)$ does not depend on the relative position of q_2 with respect to q_1.

The approximation scheme that we propose consists in choosing for n a power of 2, i.e., $n=2^m$; then using the appropriate base points and weights for the numerical integration, we repeatedly square the matrix $A_j^1(\beta/n, q_1, q_2)$:

$$A_j^{1(2^m)} = (\cdots((A_j^1)^2)\cdots)^2 \simeq A_j(\beta, q_1, q_2). \tag{23}$$

Finally, we choose $\psi \in L^2(R^{3M})$ and we compute the moments μ_r [see Eq. (4)] and the zeros of Q_N [see Eq. (6)]. By Trotter's formula and by inequality (15) we see that in principle any desired accuracy can be reached by taking m and N sufficiently large. Of course, the eigenvalue λ_i can be obtained only if $a_i \neq 0$. Approximate eigenvectors of $e^{-\beta H}$, i.e., eigenvectors of A_N, are easily determined by the usual methods (see, e.g., Ref. 6, p. 56).

Finally, we add two remarks. (1) Our approximation of Trotter's formula is equivalent to replacing the functional integral which appears in the Feynman–Kac formula by an integral over R^n with $n=2^m$. (2) Singular "hard-core" potentials can be treated without difficulty since the potential appears only through the factor $e^{-\beta V/2n}$.

Numerical calculations, using the IBM 370/165 of the Centro di Calcolo del CNEN, Bologna, and the CDC 6600 of the University of Bologna, have been performed in the two- and three-body cases. The results confirm the good features of the method and will be published elsewhere. See also Refs. 9 and 10.

[1] E. Nelson, J. Math. Phys. 5, 332 (1964).
[2] R.C. Grimm and R.G. Storer, J. Comp. Phys. 4, 230 (1969).
[3] R.C. Grimm and R.G. Storer, J. Comp. Phys. 5, 350 (1970).
[4] R.C. Grimm and R.G. Storer, J. Comp. Phys. 7, 134 (1971).
[5] R.C. Grimm, R.G. Storer, and B. Davies, J. Comp. Phys. 9, 538 (1972).
[6] Yu.V. Vorobyev, *Method of Moments in Applied Mathematics* (Gordon and Breach, New York 1965), p. 64.
[7] D. Bessis and M. Villani, "Perturbative variational approximations," Preprint DPh-T/74/27, Service de Physique Théorique, Saclay.
[8] D. Masson, "Hilbert space and the Padé approximant" in *The Padé Approximant in Theoretical Physics*, edited by G.A. Baker and J.L. Gammel (Academic, New York, 1970).
[9] A. Bove, G. Fano, and A.G. Teolis, "Functional integration for solving the Schroedinger equation," Technical Report RT/FIMA (73)1, CNEN, Bologna.
[10] G. Fano, "An application of Trotter's formula to a few-nucleon problem," in *The Nuclear Many-Body Problem*, edited by F. Calogero and C. Ciofi (Degli Atti, Roma, 1973).

Particles and simple scattering theory

J. V. Corbett*

Mathematics Department, Indiana University, Bloomington, Indiana
(Received 18 July 1974)

A necessary and sufficient condition for a pair of subspaces to be the in and out scattering subspaces for a simple scattering system is obtained. It involves the existence of certain representations of the Galilean presymmetry group on these subspaces. The physical interpretation is that, in scattering, particles remain as particles even in the presence of the interaction.

The purpose of this note is to give a necessary and sufficient condition for a pair of subspaces to be the in and out scattering subspaces for a simple scattering process.

Consider a system of two massive spinless particles whose interaction can be represented by a potential function V that depends only on the relative separation of the particles. The relative motion of the particles is governed by a Hamiltonian $H=H_0+V$ with domain dense in $L^2(\mathbf{R}^3)$. H_0 is the relative kinetic energy of the pair of particles. V is taken to be such that H_0+V is self-adjoint. Suppose that we are given a pair of subspaces $M^{\pm}(H)\subset L^2(\mathbf{R}^3)$ that reduce H; under what conditions are they scattering subspaces associated with a unitary S matrix?

In order to have a scattering theory we must identify the following structures;

(S_1) A pair of subspaces $\mathcal{H}_{in}(H)$, $\mathcal{H}_{out}(H)$, the subspaces of incoming and outgoing scattering states for the dynamics given by the Hamiltonian H. Both $\mathcal{H}_{in}(H)$ and $\mathcal{H}_{out}(H)$ reduce H.

(S_2) An S matrix $S(H)$ that is a unitary map from $\mathcal{H}_{out}(H)$ to $\mathcal{H}_{in}(H)$ and that intertwines the restrictions of the Hamiltonian H to $\mathcal{H}_{out}(H)$ and to $\mathcal{H}_{in}(H)$,

$$S(H):\mathcal{H}_{out}(H)\to\mathcal{H}_{in}(H);\quad S(H)H_+=H_-S(H). \tag{1}$$

In general these two conditions are not sufficient conditions for $S(H)$ to be a physical S matrix in the sense that observable quantities, such as cross sections, can be deduced from the properties of $S(H)$. The additional structure that gives a physical S matrix can be presented in many ways; the usual method is the following:

(S_3) Define a standard dynamics through a standard Hamiltonian H_0 for which the scattering subspaces are equal, $\mathcal{H}_{in}(H_0)=\mathcal{H}_{out}(H_0)=\mathcal{H}_{scat}(H_0)$, and the scattering matrix is trivial, $S(H_0)=I$. Then $S(H)$ for the nonstandard Hamiltonian H will be a physical S matrix if there exist isometries $\Omega_{\pm}$ from $\mathcal{H}_{scat}(H_0)$ to $\mathcal{H}_{in}(H)$ such that

$$\begin{aligned}\Omega_+H_0\big|_{\mathcal{H}_{scat}(H_0)}&=H\big|_{\mathcal{H}_{out}(H)}\Omega_+,\\ \Omega_-H_0\big|_{\mathcal{H}_{scat}(H_0)}&=H\big|_{\mathcal{H}_{in}(H)}\Omega_-,\end{aligned} \tag{2}$$

and $S(H)$ will be given by

$$S(H)=\Omega_-\Omega_+^*. \tag{3}$$

It should be noted that under these conditions both $\Omega_{\pm}$ are only defined up to a multiplicative constant of modulus 1. This nonuniqueness is usually absorbed by taking $S(H)$ to have the form

$$S(H)=I+T(H).$$

The usual standard dynamics for a nonrelativistic system of two particles is that governed by the free Hamiltonian H_0. In the relative coordinates $H_0=-\nabla^2/2\mu$. This is not the only candidate. It may be argued that the dynamics given by an external gravitational field should be employed. In this case the apparent effectiveness of the standard of free dynamics would be due to the homogeneity of the external gravitational field over the dimensions of the apparatus that is used in scattering experiments. It is easy to think of situations in which this effective homogeneity would not be a good approximation. It is also possible that other standards are necessary in other studies. For example we may wish to study the nuclear forces between a pair of protons by comparing the unbound motion of such a system with the unbound motion of a pair of particles of the same mass where only interaction is coulombic. In the following we take the free dynamics as the standard.

The standard model given by the free dynamics has the following characteristic:

(a) $\mathcal{H}_{scat}(H_0)=L^2(\mathbf{R}^3)$.

(b) The physical kinematic observables associated with the free particle of mass μ and spin 0 are given by the self-adjoint elements of the Lie algebra of the unitary irreducible projective representation, labelled by $(\mu,0)$, of the group G_t. We call this representation $U_\mu(g)$, $g\in G_t$.

The group G_t is the subgroup of the Galilean group obtained by deleting the temporal displacements. If $g\in G_t$,

$$g=(\mathbf{a},\mathbf{v},R)$$

where $\mathbf{a}$ is a spatial displacement, $\mathbf{v}$ is a boost (displacement in velocity space), and $R\in SU(2)$, the covering group of the rotations. It is the subgroup of the presymmetry group, called the kinematic subgroup, that maps a fixed-time hypersurface, $t=\text{const}$, onto itself.[1]

(c) The generators of the spatial displacements in the representation $U_\mu(g)$ of G_t are $\mathbf{P}$ (the triplet of momentum operators) and are related to the Hamiltonian H_0 by

$$H_0=\mathbf{P}\cdot\mathbf{P}/2\mu \tag{4}$$

on a dense subset of $L^2(\mathbf{R}^3)$.

The canonical irreducible unitary projective representation $U_\mu(g)$ of G_t is obtained from a true represen-

tation $U(\theta, g)$ of the covering group $G_{t,\mu}$ of G_t. $G_{t,\mu}$ is the central extension of G_t by the Abelian group of real numbers. Members of $G_{t,\mu}$ are ordered ordered pairs (θ, g), $\theta \in \mathbb{R}$, $g \in G_t$. The group law of multiplication is

$$(\theta', g')(\theta, g) = (\theta' + \theta + \xi_\mu(g', g), g'g), \tag{5}$$

where $w_\mu(g', g) = \exp[i\xi_\mu(g', g)] = \exp(i\mu \mathbf{v} \cdot R'\mathbf{a})$. The choice of $w_\mu(g', g)$ is not unique; equivalent ray representations can be obtained if $w'_\mu(g', g) = [\Phi(g)\Phi(g')/\Phi(g'g)]w_\mu(g', g)$ replaces the factor $w_\mu(g', g)$, where $\Phi(g)$ is a complex function of modulus 1 on the group G_t.[2] The choice given here leads to the usual projective representation $U_\mu(g)$ of G_t through the formula

$$U_\mu(g) = \exp(-i\theta)U(\theta, g),$$

where $U(\theta, g)$ is the true representation of $G_{t,\mu}$ with group multiplication given by Eq. (5):

$$(U_\mu(g)\Psi)(\mathbf{x}) = \exp(i\mu \mathbf{a} \cdot \mathbf{V}/2 - i\mu \mathbf{V} \cdot \mathbf{x}')\Psi(\mathbf{x}'),$$

where $\mathbf{x}' = R\mathbf{x} + \mathbf{a}$.[3]

A different equivalent choice of the factor $w_\mu(g', g)$ would lead to a representation

$$U'_\mu(g) = \Phi(g)U_\mu(g).$$

It is usual to consider not G_t but G, the Galilean group, as the group whose representations are associated with a free particle, and the representation of G_t given here can be extended to a representation of G on account of the relation between P and H_0, given under (c). However, in the presence of an external field we can, in general, only assume the existence of a representation of G_t. Even if the external field can be represented by a potential function, the corresponding Hamiltonian need not be the generator of time translations for the representation of G induced from that of G_t. In fact, the next theorem can be viewed as a statement of the fact that it is only when the representation of G_t is carried by a space of scattering states that this extension of the representation of G_t to a representation of G is possible. The possible importance of representation of G_t as distinct from representations of G was first discussed in Ekstein's articles on presymmetry.[1,4]

We will now obtain necessary and sufficient conditions for a pair of subspaces to be scattering subspaces that are associated with a unitary S matrix. In the following $U_\mu(g)$ will stand for an irreducible representation labelled by $(\mu, 0)$. $H_+ = H|_{M^+(H)}$, $H_- = H|_{M^-(H)}$ are the restrictions of the Hamiltonian H to the subspaces $M^\pm(H)$ that reduce H.

Theorem 1: $M^\pm(H)$ are subspaces of scattering states if and only if there exists a pair of unitary irreducible projective representations of G_t, $U_\mu(g, \text{out})$ on $M^+(H)$, $U_\mu(g, \text{in})$ on $M^-(H)$ such that

$$H_+ = \mathbf{P}(\text{out}) \cdot \mathbf{P}(\text{out})/2\mu, \qquad H_- = \mathbf{P}(\text{in}) \cdot \mathbf{P}(\text{in})/2\mu, \tag{6}$$

where $\mathbf{P}(\text{out})$, $\mathbf{P}(\text{in})$ are the generators of the spatial translations in the corresponding representations $U_\mu(g, \text{out})$, $U_\mu(g, \text{in})$ of G_t.

Proof: The sufficiency is evident from the equivalence of the unitary irreducible representations of $G_{t,\mu}$ labelled by $(\mu, 0)$. Let W_+ be the unitary that intertwines $U_\mu(g, \text{out})$ and $U_\mu(g)$ and W_- that which intertwines $U_\mu(g, \text{in})$ and $U_\mu(g)$:

$$W_+ : L^2(\mathbb{R}^3) \to M^+(H), \qquad W_- : L^2(\mathbb{R}^3) \to M^-(H).$$

If $S(H) = W_- W_+^*$, it is a unitary map from $M^+(H)$ to $M^-(H)$ that intertwines $U_\mu(g, \text{out})$ and $U_\mu(g, \text{in})$. In particular

$$\mathbf{P}(\text{in})\, S(H) = S(H)\mathbf{P}(\text{out})$$

on a dense subset of $M^+(H)$, because $\mathbf{P}(\text{in})$, $\mathbf{P}(\text{out})$ are self-adjoint representatives of elements in the Lie algebra of G_t. But by Eq. (4) and the unitarity of $S(H)$,

$$H_- S(H) = S(H) H_+$$

on a dense subset of $M^+(H)$.

The necessity goes as follows. If $S(H) = \Omega_- \Omega_+^*$ is a unitary map from $M^+(H)$ to $M^-(H)$ with the $\Omega_\pm$ isometrics from $L^2(\mathbb{R}^3)$ to $M^\pm(H)$ such that

$$\Omega_- H_0 = H_- \Omega_- \quad \text{on } \mathcal{D}(H_0) \subseteq L^2(\mathbb{R}^3),$$

$$\Omega_+ H_0 = H_+ \Omega_- \quad \text{on } \mathcal{D}(H_0) \subseteq L^2(\mathbb{R}^3),$$

then we can define representations $U_\mu(g, \text{out})$, $U_\mu(g, \text{in})$ on $M^+(H)$ and $M^-(H)$ by the equations

$$U_\mu(g, \text{out}) = \Omega_+ U_\mu(g) \Omega_+^*, \qquad U_\mu(g, \text{in}) = \Omega_- U_\mu(g) \Omega_-^*.$$

These representations of G_t are unitary, irreducible projective representations on account of the isometric property of $\Omega_\pm$, and the fact that

$$U_\mu\left(g, \begin{matrix}\text{out} \\ \text{in}\end{matrix}\right) = e \exp(-i\theta)\Omega_\pm U(\theta, g)\Omega_\pm^*,$$

where $\Omega_\pm U(\theta, g)\Omega_\pm^*$ is a unitary irreducible representation of $G_{t,\mu}$. The representatives of the Lie algebra of G_t are mapped onto each other in such a way that

$$\mathbf{P}(\text{in}) = \Omega_- \mathbf{P} \Omega_-^* \quad \text{and} \quad \mathbf{P}(\text{out}) = \Omega_+ \mathbf{P} \Omega_+^*$$

are, respectively, the generators of the spatial translations for the representations $U_\mu(g, \text{in})$ and $U_\mu(g, \text{out})$. But $H_0 = \mathbf{P} \cdot \mathbf{P}/2\mu$ and therefore

$$H_- = \mathbf{P}(\text{in}) \cdot \mathbf{P}(\text{in})/2\mu \quad \text{and} \quad H_+ = \mathbf{P}(\text{out}) \cdot \mathbf{P}(\text{out})/2\mu.$$

As was mentioned in the discussion preceding the statement of this theorem, the theorem can be given the following form.

Corollary: The subspace $M^\pm(H)$ are subspaces of scattering states if and only if they each are the carrier spaces for a unitary irreducible projective representation $(\mu, 0, 0)$ of the Galilean group G, in which the representative of the generator of temporal translations is the restriction of H to $M^\pm(H)$, respectively.

If the motion-reversal invariance of the Hamiltonian H is used, a relation between $\mathbf{P}(H, \text{in})$ and $\mathbf{P}(H, \text{out})$ is obtained that permits a development of the correspondence between the previous conditions for the existence of a unitary S matrix and spectral properties of the Hamiltonian H.

We begin by assuming that a unitary S matrix $S(H)$ of

the form $S(H)=\Omega_-\Omega_+^*$ exists as a linear map from $\mathcal{H}_{\rm out}(H)$ to $\mathcal{H}_{\rm in}(H)$. The motion reversal operators Θ is an antiunitary map from $\mathcal{H}_{\rm out}(H)$ to $\mathcal{H}_{\rm in}(H)$ with the following property: If $\Phi\in\mathcal{H}_{\rm out}(H)$ is an outgoing scattering state with mean outgoing momentum $\mathbf{k}$,

$$(\Phi, \mathbf{P}(H,\text{out})\Phi)=\mathbf{k},$$

then $\Theta\Phi\in\mathcal{H}_{\rm in}(H)$ is an incoming scattering state with mean incoming momentum $-\mathbf{k}$,

$$(\Theta\Phi, \mathbf{P}(H,\text{in})\Theta\Phi)=-\mathbf{k}.$$

Since this holds for all $\Phi\in\mathcal{D}(\mathbf{P}(H,\text{out}))$,

$$\Theta^{-1}\mathbf{P}(H,\text{in})\Theta=-\mathbf{P}(H,\text{out}) \tag{7}$$

holds as an operator equality on $\mathcal{D}(\mathbf{P}(H,\text{out}))$. Moreover, $\mathcal{H}_{\rm out}(H)=\mathcal{H}_{\rm in}(H)=\mathcal{H}_{\rm scat}(H)$.

Theorem 2: If H is motion reversal invariant,

$$\Omega_+=\Theta^{-1}\Omega_-\Theta, \tag{8}$$

Proof: By definition of Ω_-, for all $\Psi\in\mathcal{D}(\mathbf{P})$

$$\Omega_-\mathbf{P}\Psi=-\Theta\mathbf{P}(H,\text{out})\Theta^{-1}\Omega_-\Psi,$$

but $-\mathbf{P}=\Theta\mathbf{P}\Theta^{-1}$ and $\Theta\mathcal{D}(\mathbf{P})=\mathcal{D}(\mathbf{P})$;

$$\therefore\quad \Omega_-\Theta\mathbf{P}\Theta^{-1}\Psi=\Theta\mathbf{P}(H,\text{out})\Theta^{-1}\Omega_-\Psi.$$

Let $x=\Theta^{-1}\Psi\in\mathcal{D}(\mathbf{P})$; then $\Theta^{-1}\Omega_-\Theta\mathbf{P}x=\mathbf{P}(H,\text{out})\Theta^{-1}\Omega_-\Theta x$.

Furthermore, $\Theta^{-1}\Omega_-\Theta$ is a linear unitary map from $L^2(\mathbb{R}^3)$ to $\mathcal{H}_{\rm scat}(H)$ and

$$\Theta^{-1}\Omega_-\Theta H_0\xi=H\Big|_{\mathcal{H}_{\rm scat}(H)}\Theta^{-1}\Omega_-\Theta\xi \tag{9}$$

for all $\xi\in\mathcal{D}(H_0)$. Since

$$H\Big|_{\mathcal{H}_{\rm scat}(H)}=H\Big|_{\mathcal{H}_{\rm out}(H)}=H_+,$$

we have the result that

$\Theta^{-1}\Omega_-\Theta=\exp(i\alpha)\Omega_+$ for some real numbers α. Since $\Omega_\pm$ are only defined up to such a multiplicative constant, we can take $\alpha=0$ without loss of generality.

The following theorem relates the existence of a physical S matrix to the spectral properties of H.

Theorem 3: If H is motion-reversal invariant and such that $M(H)=\mathcal{H}_{\rm a.c.}$, the subspace of absolute continuity of H, and $H_{\rm a.c.}$ has uniformly infinite spectral multiplicity, $\sigma_{\rm a.c.}(H)=(0,\infty)$, then there exists a unitary S matrix $S(H)$ that maps $M(H)$ onto $M(H)$.

Proof: Under the stated conditions $H|_{M(H)}=H_{\rm a.c.}$ is unitarily equivalent to H_0. Let W denote the unitary map from $L^2(\mathbb{R}^3)$ to $M(H)$ that intertwines H_0 and $H_{\rm a.c.}$. Since $H_0=\mathbf{P}\cdot\mathbf{P}/2\mu$, where $\mathbf{P}$ is the generator of spatial translations for the representation $U_\mu(g)$ of G_t on $L^2(\mathbb{R}^3)$, W maps this structure onto a representation $U_\mu(g,M)$ of G_t on $\mathcal{H}_{\rm a.c.}(H)$, for which $\mathbf{P}_M$ is the generator of spatial translations and $H_{\rm a.c.}=\mathbf{P}_M\cdot\mathbf{P}_M/2\mu$.

Therefore, we can take W as Ω_+, without loss of generality. Ω_- will then be given by Theorem 2,

$$\Omega_+=W,\quad \Omega_-=\Theta W\Theta^{-1},$$

and therefore, taking

$$S(H)=\Theta W\Theta^{-1}W^*,$$

we have constructed a unitary S matrix, and $M(H)=\mathcal{H}_{\rm a.c.}(H)$ is the subspace of scattering states.

It is worthwhile to compare this result with the usual formulation of scattering. We first note that the demand for asymptotic freedom is satisfied in the sense that the scattering states carry an irreducible representation of G_+, and hence, of the canonical commutation relations, that is unitarily equivalent to the free representation. The fact that the irreducible representation carried by these states is unitarily equivalent to the free representation is an accident due to the lack of inequivalent irreducible representations of the commutation relations for finite numbers of particles. In an analogous field theoretic model we could only expect that scattering states carry an irreducible representation of the commutation relations, the in and out representations would be unitarily equivalent to each other but inequivalent to the Fock space representation.

This analysis also illustrates the care that must be taken when describing the symmetry properties of the Hamiltonian. The Hamiltonian H may break a symmetry in one representation of the symmetry group, and yet the restriction of H will not break the symmetry in a representation that is carried on the reducing subspace. For example, the Hamiltonian need not be spherically symmetric for Theorem 1 to hold but the restriction of H to $M^\pm(H)$ will be rotationally invariant under the action of the representation $U_\mu(R,{}^{\rm out}_{\rm in})$ of the rotation group obtained from the representation of $U_\mu(g,{}^{\rm out}_{\rm in})$ on $M^\pm(H)$. Similarly the Hamiltonian H is not dilation invariant in general with respect to the representation of the dilation group generated by the self-adjoint operator $A=\frac{1}{2}(\mathbf{x}\cdot\mathbf{P}+\mathbf{P}\cdot\mathbf{x})$, which is contained in the enveloping algebra of the Lie algebra of G_+.[5] Nevertheless, $H/M^\pm(H)$ is dilation invariant relative to the dilation group generated by $A({}^{\rm out}_{\rm in})$ which is contained in the enveloping algebra of G_+ in the representation $U_\mu(g,{}^{\rm out}_{\rm in})$. The question then becomes which representation of G_+ is the physically correct representation, and this work indicates that the representation best suited to scattering states is that which is carried by the manifold of scattering states.

It is of some interest to note that a result similar to this has been obtained by Thomas,[6] using the algebraic approach to simple scattering. In his work Thomas uses representations of the Euclidean group to obtain the generalized wave operators that correspond to our $\Omega_\pm$.

As is indicated by Theorem 3, the model of scattering discussed in this paper is similar to the Kato–Kuroda model in which asymptotic completeness is taken to mean $M^+(H)=M^-(H)=\mathcal{H}_{\rm a.c.}(H)$.[7] In effect the content of this paper is merely the natural remark that scattering has to do with particles and particles are defined, in a nonrelativistic theory, by unitary irreducible representations of the Galilean group labelled by their asymptotic particles, that is, in the language of this paper, particles existing in the presence of the external field. It is

interesting that such a demand leads to the identification of the subspaces of scattering states with the states in the absolutely continuous subspace of H.

The extension of this result to particles with nonzero spin can also be handled by suitably changing the representation of the presymmetry group. The extension to many-particle systems is more complicated and will be discussed in a subsequent article.

*Present address: Department of Mathematical Physics, University of Adelaide, Adelaide, South Australia 5001.

[1]H. Ekstein, Phys. Rev. 153, 1937 (1967).
[2]J. M. Levy-Leblond, J. Math. Phys. 4, 776 (1963).
[3]J. V. Corbett, Phys. Rev. D 1, 3331 (1970).
[4]H. Ekstein, Phys. Rev. 184, 1315 (1969).
[5]H. A. Kastrup, Nucl. Phys. B 7, 574 (1968). For a comparison of some of approaches to simple scattering theory see B. Simon, *Quantum Mechanics for Hamiltonians Defined as Quadratic Forms* (Princeton U. P., Princeton, N.J., 1971), Chap. IV.
[6]L. E. Thomas, "On the algebraic theory of scattering," Geneva preprint, May 1973.
[7]S. T. Kuroda, Nuovo Cimento 12, 431 (1959).

Variational methods and nonlinear forced oscillations

D. Y. Hsieh

Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912
(Received 12 August 1974)

An approximate, direct variational method, simple in concept, and straightforward in application is presented to deal with the problem of forced oscillation of nonlinear systems. The general procedure is illustrated in detail by treating a particular example, i.e., the Duffing's equation. The same procedure is also applied to some other examples in mathematical physics.

I. INTRODUCTION

Direct methods in the calculus of variation are often useful for obtaining approximate solutions of differential equations. We usually start with the formulation of an equivalent variational problem, and then substitute in the functional to be varied some trial functions with adjustable parameters. The adjustable parameters are then determined by the condition that the functional be stationary. From the practical point of view, the success of this approach depends very much on the trial function we choose.

Literature abounds with the use of variational methods for the solution of both linear and nonlinear differential equations of various kinds.[1,2] However, to our knowledge, it has scarcely been applied to the problems of forced oscillation, especially the nonlinear forced oscillation. It is well known that the effect of an external forcing term on a nonlinear system is one of the most complicated problems in the theory of nonlinear oscillations.[3] Many interesting and challenging questions are raised from the study of these problems. We may mention, among others, the existence of asymptotic periodic solutions, the generation of subharmonic solutions, and the jump phenomenon as observed from numerical studies.[4]

Besides the numerical studies, most works on nonlinear oscillations adopt in essence the procedure of perturbation expansions. Therefore, the practical usefulness is limited to areas where the nonlinearity is not large. The variational scheme to be developed in this paper, however, is capable of dealing with the general nonlinear problem, although it has other defects. It is also an analytical method simple in concept, therefore we should be able to gain some insight from the approximate solution on this class of difficult problems.

In the following, we shall first briefly sketch the general procedure, then treat one example, i.e., the Duffing's equation, in detail to illustrate the procedure. Then we indicate how the same procedure can also be applied to some other well-known equations in mathematical physics.

II. GENERAL SCHEME OF THE VARIATIONAL METHOD

Consider an ordinary differential equation of the form

$$L(x^{(n)}, x^{(n-1)}, \ldots, x', x, \sin\omega t) = 0, \tag{II.1}$$

where we have used the notation $x' = dx/dt$, $x^{(n)}(t) = d^n x/dt^n$ and so on. A special case of (II.1) is the equation

$$L_0(x^{(n)}, x^{(n-1)}, \ldots, x', x) = K\sin\omega t. \tag{II.2}$$

When $K=0$, Eq. (II.2) may represent the free oscillation of a dynamical system; and when $K\neq 0$, it represents the forced oscillation of the system. Let us take the case that the solution (II.1) is such that the functional

$$J = \int_0^t F(x^{(n)}, x^{(n-1)}, \ldots, x', x, \sin\omega t)\, dt \tag{II.3}$$

is stationary. Thus a direct substitution of some trial solution $x(t)$ in (II.3) and the minimization of J will yield an approximate solution. However, in this case the indefinite upper limit t in the functional integral presents a serious difficulty. What we propose to do for finding an approximate solution is to expect that asymptotically a solution of the form

$$x(t) = A(t)\sin[\omega t + \delta(t)] + B(t) \tag{II.4}$$

may prevail. In the expression (II.4), we require that $A(t)$ and $\delta(t)$ are both slowly varying functions of t, and $B(t)$ is a slowly varying function of t, or a periodic function of t with periods different from $2\pi/\omega$. The first term will hopefully take care of the direct effect of the forcing terms, and the second term the free oscillations when the forcing term is absent. After substitution of (II.4) into (II.3), the evaluation of various integral terms in J is to be carried out in the following manner. We first express

$$\int_0^t = \sum_{j=1}^{m} \int_{(j-1)\Delta}^{j\Delta} dt + \int_{m\Delta}^{t} dt,$$

where Δ is either $2\pi/\omega$ or the period of $B(t)$, and $|t - m\Delta| < \Delta$. In each integral over a period, we take A and δ as constants, say $A(j\Delta)$ and $\delta(j\Delta)$. $B(t)$ and $\sin[\omega t + \delta(t)]$ are assumed to have an annihilative interference between them, thus terms with odd powers of $\sin[\omega t + \delta]$ will have vanishingly small contribution to the integral. When t is large, the contribution from $\int_{m\Delta}^{t} dt$ is insignificant compared with the rest. In this way, the effect of the factors $\sin[\omega t + \delta]$ and $\sin\omega t$ is explicitly integrated out, and then we revert the sums over functions of $A(j\Delta)$ and $\delta(j\Delta)$ to the original integral $\int_0^t dt$. Thus the functional J would be approximated by an expression as follows:

$$J = \int_0^t G(A^{(n)}, \ldots, A; \delta^{(n)}, \ldots, \delta; B^{(n)}, \ldots, B)\, dt. \tag{II.5}$$

From the expression (II.5), we can deduce three Euler–Langrange equations from the independent variation of A, δ, and B.

It may appear that the problem is getting even more

complicated than before since we now have three coupled nonlinear equations instead of one. In fact, the problem is greatly simplified since we have taken $A(t)$ and $\delta(t)$ as slowly varying functions of time. Often we may indeed take A and δ as constants to be our first approximation. Also, the forcing term which causes most of the difficulties is no longer there.

The form of trial solution (II.4) may not be adequate for certain situations. For instance, when subharmonic resonance is expected to occur, a trial function of the form

$$x(t)=A(t)\sin[\omega t+\delta(t)]+B(t)\sin[p\omega t+\nu(t)], \tag{II.6}$$

may be a better choice.

If the system is dissipative, and we cannot find a functional J whose variation will lead to the differential Eq. (II.1), we can still formulate the problem to be equivalent of

$$\Delta J+\Delta I=0, \tag{II.7}$$

where J is again of the same form as (3) and

$$\Delta I=\int_0^t E(x^{(n)},\ldots,x)\,\Delta x\cdot dt. \tag{II.8}$$

Then we proceed as before.

In what follows, we shall treat an example in considerable detail, and how the scheme just outlined works will become quite clear.

III. APPROXIMATE ASYMPTOTIC SOLUTIONS OF DUFFING'S EQUATION

Consider the Duffing equation

$$x''+x+\gamma x^3=K\sin\omega t, \tag{III.1}$$

where γ and K are real constants. When $\gamma=0$, the natural frequency of the system is 1. When $K=0$, the solution of the nonlinear differential equation can be expressed in terms of elliptic finctions. From (III.1) we can construct the functional

$$J=\int_0^t[\tfrac{1}{2}(x'^2-x^2-\tfrac{1}{2}\gamma x^4)+Kx\sin\omega t]\,dt. \tag{III.2}$$

Now let us take

$$x=A(t)\sin[\omega t+\delta(t)]+B(t), \tag{III.3}$$

where $A(t)$ and $\delta(t)$ are slowly varying functions of t, and $B(t)$ is such that

$$\int_0^t B^m\sin^n(\omega t+\delta)\,dt=O(t) \text{ for } m \text{ integers}, \ n \text{ odd integers}. \tag{III.4}$$

Thus $B(t)$ could be a slowly varying function of t, or a periodic function with periods different from $2\pi/\omega$, certain multiples or rational fractions of $2\pi/\omega$, or the product of such periodic functions and a slowly varying function of t. From (III.3) we obtain, for instance,

$$x'=A'\sin(\omega t+\delta)+(\omega+\delta')A\cos(\omega t+\delta)+B',$$

$$x^4=A^4\sin^4(\omega t+\delta)+4A^3B\sin^3(\omega t+\delta)+6A^2B^2\sin^2(\omega t+\delta)+4AB^3\sin(\omega t+\delta)+B^4.$$

Let us substitute these expressions in (III.2). Take a typical term $\int_0^t A^2B^2\sin^2(\omega t+\delta)\,dt$ as an example. Now, if $P(t)$ and $Q(t)$ are both periodic functions with period τ_P and τ_Q, respectively, let

$$\langle P^2\rangle=\frac{1}{\tau_P}\int_\tau^{\tau+\tau_P}P^2\,dt,\quad \langle Q^2\rangle=\frac{1}{\tau_Q}\int_\tau^{\tau+\tau_Q}Q^2\,dt.$$

Then $[P^2-\langle P^2\rangle]$ and $[Q^2-\langle Q^2\rangle]$ are both periodic functions with zero mean. If τ_P and τ_Q are different and one is not a rational multiple of the other, we obtain, when t is large,

$$\int_0^t[P^2Q^2-\langle P^2\rangle\langle Q^2\rangle]\,dt\approx\int_0^t[P^2-\langle P^2\rangle][Q^2-\langle Q^2\rangle]\,dt=o(t).$$

Thus we obtain

$$\int_0^t A^2B^2\sin^2(\omega t+\delta)\,dt\approx\int_0^t A^2\langle B^2\rangle\langle\sin^2(\omega t+\delta)\rangle\,dt$$
$$=\tfrac{1}{2}\int_0^t A^2\langle B^2\rangle\,dt.$$

We shall use the above result as far as A-variation is concerned. When B-variation is applied, we shall use $\frac{1}{2}\int_0^t A^2B^2\,dt$ instead. Thus we obtain after some straightforward computations

$$J\approx\int_0^t\{\tfrac{1}{4}A'^2+\tfrac{1}{4}[(\omega+\delta')^2-1]A^2-\tfrac{3}{32}\gamma A^4-\tfrac{3}{4}\gamma A^2\langle B^2\rangle+\tfrac{1}{2}(B'^2-B^2-\tfrac{1}{2}\gamma B^4)+\tfrac{1}{2}KA\cos\delta\}\,dt. \tag{III.5}$$

The Euler equations obtained from variations of A, δ, and B are

$$A''+[1-(\omega+\delta')^2]A+\tfrac{3}{4}\gamma A^3+3\gamma\langle B^2\rangle A=K\cos\delta, \tag{III.6}$$

$$\frac{d}{dt}[A^2(\omega+\delta')]=-KA\sin\delta, \tag{III.7}$$

$$B''+B+\gamma B^3+\tfrac{3}{2}\gamma A^2B=0. \tag{III.8}$$

Since $A(t)$ and $\delta(t)$ are supposed to be slowly varying functions of t. Thus as a first approximation, we may take both A and δ as real constants. Then we obtain

$$\sin\delta=0 \text{ or } \delta=0 \text{ or } \pi, \tag{III.9}$$

and

$$[1-\omega^2+3\gamma\langle B^2\rangle]A+\tfrac{3}{4}\gamma A^3=\pm K. \tag{III.10}$$

When A is constant, Eq. (III.8) can be solved and its solution can be expressed in terms of elliptical functions with A^2 as a parameter. Since $\langle B^2\rangle$ enters in the expression of A as a parameter, there is a coupling between Eqs. (III.8) and (III.10). The coupling is insignificant if γ is small.

For $\gamma=0$, i.e., the linear case, Eqs. (10) and (8) become, if we take $\delta=0$,

$$A=K/(1-\omega^2), \tag{III.11}$$

and

$$\ddot{B}+B=0, \tag{III.12}$$

and we recover the familiar result.

For the next approximation, let us take $\delta'\ll\omega$, and retain the A' term only; thus, we obtain

$$[1-\omega^2+3\gamma\langle B^2\rangle]A+\tfrac{3}{4}\gamma A^3=K\cos\delta, \tag{III.13}$$

and

$$2\omega\frac{dA}{dt}=-K\sin\delta. \tag{III.14}$$

Thus, we obtain

$$4\omega^2\left(\frac{dA}{dt}\right)^2=K^2-\{[1-\omega^2+3r\langle B^2\rangle]A+\tfrac{3}{4}rA^3\}^2. \tag{III.15}$$

It is again illuminating to consider the linear case for which $r=0$. Then the solution is readily found to be

$$A=\pm\frac{K}{1-\omega^2}\sin[(1-\omega^2)/2\omega]t+\beta\}. \tag{III.16}$$

We should recall that $A(t)$ is supposed to be a slowly varying function of time compared with $\sin\omega t$. It is evident unless $1-\omega^2\approx 0$, $A(t)$ will not be so. We should also note that A and δ both constants are solutions for Eqs. (III.6) and (III.7). Thus they represent more reasonable solutions, than (III.16) if $(1-\omega^2)$ is not small. For $1-\omega^2\approx 0$, to keep A finite, we shall take $\beta=0$; thus, as $1-\omega^2\to 0$, we obtain

$$A=\pm(K/2\omega)t, \tag{III.17}$$

while

$$\sin\delta=\mp 1 \quad \text{or} \quad \delta=\mp\pi/2.$$

Again we recover the familiar results for the linear case.

For the nonlinear case, the situation may not quite be the same. It is easy to see that the solution of (III.15) will be a periodic function of time. Thus unless tha period is large compared with $2\pi/\omega$, we cannot consider $A(t)$ as a slowly varying function of time. Hence the constant solutions of (III.9) and (III.10) are more appropriate. Even at the linear resonance, i.e., when $\omega^2=1$, the constant solutions may still be adequate, unless r is very small. When r is very small, we have

$$4\omega^2\left(\frac{dA}{dt}\right)^2=K^2-\frac{9r^2}{16}A^6, \tag{III.18}$$

where we have ignored the term with $\langle B^2\rangle$, since the inclusion of B when $\omega^2\approx 1$ is unnecessary. Equation (III.18) may be readily put in the form,

$$\frac{3r}{8\omega}dt=\frac{dA}{\pm[(4K/3r)^2-A^6]^{1/2}}. \tag{III.19}$$

Thus $A(t)$ will be a periodic function of t, with maximum amplitude $|4K/3r|^{1/3}$ and a rough estimate of the period $\tau\approx 10/K^{2/3}r^{1/3}$.

The plausibility of the constant solution given by (III.9) and (III.10) may be checked by a stability analysis based on Eqs. (III.6) and (III.7). Let the solution of (III.9) and (III.10) be denoted as δ_0 and A_0, and write $\delta=\delta_0+\delta_1$, $A=A_0+A_1$, then for small δ_1 and A_1, we have, to the first order of δ_1 and A_1

$$A_1''+[1-\omega^2+3r\langle B^2\rangle+\tfrac{9}{4}rA_0^2]A_1-2\omega A_0\delta_1'=0, \tag{III.20}$$

$$A_0^2\delta_1''\pm KA_0\delta_1+2\omega A_0A_1'=0. \tag{III.21}$$

There is no loss of generality to take $\delta_0=0$ and hence $+K$ in Eq. (III.21). Then we obtain

$$A_1^{iv}+[1+3\omega^2+\tfrac{9}{4}rA_0^2+3r\langle B^2\rangle+K/A_0]A_1''$$
$$+(K/A_0)[1-\omega^2+\tfrac{9}{4}rA_0^2+3r\langle B^2\rangle]A_1=0. \tag{III.22}$$

Let us put $A_1=\exp(\nu t)$; then after making use of (III.10), we obtain

$$\nu^2=\tfrac{1}{2}[-(4\omega^2+2K/A_0+\tfrac{3}{2}rA_0^2)\pm\{(\tfrac{3}{2}rA_0^2+4\omega^2)^2+16\omega^2K/A_0\}^{1/2}]. \tag{III.23}$$

Thus if $K/A_0>0$ and $r>0$, ν^2 will be real and negative and the constant solutions are stable. Otherwise, the constant solution may be unstable. We should note Eq. (III.10) may admit more than one real solution, the stability consideration can help to determine which branch to take.

IV. SUBHARMONIC OSCILLATIONS IN DUFFING'S PROBLEM

In order to deal with the phenomenon of subharmonic oscillations which is a common feature in nonlinear oscillation, we shall take the trial function, instead of (III.3):

$$x=A(t)\sin[\omega t+\delta(t)]+B(t)\sin[p\omega t+\nu(t)] \tag{IV.1}$$

where A, B, δ, and ν are all slowly varying functions of t and $p\neq 1$ is some positive real number. If $p=1$, then the second term can be absorbed into the first term.

Substitute (IV.1) into (III.2), and carry out the same procedure as we have done in the Sec. III, the Euler equations obtained from variations of A, δ, B, and ν become

$$A''+[1-(\omega+\delta')^2]A+\tfrac{3}{4}rA^3+\tfrac{3}{2}rB^2A=K\cos\delta+R_A, \tag{IV.2}$$

$$\frac{d}{dt}[A^2(\omega+\delta')]=-KA\sin\delta+R_\delta, \tag{IV.3}$$

$$B''+[1-(p\omega+\nu')^2]B+\tfrac{3}{4}rB^3+\tfrac{3}{2}rA^2B=R_B, \tag{IV.4}$$

and

$$\frac{d}{dt}[B^2(p\omega+\nu')]=R_\nu, \tag{IV.5}$$

where

$$\begin{aligned}&R_A=\tfrac{1}{4}rB^3\cos(3\nu-\delta)\\ &R_\delta=\tfrac{1}{4}rAB^3\sin(3\nu-\delta)\\ &R_B=\tfrac{3}{4}rAB^2\cos(3\nu-\delta)\\ &R_\nu=-\tfrac{3}{4}rAB^3\sin(3\nu-\delta)\end{aligned}\quad \text{for } p=\tfrac{1}{3}, \tag{IV.6}$$

$$\begin{aligned}&R_A=\tfrac{3}{4}rA^2B\cos(3\delta-\nu)\\ &R_\delta=-\tfrac{3}{4}rA^3B\sin(3\delta-\nu)\\ &R_B=\tfrac{1}{4}rA^3\cos(3\delta-\nu)\\ &R_\nu=\tfrac{1}{4}rA^3B\sin(3\delta-\nu)\end{aligned}\quad \text{for } p=3, \tag{IV.7}$$

and

$$R_A=R_\delta=R_B=R_\nu=0 \quad \text{otherwise}. \tag{IV.8}$$

Let us consider only the constant solutions such that $A'=\delta'=B'=\nu'=0$. Thus, for $p\neq\frac{1}{3}$ and $p\neq 3$, we obtain from (IV.3)

$$\sin\delta=0 \quad \text{or} \quad \delta=0, \tag{IV.9}$$

with no loss of generality. Equations (IV.2) and (IV.4) then lead to

$$A(1-\omega^2)+\tfrac{3}{4}r(A^3+B^2A)=K, \tag{IV.10}$$

$$1-(p\omega)^2+\tfrac{3}{4}r(B^2+2A^2)=0. \tag{IV.11}$$

For $p=\frac{1}{3}$, Eqs. (IV.3), (IV.5), and (IV.6) lead to

$$\delta=0 \text{ and } \sin(3\nu-\delta)=0. \tag{IV.12}$$

Then we obtain from (IV.2), (IV.4), and (IV.6)

$$A(1-\omega^2)+\tfrac{1}{4}r(3A^3+6B^2A\pm B^3)=K, \tag{IV.13}$$

$$1-(\omega/3)^2+\tfrac{3}{4}r(B^2+2A^2\pm AB)=0. \tag{IV.14}$$

Similarly, for $p=3$, we obtain from (IV.2)—(IV.5), and (IV.7)

$$A(1-\omega^2)+\tfrac{3}{4}r(A^3+2B^2A\pm A^2B)=K \tag{IV.15}$$

and

$$1-(3\omega)^2+\tfrac{3}{4}r(B^2+2A^2\pm A^3/3B)=0. \tag{IV.16}$$

These results are the so-called frequency response relations for subharmonic oscillations.[5] When the nonlinear effects are small, these simple harmonic solutions with constant amplitudes and phases should be very good approximation solutions.

V. DUFFING EQUATION WITH DAMPING

Let us now consider the Duffing equation with damping:

$$x''+\alpha x'+x+rx^3=K\sin\omega t, \tag{V.1}$$

where α is a real and positive constant. The equivalent variational formulation will be

$$\Delta J+\Delta I=0, \tag{V.2}$$

where

$$\Delta I=-\int_0^t[\alpha x'\Delta x]\,dt, \tag{V.3}$$

and J is again given by (III.2). Let us again take the trial function as (IV.1):

$$x=A(t)\sin[\omega t+\delta(t)]+B(t)\sin[p\omega t+\nu(t)], \tag{V.4}$$

where $p\neq 1$ is a positive constant.

Thus, for instance,

$$\begin{aligned}x'\Delta x=&[A'\sin(\omega t+\delta)+(\omega+\delta')A\cos(\omega t+\delta)\\&+B'\sin(p\omega t+\nu)+B(p\omega+\nu')\cos(p\omega t+\nu)]\\&\times[\sin(\omega t+\delta)\Delta A+A\cos(\omega t+\delta)\Delta\delta+\sin(p\omega t+\nu)\Delta B\\&+B\cos(p\omega t+\nu)\Delta\nu].\end{aligned}$$

Since A, B, δ, and ν as well as their variations are again supposed to be slowly varying functions of time, for large t, the approximate expression of ΔI can be evaluated by the same procedure as before. The Euler equation then becomes

$$A''+\alpha A'+[1-(\omega+\delta')^2]A+\tfrac{3}{4}rA^3+\tfrac{3}{2}rAB^2=K\cos\delta+R_A, \tag{V.5}$$

$$\frac{d}{dt}[A^2(\omega+\delta')]+\alpha(\omega+\delta')A^2=-KA\sin\delta+R_\delta, \tag{V.6}$$

$$B''+\alpha B'+[1-(p\omega+\nu')^2]B+\tfrac{3}{4}rB^3+\tfrac{3}{2}rA^2B=R_B. \tag{V.7}$$

and

$$\frac{d}{dt}[B^2(p\omega+\nu')]+\alpha(\omega+\nu')B^2=R_\nu, \tag{V.8}$$

where R_A, R_δ, R_B, and R_ν are again given by (IV.6)—(IV.8). If we again look for constant solutions, then some very interesting results are obtained. When A, B, δ, and ν are all independent of t, Eqs. (V.5)—(V.8) become

$$(1-\omega^2)A+\tfrac{3}{4}r(A^3+2AB^2)=K\cos\delta+R_A, \tag{V.9}$$

$$\alpha\omega A^2=-KA\sin\delta+R_\delta, \tag{V.10}$$

$$(1-p^2\omega^2)B+\tfrac{3}{4}r(B^3+2A^2B)=R_B, \tag{V.11}$$

and

$$\alpha\omega B^2=R_\nu. \tag{V.12}$$

From (V.12), we obtain immediately for $\alpha\neq 0$,

$$B=0, \tag{V.13}$$

unless $p=1/3$ or $p=3$. This means, if the system is dissipative, only the subharmonics and the harmonics of order 3 can persist.

When the nonlinear effect is small, i.e., when r is small, we also obtain readily from (V.9) and (V.10) the familiar result

$$A=\frac{K}{[(1-\omega^2)^2+(\alpha\omega)^2]^{1/2}} \tag{V.14}$$

and

$$\sin\delta=\frac{\alpha\omega/2}{[(1-\omega^2)^2+(\alpha\omega)^2]^{1/2}}. \tag{V.15}$$

Let us now investigate in more detail the case $p=1/3$, i.e., the subharmonic oscillation of order 3. For this case, Eqs. (V.11) and (V.12) have the form

$$1-\tfrac{1}{9}\omega^2+\tfrac{3}{4}r(B^2+2A^2)=\tfrac{3}{4}rAB\cos(3\nu-\delta), \tag{V.16}$$

$$\alpha\omega=-\tfrac{3}{4}rAB\sin(3\nu-\delta). \tag{V.17}$$

Eliminating the $(3\nu-\delta)$ term, we obtain

$$[1-\tfrac{1}{9}\omega^2+\tfrac{3}{4}r(B^2+2A^2)]^2+\alpha^2\omega^2=\tfrac{9}{16}r^2A^2B^2. \tag{V.18}$$

The last equation is of the form

$$(a+bB^2)^2+c=dB^2, \tag{V.19}$$

where

$$a=1-\tfrac{1}{9}\omega^2+\tfrac{3}{2}rA^2,\quad b=\tfrac{3}{4}r,$$

$$c=\alpha^2\omega^2,\quad d=\tfrac{9}{16}r^2A^2.$$

Thus we obtain

$$B^2=\frac{1}{2b^2}[d-2ab\pm\{d(d-4ab)-4b^2c\}^{1/2}]. \tag{V.20}$$

The solution is meaningful only when B^2 is real and positive. Since both c and d are positive, it is necessary that

$$d-4ab>0,$$

or

$$-3r(1-\tfrac{1}{9}\omega^2)-\tfrac{63}{16}r^2A^2>0 \tag{V.21}$$

Thus for $r<0$, the subharmonic oscillation can be

established only for $\omega < 3$. Let us keep in mind the expression of A for the linear case (V.14) as a background guide. Thus as ω decreases to toward the value 1, A may become rather large. This will offer a lower bound for the permissible value of ω. Further, since it is also necessary that

$$d(d-4ab) > 4b^2c.$$

Therefore, a minimum threshold value of A^2 or K^2 is required to overcome the interaction of damping and nonlinearity as expressed by $4b^2c = \frac{9}{4}\gamma^2\alpha^2\omega^2$.

This term $\frac{9}{4}\gamma^2\alpha^2\omega^2$ will also furnish an upper bound for the permissible value of ω if $\gamma > 0$. For the case $\gamma > 0$, the criterion (V.21) will require that $\omega > 3$.

The above analysis may again be best only when the nonlinear effect is not too large, since we have been using the simple harmonic solutions with constant amplitudes and phases. For the highly nonlinear system, the general Eqs. (V.5)–(V.8) need to be explored more fully.

VI. OTHER EXAMPLES

We now briefly describe the application of this variational method to a few other equations.

A. The Mathieu equation

For the Mathieu equation

$$x'' + (\alpha + \beta\cos 2t)x = 0, \tag{VI.1}$$

we take

$$J = \int_0^t \tfrac{1}{2}[x'^2 - \alpha x^2 - \beta x^2\cos 2t]\,dt. \tag{VI.2}$$

The Mathieu equation is a linear equation, and we shall let

$$x = \frac{A_0}{2} + \sum_{k=1} A_k\cos kt + \sum_{k=1} B_k\sin kt, \tag{VI.3}$$

where A_k, B_k's are all constants.

Substitute (VI.3) into (VI.2), and use the same procedure as we used in previous sections; we obtain

$$J \approx \tfrac{1}{2}\int_0^t dt\Big[-\tfrac{1}{4}\alpha A_0^2 - \tfrac{1}{4}\beta A_1^2 + \tfrac{1}{2}\sum_{m=1}(m^2-\alpha)A_m^2 - \tfrac{1}{2}\beta\sum_{m=0} A_m A_{m+2} + \tfrac{1}{4}\beta B_1^2 + \tfrac{1}{2}\sum_{m=1}(m^2-\alpha)B_m^2 - \tfrac{1}{2}\beta\sum_{m=0} B_m B_{m+2}\Big]. \tag{VI.4}$$

The variations of A_m and B_m lead to

$$\alpha A_0 + \beta A_2 = 0, \tag{VI.5}$$

$$(1-\alpha)A_1 - \tfrac{1}{2}\beta(A_1 + A_3) = 0, \tag{VI.6}$$

$$(m^2-\alpha)A_m - \tfrac{1}{2}\beta(A_{m-2} + A_{m+2}) = 0 \quad \text{for } m \geq 2, \tag{VI.7}$$

$$(1-\alpha)B_1 + \tfrac{1}{2}\beta(B_1 - B_3) = 0, \tag{VI.8}$$

and

$$(m^2-\alpha)B_m - \tfrac{1}{2}\beta(B_{m-2} + B_{m+2}) = 0. \tag{VI.9}$$

Equations (VI.5)–(VI.9) are the same as those may be found in standard tests.[6]

We may also apply the variational scheme to the nonlinear Mathieu equation

$$x'' + (\alpha + \beta\cos 2t)x + \gamma x^3 = 0. \tag{VI.10}$$

The corresponding functional is now

$$J = \int_0^t \tfrac{1}{2}(x'^2 - \alpha x^2 - \beta x^2\cos 2t - \tfrac{1}{2}\gamma x^4)\,dt. \tag{VI.11}$$

If we take as trial function

$$x = A\cos mt, \quad m \text{ integers}, \tag{VI.12}$$

then we obtain

$$J \approx \int_0^t \tfrac{1}{4}(m^2A^2 - \alpha A^2 - \tfrac{1}{2}\beta\delta_{m1}A^2 - \tfrac{3}{8}\gamma A^4)\,dt. \tag{VI.13}$$

The variation of J with respect to A leads to

$$(m^2-\alpha) - \tfrac{3}{4}\gamma A^2 = 0 \quad \text{for } m \neq 1 \tag{VI.14}$$

and

$$1 - \alpha - \tfrac{1}{2}\beta - \tfrac{3}{4}\gamma A^2 = 0 \quad \text{for } m = 1, \tag{VI.15}$$

which is the same as given elsewhere.[5]

B. The Van der Pol equations

The Van der Pol equation

$$x'' - \alpha(1-x^2)x' + x = K\sin\omega t, \tag{VI.16}$$

contains a dissipative term. Hence the corresponding variational problem is

$$\Delta J + \Delta I = 0, \tag{VI.17}$$

where

$$J = \int_0^t \left(\frac{x'^2}{2} - \frac{x^2}{2} + Kx\sin\omega t\right)dt, \tag{VI.18}$$

and

$$\Delta I = \int_0^t [\alpha(1-x^2)x'\cdot\Delta x]\,dt. \tag{VI.19}$$

Consider a trial function

$$x = A(t)\sin(\omega t + \delta(t)) + B(t), \tag{VI.20}$$

where $A(t)$, $\delta(t)$, and $B(t)$, their derivatives and variations are all slowly varying functions t or those functions which will not yield secular terms as we discussed in Sec. III. Then the application of our approximation scheme, after some computations, lead to

$$A'' - \alpha[1 - \tfrac{3}{4}A^2 - \langle B^2\rangle]A' + [1 - (\omega+\delta')^2]A + \alpha A\langle BB'\rangle = K\cos\delta, \tag{VI.21}$$

$$\frac{d}{dt}[A^2(\omega+\delta')] - \alpha(\omega+\delta')(A^2 - \tfrac{1}{4}A^4 - A^2\langle B^2\rangle) = -KA\sin\delta \tag{VI.22}$$

and

$$B'' - \alpha(1 - \tfrac{1}{2}A^2 - B^2)B' + \tfrac{1}{2}\alpha ABA' + B = 0. \tag{VI.23}$$

When constant solutions are sought, we found

$$B = 0, \tag{VI.24}$$

and

$$[(1-\omega^2)^2 + \alpha^2\omega^2(1 - \tfrac{1}{4}A^2)]A^2 = K^2. \tag{VI.25}$$

C. Nonlinear oscillation of a gas bubble in a liquid

The variational method has also been applied to the following equation:

$$RR'' + \tfrac{3}{2}R'^2 + \alpha RR' + AR - B/R^{3\gamma} + C = K\sin\omega t, \qquad \text{(VI.26)}$$

which describes the oscillation of a spherical bubble in a liquid in an external sinusoidal pressure field. The threshold subharmonic oscillations of order 2 and 3 can be established by the schemes developed in this paper. The detailed analysis of this problem is reported elsewhere.[7]

The variational method developed here is simple in concept and straightforward in application. We have demonstrated its usefulness by applying the method to some familiar equations as well as to not so familiar ones. For cases when the nonlinear effects are small, usually the only cases with available results since perturbation approach is ordinarily adopted, the established results are recovered. The variational method should be useful also when nonlinear effects are not small, and we have indicated the direction of the study one should follow. As is common for all variational approaches, the defect of this method is the difficulty to assess the accuracy of the solution. Hence a good trial solution either from experience or ingenuity is often essential for the successful application of this method. Therefore, much more work is needed to explore various aspects of this scheme. But let us emphasize one point here, i.e., this is an analytical method. Thus this method could offer much needed insight to the problem we study which purely numerical solutions often fail to accomplish.

ACKNOWLEDGMENTS

The study is supported by the Brown University MRL funded by National Science Foundation.

[1]S. H. Gould, *Variation Method for Eigenvalue Problem* (University of Toronto Press, Toronto, 1966).
[2]I. S. Sokolnifkoff, *Mathematical Theory of Elasticity* (McGraw-Hill, New York, 1956).
[3]N. V. Butenin, *Elements of the Theory of Nonlinear Oscillations* (Blaisdell, New York, 1965).
[4]H. T. Davis, *Introduction to Nonlinear Differential and Integral Equations* (Dover, New York, 1962).
[5]J. K. Hale, *Oscillations in Nonlinear Systems* (McGraw-Hill, New York, 1963).
[6]J. Mathews and R. L. Walker, *Mathematical Methods of Physics* (Benjamin, New York, 1970).
[7]D. Y. Hsieh, "Variational Methods and Nonlinear Oscillation of Bubbles" (to be published).

Degeneracies in energy levels of quantum systems of variable dimensionality

D. R. Herrick

Bell Laboratories, Murray Hill, New Jersey 07974
(Received 5 August 1974)

Introduction of variable dimensionality to the Schrödinger equation gives rise to interdimensional degeneracies in the energy levels of one-, two-, and three-electron atoms and molecules. In all cases the degeneracies result from a factorization of the wavefunction into a product of a "radial" type function times an "angular" type function. Scaling of the orbital angular momentum quantum number in the one-electron radial equation to obtain "excess angular momentum" is shown to be equivalent to a variation in dimensionality.

I. INTRODUCTION

This paper demonstrates, for the first time, the existence of degeneracies in the energy spectra of one-, two-, and three-electron atoms and molecules of variable dimensionality ($\equiv D$). Our interest in the D dependence of energy levels arose in an earlier investigation[1] in which the ground state energy of heliumlike ions was determined variationally as a continuous function of D. Comparison of these approximate energies with the exact ground state result for $D=1$ and 5 ions was made possible by the fact that the $D=1$ state is equivalent (by suitable scaling of length and energy) to the simple delta function model atom,[2] and the $D=5$ state is identical (to within trivial factors in the wavefunction) to the doubly excited $2p^2\,{}^3P^e$ state at $D=3$. Our current investigation shows that many such "interdimensional degeneracies" exist, and a study of the examples given may eventually lead to a formulation of methods for predicting all degeneracies of this type.

II. ONE-ELECTRON SYSTEMS

The single-electron Schrödinger equation in a D-dimensional Cartesian coordinate system x_k, $k=1,2,\ldots,D$, is

$$(\Delta^{(D)}+2(E-V))\Psi^{(D)}=0, \tag{1}$$

where $\Delta^{(D)}$ is the Laplacian

$$\Delta^{(D)}=\sum_{k=1}^{D}\frac{\partial^2}{\partial x_k^2} \tag{2}$$

and $V=V(\mathbf{r})$ is an arbitrary potential function of the coordinates, $\mathbf{r}=(x_1,x_2,\ldots,x_D)$. Whenever $V(\mathbf{r})$ is such that the wavefunction $\Psi^{(D)}$ factors into a product of two functions, one of which is a spherical harmonic for a subspace $D'<D$, then there is an effective reduction in the number of coordinates necessary to describe the energy spectrum.

In the case of a central field potential $V=V(r)$ the factorization is simply

$$\Psi^{(D)}=r^c R_\Lambda^{(D)}(r)\cdot Y_\Lambda^{(D)}, \tag{3}$$

where $Y_\Lambda^{(D)}$ is a D-dimensional spherical harmonic of characteristic value $\Lambda(\Lambda+D-2)$, $\Lambda=0,1,\ldots$, and $R_\Lambda^{(D)}$ is a radial function. The power c is arbitrary, but for convenience we set $c=\Lambda$.[3] The corresponding radial Schrödinger equation is

$$\left(\frac{\partial^2}{\partial r^2}+\frac{2\Lambda+D-1}{r}\frac{\partial}{\partial r}+2(E-V)\right)R_\Lambda^{(D)}=0, \tag{4}$$

which clearly is invariant to the transformation

$$(\Lambda,D)\rightarrow(\Lambda\pm1,D\mp2). \tag{5}$$

This invariance establishes the existence of interdimensional degeneracies in the energy spectrum. In fact repeated application of (5) shows that every state with $\Lambda\neq0$ is degenerate with a state having $\Lambda=0$ in a higher dimensionality.

Example: *Bound state hydrogen atom*: For $V=-1/r$ the radial function is[4]

$$R_{n\Lambda}^{(D)}=\exp(-kr)\,F(-n+\Lambda+1;2\Lambda+D-1;2kr), \tag{6}$$

where n is the principal quantum number $(n=\Lambda+1,\Lambda+2,\cdots)$ and $k=2/(2n+D-3)$. The energy is independent of Λ (an *intradimensional* degeneracy), with $E_{n\Lambda}=-k^2/2$. Application of transform (5) to the $R_{n\Lambda}$ shows that in order to preserve the number of nodes $(=n-\Lambda-1)$ in a particular radial solution we must simultaneously transform $n\rightarrow n\pm1$. For instance the ground state solution $R_{10}^{(5)}$ is identical to the excited $R_{21}^{(3)}$ state, and both have energy $=-1/8$.[5]

Identical results are obtained for less symmetric potentials $V(\mathbf{r})$. For instance, a cylindrically symmetric potential $V(R,z)$, where $z=x_D$ and $R=[\sum_{k=1}^{D-1}x_k^2]^{1/2}$, allows the factorization

$$\Psi^{(D)}=R^\Lambda F_\Lambda(R,z)\cdot Y_\Lambda^{(D-1)}. \tag{7}$$

F_Λ is the cylindrical analogue of the radial function and satisfies

$$\left(\frac{\partial^2}{\partial z^2}+\frac{\partial^2}{\partial R^2}+\frac{(2\Lambda+D-2)}{R}\frac{\partial}{\partial R}+2(E-V)\right)F_\Lambda=0, \tag{8}$$

which is invariant under transformation (5). A special case of a cylindrically symmetric potential is that of the H_2^+ molecular ion, $V=-1/r_A-1/r_B$. Here r_A and r_B are the distances of the particle from the fixed attractive nuclei A and B located at $+R/2$ and $-R/2$ respectively on the x_D axis. In terms of hyperelliptic coordinates $\rho=(r_B+r_A)/R$ and $\mu=(r_B-r_A)/R$ the factorization is

$$\Psi^{(D)}=X(\rho)\,Y(\mu)\cdot Y_\Lambda^{(D-1)}. \tag{9}$$

Invariance of the corresponding Schrödinger equation under transformation (5) shows for instance that the lowest energy σ_g state[6] at $D=5$ is degenerate with the lowest energy π_u state at $D=3$ for all R. The united atom limits ($R\to 0$) of these molecular states are the $1s(D=5)$ and $2p(D=3)$ atomic states respectively. At infinite nuclear separation the $D=5$ σ_g state correlates to the $1s_A+1s_B$ linear combination, while the π_u state goes to $2p_A+2p_B$. In analogy to our previous remark for the atom, each potential energy curve with $\Lambda>0$ for this one-electron problem is degenerate with a curve having $\Lambda=0$ in a higher dimensionality.

These one-electron results have a very general nature, being independent of the precise form of the potential. It is clear that interdimensional degeneracies exist for other simple molecular systems, such as H_3^{++}.

III. MANY-ELECTRON SYSTEMS

The N-electron Schrödinger equation is

$$\left(\sum_{j=1}^{N}\Delta_j^{(D)}+2(E-V)\right)\varphi^{(D)}=0, \tag{10}$$

where $V=V(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N)$. It is not obvious at first sight that interdimensional degeneracies exist for systems with two or more interacting electrons, although the one-electron results indicate that degeneracies may exist for states which can be factored into a product of a "radial" function and an "angular" function. Such a separation was possible in Ref. 1 for the $2p^2\,{}^3P^e$ state ($D=3$) of helium, for which the wavefunction is

$$\Psi^{(3)}=P\cdot\Phi(r_1,r_2,r_{12}), \tag{11a}$$

$$P=x_1y_2-y_1x_2, \tag{11b}$$

x_i and y_i being any two different Cartesian coordinates. The function Φ was shown to satisfy the same Schrödinger equation as the exact ground state wavefunction at $D=5$. For convenience in what follows, we shall assume all potentials to be Coulombic, although generalization to other forms is equally possible.

A. Hydrogen molecule

In direct analogy with the helium results, the ground state wavefunction of H_2 in D dimensions may be written as $\Phi^{(D)}(r_{1A},r_{2A},r_{1B},r_{2B},r_{12})$, where r_{iA} and r_{iB} were defined for the H_2^+ molecule. The explicit form of the Laplacian in this coordinate system, $\Delta_{12,AB}^{(D)}$, is omitted for brevity. It is easily seen that if χ is a function only of $r_{1A},r_{2A},r_{1B},r_{2B}$, and r_{12}, then

$$(\Delta_1^{(D-2)}+\Delta_2^{(D-2)})\,Q\chi=Q(\Delta_{12,AB}^{(D)}\chi). \tag{12}$$

Here $Q=x_1y_2-y_1x_2$, where x_i and y_i are Cartesian coordinates perpendicular to the nuclear axis. Q has zero angular momentum about the internuclear axis. If χ is taken to be $\Phi^{(D)}$, Eq. (12) shows that an interdimensional degeneracy exists, with $Q\Phi^{(D)}$ and $\Phi^{(D)}$ having the same energy eigenvalue in $(D-2)$ and D dimensions respectively. Thus the ground state ${}^1\Sigma_g^+$ energy at $D=5$ is the same as the lowest energy doubly excited ${}^3\Sigma_g^-$ state at $D=3$. The united atom limits of the states are $1s^2\,{}^1S^e$ and $2p^2\,{}^3P^e$, in agreement with the helium results.

It is evident from Eq. (12) that similar degeneracies exist for other states in the $\Sigma^+(D=5)$ and $\Sigma^-(D=3)$ spectra.

B. Helium atom, excited states

The singly excited $D=5$ states having one unit of angular momentum and odd parity (i.e., $1snp$, $n=2,3,\cdots$) are exactly degenerate with the doubly excited states $2pnd$, P^0 ($n=3,4,\cdots$) at $D=3$. The $D=5$ states can be written generally as

$$\Psi_\pm=z_2F_{12}^{\pm}\pm z_1F_{21}^{\pm}, \tag{13}$$

where F_{12}^{+} and F_{12}^{-} are functions only of r_1, r_2, and r_{12} ($\equiv u$), and z_k is one of the Cartesian coordinates. Ψ_+ and Ψ_- are symmetric and antisymmetric with respect to electron exchange. By defining the operator

$$\begin{aligned}\Delta_{D,12}=&\frac{\partial^2}{\partial r_1^2}+\frac{\partial^2}{\partial r_2^2}+2\frac{\partial^2}{\partial u^2}\\&+\frac{u^2+r_1^2-r_2^2}{ur_1}\frac{\partial^2}{\partial r_1\partial u}+\frac{u^2+r_2^2-r_1^2}{ur_2}\frac{\partial^2}{\partial r_2\partial u}\\&+(D-1)\left(\frac{1}{r_1}\frac{\partial}{\partial r_1}+\frac{1}{r_2}\frac{\partial}{\partial r_2}+\frac{2}{u}\frac{\partial}{\partial u}\right),\end{aligned} \tag{14}$$

F_{12}^{+} and F_{12}^{-} are seen to satisfy the exchange coupled equations

$$\left(\Delta_{5,12}+\frac{2}{r_2}\frac{\partial}{\partial r_2}+\frac{2}{u}\frac{\partial}{\partial u}+2(E-V)\right)F_{12}^{\pm}=\pm\left(\frac{2}{u}\frac{\partial}{\partial u}\right)F_{21}^{\pm}. \tag{15}$$

Note that $F_{12}^{\pm}\neq F_{21}^{\pm}$.

Making use of the one-electron spherical harmonics p_m and d_m in three dimensions, the function

$$Y_{12}=\frac{r_1r_2}{\sqrt{2}}\,(p_{-1}(1)\,d_{+1}(2)-p_{+1}(1)\,d_{-1}(2)) \tag{16}$$

will be recognized as having two units of angular momentum (i.e., a pd, D state). It is straightforward to show that if G_{12} is any function of only r_1,r_2, and u, then

$$\begin{aligned}&(\Delta_1^{(3)}+\Delta_2^{(3)})\,Y_{12}G_{12}\\&=Y_{12}\left(\Delta_{3,12}+\frac{2}{r_1}\frac{\partial}{\partial r_1}+\frac{4}{r_2}\frac{\partial}{\partial r_2}+\frac{6}{u}\frac{\partial}{\partial u}\right)G_{12}\\&\quad+Y_{21}\left(\frac{2}{u}\frac{\partial}{\partial u}\right)G_{12}.\end{aligned} \tag{17}$$

The corresponding exchange symmetrized functions

$$\chi_\pm=Y_{12}G_{12}^{\pm}\pm Y_{21}G_{21}^{\pm}, \tag{18}$$

which are eigenfunctions of the Schrödinger equation, lead to the coupled equations

$$\left(\Delta_{5,12}+\frac{2}{r_2}\frac{\partial}{\partial r_2}+\frac{2}{u}\frac{\partial}{\partial u}+2(E-V)\right)G_{12}^{\pm}=\mp\left(\frac{2}{u}\frac{\partial}{\partial u}\right)G_{21}^{\pm}. \tag{19}$$

The $\pm$ superscripts have been added to indicate that $G_{12}^{+}\neq G_{12}^{-}$. Comparison of Eqs. (19) and (15) shows that each singly excited P eigenfunction in five dimensions can be used to construct a doubly excited D eigenfunction in three dimensions by the correspondence $G_{12}^{\pm}\equiv F_{12}^{\mp}$.

Since both states satisfy the same Schrödinger equation, they are necessarily degenerate.[7]

C. The lithium atom

In all of the preceding examples it was not necessary to discuss spin explicitly, since in each case the total wavefunction factors into a product of a spatial function times a spin function. By analogy with the $D=3$ spin formalism, one may define single particle spin representations of the D-dimensional angular momentum group which can be coupled to give spin eigenstates. We shall not discuss multidimensional spin further since the only example of an interdimensional degeneracy we have found for the Li atom is not spin dependent.

Consider for $D=3$ the triply excited $2p^3\,{}^4S$ fermion state of Li. The spin function is $\alpha(1)\alpha(2)\alpha(3)$ for the $M_s=3/2$ component, and the corresponding spatial wavefunction may be written

$$\Psi_{123}=W_{123}\circ\Phi_{123}, \tag{20}$$

where Φ_{123} is a function only of the six coordinates $r_1,r_2,r_3,r_{12},r_{23},r_{31}$. W_{123} is the totally antisymmetric (with respect to two-electron exchange) function of the three coordinates x,y,z:

$$W_{123}=x_1y_2z_3+z_1x_2y_3+y_1z_2x_3 - z_1y_2x_3-x_1z_2y_3-y_1x_2z_3. \tag{21}$$

The condition of antisymmetry for the total wavefunction $W_{123}\Phi_{123}\alpha(1)\alpha(2)\alpha(3)$ leads to the requirement that Φ_{123} be a totally symmetric function with respect to exchange.

If $\Delta_{D,123}$ is the three-electron representation of the D-dimensional Laplacian in the coordinates $r_1,r_2,r_3,r_{12},r_{23},r_{31}$, then one easily verifies that

$$(\Delta_1^{(3)}+\Delta_2^{(3)}+\Delta_3^{(3)})\,W_{123}\Phi_{123} = W_{123}\left(\Delta_{3,123}+\frac{2}{r_1}\frac{\partial}{\partial r_1}+\frac{2}{r_2}\frac{\partial}{\partial r_2}+\frac{2}{r_3}\frac{\partial}{\partial r_3} + \frac{4}{r_{12}}\frac{\partial}{\partial r_{12}}+\frac{4}{r_{23}}\frac{\partial}{\partial r_{23}}+\frac{4}{r_{31}}\frac{\partial}{\partial r_{31}}\right)\Phi_{123} \tag{22a}$$

$$=W_{123}(\Delta_{5,123}\Phi_{123}). \tag{22b}$$

We thus have the interesting result that any *totally symmetric* eigenfunction Φ_{123} of the five-dimensional Schrödinger equation can be used to define a (degenerate) solution Ψ_{123} of the three-dimensional Schrödinger equation by means of Eq. (20). In particular the exact $2p^3\,{}^4S(D=3)$ fermion state is degenerate with the exact spinless boson $1s^3$ $(D=3)$ ground state.

IV. RELATIONSHIP OF VARIABLE DIMENSIONALITY TO "EXCESS ANGULAR MOMENTUM"

Recent investigations[8,9] have used continuous scaling of the angular momentum (i.e., excess angular momentum) of one-particle quantum systems for purposes of determining degeneracies. It is clear from our present work (and also Ref. 1) that each variation in angular momentum is equivalent to a variation in D. This is most easily seen if the power c in Eq. (3) is chosen to be $c=(D-1)/2$, so that the corresponding radial Schrödinger equation effectively describes a one-dimensional motion (by removal of r^{D-1} from the integration volume element) along the r coordinate:

$$\left(-\frac{\partial^2}{\partial r^2}+\frac{K(K+1)}{r^2}-2(E-V)\right)R_\Lambda^{(D)}=0, \tag{23a}$$

with

$$K=\Lambda+((D-3)/2). \tag{23b}$$

Changes in the magnitude of $K(K+1)$ can be effected by variations of either Λ or D, or both. For the hydrogen atom [Eq. (6)] the shift $\Lambda\to\Lambda+((D-3)/2)=K$ is accompanied by a similar shift of the principal quantum number: $n\to n+((D-3)/2)\equiv N$, which preserves the degeneracy of radial states having the same n.[10] In the present paper we have considered only transformations of Λ and D which leave $K(K+1)$ unchanged.

V. CONCLUDING REMARKS

The existence of interdimensional degeneracies in the spectra of one-, two-, and three-particle systems is quite interesting, although at present no general criteria exist for predicting the occurrence of the degeneracies. In all cases the proof of degeneracy was made by investigation of the kinetic energy portion of the energy, and identical results may be obtained in the corresponding momentum space representation of the wavefunction . Separability of the wavefunctions into a product of "angular" and "radial" functions is found in all of our examples.

ACKNOWLEDGMENT

The author wishes to thank F.H. Stillinger for helpful discussions in regard to this work.

[1] D.R. Herrick and F.H. Stillinger, Phys. Rev. A (in press).

[2] Exact energies for this model potential were obtained by C.M. Rosenthal, J. Chem. Phys. 55, 2474 (1971).

[3] This leads naturally to solutions $R_\Lambda^{(D)}$ which are regular at the origin. Irregular solutions follow from setting $c=-\Lambda-D+2$. Our results are valid for either case, and in all subsequent examples only regular solutions are considered.

[4] S.P. Alliluev, Zh. Eksp. Teor. Fis. 33, 2000 (1957) [Sov. Phys. JETP 6, 156 (1958)].

[5] In more conventional notation atomic states with $\Lambda=0, 1, 2, \cdots$ are labeled $s, p, d, \cdots$. Thus the $1s$ $(D=5)$ ground state is identical (to within angular factors) to the $2p$ $(D=3)$ excited state.

[6] For diatomics we use the convention that $\Lambda=0, 1, 2, \cdots$ is denoted by $\sigma, \pi, \delta, \cdots$. Subscripts g and u indicate respectively symmetry and antisymmetry with respect to parity.

[7] An alternate form for $\chi\pm$ (within a normalization constant) which displays the relationship between the $D=3$ and 5 states is $P\cdot(z_2G_{12}^{\pm}\mp z_1G_{21}^{\pm})$, where $P=x_1y_2-y_1x_2$.

[8] A. Cisneros and H.V. McIntosh, J. Math. Phys. 11, 870 (1970).

[9] G.M. King, J. Phys. A 7, 202 (1974).

[10] For $D=3$ the degeneracy of radial states having the same principal quantum number is described by a radial Lie algebra locally isomorphic to $O(3)$. [See D.R. Herrick and O. Sinanoğlu, Phys. Rev. A 5, 2309 (1972)]. These results are essentially unchanged by transformation to the shifted quantum numbers N, K.

Electrostatic screening*

Jeffrey Rauch and Michael Taylor

Department of Mathematics, The University of Michigan, Ann Arbor, Michigan 48104
(Received 7 August 1974)

Using the methods of partial differential equations and functional analysis, we investigate the electric field in the presence of a screen composed of wires of radius r spaced at distance R spread over a surface S. In the limit as r and R converge to zero if $[R\ \ln r]^{-1} \to -\infty$, the field in the presence of the screen converges to the field with a conducting sheet spread over S. If $[R\ \ln r]^{-1} \to 0$, the field converges to the field with no conductors.

1. INTRODUCTION

It is well known that a region enclosed by a mesh of conducting wire is shielded from external static electric fields. In this sense the mesh acts like a solid sheet of conductor. On the other hand, it is clear that if the wires of the mesh are sufficiently narrow (for fixed mesh width), then they will have a negligible effect on the electric field. In this note we will study the problem of determining what range of physical parameters correspond to these two types of behavior. If the screen consists of wires of radius r whose axes are spaced at approximately distance R from each other, the critical parameter is $[R \ln r]^{-1} \equiv \delta$. We consider screens spread over a surface s in the limit as r and R approach zero and prove that for any charge distribution, if $\delta \to +\infty$, then the field in the presence of the screens converges to the field in the presence of a sheet of conductor spread over S (Theorem 2). In the opposite extreme case, if $\delta \to 0$, then the field converges to the field without any conductors present, that is, the screen becomes negligible.

2. VARIATIONAL FORMULATION OF THE BASIC BOUNDARY VALUE PROBLEM

We seek the electrostatic potential u in the exterior of a finite number of conductors $\kappa_1, \kappa_2, \cdots, \kappa_j$, arising from a charge distribution with density $\rho(x)$. For convenience we suppose that the whole system lies inside a very large but bounded region $\mathcal{R}$ whose boundary is kept at potential zero and is assumed to be smooth. With a little extra effort the problem in unbounded regions can also be handled by our methods. The boundary value problem for u is

$$\Delta u = -4\pi\rho \ \text{ in } \ \mathcal{R}\backslash\cup\kappa_i, \tag{1}$$

$$u = \text{const} \ \text{ on each } \kappa_i, \ \ i, \cdots, j, \tag{2}$$

$$\int_{\partial\kappa_i} \frac{\partial u}{\partial \nu} = 0, \ \ i = 1, \cdots, j, \tag{3}$$

$$u = 0 \ \text{ on } \partial\mathcal{R}. \tag{4}$$

From a mathematical standpoint the condition (3) which asserts that the conductors carry no charge is the most troublesome, and we will give a weak or variational formulation in which (3) becomes a natural boundary condition. Let $K = \cup\, \kappa_i$, $\Omega = \mathcal{R}\backslash K$, and $H_1(\Omega)$ the Sobolev space of functions on Ω which are square integrable together with their partial derivatives of order one.

Definition 1: $\mathcal{B}$ is the closed subspace of $H_1(\Omega)$ consisting of functions u which vanish on $\partial\mathcal{R}$ and in addition are constant on $\partial\kappa_i$, $i = 1, 2, \cdots, j$. For $u, v \in H_1(\Omega)$ let

$$a(u,v) = -\int_\Omega (\text{grad}u \cdot \text{grad}v).$$

It is not hard to show that u is a solution of (1)–(4) if and only if $u \in \mathcal{B}$ and

$$a(u,v) = 4\pi \int_\Omega \rho(x) v(x) dx \ \ \forall\ v \in \mathcal{B}. \tag{5}$$

Equation (5) is just the Euler–Lagrange equation associated with Thompson's principle: u minimizes $-a(u,u)/2 + \int_\Omega \rho u$ over all $u \in \mathcal{B}$. Note that (3) is a natural boundary condition. It is useful to notice that if $u \in \mathcal{B}$ satisfies (5), then $\Delta u = \rho$ in the sense of distributions and $u = \text{const}$ on $\partial\kappa_i$ so that the regularity theorems for the Dirichlet problem can be applied to show that u is smooth provided that the boundaries of the κ_i are smooth, which we will assume henceforth.

The quadratic form a on $L^2(\Omega)$ with domain $D(a) = \mathcal{B}$ is a closed symmetric and nonpositive. It is well known[1] that there is a self-adjoint operator Δ defined by the recipe

$$D(\Delta) = \{u \in \mathcal{B}: \ (\exists f \in L^2(\Omega)) \ \text{ such that } a(u,v) = (f,v)_{L^2(\Omega)}\},$$

$$\Delta u = f \ \text{ for } u \in D(\Delta).$$

With the aid of the regularity theorems mentioned above one can show that

$$D(\Delta) = \{u \in H^2(\Omega): u \in \mathcal{B} \text{ and } \int_{\partial\kappa_i} \frac{\partial u}{\partial \nu} = 0, \ i = 1, \cdots, j\},$$

$$\Delta u = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} \ \text{ for } u \in D(\Delta).$$

The solution to the electrostatics problem (1)–(4) is therefore $\Delta^{-1}(-4\pi\rho)$, the inverse of Δ applied to $-4\pi\rho$.

3. A THEOREM ON VANISHING SCREENS

We now pose the basic problem. For each integer n we consider the electrostatics problem in the presence of conductors $\kappa_1^n, \kappa_2^n, \cdots, \kappa_{j_n}^n$ and we ask whether the effect of the conductors has some limiting behavior as $n \to \infty$. In this section we prove a theorem which asserts that the effect of the conductors disappears as $n \to \infty$ provided they are sufficiently small. As an application we obtain the result on vanishing screens mentioned in the Introduction.

The appropriate measure of smallness turns out to be electrostatic capacity. Recall that for reasonable sub-

sets Λ of $\mathbb{R}^3$, cap(Λ) is defined as follows: Let v be the solution of the boundary value problems

$$\Delta v = 0 \text{ on } \mathbb{R}^3 \setminus \Lambda,$$
$$v = O(1/|x|) \text{ as } |x| \to \infty,$$
$$v = 1 \text{ on } \partial\Lambda.$$

Then $-\int_{|x|=L} \partial v/\partial r$ is independent of L for L large and is the total charge on a conductor occupying the region Λ and raised to potential one. This quantity is cap(Λ), the capacity of Λ.

Notations: Let Δ_n, a_n, $\mathcal{B}_n$ be the operator, form, and form domain on $\Omega_n \equiv R \setminus \cup_i \kappa_i^n$ as defined in Sec. 2. In addition, for $v \in L^2(R)$ let $P_n v \in L^2(\Omega_n)$ be the restriction of v to Ω_n. Any element of $L^2(\Omega_n)$ is considered as an element of $L^2(R)$ by extending it to vanish on the union of the κ_i^n. Let $K(n) = \cup_i \kappa_i^n$ denote this union. We suppose all $K(n)$ are contained in some compact set $\Gamma \subset R$.

The main tool we use to show that $K(n)$ vanishes is Theorem 3.1 of Ref. 2. This asserts that $f(\Delta_n)P_n u \to f(\Delta)u$ in $L^2(R)$ for all $u \in L^2(R)$ and any f bounded and continuous on $(-\infty, 0]$ provided the Ω_n satisfy mild regularity conditions, that the quadratic form $a(u,u)$ satisfies the coerciveness hypothesis $-a(u,u) \geq \int_{\Omega_n} |\text{grad} u|^2$ for all $u \in \mathcal{B}_n$, and the following two special assumptions:

(A) There exist extension operators $E_n: \mathcal{B}_n \to \mathcal{B}$ [the domain of the form $a(u,v)$ on R without conductors] with the properties

(i) $E_n u = u$ on Ω_n for all $u \in \mathcal{B}_n$,

(ii) there is a constant M such that for all n and $u \in \mathcal{B}_n$,

$$\|E_n u\|_{H_1(R)} \leq M \|u\|_{H_1(\Omega_n)};$$

and either

(B) Meas$(K(n)) \to 0$, and if $u \in \mathcal{B}$, there exist $u_j \to u$ in $\mathcal{B}$ such that $u_j|_{\Omega_j} \in \mathcal{B}_j$,

or

(B′) cap$(K(n)) \to 0$ as $n \to \infty$.

That (A), (B′) imply operator convergence is stated in Theorem 4.2 of Ref. 2; alternatively, condition (B′) implies condition (B).

Theorem 1: Suppose there is a compact set $\Gamma \subset R$ with $K(n) \subset \Gamma$ for all n and that cap$(K(n)) \to 0$ as $n \to \infty$. Then for any continuous function f on $(-\infty, 0)$ bounded at $-\infty$ and any $u \in L^2(R)$ we have

$$f(\Delta_n)P_n u \to f(\Delta)u \text{ in } L^2(R),$$

where Δ is the operator on R without any conductors.

As a particular example, for $\rho \in L^2(R)$ with ρ supported in the exterior of all conductors, we can take $f(x) = 1/x$ to get $\Delta_n^{-1}(\rho) \to \Delta^{-1}(\rho)$ in $L^2(R)$. Thus the solutions of the electrostatics problems converge to the solution to the problem with no conductors at all.

Proof: Note that $\sigma(\Delta)$ and $\sigma(\Delta_n) \subset (-\infty, \delta)$ for some $\delta < 0$, so f can be altered to be bounded and continuous on $(-\infty, 0]$ without changing $f(\Delta_n)$ or $f(\Delta)$. To complete the proof, it is only necessary to verify hypothesis (A).

To describe E_n notice that if $u \in \mathcal{B}_n$, then u is constant on $\partial\kappa_i^n$, $i = 1, 2, \cdots, j_n$, say $u = c_i$ on $\partial\kappa_i^n$. Define $E_n u = c_i$ on κ_i^n. It is clear that $\int_R |\text{grad} E_n u|^2 = \int_{\Omega_n} |\text{grad} u|^2$. Furthermore, since $E_n u = 0$ on ∂R, we have

$$\int_R |E_n u|^2 \leq \int_R |\text{grad} E_n u|^2,$$

where $\lambda < 0$ is the largest eigenvalue of the Laplacian on R with Dirichlet boundary conditions on R. Thus (ii) is satisfied with $M = 1 + \lambda^{-1}$ and the proof is complete. □

It is quite easy to apply this result to screens. The basic fact that is needed is that the capacity of a solid circular cylinder of length L and radius r is proportional to $-L/\ln r$. Similarly a not excessively curved piece of wire of length L and radius R has capacity $O(-L/\ln r)$. In addition, capacity is a subadditive set function, that is, cap$(\cup A_i) \leq \Sigma_i$ cap(A_i) for any countable union of sets. Thus the capacity of a curved screen of fixed area with wires of radius r and spacing R between axes of the wires is $O(-1/R\ln r)$. Thus if $K(n)$ is a screen as above with r and R approaching zero as $n \to \infty$ in such a way that $1/R\ln r \to 0$, then the effect of the screen is negligible for n large.

For the electrostatic problem, cap$K(n) \to 0$ is by no means a necessary condition for the $K(n)$ to have a negligible effect. Suppose, for example, that $K(n)$ consists of n balls, of radius r_n, and say their center ξ_{jn} are spaced at a distance at least $4r_n$. By defining extension operators E_n as in the proof of Theorem 1, it is easy to see that hypothesis (A) is satisfied. We show that hypothesis (B) is verified, assuming vol$K(n) = (4/3\pi)nr_n^3 \to 0$.

Define a continuous linear map $Q: H_1(B_2) \to H_1(B_2)$ $(B_2 = \{x: |x| \leq 2\})$ such that

(i) $Qu(x) = u(x)$ for $3/2 \leq |x| \leq 2$

(ii) $Qu(x)$ is constant for $|x| \leq 1$

(iii) $\int_{B_2} |Qu|^2 \leq C_0 \int_{B_2} |u|^2$

(iv) $\int_{B_2} |\text{grad} Qu|^2 \leq C_0 \int_{B_2} |\text{grad} u|^2$.

This is easy to arrange. Given this, you can scale B_2 to $B_{2r_n}(\xi_{jn}) = \{x: |x - \xi_{jn}| \leq 2r_n\}$ and get maps with the same properties as i–iv (same constant C_0). Thus you get maps $Q_n: \mathcal{B}_n \to \mathcal{B}$ such that

(i) $Q_n u(x) = u(x)$, $x \notin \cup_j B_{2r_n}(\xi_{jn})$,

(ii) $\|Q_n u\|^2_{H_1(B_{2r_n}(\xi_{jn}))} \leq C_0 \|u\|^2_{H_1(B_{2r_n}(\xi_{jn}))}$,

(iii) $Q_n u|_{\Omega_n} \in \mathcal{B}_n$.

Now with $u_n = Q_n u$ you get

$$\|u_n - u\|^2_{H_1(R)} = \sum_j \|u_n - u\|^2_{H_1(B_{2r_n}(\xi_{jn}))}$$
$$\leq 4C_0 \sum_j \|u\|^2_{H_1(B_{2r_n}(\xi_{jn}))} \to 0$$
$$\text{as } n \to \infty \text{ since meas } \cup_j B_{2r_n}(\xi_{jn}) \to 0.$$

This verifies hypothesis (B).

The conclusion is that if $K(n)$ consists of n "well spaced" balls of radius r_n, then $K(n)$ disappears as $n \to \infty$, *assuming only that vol* $K(n) \to 0$.

4. THE CASE OF ELECTROSTATIC SCREENING

In this section we will investigate the observed phe-

nomenon of screens behaving like solid barriers. To be more precise, suppose that $K(n)$ is a conducting screen, with wires of radius r and spacing R, spread smoothly over the surface S and that r and R tend to zero as $n \to \infty$. If $(-R\ln r)^{-1} \to +\infty$ as $n \to \infty$, then for any charge distribution ρ on $\mathcal{R}$ the solutions, $\Delta_n^{-1}(\rho)$, of the electrostatics problems in $\mathcal{R}\backslash K(n)$ converge to the solution u of the problem where S is covered by a sheet of perfect conductor, that is,

$$\Delta u = -4\pi\rho \quad \text{in } \mathcal{R}\backslash s, \tag{6}$$

$$u = \text{const} \quad \text{on } S, \tag{7}$$

$$\int_S \left[\frac{\partial u}{\partial v}\right] = 0 \quad ([\ ] \text{ denotes jump on crossing } S) \tag{8}$$

$$u = 0 \quad \text{on } \partial\mathcal{R}. \tag{9}$$

This result complements the result of Sec. 3 and confirms the idea that the parameter $(-R\ln r)^{-1}$ is a reasonable measure of the solidity of a screen. It is interesting to note the same parameter occurs in the clever special problem treated in §203 of Maxwell's treatise.[3] In addition, as Maxwell observed, a complete screen is not needed, just one family of parallel wires which are connected to each other in any way at all will suffice.

We must make precise the notion of a screen spread smoothly over S, where S is an open subset of a compact surface in the interior of $\mathcal{R}$. The intuitive idea is to take a piece of planar screen and give a mapping of the planar region to the surface. Precisely, if $s \in S$ and $\mathcal{O}$ is an open neighborhood of s in $\mathbb{R}^3$, then a mapping ψ: $\mathcal{U} \to \mathcal{O}$ is called a δ *bending* if

(i) $\mathcal{U}$ is a cube $|x_i| < \alpha$, $i = 1, 2, 3$,

(ii) $\psi[\mathcal{U} \cap \{x_3 = 0\}] = S \cap \mathcal{O}$,

(iii) ψ is a diffeomorphism with $\|J_\psi\|$ and $\|J_{\psi^{-1}}\|$ less than δ where J is the Jacobian matrix.

Screens are laid on S by placing a screen in the $x_3 = 0$ plane of $\mathcal{U}$ and carrying it to S by the map ψ.

Definition 2: A *patch of* δ *bent screen* on S consisting of wires of radius r and spacing R is the set $\psi[\Sigma]$, where ψ: $\mathcal{U} \to \mathcal{O}$ is a δ bending and

$$\Sigma = \{x \in \mathcal{U}: (x_1 - jR)^2 + x_2^2 \leq r^2 \text{ for some } j\}.$$

In addition we require $R > 3r$.

To form a picture, notice that the wires in Σ are parallel to the x_2 axis. The only interesting case of screening is when the screen has large gaps, that is, $R \gg r$.

Definition 3: A sequence of systems of conductors will be called *screens smoothly covering* S if there is a $\delta > 0$, an $\alpha > 0$, and an integer M such that (1) each system consists of at most M patches of δ bent screen on S, (2) the sets $\psi(\mathcal{U}_i)$, $i = 1, \cdots, M$, cover S for each system, and (3) the lengths of the sides of the cubes are all greater than α.

It is important that the electrostatic potential be constant on the screen, not just on the individual wires from which it is constructed. There are two ways we could arrange this. In one approach, we could suppose that a few wires are added to the screen so that it becomes a connected set. In the second we just prescribe the constancy of the potential on the screen as a boundary condition. Both methods yield the same results and we will adopt the second so that the basic boundary value problem becomes (1)—(4) with $j = 1$ and κ_1 the screen on S.

As in Sec. 2: the boundary value problem (6)—(9) can be given a variational formulation in which $u = \Delta_\infty^{-1}(-4\pi\rho)$, where Δ_∞ is the operator on $L^2(R)$ defined by the quadratic form

$$a(u,v) = \int_R \text{grad}u \cdot \text{grad}v,$$

$$D(a_\infty) = \{u \in H_1(R): u = 0 \text{ on } \partial\mathcal{R} \text{ and } u \text{ is constant on } S\}.$$

Theorem 2: Suppose that $K(n)$, $n = 1, 2, \cdots$, are screens smoothly placed on S, where $K(n)$ consists of wires of radius r_n and spacing R_n. Let Δ_n be the operator on $L^2(\mathcal{R}\backslash K(n))$ as in Sec. 2 and P_n: $L^2(R) \to L^2(\mathcal{R}\backslash K(n))$ be the restriction mapping. If $(-R_n \ln r_n)^{-1} \to \infty$, then for any continuous function f on $(-\infty, 0)$ bounded at $-\infty$, $f(\Delta_n)P_n\rho \to f(\Delta_\infty)\rho$ in $L^2(\mathcal{R})$ for any $\rho \in L^2(\mathcal{R})$.

Proof: We describe the modifications that are required to adopt the methods of our paper on wild perturbations[2] to this setting. For the remainder of the proof this paper is referred to as PSWPD. First we define uniformly bounded extension operators E_n: $\mathcal{B}_n \to \mathcal{B}_\infty \equiv D(a_\infty)$ by extending functions to be constant inside κ_i^n. As in our previous work (see the proof of Theorem 1.2 of PSWPD) it suffices to prove the result for $f = (1 - x)^{-1}$. Imitating the proof of Theorem 4.4 of PSWPD, we notice that for $g \in L^2(\mathcal{R})$

$$\begin{aligned}\|(1 - \Delta_n)^{-1}P_n g\|^2_{H_1(\mathcal{R}\backslash K(n))} &= ((1 - \Delta_n)(1 - \Delta_n)^{-1}P_n g, (1 - \Delta_n)^{-1}P_n g)_{\mathcal{R}\backslash K(n)} \\ &= (P_n g, (1 - \Delta_n)^{-1}P_n g)_{\mathcal{R}\backslash K(n)} \\ &\leq \|g\|^2_{L^2(\mathcal{R})}\end{aligned}$$

so that $w_n \equiv E_n(1 - \Delta_n)^{-1}P_n g$ is a bounded sequence in $H_1(\mathcal{R})$. By using Eq. (5) on $\Omega = \mathcal{R}\backslash K(n)$ for the function w_n it is easy to show that if w is a limit point of the sequence $\{w_n\}$ in the weak topology for $H_1(\mathcal{R})$, then

$$\int_{\mathcal{R}} (wu - \text{grad}w \cdot \text{grad}u) = \int_{\mathcal{R}} gu \tag{10}$$

for all $u \in H_1(\mathcal{R})$ such that u is constant on a neighborhood of S. Since these u are dense in $\mathcal{B}_\infty = D(a)$, (10) holds for all $u \in \mathcal{B}_\infty$. To show that $w = (1 - \Delta_\infty)^{-1}g$, it therefore suffices to prove that $w \in \mathcal{B}_\infty$, that is, $w = \text{const}$ on S and $w = 0$ on $\partial\mathcal{R}$. The latter is true since $\{v \in H_1(\mathcal{R}) | v = 0 \text{ on } \partial\mathcal{R}\}$ is a closed linear subspace, hence weakly closed. That w is constant on S lies considerably deeper. The crucial inequality is the following:

Let $\mathcal{U}$, Σ, r, R be as in Definition 2 and let $\mathcal{U}_H = \mathcal{U} \cap \{|x_3| \leq H\}$.

There is a constant b independent of H such that for all $v \in H_1(\mathcal{U})$ with $v|_\Sigma = 0$,

$$\int_{\mathcal{U}_H} |\text{grad}v|^2 \Big/ \int_H |v|^2 \geq \frac{b}{H^2 - HR\ln(r/R)} \tag{11}$$

provided $H > R > 3r$.

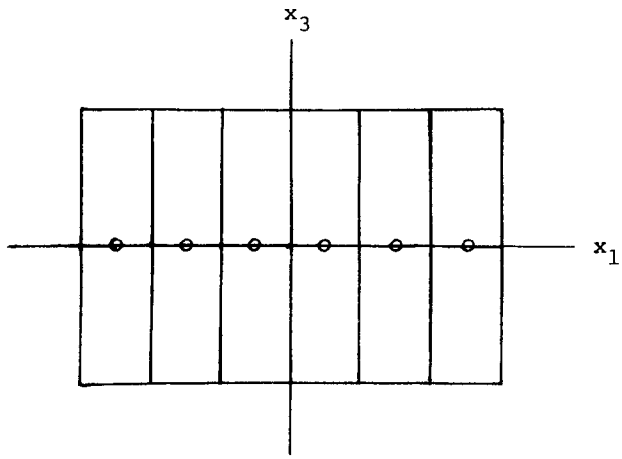


FIG. 1.

The verification of inequality (11) is postponed to the end of the proof. Let $U_1^n, \cdots, U_{j_n}^n$ be cubes with $\cup_i \psi_i^n(U_i^n) \supset S$, the screen $K(n) = \cup_i \psi_i^n(\Sigma_i^n)$. Let $c_n = w_n|_{\text{screen}}$ and apply the inequality to $w_n \circ \psi_i^n - c_n$. In this case, $R_n \ln(r_n/R_n) \to 0$ as $n \to \infty$ so the right-hand side of (11) behaves like const H^{-2} for n large. Letting $S_H^n = \cup_i \psi_i^n(U_H)$, we get

$$\int_{S_H^n} |w_n - c_n|^2 \leq \text{const } H^2. \tag{12}$$

Let $S_H = \{x: \text{dist}(x, S) \leq H\}$; then, since $|c_n - c_m| \leq |w_n - c_n| + |w_n - c_m|$, it follows that $\int_{S_H} |c_n - c_m|^2 \leq \text{const } H^2$ provided $\delta H > R_n, R_m$ for some δ independent of n, m. Since vol (S_H) approaches zero like a multiple of H, we get $|c_n - c_m|^2 = O(H)$ for $H > R_n, R_m$. Letting n, m tend to infinity, we see that $\{c_n\}$ is a Cauchy sequence so that $c_n \to c$ for some c. Passing to the limit in (12) yields

$$\frac{1}{H} \int_{S_H} |w - c|^2 = O(H), \tag{13}$$

and it follows that $w = c$ on S, since

$$\int_S (w - c)^2 \leq \text{const} \lim_{H \to 0} H^{-1} \int_{S_H} (w - c)^2.$$

We have now shown that w_n converges weakly in $H_1(R)$ to $w \equiv (1 - \Delta_\infty)^{-1} g$. Since $\|w_n\|_{H_1(R)}$ is bounded independent of n, it follows by the Rellich compactness theorem that $\{w_n\}$ is precompact in $L^2(R)$. Since w_n converges weakly to w in $L^2(R)$ it follows that $w_n \to w$ in $L^2(R)$ which is the desired result.

We now return to the proof of inequality (11). This is reduced to a two-dimensional problem by considering the $x_2 = \text{const}$ cross sections of U_H. For these cross sections we prove that

$$\int_{\text{cross section}} \left(\frac{\partial v}{\partial x_1}\right)^2 + \left(\frac{\partial v}{\partial x_2}\right)^2 dx_1 dx_3 \geq \frac{\text{const}}{H^2 - HR \log(r/R)} \int_{\text{cross section}} v^2 \tag{14}$$

for v which vanish on Σ. This in turn can be proven by chopping the cross section into punctured rectangles as in Fig. 1. It suffices to prove (14) where the integration is only over one of the punctured rectangles. The lower bound for the punctured rectangles is proved exactly as inequality (4.1) of PSWPD and the argument is not reproduced here. This completes the proof of Theorem 2. □

The phenomenon just considered has a great deal in common with the behavior of the Dirichlet problem, although the proof in the case of the electrostatic boundary problem is a little more involved. It is interesting to note that the electrostatic problem can exhibit behavior markedly different from that of the Dirichlet problem. For example, suppose the wire screen described above consists of wires which are not connected, that is, not at a common potential. If the wires are parallel to a vector field X on the surface S, and if $-(R_n \log r_n)^{-1} \to \infty$, then the u_n converge to a solution to the problem

$$\Delta u = -4\pi\rho \quad \text{on } R \setminus S, \tag{14}$$

$$[u] = 0 \quad \text{on } S, \tag{15}$$

$$Xu = 0 \quad \text{on } S, \tag{16}$$

$$\int_S \left[\frac{\partial u}{\partial v}\right] v = 0 \quad \text{for all } v \in C^\infty(S) \text{ with } Xv = 0, \tag{17}$$

$$u = 0 \quad \text{on } \partial R. \tag{18}$$

Since this is not a straightforward application of previously stated results, we indicate a proof. Let $u_n = \Delta_n^{-1}(-4\pi\rho)$, where Δ_n is defined on R with electrostatic boundary conditions on the wires $K(n)$, and u_n is extended as a constant on each wire. As usual, $\{u_n\}$ is bounded in $\mathring{H}_1(R)$, and so has a weak limit point $u \in \mathring{H}_1(R)$. Clearly u satisfies (14), (15), and (18) above so that we need to prove (16) and (17). Furthermore, we need only consider those ρ which vanish in a neighborhood of S since these are dense in $L^2(R)$. If we prove

$$-a(u, v) = (-4\pi\rho, v) \text{ for all } v \in B \quad \text{where } B = \{v \in H_1(R): Xv = 0 \text{ on } S\}, \tag{19}$$

then (17) will arise as a natural boundary condition.

To prove (19), we need only observe that for each $v \in B$ there exist $v_n \in B$ such that v_n is constant on each wire of $K(n)$ and $v_n \to v$ in B as $n \to \infty$. Then (19) holds for v_n and we may pass to the limit. The existence of such v_n is proven by constructing operators analogous to the Q's at the end of Sec. 3.

It only remains to prove that $Xu = 0$ on S, i.e., that $u \in B$. Indeed, by previous calculations

$$\frac{1}{H} \int_{S_H} |u_n - c_n|^2 \leq b\left(H - R_n \ln \frac{r_n}{R_n}\right).$$

This time, c_n is not a constant, but it is constant on the wires of $K(n)$, and in the direction normal to S. It merely varies from wire to wire. Thus $c_n \in L^2(S)$, $Xc_n = 0$. A trivial estimate is

$$\frac{1}{H} \int_{S_H} |u_n - \tilde{u}_n|^2 \leq \beta(H) \to 0 \quad \text{as } H \to 0,$$

where $\tilde{u}_n = u_n|_S$ extended to S_H as a function independent of the normal variable. Putting these together and letting $H \to 0$ yields

$$\int_S |\tilde{u}_n - c_n|^2 \to 0 \quad \text{as } n \to \infty.$$

Since $\tilde{u}_n \in H^{1/2}(S)$ is bounded, passing to a subsequence you get $\tilde{u}_n \to u|_S$ in $L^2(S)$. Hence $c_n \to u|_S$ in $L^2(S)$, so that $Xu = 0$ on S, as desired.

An even greater disparity is observed if $K(n)$ consists of n balls of radius r_n, with center ξ_{jn} lying on S and spaced apart a distance at least $4r_n$ [or $K(n)$ could consist of discs, the intersection of S with these balls]. If these obstacles are connected, say by arbitrarily thin wires, arguments as in the proof of Theorem 2 show that $K(n)$ behaves in the limit as a solid screen S, provided $nr_n \to \infty$. For this proof Lemma 4.5 of Ref. 2 is needed in place of (11). On the other hand, surely $\mathrm{vol}K(n) \to 0$, so if $K(n)$ is not connected, as we have seen at the end of Sec. 3, the obstacles disappear as $n \to \infty$.

*This work was partially supported by the National Science Foundation under Grant NSF GP-34260.

[1]T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966).
[2]J. Rauch and M. Taylor, "Potential and Scattering Theory on Wildly Perturbed Domains," J. Funct. Anal. (to appear).
[3]J. C. Maxwell, *A Treatise on Electricity and Magnetism* (Dover, New York, 1954).

A generalized theory of multiplicative stochastic processes using cumulant techniques

Ronald Forrest Fox*

School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332
(Received 4 October 1974)

The rules for the construction of the nth order cumulant for time-dependent, stochastic, matrices or operators which do not commute with themselves at unequal times are derived. The results are identical with van Kampen's rules. In the Gaussian case, Kubo's concept of a generalized Gaussian process is criticized. Under certain conditions Kubo's idea becomes asymptotically valid, while the same conditions justify use of the author's earlier delta function theory. A generalized density matrix equation is presented and its behavior during the approach to equilibrium is discussed. A finite correlation time, τ_c, does not necessarily invalidate a monotonic approach to equilibrium.

1. INTRODUCTION

In this paper, the author's earlier theory of multiplicative stochastic processes,[1] with its application to the treatment of the time development of the quantum mechanical density matrix,[2,3] is generalized. The earlier theory was restricted to the consideration of purely random, Gaussian, stochastic matrices, whereas the theory presented here applies to any class of stochastic matrices.

Following earlier work by Kubo[4,5] and more recent work by van Kampen,[6,7] cumulant techniques are used to achieve the generalization. General rules for the construction of the nth order cumulant expressions have been published by van Kampen. In this paper a derivation of those rules is presented, and the derivation requires the introduction of the concept of a time-ordered logarithm. The Gaussian property is introduced, but without the added restriction of "purely randomness," and it is shown that, for stochastic matrices which do not commute with themselves at unequal times, the cumulant expansion does not truncate to the first two cumulants only. While this point has been published before,[8] here it is pointed out that this circumstance invalidates Kubo's concept of a "generalized Gaussian" stochastic, matrix process.[4,5] By studying a special case in which the correlation matrices dampen out exponentially, it is demonstrated that the Gaussian property does lead to the desired truncation, asymptotically in time, after all.

The application of these mathematical results to the density matrix is made in the last section. A generalized density matrix equation for the approach to the microcanonical density matrix is presented. With the Gaussian property imposed, the generalized equation asymptotically approaches the Redfield equation presented earlier.[1] Moreover, the H-theorem, proved within context of the earlier, Gaussian, purely random, theory,[2] can be generalized in some situations which are discussed.

Because the asymptotic behavior of the generalized equations is identical with the earlier, more restricted theory, the general theory has as one of its virtues the property that it justifies the earlier theory, rather than replaces it.

2. MATHEMATICAL FOUNDATIONS

It is the purpose of the following calculations to investigate the averaged behavior of the equation:

$$\frac{d}{dt}a_\alpha(t)=\tilde{A}_{\alpha\alpha'}(t)a_{\alpha'}(t) \tag{1}$$

when

$$\langle\tilde{A}_{\alpha\alpha'}(t)\rangle=M_{\alpha\alpha'}, \tag{2}$$

$$\langle\tilde{A}_{\alpha\alpha'}(t)\tilde{A}_{\beta\beta'}(s)\rangle=2Q_{\alpha\alpha'\beta\beta'}(t-s)+M_{\alpha\alpha'}M_{\beta\beta'}; \tag{3}$$

and there are nonvanishing higher order moments of $\tilde{\mathbf{A}}(t)$ which will not be explicitly indicated at this point. Both $\tilde{\mathbf{A}}(t)$ and $\mathbf{M}$ are antisymmetric matrices. Repeated indices imply summation, and the variance tetratic, $Q_{\alpha\alpha'\beta\beta'}(t-s)$, is a function of the time interval $t-s$. In the earlier, purely random, theory, the time dependence was always given by a delta function times a time-independent, tetratic variance, $Q_{\alpha\alpha'\beta\beta'}\times\delta(t-s)^1$. That restriction is now removed.

The study of Eq. (1) is equivalent to the study of the averaged behavior of

$$\frac{d}{dt}a_\alpha(t)=M_{\alpha\alpha'}a_{\alpha'}(t)+\tilde{A}'_{\alpha\alpha'}(t)a_{\alpha'}(t) \tag{4}$$

when

$$\langle\tilde{A}'_{\alpha\alpha'}(t)\rangle=0, \tag{5}$$

$$\langle\tilde{A}'_{\alpha\alpha'}(t)\tilde{A}'_{\beta\beta'}(s)\rangle=2Q_{\alpha\alpha'\beta\beta'}(t-s), \tag{6}$$

and again the higher order moments will not yet be explicity indicated.

It is also equivalent, and convenient, to perform a linear transformation and work with the equation

$$\frac{d}{dt}b_\alpha(t)=\tilde{B}_{\alpha\alpha'}(t)b_{\alpha'}(t) \tag{7}$$

wherein $b_\alpha(t)$ and $\tilde{B}_{\alpha\alpha'}(t)$ are defined by

$$a_\alpha(t)\equiv[\exp(t\mathbf{M})]_{\alpha\alpha'}b_{\alpha'}(t), \tag{8}$$

$$\tilde{B}_{\alpha\alpha'}(t)\equiv[\exp(-t\mathbf{M})]_{\alpha\beta}\tilde{A}_{\beta\beta'}(t)[\exp(t\mathbf{M})]_{\beta'\alpha'}. \tag{9}$$

It follows from (5) and (6) that

$$\langle\tilde{B}_{\alpha\alpha'}(t)\rangle=0, \tag{10}$$

$$\langle\tilde{B}_{\alpha\alpha'}(t)\tilde{B}_{\beta\beta'}(s)\rangle = 2Q_{\mu\mu'\nu\nu'}(t-s)[\exp(-t\mathbf{M})]_{\alpha\mu}[\exp(t\mathbf{M})]_{\mu'\alpha'} \times[\exp(-s\mathbf{M})]_{\beta\nu}[\exp(s\mathbf{M})]_{\nu'\beta'}. \tag{11}$$

Note also that the antisymmetry of $\tilde{\mathbf{A}}(t)$ and $\mathbf{M}$ imply through (9) that

$$\tilde{B}_{\alpha'\alpha}(t) = -\tilde{B}_{\alpha\alpha'}(t) \tag{12}$$

The solution to (7) is expressible as a series of time ordered integrals:

$$\begin{aligned} b_\alpha(t) &= \sum_{n=0}^{\infty} \int_0^t \int_0^{t_1} \int_0^{t_2} \cdots \int_0^{t_{n-2}} \int_0^{t_{n-1}} [\tilde{B}_{\alpha\mu_1}(t_1)\tilde{B}_{\mu_1\mu_2}(t_2)\cdots \\ &\quad \times \tilde{B}_{\mu_{n-2}\mu_{n-1}}(t_{n-1})\tilde{B}_{\mu_{n-1}\alpha'}(t_n)]dt_n \cdots dt_1 b_{\alpha'}(0) \\ &= [\underset{\sim}{T}\exp(\int_0^t \tilde{\mathbf{B}}(s)ds)]_{\alpha\alpha'} b_{\alpha'}(0) \end{aligned} \tag{13}$$

wherein the $n=0$ term of the sum is defined to be $\delta_{\alpha\alpha'} b_{\alpha'}(0) \equiv b_\alpha(0)$, and the last equivalence defines the time ordering symbol, $\underset{\sim}{T}$.

The average of (13) may be expressed using time ordered cumulants[4–8]:

$$\begin{aligned} &\langle \underset{\sim}{T}\exp(\int_0^t \tilde{\mathbf{B}}(s)ds)\rangle \\ &\quad \equiv \underset{\sim}{T}\exp\Big(\sum_{n=1}^{\infty}\int_0^t \int_0^{t_2}\cdots\int_0^{t_{n-2}}\int_0^{t_{n-1}} \langle \tilde{\mathbf{B}}(t_1)\tilde{\mathbf{B}}(t_2)\cdots \\ &\qquad \times \tilde{\mathbf{B}}(t_{n-1})\tilde{\mathbf{B}}(t_n)\rangle_c dt_n \cdots dt_1\Big) \\ &\quad \equiv \underset{\sim}{T}\exp\Big(\sum_{n=1}^{\infty}\int_0^t \mathbf{G}^{(n)}(s)ds\Big), \end{aligned} \tag{14}$$

where $\mathbf{G}^{(n)}(s)$ is defined by

$$\begin{aligned} \mathbf{G}^{(n)}(s) \equiv \int_0^s \int_0^{t_2}\cdots\int_0^{t_{n-2}}\int_0^{t_{n-1}} &\langle \tilde{\mathbf{B}}(s)\tilde{\mathbf{B}}(t_2)\cdots \\ &\times \tilde{\mathbf{B}}(t_{n-1})\tilde{\mathbf{B}}(t_n)\rangle_c dt_n \cdots dt_2. \end{aligned} \tag{15}$$

So far, all that has been done is to define the cumulant averages through (14) and (15), and it is necessary to find expressions for $\mathbf{G}^{(n)}(s)$ solely in terms of ordinary averages of products of $\tilde{\mathbf{B}}(t)$'s if (14) is to be useful. This task is the heart of the difficulty with the cumulant method, and the purpose of this section of this paper will be to demonstrate a general method of solution.

The method of solution begins with the replacement of $\tilde{\mathbf{B}}(t)$ by $\lambda\tilde{\mathbf{B}}(t)$ everywhere in (14) and (15). The parameter λ is an ordinary real number, with respect to which derivatives may be performed. Using (13), (14), and (15) gives

$$\begin{aligned} &\langle \underset{\sim}{T}\exp(\int_0^t \lambda\tilde{\mathbf{B}}(s)ds)\rangle \\ &\quad = \sum_{n=0}^{\infty}\int_0^t \int_0^{t_1}\int_0^{t_2}\cdots\int_0^{t_{n-2}}\int_0^{t_{n-1}} \lambda^n \langle \tilde{\mathbf{B}}(t_1)\tilde{\mathbf{B}}(t_2)\cdots \\ &\qquad \times \tilde{\mathbf{B}}(t_{n-1})\tilde{\mathbf{B}}(t_n)\rangle dt_n \cdots dt_1 \\ &\quad \equiv \sum_{n=0}^{\infty}\lambda^n \int_0^t \mathbf{A}^{(n)}(s)ds \\ &\quad = \underset{\sim}{T}\exp\Big(\sum_{n=1}^{\infty}\lambda^n \int_0^t \mathbf{G}^{(n)}(s)ds\Big), \end{aligned} \tag{16}$$

where $\mathbf{A}^{(n)}(s)$ is defined by

$$\begin{aligned} \mathbf{A}^{(n)}(s) \equiv \int_0^s \int_0^{t_2}\cdots\int_0^{t_{n-2}}\int_0^{t_{n-1}} &\langle \tilde{\mathbf{B}}(s)\tilde{\mathbf{B}}(t_2)\cdots\tilde{\mathbf{B}}(t_{n-1})\tilde{\mathbf{B}}(t_n)\rangle \\ &\times dt_n \cdots dt_2. \end{aligned} \tag{17}$$

Clearly, $\mathbf{A}^{(n)}(s)$ is the analog, in terms of ordinary averages, of $\mathbf{G}^{(n)}(s)$ which is given in (15) by cumulant averages. Now, notice that if (16) is differentiated with respect to λ, n times, and then λ is set equal to 0, the first and third terms in (16) give

$$\frac{d^n}{d\lambda^n}\Big\langle \underset{\sim}{T}\exp\Big(\int_0^t \lambda\tilde{\mathbf{B}}(s)d\Big)\Big\rangle\Big|_{\lambda=0} = n!\int_0^t \mathbf{A}^{(n)}(s)ds. \tag{18}$$

The corresponding quantity which comes from the fourth term in (16) is

$$\frac{d^n}{d\lambda^n}\underset{\sim}{T}\exp\Big(\sum_{n=1}^{\infty}\lambda^n \int_0^t \mathbf{G}^{(n)}(s)ds\Big)\Big|_{\lambda=0} = ?. \tag{19}$$

The right-hand side of (19) is sufficiently complicated that a special method is required to arrive at an explicit formula.

The evaluation of (19) is a special case of the general problem of evaluating the nth derivative of a composite function which is time ordered. The method to be employed here is suggested by Riordan.[9] Let f be a time ordered composite function of the form

$$f(\lambda) = \underset{\sim}{T}y(w(\lambda)). \tag{20}$$

In (19), $y(u)$ is $\exp(u)$ and $w(\lambda)$ is $\sum_{n=1}^{\infty}\lambda^n \int_0^t \mathbf{G}^{(n)}(s)ds$. Denote $(d^n/d\lambda^n)f(\lambda)$ by f_n, $(d^n/d\lambda^n)w(\lambda)$ by w_n, and $(d^n/du^n)/y(u)$ by y_n. From (20) it follows that

$$\begin{aligned} f_1 &= \underset{\sim}{T}(w_1 y_1), \\ f_2 &= \underset{\sim}{T}(w_2 y_1 + w_1^2 y_2), \\ f_3 &= \underset{\sim}{T}(w_3 y_1 + 2w_2 w_1 y_2 + w_1 w_2 y_2 + w_1^3 y_3), \\ f_4 &= \underset{\sim}{T}(w_4 y_1 + 3w_3 w_1 y_2 + 3w_2^2 y_2 + w_1 w_3 y_2 + 3w_2 w_1^2 y_3 \\ &\quad + 2w_1 w_2 w_1 y_3 + w_1^2 w_2 y_3 + w_1^4 y_4) \\ &\ \ \vdots \end{aligned} \tag{21}$$

So far, careful preservation of the order of terms in these derivatives has been observed because noncommutivity is in general possible. The presence of the time ordering operator, $\underset{\sim}{T}$, however, makes this scrupulous observance of order unnecessary because $\underset{\sim}{T}$ permits treating the noncommutative quantities as if they commuted up until the time ordering is applied.[10] For example, $\underset{\sim}{T}(2w_2 w_1 y_2 + w_1 w_2 y_2) = \underset{\sim}{T}(3w_2 w_1 y_2)$. Therefore, Eqs. (21) may be rewritten as

$$\begin{aligned} f_1 &= \underset{\sim}{T}(w_1 y_1), \\ f_2 &= \underset{\sim}{T}(w_2 y_1 + w_1^2 y_2), \\ f_3 &= \underset{\sim}{T}(w_3 y_1 + 3w_2 w_1 y_2 + w_1^3 y_3), \\ f_4 &= \underset{\sim}{T}(w_4 y_1 + 4w_3 w_1 y_2 + 3w_2^2 y_2 + 6w_2 w_1^2 y_3 + w_1^4 y_4). \end{aligned} \tag{22}$$

These expressions are sufficient to suggest the conjecture that the nth term is

$$f_n = \sum_{\sum_{l=1}^{\infty} l m_l = n} \underset{\sim}{T}\Big\{\prod_{l=1}^{\infty}\frac{n!}{(l!)^{m_l} m_l!}(w_l)^{m_l} y_p\Big\}, \quad \text{where } p \equiv \sum_{l=1}^{\infty} m_l, \tag{23}$$

where $\sum_{l=1}^{\infty} l m_l = n$ defines a partition of n, and the summation in (23) is over all such partitions. The values of the multiplicities, m_l, are 0, 1, 2, $\cdots$. Therefore, $p \equiv \sum_{l=1}^{\infty} m_l$ is always finite because all but a finite number of the m_l's are zero. The formula given by (23) is the time ordered generalization of di Brunno's formula.[11] It may be proved by induction as follows.

Clearly, (23) agrees with (22) for $n=1$, 2, 3, and 4. Therefore, assume that (23) is true for n, and it shall be proved that it is true for $n+1$. Then, by induction, (23) is true for all positive n. Using (23), f_{n+1} can be directly obtained and is given by

$$f_{n+1} = \sum_{\sum_{l=1}^{\infty} l m_l = n} \underset{\sim}{T}\Big\{\prod_{l=2}^{\infty}\frac{n!}{(l!)^{m_l} m_l!}\frac{1}{m_1!}(w_l)^{m_l}(w_1)^{m_1+1} y_{p+1}\Big\} \tag{24}$$

$$+\sum_{\sum_{l=1}^{\infty} l m_l = n} \sum_{j=1}^{\infty} T\left\{ \prod_{\substack{l=1 \\ l\neq j \\ l\neq j+1}}^{\infty} \frac{n!}{(l!)^{m_l} m_l!} \frac{m_j}{(j!)^{m_j} m_j!} \times \frac{1}{((j+1)!)^{m_{j+1}} m_{j+1}!} (w_l)^{m_l} (w_j)^{m_j - 1} (w_{j+1})^{m_{j+1}+1} y_p \right\}$$

where $p \equiv \sum_{l=1}^{\infty} m_l$.

The first summation comes from one more derivative of y which generates another factor of w_1, whereas the second summation comes from the derivative of all other factors in (23) other than the y factor. It must be shown that (24) is of the form given by (23) for $n+1$, which is

$$f_{n+1} = \sum_{\sum_{l=1}^{\infty} l \bar{m}_l = n+1} T\left\{ \prod_{l=1}^{\infty} \frac{(n+1)!}{(l!)^{\bar{m}_l} \bar{m}_l!} (w_l)^{\bar{m}_l} y_{\bar{p}} \right\}, \quad \text{where } \bar{p} \equiv \sum_{l=1}^{\infty} \bar{m}_l. \tag{25}$$

To show the equivalence of (24) and (25) the following replacements are needed:

$$\frac{1}{m_1!} \to \frac{m_1+1}{(m_1+1)!}, \quad \frac{1}{(j!)^{m_j}} \to \frac{1}{(j!)^{m_j - 1} j!}, \quad \frac{1}{((j+1)!)^{m_{j+1}} m_{j+1}!} \to \frac{(j+1)!\,(m_{j+1}+1)}{((j+1)!)^{m_{j+1}+1} (m_{j+1}+1)!}. \tag{26}$$

Using (26) in (24) gives

$$f_{n+1} = \sum_{\sum_{l=1}^{\infty} l m_l = n} T\left\{ \prod_{l=2}^{\infty} \frac{n!\,(m_1+1)}{(l!)^{m_l} m_l!\,(m_1+1)!} (w_l)^{m_l} (w_1)^{m_1+1} y_{p+1} \right\}$$

$$+\sum_{\sum_{l=1}^{\infty} l m_l = n} \sum_{j=1}^{\infty} T\left\{ \prod_{\substack{l=1 \\ l\neq j \\ l\neq j+1}}^{\infty} \frac{n!\,(j+1)(m_{j+1}+1)}{(l!)^{m_l} m_l!\,(j!)^{m_j - 1} (m_j - 1)!\,[(j+1)!]^{m_{j+1}+1} (m_{j+1}+1)!} (w_l)^{m_l} (w_j)^{m_j - 1} (w_{j+1})^{m_{j+1}+1} y_p \right\} \quad \text{where } p \equiv \sum_{l=1}^{\infty} m_l. \tag{27}$$

In arriving at (27), $m_j/m_j!$ has been replaced by $1/(m_j - 1)!$, and this does not lead to difficulty when $m_j = 0$ if it is simply recalled that

$$\frac{1}{(m_j - 1)!} \equiv \frac{m_j}{m_j!} = \frac{0}{0!} = \frac{0}{1} = 0.$$

In the first summation in (27) let $\bar{m}_l = m_l$ for $l \geq 2$ and let $\bar{m}_1 = m_1 + 1$. Then it follows that

$$\sum_{l=1}^{\infty} l \bar{m}_l = (m_1 + 1) + \sum_{l=2}^{\infty} l m_l = 1 + \sum_{l=1}^{\infty} l m_l = 1 + n.$$

Similarly, in the second summation in (27) let $\bar{m}_l = m_l$ for $l = 1, 2, \cdots$, but $l \neq j$ and $l \neq j+1$, and let $\bar{m}_j = m_j - 1$ while letting $\bar{m}_{j+1} = m_{j+1} + 1$. Then it follows that

$$\sum_{l=1}^{\infty} l \bar{m}_l = \sum_{\substack{l=1 \\ l\neq j \\ l\neq j+1}}^{\infty} l m_l + j(m_j - 1) + (j+1)(m_{j+1} + 1) = \sum_{l=1}^{\infty} l m_l - j + (j+1) = n + 1.$$

In addition, in the first summation in (27)

$$p + 1 = \left(\sum_{l=1}^{\infty} m_l\right) + 1 = \sum_{l=1}^{\infty} \bar{m}_l,$$

whereas in the second summation in (27) $p = \sum_{l=1}^{\infty} m_l = \sum_{l=1}^{\infty} \bar{m}_l = \bar{p}$.

These considerations permit rewriting (27) as

$$f_{n+1} = \sum_{\substack{\sum_{l=1}^{\infty} l \bar{m}_l = n+1 \\ m_1 \geq 1}} T\left\{ \prod_{l=1}^{\infty} \frac{n!\,\bar{m}_1}{(l!)^{\bar{m}_l} \bar{m}_l!} (w_l)^{\bar{m}_l} y_{\bar{p}} \right\} + \sum_{\sum_{l=1}^{\infty} l \bar{m}_l = n+1} \sum_{\substack{j=1 \\ \bar{m}_1 \neq n+1}}^{\infty} T\left\{ \prod_{l=1}^{\infty} \frac{n!\,(j+1)\bar{m}_{j+1}}{(l!)^{\bar{m}_l} \bar{m}_l!} (w_l)^{\bar{m}_l} y_{\bar{p}} \right\}$$

$$\text{where } \bar{p} \equiv \sum_{l=1}^{\infty} \bar{m}_l. \tag{28}$$

Notice that the summations over partitions of $n+1$ are not unrestricted, and that the numerators still require simplification. Reductions of (28) to (25) requires consideration of two cases, $\bar{m}_1 = 0$ and $\bar{m}_1 \neq 0$. In the case in which $\bar{m}_1 = 0$ only the second summation can contribute and the summation over j gives

$$\sum_{j=1}^{\infty} (j+1)(\bar{m}_{j+1}) = \sum_{j=2}^{\infty} j \bar{m}_j = n + 1 \quad \text{when } \bar{m}_1 = 0. \tag{29}$$

Therefore, the numerator becomes $n!(n+1) = (n+1)!$ and the j summation is gone. In the case in which $\bar{m}_1 \neq 0$, both summations in (28) contribute, and the combination of the first summation with the j summation in the second summation yields

$$n!\left(\bar{m}_1 + \sum_{j=1}^{\infty} (j+1)\bar{m}_{j+1}\right) = n!\left(\bar{m}_1 + \sum_{j=2}^{\infty} j \bar{m}_j\right) = n!\left(\sum_{j=1}^{\infty} j \bar{m}_j\right) = n!\,(n+1) = (n+1)!. \tag{30}$$

Consequently, (28) is identical with (25). This completes the proof of (23).

The original objective was to find the right-hand side of (19) and equate it with the right-hand side of (18). Equation (19) may be treated as an example of (23) in which $y(u) \equiv \exp(u)$ and $w(\lambda) \equiv \sum_{n=1}^{\infty} \lambda^n \int_0^t \mathbf{G}^{(n)}(s)ds$ as was previously indicated in the discussion of (20). In this case, it follows that

$$y_n\big|_{\lambda=0} = y_n\big|_{u=0} = 1 \quad \text{for all } n$$

and

$$w_n\big|_{\lambda=0} = n! \int_0^t \mathbf{G}^{(n)}(s)ds. \tag{31}$$

Therefore, (23) becomes

$$f_n\big|_{\lambda=0} = \sum_{\sum_{l=1}^{\infty} l m_l = n} T\left\{ \prod_{l=1}^{\infty} \frac{n!}{(l!)^{m_l} m_l!} (l!)^{m_l} \left(\int_0^t \mathbf{G}^{(l)}(s)ds \right)^{m_l} \right\} = \sum_{\sum_{l=1}^{\infty} l m_l = n} T\left\{ \prod_{l=1}^{\infty} \frac{n!}{m_l!} \left(\int_0^t \mathbf{G}^{(l)}(s)ds \right)^{m_l} \right\}. \tag{32}$$

However, from (16) and (18) it is also true that

$$f_n\big|_{\lambda=0} = n! \int_0^t \mathbf{A}^{(n)}(s)ds. \tag{33}$$

Therefore, it follows that

$$\int_0^t \mathbf{A}^{(n)}(s)ds = \sum_{\sum_{l=1}^{\infty} l m_l = n} T\left\{ \prod_{l=1}^{\infty} \frac{1}{m_l!} \left(\int_0^t \mathbf{G}^{(l)}(s)ds \right)^{m_l} \right\}. \tag{34}$$

By explicitly working out (34) through $n = 4$, agreement with previously published expressions, derived by a more laborious method, will be reached.[8]

The problem of inverting (34) in order to obtain expressions for $\int_0^t \mathbf{G}^{(n)}(s)ds$ in terms of appropriate pro-

ducts of $\int_0^t \mathbf{A}^{(l)}(s)ds$ is not simple because (34) is nonlinear in $\int_0^t \mathbf{G}^{(l)}(s)ds$. This problem is the original objective and it has so far lead to (34). Its solution is obtained by applying the method used to obtain (34) in a different manner, as was also suggested by Riordan.[9] Use of (23) will again be required, although this time w and y will be different functions than those used to arrive at (34). It is necessary to find the appropriate new functions.

To get (34), use of (16) has been made; particularly

$$\sum_{n=0}^{\infty}\lambda^n \int_0^t \mathbf{A}^{(n)}(s)ds = \underset{\sim}{T}\exp\left(\sum_{n=0}^{\infty}\lambda^n \int_0^t \mathbf{G}^{(n)}(s)ds\right) = \left\langle \underset{\sim}{T}\exp\left(\lambda \int_0^t \bar{\mathbf{B}}(s)ds\right)\right\rangle. \tag{35}$$

Define $\mathbf{W}(\lambda,t)$ and $\mathbf{R}(\lambda,t)$ by

$$\mathbf{W}(\lambda,t)\equiv\sum_{n=1}^{\infty}\lambda^n\int_0^t \mathbf{G}^{(n)}(s)ds, \qquad \mathbf{R}(\lambda,t)\equiv\sum_{n=1}^{\infty}\lambda^n\int_0^t \mathbf{A}^{(n)}(s)ds. \tag{36}$$

Using (36) in (35) gives

$$\mathbf{1}+\mathbf{R}(\lambda,t)=\mathbf{1}+\sum_{n=1}^{\infty}\frac{1}{n!}\underset{\sim}{T}\{(\mathbf{W}(\lambda,t))^n\}. \tag{37}$$

Finally, define $\mathbf{R}'(\lambda,t)$ by

$$\mathbf{R}'(\lambda,t)\equiv\sum_{n=1}^{\infty}\frac{1}{n!}\underset{\sim}{T}\{(\mathbf{W}(\lambda,t))^n\}. \tag{38}$$

Consequently, (37) is simply

$$\mathbf{1}+\mathbf{R}(\lambda,t)=\mathbf{1}+\mathbf{R}'(\lambda,t). \tag{39}$$

For sufficiently small λ, both $\mathbf{R}(\lambda,t)$ and $\mathbf{R}'(\lambda,t)$ can be made to approach the zero matrix as closely as is desired. Therefore, it is possible to define a time-ordered logarithm by the formula

$$\underset{\sim}{\ln}[\mathbf{1}+\mathbf{R}(\lambda,t)]\equiv\sum_{l=1}^{\infty}(-1)^{l-1}\frac{1}{l}\underset{\sim}{T}\{(\mathbf{R}(\lambda,t))^l\} = \underset{\sim}{\ln}[\mathbf{1}+\mathbf{R}'(\lambda,t)]\equiv\sum_{l=1}^{\infty}(-1)^{l-1}\frac{1}{l}\underset{\sim}{T}\{(\mathbf{R}(\lambda,t))^l\} \tag{40}$$

If the matrices in (40) are replaced by commuting quantities, then the time ordering operations become unnecessary and the series correspond with the series expansion for the ordinary logarithm, $\ln(1+\lambda x)$, which is valid for sufficiently small λ. Because the λ derivatives to be performed will be evaluated at $\lambda=0$, (40) is entirely sufficient, even though it only makes sense for sufficiently small λ. To use (23) in this case, identify f with $\underset{\sim}{\ln}[\mathbf{1}+\mathbf{R}'(\lambda,t)]=\underset{\sim}{\ln}[\mathbf{1}+\mathbf{R}(\lambda,t)]$. For $y(u)$ in this case take: $y(u)\equiv\sum_{l=1}^{\infty}[(-1)^{l-1}/l]u^l$, and for $w(\lambda)$ take: $w(\lambda)\equiv\mathbf{R}(\lambda,t)$.

Observe that

$$\underset{\sim}{\ln}[\mathbf{1}+\mathbf{R}'(\lambda,t)]=\mathbf{W}(\lambda,t) \tag{41}$$

This remarkable result can be written as

$$\underset{\sim}{\ln}\{\underset{\sim}{T}\exp[\mathbf{W}(\lambda,t)]\}=\mathbf{W}(\lambda,t), \tag{42}$$

where use of (35) and (36) has been made. The proof of (41) and (42) goes as follows:

$$\underset{\sim}{\ln}[\mathbf{1}+\mathbf{R}'(\lambda,t)] =\sum_{l=1}^{\infty}(-1)^{l-1}\frac{1}{l}\underset{\sim}{T}\{(\mathbf{R}'(\lambda,t))^l\} =\sum_{l=1}^{\infty}(-1)^{l-1}\frac{1}{l}\underset{\sim}{T}\left\{\left(\sum_{n=1}^{\infty}\frac{1}{n!}\underset{\sim}{T}\{(\mathbf{W}(\lambda,t))^n\}\right)^l\right\} =\sum_{l=1}^{\infty}(-1)^{l-1}\frac{1}{l}\underset{\sim}{T}\left\{\left(\sum_{n=1}^{\infty}\frac{1}{n!}(\mathbf{W}(\lambda,t))^n\right)^l\right\}. \tag{43}$$

The last equality follows from the definition of the time ordering operator.[10] It may be expressed by: the time-ordered product of factors which are themselves time-ordered is equal to the time-ordered product of those same factors without internal time ordering.[12] The easiest way to see this point is to use the definition of $\underset{\sim}{T}$ which involves θ factions.[10] The last line of (43) is most easily evaluated after taking the time-ordering operator $\underset{\sim}{T}$ out in front of the sum over l. The result is

$$\underset{\sim}{\ln}[\mathbf{1}+\mathbf{R}'(\lambda,t)]=\underset{\sim}{T}\left\{\sum_{l=1}^{\infty}(-1)^{l-1}\frac{1}{l}\left(\sum_{n=1}^{\infty}\frac{1}{n!}(\mathbf{W}(\lambda,t))^n\right)^l\right\}. \tag{44}$$

Now, consider e^x for x positive but less than ln2. Then it is so that

$$e^x=1+\sum_{n=1}^{\infty}\frac{1}{n!}X^n \quad\text{and}\quad \sum_{n=1}^{\infty}\frac{1}{n!}X^n<1. \tag{45}$$

Consequently,

$$\ln\left(1+\sum_{n=1}^{\infty}\frac{1}{n!}X^n\right)=\sum_{l=1}^{\infty}(-1)^{l-1}\frac{1}{l}\left(\sum_{n=1}^{\infty}\frac{1}{n!}X^n\right)^l. \tag{46}$$

The form of (46) is precisely the same as the form of (44) provided λ is small enough. Consequently, because $\ln[1+\sum_{n=1}^{\infty}(1/n!)X^n]=\ln(e^X)=X$, it follows from (44) that

$$\underset{\sim}{\ln}[\mathbf{1}+\mathbf{R}'(\lambda,t)]=\mathbf{W}(\lambda,t). \tag{47}$$

The time-ordering operator in front of the right-hand side of (44) is not present explicitly in (47) because $\mathbf{W}(\lambda,t)$ is already fully time-ordered according to (36) and (15). This completes the proof of (41) and (42), but it must be remembered that the attendant analysis is only valid for sufficiently small λ, and (41) and (42) have not been proved for arbitrary λ.

Returning to the discussion between Eqs. (40) and (41), it follows that

$$f_n\big|_{\lambda=0}=\frac{d^n}{d\lambda^n}\underset{\sim}{\ln}[\mathbf{1}+\mathbf{R}'(\lambda,t)]\big|_{\lambda=0}=\frac{d^n}{d\lambda^n}\mathbf{W}(\lambda,t)\Big|_{\lambda=0} =n!\int_0^t \mathbf{G}^{(n)}(s)ds. \tag{48}$$

On the other hand, the equivalence of $\underset{\sim}{\ln}[\mathbf{1}+\mathbf{R}'(\lambda,t)]$ and $\underset{\sim}{\ln}[\mathbf{1}+\mathbf{R}(\lambda,t)]$, along with the choices for $y(u)$ and $w(\lambda)$ given in the discussion between Eqs. (40) and (41), imply through (23) that

$$f_n\Big|_{\lambda=0}=\sum_{\Sigma_{l=1}^{\infty}lm_l=n}\underset{\sim}{T}\left\{\prod_{l=1}^{\infty}\frac{n!}{(l!)^{m_l}m_l!}(w_l)^{m_l}y_p\Big|_{\lambda=0}\right\} =\sum_{\Sigma_{l=1}^{\infty}lm_l=n}\underset{\sim}{T}\left\{\prod_{l=1}^{\infty}\frac{n!}{(l!)^{m_l}m_l!}\left(l!\int_0^t \mathbf{A}^{(l)}(s)ds\right)^{m_l}(-1)^{p-1}(p-1)!\right\}, \tag{49}$$

where $p\equiv\sum_{l=1}^{\infty}m_l$. To get (49), the following two identities were used:

$$w_l\Big|_{\lambda=0}=\frac{d^l}{d\lambda^l}\left(\sum_{n=1}^{\infty}\lambda^n\int_0^t \mathbf{A}^{(n)}(s)ds\right)\Big|_{\lambda=0} =l!\int_0^t A^{(n)}(s)ds \tag{50}$$

and

$$y_p\Big|_{\lambda=0} = \frac{d^p}{du^p}\left(\sum_{l=1}^{\infty}(-1)^{l-1}\frac{1}{l}u^l\right)_{u=0}$$
$$= (-1)^{p-1}(p-1)!, \tag{51}$$

wherein $u=0$ corresponds with $\lambda=0$.

Combining (48) and (49) provides the desired inversion formula:

$$\int_0^t \mathbf{G}^{(n)}(s)ds$$
$$= \sum_{\sum_{l=1}^{\infty} l m_l = n} (-1)^{p-1}(p-1)!\,\underline{T}\left\{\prod_{l=1}^{n}\frac{1}{m_l!}\left(\int_0^t \mathbf{A}^{(l)}(s)ds\right)^{m_l}\right\}, \tag{52}$$

where $p \equiv \sum_{l=1}^{\infty} m_l$ in each summand of the partition summation. Formula (52) is identical with the results obtained from van Kampen's rules.[6,7] Together, (19), (17), and (52) provide a general expression for the nth order, time-ordered cumulant averages in terms of ordinary, time-ordered averages.

The averaged solution to (7) can be written using (13) and (14) as

$$\langle b_\alpha(t)\rangle = \left[\underline{T}\exp\left(\sum_{n=1}^{\infty}\int_0^t \mathbf{G}^{(n)}(s)ds\right)\right]_{\alpha\alpha'} b_{\alpha'}(0). \tag{53}$$

The corresponding differential equation is

$$\frac{d}{dt}\langle b_\alpha(t)\rangle = \sum_{n=1}^{\infty}[\mathbf{G}^{(n)}(t)]_{\alpha\alpha'}\langle b_{\alpha'}(t)\rangle. \tag{54}$$

By using (8), this corresponds with

$$\frac{d}{dt}\langle a_\alpha(t)\rangle = M_{\alpha\alpha'}\langle a_{\alpha'}(t)\rangle + [\exp(t\mathbf{M})]_{\alpha\beta}\sum_{n=1}^{\infty}[\mathbf{G}^{(n)}(t)]_{\beta\beta'} \times [\exp(-t\mathbf{M})]_{\beta'\alpha'}\langle a_{\alpha'}(t)\rangle. \tag{55}$$

Because only the first two moments of $\tilde{\mathbf{B}}(t)$ have been explicitly indicated in (10) and (11), only for $n=1$ and $n=2$ in (55) will explicit expressions be given at this point. From (52) it follows that

$$\mathbf{G}^{(1)}(t) = \mathbf{A}^{(1)}(t) = \langle\tilde{\mathbf{B}}(t)\rangle = 0. \tag{56}$$

The second and third equalities follow from (17) and (10) respectively. Again from (52), it follows that

$$\mathbf{G}^{(2)}(t) = \mathbf{A}^{(2)}(t) = \int_0^t \langle\tilde{\mathbf{B}}(t)\tilde{\mathbf{B}}(s)\rangle ds, \tag{57}$$

where the second equality follows from (17). Therefore, the $n=2$ term in (55) becomes

$$[\exp(t\mathbf{M})]_{\alpha\beta}\mathbf{G}^{(2)}_{\beta\beta'}(t)[\exp(-t\mathbf{M})]_{\beta'\alpha'}$$
$$= \int_0^t \langle\tilde{A}'_{\alpha\mu}(t)[\exp((t-s)\mathbf{M}]_{\mu\nu}\tilde{A}'_{\nu\theta}(s)[\exp((s-t)\mathbf{M})]_{\theta\alpha'}\rangle ds$$
$$= 2\int_0^t Q_{\alpha\mu\nu\theta}(t-s)[\exp((s-t)\mathbf{M}]_{\mu\nu}[\exp((s-t)\mathbf{M})]_{\theta\alpha'}ds, \tag{58}$$

where (9) and (6) have been used to get the last two lines. lines. Without explicit expressions for the higher order moments of $\tilde{\mathbf{A}}'(t)$, the higher order terms in (55) can be rewritten part way only, giving

$$\exp(t_1\mathbf{M})\mathbf{G}^{(n)}(t_1)\exp(-t_1\mathbf{M})$$
$$= \int_0^{t_1}\int_0^{t_2}\cdots\int_0^{t_{n-2}}\int_0^{t_{n-1}}\langle\tilde{\mathbf{A}}'(t_1)\prod_{l=2}^{n}\exp[(t_{l-1}-t_l)\mathbf{M}]\tilde{\mathbf{A}}'(t_l) \times\exp[(t_l-t_1)\mathbf{M}]\rangle_c dt_n\, dt_{n-1}\cdots dt_2. \tag{59}$$

This expression contains cumulant averaging in the integrand, and (52) must be used to get an expression in terms of ordinary averaging. The explicit, very complicated expression will not be written down here.

3. THE GAUSSIAN PROPERTY

When dealing with stochastic processes which do not involve the difficulties of noncommutivity, the stipulation that the stochastic process is Gaussian leads to great simplifications in the cumulant expression.[8] All cumulants of order higher than two vanish for Gaussian, commutative, stochastic processes.[8] However, the presence of noncommutivity leads to nonvanishing higher than second order cumulants.[8] It is for this reason that Kubo's concept of a generalized Gaussian process in the noncommutative case is invalid.[4,5] To further clarify this situation, the following brief review of the principal points of relevance are presented.

For convenience assume that $\mathbf{M}=0$ in the preceding section. It then follows that $\tilde{\mathbf{A}}(t) \equiv \tilde{\mathbf{A}}'(t) \equiv \tilde{\mathbf{B}}(t)$, and

$$\langle\tilde{\mathbf{A}}(t)\rangle = 0, \tag{60}$$

$$\langle\tilde{A}_{\alpha\beta}(t)\tilde{A}_{\alpha'\beta'}(s)\rangle = 2Q_{\alpha\beta\alpha'\beta'}(t-s). \tag{61}$$

The Gaussian property leads to explicit expressions for all higher order moments:

$$\langle\tilde{A}_{\alpha_1\beta_1}(t_1)\cdots A_{\alpha_{2n-1}\beta_{2n-1}}(t_{2n-1})\rangle = 0, \tag{62}$$

$$\langle\tilde{A}_{\alpha_1\beta_1}(t_1)\cdots\tilde{A}_{\alpha_{2n}\beta_{2n}}(t_{2n})\rangle$$
$$= \frac{1}{2^n n!}\sum_{p\in S_{2n}}\prod_{j=1}^{n}\langle\tilde{A}_{\alpha_{p(2j-1)}\beta_{p(2j-1)}}(t_{p(2j-1)})\tilde{A}_{\alpha_{p(2j)}\beta_{p(2j)}}(t_{p(2j)})\rangle, \tag{63}$$

where p is a permutation in the symmetric group of order $(2n)!$, S_{2n}. The easiest way to understand the origin of these formulas is to use the method of S. O. Rice.[13,14] $\tilde{A}(t)$ is represented by a Fourier series:

$$\tilde{A}(t) = \left(\frac{2}{T}\right)^{1/2}\sum_{n=1}^{N}\left(\tilde{\mathbf{A}}_n\cos\frac{2\pi nt}{T} + \tilde{\mathbf{B}}_n\sin\frac{2\pi nt}{T}\right), \tag{64}$$

in which both N and T ultimately are allowed to go to infinity and the coefficients $\tilde{\mathbf{A}}_n$ and $\tilde{\mathbf{B}}_n$ are time-independent, statistically independent, stochastic matrices with Gaussian distributions. In particular, the distribution for $\tilde{A}_n$ is of the form

$$D(\mathbf{A}_n) = \left(\frac{\|\mathbf{G}_n\|}{(2\pi)^n}\right)^{1/2}\exp(-\tfrac{1}{2}\mathbf{A}_n\circ\mathbf{G}_n\circ\mathbf{A}_n), \tag{65}$$

where $\mathbf{G}_n$ is a positive definite tetratic correlation with determinant $\|\mathbf{G}_n\|$. Using (64) in the left-hand side of (63) leads to all the different products of pairs on the right-hand side. In the special case in which all the correlation tetratics for all the $\mathbf{A}_n$'s and $\mathbf{B}_n$'s are identical, the time dependence in (61) will be $\delta(t-s)$, and $2\mathbf{Q} \equiv \mathbf{G}^{-1}$. This is seen by using (64) in (61) which gives

$$\langle A_{\alpha\beta}(t)A_{\alpha'\beta'}(s)\rangle$$
$$= \frac{2}{T}\sum_{n=1}^{N}\left(\langle\tilde{A}_{n\alpha\beta}\tilde{A}_{n\alpha'\beta'}\rangle\cos\frac{2\pi nt}{T}\cos\frac{2\pi ns}{T}\right.$$
$$\left. + \langle\tilde{B}_{n\alpha\beta}\tilde{B}_{n\alpha'\beta'}\rangle\sin\frac{2\pi nt}{T}\sin\frac{2\pi ns}{T}\right) \tag{66}$$
$$= \frac{2}{T}\sum_{n=1}^{N}\left(G^{-1}_{(\alpha\beta)(\alpha'\beta')}\cos\frac{2\pi nt}{T}\cos\frac{2\pi ns}{T}\right.$$

$$+G^{-1}_{(\alpha\beta)(\alpha'\beta')}\sin\frac{2\pi nt}{T}\sin\frac{2\pi ns}{T}\Big)$$

$$=\frac{2}{T}G^{-1}_{(\alpha\beta)(\alpha'\beta')}\sum_{n=1}^{N}\cos\frac{2\pi n(t-s)}{T}$$

$$\to G^{-1}_{(\alpha\beta)(\alpha'\beta')}\delta(t-s)\quad\text{as } T\to\infty \text{ and } N\to\infty.$$

The first equality follows from the postulated statistical independence of all the stochastic, matrix coefficients in (64). The second equality follows from the assumption for this special case that all correlation tetratics are the same, and the fact that

$$\langle\tilde{A}_{n_{\mu\nu}}\tilde{A}_{n_{\mu'\nu'}}\rangle\equiv\int\cdots\int\left(\frac{\|\mathbf{G}\|}{(2\pi)^n}\right)^{1/2}\exp[-\tfrac{1}{2}A_{n_{\alpha\beta}}G_{(\alpha\beta)(\alpha'\beta')}A_{n_{\alpha'\beta'}}]$$

$$\times A_{n_{\mu\nu}}A_{n_{\mu'\nu'}}d\mathbf{A}_n=G^{-1}_{(\mu\nu)(\mu'\nu')}. \tag{67}$$

The validity of (67) is perhaps best seen in a more familiar matrix notation in the case of a vector process, $\tilde{X}_i$, where $i=1,\ 2,\ \cdots,\ N$ with a Gaussian distribution given by

$$D(X)=[\|\mathbf{E}\|/(2\pi)^n]^{1/2}\exp(-\tfrac{1}{2}X_iE_{ij}X_j), \tag{68}$$

where $\|\mathbf{E}\|$ is the determinant of E_{ij} which is positive definite and symmetric. Therefore,

$$\langle X_lX_k\rangle=\int\cdots\int\left(\frac{\|\mathbf{E}\|}{(2\pi)^n}\right)^{1/2}\exp(-\tfrac{1}{2}X_iE_{ij}X_j)X_lX_k d\mathbf{X}$$

$$=E^{-1}_{ij}. \tag{69}$$

Equation (67) is the natural generalization of (69) when vectors become matrices, and matrices become tetratics. Other special cases will involve G_n's which are not all equal, and as a consequence the time dependence in (61) will not be simply $\delta(t-s)$ as in (66). Of particular interest is the special case in which G_n^{-1} has the form

$$\mathbf{G}_n^{-1}=2\mathbf{Q}\frac{\tau_c}{1+(2\pi n/T)^2\tau_c^2}[1-\exp(-T/\tau_c)]. \tag{70}$$

Following the procedure used in (66) leads to

$$\langle\tilde{A}_{\alpha\beta}(t)\tilde{A}_{\alpha'\beta'}(s)\rangle=2Q_{(\alpha\beta)(\alpha'\beta')}\frac{2}{T}\sum_{n=1}^{N}\cos\frac{2\pi n}{T}(t-s)$$

$$\times\frac{\tau_c}{1+(2\pi n/T)^2\tau_c^2}[1-\exp(-T/\tau_c)]$$

$$\to 2Q_{(\alpha\beta)(\alpha'\beta')}\exp(-|t-s|/\tau_c)$$

$$\text{as } N\to\infty \text{ and } T\to\infty \tag{71}$$

This special case will be of interest when Kubo's concept of a generalized Gaussian process is considered.

With the stipulation $\mathbf{M}\equiv 0$, introduced at the beginning of this section, the fourth-order cumulant, according to (52), (17), (15), and (9), is

$$\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\langle\tilde{\mathbf{A}}(t_1)\tilde{\mathbf{A}}(t_2)\tilde{\mathbf{A}}(t_3)\tilde{\mathbf{A}}(t_4)\rangle_c dt_4dt_3dt_2dt_1$$

$$=\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\langle\tilde{\mathbf{A}}(t_1)\tilde{\mathbf{A}}(t_2)\tilde{\mathbf{A}}(t_3)\tilde{\mathbf{A}}(t_4)\rangle dt_4dt_3dt_2dt_1$$

$$-\int_0^t\int_0^{t_1}\int_0^{t_1}\int_0^{t_3}\langle\tilde{\mathbf{A}}(t_1)\tilde{\mathbf{A}}(t_2)\rangle\langle\tilde{\mathbf{A}}(t_3)\tilde{\mathbf{A}}(t_4)\rangle dt_4dt_3dt_2dt_1. \tag{72}$$

Notice the upper limits on the second multiple integral of the right-hand side of (72). These upper limits follow from the time ordering in (52). Using (63) yields for the first multiple integral of the right-hand side of (72) the identity

$$\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_1\mu_2}(t_2)\tilde{A}_{\mu_2\mu_3}(t_3)\tilde{A}_{\mu_3\beta}(t_4)\rangle dt_4dt_3dt_2dt_1$$

$$=\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\{\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_1\mu_2}(t_2)\rangle\langle\tilde{A}_{\mu_2\mu_3}(t_3)\tilde{A}_{\mu_3\beta}(t_4)\rangle$$

$$+\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_2\mu_3}(t_3)\rangle\langle\tilde{A}_{\mu_1\mu_2}(t_2)\tilde{A}_{\mu_3\beta}(t_4)\rangle$$

$$+\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_3\beta}(t_4)\rangle\langle\tilde{A}_{\mu_1\mu_2}(t_2)\tilde{A}_{\mu_2\mu_3}(t_3)\rangle\}dt_4dt_3dt_2dt_1. \tag{73}$$

The second multiple integral of the right-hand side of (72) may be simplified also by noting the identity

$$\int_0^t\int_0^{t_1}\int_0^{t_1}\int_0^{t_3}g(t_1t_2t_3t_4)dt_4dt_3dt_2dt_1$$

$$=\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}g(t_1t_2t_3t_4)dt_4dt_3dt_2dt_1$$

$$+\int_0^t\int_0^{t_1}\int_0^{t_1}\int_0^{t_3}g(t_1t_2t_3t_4)dt_4dt_3dt_2dt_1$$

$$=\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}g(t_1t_2t_3t_4)dt_4dt_3dt_2dt_1$$

$$+\int_0^t\int_0^{t_1}\int_0^{t_3}\int_0^{t_3}g(t_1t_2t_3t_4)dt_4dt_2dt_3dt_1$$

$$=\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}g(t_1t_2t_3t_4)dt_4dt_3dt_2dt_1$$

$$+\int_0^t\int_0^{t_1}\int_0^{t_3}\int_0^{t_2}g(t_1t_2t_3t_4)dt_4dt_2dt_3dt_1$$

$$+\int_0^t\int_0^{t_1}\int_0^{t_3}\int_0^{t_3}g(t_1t_2t_3t_4)dt_4dt_2dt_3dt_1$$

$$=\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}g(t_1t_2t_3t_4)dt_4dt_3dt_2dt_1$$

$$+\int_0^t\int_0^{t_1}\int_0^{t_3}\int_0^{t_2}g(t_1t_2t_3t_4)dt_4dt_2dt_3dt_1$$

$$+\int_0^t\int_0^{t_1}\int_0^{t_3}\int_0^{t_4}g(t_1t_2t_3t_4)dt_2dt_4dt_3dt_1$$

$$=\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\{g(t_1t_2t_3t_4)+g(t_1t_3t_2t_4)$$

$$+g(t_1t_4t_2t_3)\}dt_4dt_3dt_2dt_1. \tag{74}$$

The first and third equalities follow from: $\int_0^x=\int_0^y+\int_y^x$. The second and fourth equalities follow from interchange of the order of integration: $\int_0^t\int_s^t dt'ds=\int_0^t\int_0^{t'}dsdt'$. The last equality follows from a renaming of time variable indices. The arbitrary function $g(t_1t_2t_3t_4)$ respects the order of the occurrence of the different indices on the time variables which are its arguments. Applying (74) to the case at hand gives

$$-\int_0^t\int_0^{t_1}\int_0^{t_1}\int_0^{t_3}\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_1\mu_2}(t_2)\rangle\langle\tilde{A}_{\mu_2\mu_3}(t_3)\tilde{A}_{\mu_3\beta}(t_4)\rangle$$

$$\times dt_4dt_3dt_2dt_1$$

$$=-\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\{\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_1\mu_2}(t_2)\rangle\langle\tilde{A}_{\mu_2\mu_3}(t_3)\tilde{A}_{\mu_3\beta}(t_4)\rangle$$

$$+\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_1\mu_2}(t_3)\rangle\langle\tilde{A}_{\mu_2\mu_3}(t_2)\tilde{A}_{\mu_3\beta}(t_4)\rangle$$

$$+\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_1\mu_2}(t_4)\rangle\langle\tilde{A}_{\mu_2\mu_3}(t_2)\tilde{A}_{\mu_3\beta}(t_3)\rangle\}dt_4dt_3dt_2dt_1. \tag{75}$$

Putting (73) and (75) into (72) gives

$$\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_1\mu_2}(t_2)\tilde{A}_{\mu_2\mu_3}(t_3)\tilde{A}_{\mu_3\beta}(t_4)\rangle_c dt_4dt_3dt_2dt_1$$

$$=\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\{\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_2\mu_3}(t_3)\rangle\langle\tilde{A}_{\mu_1\mu_2}(t_2)\tilde{A}_{\mu_3\beta}(t_4)\rangle$$

$$-\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_1\mu_2}(t_3)\rangle\langle\tilde{A}_{\mu_2\mu_3}(t_2)\tilde{A}_{\mu_3\beta}(t_4)\rangle$$

$$+\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_3\beta}(t_4)\rangle\langle\tilde{A}_{\mu_1\mu_2}(t_2)\tilde{A}_{\mu_2\mu_3}(t_3)\rangle$$
$$-\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_1\mu_2}(t_4)\rangle\langle\tilde{A}_{\mu_2\mu_3}(t_2)\tilde{A}_{\mu_3\beta}(t_3)\rangle\}dt_4dt_3dt_2dt_1. \tag{76}$$

Only the leading terms in each of (73) and (75) cancel out in general. Therefore, Kubo's specification that the fourth order cumulant vanish for a "generalized" Gaussian, stochastic, matrix process cannot, in general, occur. However, the situation is not too bad, after all, for the following reasons.

If the time dependence for the second moments of $\tilde{\mathbf{A}}(t)$ is the delta function $\delta(t-s)$, then (76) does indeed vanish, as do all higher order cumulants.[8,15] This is seen for (76) because the time integrals become

$$\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\delta(t_1-t_3)\delta(t_2-t_4)dt_4dt_3dt_2dt_1=0$$

and

$$\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\delta(t_1-t_4)\delta(t_2-t_3)dt_4dt_3dt_2dt_1=0. \tag{77}$$

If instead the time dependence is as in (71), the integrals become

$$\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\exp\left(-\frac{t_1-t_3}{\tau_c}\right)\exp\left(-\frac{t_2-t_4}{\tau_c}\right)dt_4dt_3dt_2dt_1$$
$$=t\left[\frac{\tau_c^3}{2}+\tau_c^3\exp\left(-\frac{t}{\tau_c}\right)\right]+2\frac{\tau_c^4}{2}\left[\exp\left(-\frac{t}{\tau_c}\right)-1\right]$$
$$+\frac{\tau_c^4}{4}\left[\exp\left(-\frac{2t}{\tau_c}\right)-1\right]$$
$$=\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\exp\left(-\frac{t_1-t_4}{\tau_c}\right)\exp\left(-\frac{t_2-t_3}{\tau_c}\right)$$
$$\times dt_4dt_3dt_2dt_1. \tag{78}$$

Indeed, (78) can be used to prove (77) if it is noticed that

$$\lim_{\tau_c\to 0}(1/\tau_c)\exp(-|\tau|/\tau_c)=\tfrac{1}{2}\delta(\tau). \tag{79}$$

For $t\gg\tau_c$, (78) becomes $\frac{1}{2}t\tau_c^3$. By using all of (71) the entire fourth cumulant becomes in this case

$$\int_0^t\int_0^{t_1}\int_0^{t_2}\int_0^{t_3}\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_1\mu_2}(t_2)\tilde{A}_{\mu_2\mu_3}(t_3)\tilde{A}_{\mu_3\beta}(t_4)\rangle_c dt_4dt_3dt_2dt_1$$
$$\xrightarrow[t\gg\tau_c]{}4(Q_{\alpha\mu_1\mu_2\mu_3}Q_{\mu_1\mu_2\mu_3\beta}-Q_{\alpha\mu_1\mu_1\mu_2}Q_{\mu_2\mu_3\mu_3\beta}$$
$$+Q_{\alpha\mu_1\mu_3\beta}Q_{\mu_1\mu_2\mu_2\mu_3}-Q_{\alpha\mu_1\mu_1\mu_2}Q_{\mu_2\mu_3\mu_3\beta})\tfrac{1}{2}t\tau_c^3. \tag{80}$$

Clearly, for a nonmatrix process there would be no indices and there would be a cancellation. The second cumulant in the case of (71) is

$$\int_0^t\int_0^{t_1}\langle\tilde{A}_{\alpha\mu_1}(t_1)\tilde{A}_{\mu_1\beta}(t_2)\rangle_c dt_2dt_1\xrightarrow[t\gg\tau_c]{}2Q_{\alpha\mu_1\mu_1\beta}t\tau_c \tag{81}$$

Both (80) and (81) are linear in t. However, suppose Q is proportional to a relaxation time τ_R, taken to the minus two power:

$$\mathbf{Q}^N\approx\tau_R^{-2}. \tag{82}$$

This is dimensionally correct, and it suggests that (80) and (81) have the forms

$$2Q_{\alpha\mu_1\mu_1\beta}t\tau_c\to 2Nt\tau_c/\tau_R^2$$

and

$$4(Q_{\alpha\mu_1\mu_2\mu_3}Q_{\mu_1\mu_2\mu_3\beta}+Q_{\alpha\mu_1\mu_3\beta}Q_{\mu_1\mu_2\mu_2\mu_3}$$
$$-2Q_{\alpha\mu_1\mu_1\mu_2}Q_{\mu_2\mu_3\mu_3\beta})\tfrac{1}{2}t\tau_c^3\to 2N^3t\tau_c^3/\tau_R^4, \tag{83}$$

where N is the approximate number of terms in the sum $Q_{\alpha\mu\mu\beta}$ which are large and of order τ_R^{-2}. This means that the ratio of the fourth order cumulant and the second order cumulant is $N^2\tau_c^2/\tau_R^2$. Now, suppose there are N' states labeled by the indices α, β, μ, ν, etc. If N is comparable to N' then the effective relaxation time is τ_R/N', which for large enough N' will surely be less than τ_c. Moreover, $N^2\tau_c^2/\tau_R^2$ will then be greater than 1, and the fourth cumulant becomes more important than the second. On the other hand, if $N\gg N'$ because not all pairs of states are coupled by equally strong stochastic matrix elements, then the effective relaxation time τ_R/N can still be longer than τ_c. In this situation the fourth order cumulant becomes unimportant compared with the second cumulant when $t>\tau_c$. This situation would imply that the $\langle a_\alpha(t)\rangle$'s relax more slowly than do the fluctuation correlations. In physical applications, it is this situation which is of most interest, and for it the second cumulant provides the dominant effects when $t>\tau_c$. Thus, two conditions are required:

$$t>\tau_c \text{ and } N\tau_c/\tau_R<1 \tag{84}$$

if the dynamics is to be dominated by the second cumulant only. This is, of course, essentially equivalent to the delta function theory[1], as was indicated by (79).

In summary, it may be claimed that Kubo's "generalized" Gaussian, stochastic, matrix process only obtains for a delta function correlation, for otherwise a truly Gaussian process will involve fourth and higher order cumulants. If, however, condition (84) holds, then the asymptotic behavior is dominated by the second cumulant after all. The fully rigorous justification of these remarks requires analysis of the higher order cumulants in special cases such as the ones studied here. It is perhaps ironic that the general theory presented here using cumulants, actually justifies the earlier, more simple, delta function theory[1], rather than superseding it, if (84) is satisfied. It should also be noted that dominance by the second cumulant in this situation requires $t>\tau_c$ but t need not be very very much greater than τ_c since $t\approx 10-100\tau_c$ already reduces the higher order corrections to very small effects.

4. THE DENSITY MATRIX

The mathematical results developed in the preceding sections will be applied now to the treatment of the stochastic Schrödinger equation[2,3]

$$i\frac{d}{dt}C_\alpha(t)=M_{\alpha\alpha'}C_{\alpha'}(t)+\tilde{M}_{\alpha\alpha'}(t)C_{\alpha'}(t), \tag{85}$$

in which $M_{\alpha\alpha'}=M^*_{\alpha'\alpha}$ and $\tilde{M}_{\alpha\alpha'}(t)=\tilde{M}^*_{\alpha'\alpha}(t)$. $\tilde{M}_{\alpha\alpha'}(t)$ corresponds to a stochastic Hamiltonian which is Hermitian. Suppose that a change of basis states is made so that $M_{\alpha\alpha'}$ becomes diagonal. Suppose further that there are N states corresponding with the degenerate eigenvalue λ. If consideration is restricted to these N degenerate eigenstates of $\mathbf{M}$, then (85) becomes

$$i\frac{d}{dt}C'_\alpha(t)=\lambda C'_\alpha(t)+\tilde{M}'_{\alpha\alpha'}(t)C'_{\alpha'}(t), \tag{86}$$

where α and α' range over 1, 2, $\cdots$, N. The stochastic Hamiltonian $\tilde{M}_{\alpha\alpha'}(t)$ will not couple eigenstates of $\mathbf{M}$

which belong to different degeneracy classes. This requirement is imposed in order to insure that the description given by (86) leads to a microcanonical equilibrium density matrix and not to the uniform density matrix. The reader is especially urged to see Ref. 3 for more details concerning this point. For convenience, the primes in (86) will be dropped in the following, and confusion with (85) should be avoided.

The moments for $\tilde{M}_{\alpha\alpha'}(t)$ are given by

$$\langle \tilde{M}_{\alpha\alpha'}(t)\rangle = 0, \tag{87}$$

$$\langle \tilde{M}_{\alpha\alpha'}(t)\tilde{M}_{\beta\beta'}(s)\rangle = 2Q_{\alpha\alpha'\beta\beta'}(t-s). \tag{88}$$

In (88), the time dependence of $Q_{\alpha\alpha'\beta\beta'}(t-s)$ can change as the values of the indices change, in general. The specific case in which the time dependence for all values of the indices is a delta function has been extensively treated in earlier publications.[1-3,16,17]

Although, in general it is unnecessary to do so, the imposition that $\tilde{M}_{\alpha\alpha'}(t)$ be also Gaussian can be made and leads to

$$\langle \tilde{M}_{\mu_1\nu_1}(t_1)\cdots\tilde{M}_{\mu_{2n-1}\nu_{2n-1}}(t_{2n-1})\rangle = 0, \tag{89}$$

$$\begin{aligned}&\langle \tilde{M}_{\mu_1\nu_1}(t_1)\cdots\tilde{M}_{\mu_{2n}\nu_{2n}}(t_{2n})\rangle\\ &= \frac{1}{2^n n!}\sum_{p\in S_{2n}}\prod_{j=1}^{n}\langle \tilde{M}_{\mu_{p(2j-1)}\nu_{p(2j-1)}}(t_{p(2j-1)})\\ &\quad\times\tilde{M}_{\mu_{p(2j)}\nu_{p(2j)}}(t_{p(2j)})\rangle\\ &= \frac{1}{n!}\sum_{p\in S_{2n}}\prod_{j=1}^{n}Q_{\mu_{p(2j-1)}\nu_{p(2j-1)}\mu_{p(2j)}\nu_{p(2j)}}(t_{p(2j-1)}-t_{p(2j)}).\end{aligned} \tag{90}$$

The density matrix $\rho_{\alpha\beta}(t)$ will be defined by[18]

$$\rho_{\alpha\beta}(t) \equiv C^*_\alpha(t)C_\beta(t), \tag{91}$$

and it satisfies the equation

$$\frac{d}{dt}\rho_{\alpha\beta}(t) = -i\tilde{L}_{\alpha\beta\alpha'\beta'}(t)\rho_{\alpha'\beta'}(t), \tag{92}$$

where $\tilde{L}_{\alpha\beta\alpha'\beta'}(t)$ is defined by

$$\tilde{L}_{\alpha\beta\alpha'\beta'}(t) \equiv \delta_{\alpha\alpha'}\tilde{M}_{\beta\beta'}(t) - \delta_{\beta\beta'}\tilde{M}^*_{\alpha\alpha'}(t). \tag{93}$$

Equation (92) is a special case of (4) when the $\mathbf{M}$ in (4) is zero. Of course, one must think of $\rho_{\alpha\beta}(t)$ as a doubly indexed vector, and $\tilde{L}_{\alpha\beta\alpha'\beta'}(t)$ as a quadruply indexed matrix to see the connection between (4) and (92). Therefore, the average of (92) is a special case of (55) in which the $\mathbf{M}$ in (55) is zero. This gives

$$\frac{d}{dt}\langle\rho_{\alpha\beta}(t)\rangle = \sum_{n=1}^{\infty}R^{(n)}_{\alpha\beta\alpha'\beta'}(t)\langle\rho_{\alpha'\beta'}(t)\rangle \tag{94}$$

wherein $R^{(n)}_{\alpha\beta\alpha'\beta'}(t)$ is defined by

$$\int_0^t \mathbf{R}^{(n)}(s)ds = \sum_{\sum_{l=1}^{\infty}lm_l=n}(-1)^{p-1}(p-1)!\,\mathbf{T}\left\{\prod_{l=1}^{\infty}\frac{1}{m_l!}\left(\int_0^t\mathbf{D}^{(l)}(s)ds\right)^{m_l}\right\}, \tag{95}$$

where $p \equiv \sum_{l=1}^{\infty}m_l$ and $n \geq 2$. For $n=1$, $R^{(1)}=0$. The $\mathbf{D}^{(l)}(s)$ in (95) is defined by

$$\begin{aligned}\mathbf{D}^{(n)}(s) \equiv (i)^n\int_0^s\int_0^{t_2}\cdots\int_0^{t_{n-2}}\int_0^{t_{n-1}}&\langle\tilde{\mathbf{L}}(s)\tilde{\mathbf{L}}(t_2)\cdots\\ &\times\tilde{\mathbf{L}}(t_{n-1})\tilde{\mathbf{L}}(t_n)\rangle dt_n\cdots dt_2.\end{aligned} \tag{96}$$

Equations (95) and (96) are special cases of (52) and (15) respectively, when $\mathbf{M}$ is zero in (52) and (15). The condition, $\mathbf{R}^{(1)}=0$, follows from (87) and (67).

The explicit expression for $R^{(2)}_{\alpha\beta\alpha'\beta'}(t)$ is

$$\begin{aligned}R^{(2)}_{\alpha\beta\alpha'\beta'}(t) &= -\int_0^t\langle\tilde{L}_{\alpha\beta\mu\nu}(t)\tilde{L}_{\mu\nu\alpha'\beta'}(s)\rangle ds\\ &= -\int_0^t\{\delta_{\alpha\alpha'}Q_{\beta\nu\nu\beta'}(t-s)+\delta_{\beta\beta'}Q_{\mu\alpha\alpha'\mu}(t-s)\\ &\quad -Q_{\beta\beta'\alpha'\alpha}(t-s)-Q_{\alpha'\alpha\beta\beta'}(t-s)\}ds.\end{aligned} \tag{97}$$

If it is further assumed that the stochastic Hamiltonian is stationary, then $Q_{\alpha\beta\alpha'\beta'}(t-s)=Q_{\alpha'\beta'\alpha\beta}(s-t)=Q_{\alpha'\beta'\alpha\beta}(t-s)$. This is equivalent with microscopic reversibility.[19] Using (89) and (90) in (96) and (95) provides explicit expressions for $\mathbf{R}^{(n)}(t)$ for $n>2$ in the Gaussian case.

In general, the microcanonical density matrix

$$\rho^{\text{microcanonical}}_{\alpha\beta} \equiv (1/N)\delta_{\alpha\beta} \tag{98}$$

is a solution to Eq. (94) regardless of any special assumptions about $\tilde{\mathbf{M}}(t)$, such as Gaussianness. This is seen from (95) and (96) as follows. It is necessary to show that

$$R_{\alpha\beta(n)\alpha'\beta'}(t)(1/N)\delta\alpha'\beta' = 0 \tag{99}$$

for each n. By using (97), this is clearly true for $n=2$. The second equality in (97) leads to

$$\begin{aligned}&-\int_0^t\{\delta_{\alpha\alpha'}Q_{\beta\nu\nu\beta'}(t-s)+\delta_{\beta\beta'}Q_{\mu\alpha\alpha'\mu}(t-s)-Q_{\beta\beta'\alpha'\alpha}(t-s)\\ &\quad -Q_{\alpha'\alpha\beta\beta'}(t-s)\}ds(1/N)\delta_{\alpha'\beta'}\\ &= -\int_0^t\{Q_{\beta\nu\nu\alpha}(t-s)+Q_{\mu\alpha\beta\mu}(t-s)-Q_{\beta\theta\theta\alpha}(t-s)\\ &\quad -Q_{\theta\alpha\beta\theta}(t-s)\}(1/N)ds \equiv 0.\end{aligned} \tag{100}$$

Alternatively, the first equality in (97) may be used to obtain the same result:

$$\begin{aligned}&-\int_0^t\langle L_{\alpha\beta\mu\nu}(t)\tilde{L}_{\mu\nu\alpha'\beta'}(t)\rangle ds(1/N)\delta_{\alpha'\beta'}\\ &\quad = -\int_0^t\langle\tilde{L}_{\alpha\beta\mu\nu}(t)\tilde{L}_{\mu\nu\alpha'\beta'}(t)(1/N)\delta_{\alpha'\beta'}\rangle ds \equiv 0\end{aligned} \tag{101}$$

because $\tilde{L}_{\mu\nu\theta\theta}(t)=0$ as follows from (93) and the hermiticity of $\tilde{\mathbf{M}}(t)$. It is the proof of (99) for $n=2$ given by (101) which readily generalizes to arbitrary n. Equation (95) shows that the computation of $R^{(n)}_{\alpha\beta\alpha'\beta'}(t)(1/N)\delta_{\alpha'\beta'}$ involves a sum over partitions of n in which each summand involves a sum of time ordered products of $\int_0^t\mathbf{D}^{(l)}(s)ds$'s. Each such time ordered product, according to (96), ends with some $\mathbf{D}^{(l)}(s)$ which ends with $\tilde{L}_{\mu\nu\alpha'\beta'}(t_{l'})$ in its integrand. Consequently, as in (101), $R^{(n)}_{\alpha\beta\alpha'\beta'}(t)(1/N)\delta_{\alpha'\beta'}$ ultimately reduces to an expression containing $\tilde{L}_{\mu\nu\alpha'\beta'}(t_{l'})(1/N)\delta_{\alpha'\beta'}\equiv 0$. This completes the proof of (99).

It is also the case in general that the time dependent solution to (94) satisfies

$$\langle\rho_{\alpha\alpha}(t)\rangle = 1. \tag{102}$$

To see this, it is sufficient to note that

$$R^{(n)}_{\alpha\alpha\alpha'\beta'}(t) = 0 \tag{103}$$

for each n. The proof of (103) is very similar to the proof of (99) and depends upon $\tilde{L}_{\alpha\alpha\alpha'\beta'}(t)\equiv 0$ as follows from (93) and the hermiticity of $\tilde{\mathbf{M}}(t)$.

Whether or not the microcanonical density matrix is approached asymptotically in time starting from an arbitrary initial density matrix depends upon the struc-

ture of $\sum_{n=1}^{\infty} R^{(n)}_{\alpha\beta\alpha'\beta'}(t)$. In the special case previously considered by the author in which the correlation in (88) is a delta function in time and $\tilde{\mathbf{M}}(t)$ is also Gaussian, it was shown that the approach to the microcanonical density matrix was in fact monotone because an H theorem was provable.[2] In the present, more general setting, a greater variety of behaviors is possible, including nonmonotonic approach to the microcanonical density matrix.

In the special case in which

$$Q_{\alpha\alpha'\beta\beta'}(t-s) \equiv Q_{\alpha\alpha'\beta\beta'} \exp(-|t-s|/\tau_c), \tag{104}$$

the presence of a finite correlation time τ_c does not necessarily destroy monotonicity. To see this, note that (104) implies

$$\langle \tilde{M}_{\alpha\alpha'}(t)\tilde{M}_{\beta\beta'}(s)\rangle = \langle \tilde{M}_{\alpha\alpha'}(t)\tilde{M}_{\beta\beta'}(t)\rangle \exp(-|t-s|/\tau_c). \tag{105}$$

Therefore, if $X_{\alpha\beta}$ is an arbitrary hermitian matrix, then

$$\begin{aligned} & X^*_{\alpha\beta} R^{(2)}_{\alpha\beta\alpha'\beta'} X_{\alpha'\beta'} \\ &= -\int_0^t \langle X^*_{\alpha\beta} \tilde{L}_{\alpha\beta\mu\nu}(t) \tilde{L}_{\mu\nu\alpha'\beta'}(s) X_{\alpha'\beta'}\rangle ds \\ &= -\int_0^t \langle (\tilde{L}_{\mu\nu\alpha\beta}(t)X_{\alpha\beta})^* (\tilde{L}_{\mu\nu\alpha'\beta'}(t)X_{\alpha'\beta'})\rangle \exp(-|t-s|/\tau_c) ds \\ &= -\langle (\tilde{L}_{\mu\nu\alpha\beta}(t)X_{\alpha\beta})^* (\tilde{L}_{\mu\nu\alpha'\beta'}(t)X_{\alpha'\beta'})\rangle \tau_c [1 - \exp(-t/\tau_c)] \\ &\leq 0. \end{aligned} \tag{106}$$

The first equality in (106) follows from (97), the second follows from (105) and $L_{\alpha\beta\mu\nu}(t) = L^*_{\mu\nu\alpha\beta}(t)$ which follows from (93), and the inequality follows from the form of the third equality. Therefore, the tetratic $R^{(2)}_{\alpha\beta\alpha'\beta'}(t)$, has nonpositive eigenvalues for its "eigenmatrices." As shown in (99), it also has a zero eigenvalue for the "eigenmatrix" $(1/N)\delta_{\alpha'\beta'}$. If it has no other zero eigenvalues, and this depends upon the details of $Q_{\alpha\alpha'\beta\beta'}$ in (104), then $R^{(2)}_{\alpha\beta\alpha'\beta'}(t)$ already shows an approach to equilibrium. It is also of interest to analyze the higher order terms in $\sum_{n=1}^{\infty} R^{(n)}_{\alpha\beta\alpha'\beta'}(t)$, but it has already been argued that that $R^{(2)}_{\alpha\beta\alpha'\beta'}(t)$ will dominate the asymptotic behavior of (94) provided that $\tilde{\mathbf{M}}(t)$ is Gaussian and that conditions analogous to (84) are satisfied. These results should be compared with the delta function theory.[1,2]

5. CONCLUDING REMARKS

The paper has addressed the following points:

(1) The rules for the construction of the nth order cumulant in the case of time dependent, stochastic, matrices or operators, as given by van Kampen have been derived. The derivation generalizes di Bruno's formula to the noncommutative situation, and makes use of the time ordered logarithm.

(2) Kubo's concept of a generalized Gaussian, stochastic, matrix or operator is criticized and shown to be valid only asymptotically, if certain conditions are met.

(3) The conditions which lead to asymptotic validity of Kubo's concept of a generalized Gaussian process also justify the delta function theory presented earlier by the author.

(4) A generalized density matrix equation is given. A discussion of the monotonicity of approach to the microcanonical density matrix follows the equation.

Remaining is the problem of discussing the details of how to describe the approach to canonical equilibrium of a subsystem in contact with a thermal reservoir. This is a special case of the microcanonical problem discussed here, and its presentation will be deferred.[3]

ACKNOWLEDGMENTS

This work was partially supported by NSF Grant GP 39309. The author wishes to express his appreciation to N. G. van Kampen for providing preprints of his rules for cumulant construction, and for other useful correspondence. H. A. Gersch is similarly acknowledged for several discussions related to this work. The author is a fellow of the Alfred P. Sloan Foundation.

*Alfred P. Sloan Fellow.

[1] R. F. Fox, J. Math. Phys. 13, 1196 (1972).
[2] R. F. Fox, J. Math. Phys. 13, 1726 (1972).
[3] R. F. Fox, J. Math. Phys. 14, 20 (1973).
[4] R. Kubo, "A Stochastic Theory of Line-Shape and Relaxation", chapter in *Fluctuation, Relaxation, and Resonance in Magnetic Systems*, edited by ter Haar (Oliver and Boyd, Edinburgh, 1962), pp. 23—68.
[5] R. Kubo, J. Math. Phys. 4, 174 (1963).
[6] N. G. van Kampen, "A Cumulant Expansion for Stochastic Linear Differential Equations I," a preprint received March, 1974.
[7] N. G. van Kampen, "A Cumulant Expansion for Stochastic Linear Differential Equations II," a preprint received March, 1974.
[8] R. F. Fox, J. Math. Phys. 15, 1479 (1974).
[9] J. Riordan, *An Introduction to Combinatorial Analysis* (Wiley, New York, 1958), Chap. 2, Sec. 8.
[10] S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Harper and Row, New York, 1961), Chap. 11, Sec. f, Eq. (153a).
[11] J. Riordan, Ref. 9, pp. 36—37.
[12] Q. Bui-Duy, Phys. Rev. D 9, 2794 (1974); see Eq. (5), p. 2796.
[13] M. C. Wang and G. E. Uhlenbeck, Rev. Mod. Phys. 17, 323 (1945), see especially Ref. 5 on p. 323, and Sec. 6, and note III of the Appendix.
[14] S. O. Rice, in *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover, New York, 1954), p. 157 and Part II.
[15] R. F. Fox, see Ref. 1, the Appendix.
[16] R. F. Fox, J. Math. Phys. 14, 1187 (1973).
[17] R. F. Fox, J. Math. Phys. 15, 217 (1974).
[18] This definition is the transpose of the definition found in my texts and articles, but it is consistent with the definition used earlier by the author.
[19] L. Onsager, Phys. Rev. 37, 405 (1931), Eq. 4.10.

The unitary irreducible representations of $\overline{SL}(3,R)$*

Dj. Šijački†

Department of Physics, Duke University, Durham, North Carolina 27706
(Received 5 July 1974)

The unitary irreducible representations of the universal covering group, $\overline{SL}(3,R)$, of the $SL(3,R)$ group are analyzed by means of the methods developed by Harish-Chandra and Kihlberg. We have found a single closed expression for the matrix elements of the noncompact generators for an arbitrary unitary representation of the $\overline{SL}(3,R)$ group. The irreducibility of the representations is achieved by using the little group technique and the scalar product for each irreducible Hilbert space is explicitly given. Contraction (in the Inönü and Wigner sense) of the $\overline{SL}(3,R)$ unitary irreducible representations to the corresponding representations of the $T_5 \textcircled{o} SU(2)$ group is discussed.

1. INTRODUCTION

The conventional approach in applications of group theory to particle physics is to postulate that the particles form a multiplet which furnishes a unitary irreducible representation of a symmetry group G which is an invariance group of the interaction Hamiltonian H. In this case the group generators X_a commute with the Hamiltonian.

If the number of states in the energy spectrum of the physical system is not finite one is naturally led to the use of noncompact symmetry groups. Their unitary irreducible representations are necessarily infinite-dimensional and could account for all states of the system. Thus, we see that a symmetry group may be useful even if its generators do not leave the Hamiltonian invariant.[1,2]

The states of hadronic spectrum exhibit an approximate symmetry (Regge trajectories) of particles with the same internal quantum numbers grouped together. Particles on the same trajectory have different spin values and possess the property that $\Delta J=2$. The higher spins could be excited by making use of the angular momentum[3] $\mathbf{L}=\mathbf{J}-\mathbf{S}$, which is, loosely speaking, an internal orbital angular momentum. One can adjoin to $\mathbf{L}$ the five Hermitian components of a noncompact quadrupole operator $\mathbf{T}$ so that $\mathbf{L}$ and $\mathbf{T}$ generate the algebra of the $SL(3,R)$ group.[4] Thus the orbital excitations are due to the quadrupole operator T and the meson and baryon states of higher J differ now from the lower ones in their L value. An algebraic model which exhibits both the $\Delta J=2$ rule for the orbital angular excitations together with a daughter-like spectrum has been constructed.[5] This model makes use of the relativistic extension of $SL(3,R)$.

The $SL(3,R)$ group has been recently successfully applied in studying the nuclear rotational motion.[6] In this application to nuclear physics the assumption was made that the states of a rotational nucleus form a basis for an irreducible representation of $SL(3,R)$ which is generated by the angular transition $E2$ operators. Also the set of admissible physical states associated with strong quantum gravitational fields for the Dirac Hamiltonian theory for general relativity exhibit approximate symmetry with respect to the unitary $SL(3,R)$ transformations.[7]

Our analysis is accomplished here by obtaining the irreducible Hermitian representations for the Lie algebra of $\overline{SL}(3,R)$. Due to a result of Nelson[8] these Lie algebra representations can be exponentiated to the corresponding continuous unitary irreducible representations of the $\overline{SL}(3,R)$ group. The necessary and sufficient condition for this procedure to go through is the hermiticity of Nelson's operator Δ, which is the sum of the squares of the group generators.

2. FOCUS ON $SL(3,R)$

$SL(3,R)$ is the group of linear unimodular transformations in a three-dimensional real vector space. The group is simple and noncompact Lie group. The group multiplication law is given as a product of transformations. If the group elements are given as 3×3 matrices then the product is matrix multiplication. The space of the group parameters is not simply connected and therefore the group is not simply connected. The maximal compact subgroup of $SL(3,R)$ is $SO(3)$. The universal covering group of $SL(3,R)$ we denote by $\overline{SL}(3,R)$. This group has the same Lie algebra as $SL(3,R)$. It is a simply connected group and its maximal compact subgroup is $SU(2)$, the covering group of $SO(3)$. The center of $\overline{SL}(3,R)$ is a two element group, i.e., Z_2, and the factor group of $\overline{SL}(3,R)$ with respect to Z_2 is isomorphic to $SL(3,R)$. Explicitly, we write

$$\overline{SL}(3,R)/Z_2 \simeq SL(3,R). \tag{2.1}$$

In the following we will always consider $SL(3,R)$ as a group of matrices.

Let $sl(3,R)$ be the Lie algebra of the $SL(3,R)$ group. It is an algebra of real 3×3 traceless matrices. The Cartan decomposition[9] of $sl(3,R)$ is $\mathbf{K}\oplus\mathbf{P}$, i.e., $[\mathbf{K},\mathbf{K}]\subset\mathbf{K}$, $[\mathbf{K},\mathbf{P}]\subset\mathbf{P}$, and $[\mathbf{P},\mathbf{P}]\subset\mathbf{K}$ is the decomposition of a traceless real matrix into the sum of an antisymmetric matrix and a symmetric one, which belong to $\mathbf{K}$ and $\mathbf{P}$, respectively. The antisymmetric matrices form an $so(3)$ algebra and they are the generators of $SO(3)$. The action of the adjoint algebra of $so(3)$, i.e., $\mathrm{ad}(so(3))$, on $\mathbf{P}$ [in the Cartan decomposition of $sl(3,R)$] is irreducible. The decomposition of matrices from $SL(3,R)$, which corresponds to the Cartan decomposition of the $sl(3,R)$ algebra, is the decomposition of a real unimodular matrix into the product of an orthogonal matrix and a symmetric one.

Let us denote by J_i the generators of the maximal compact subgroup $SO(3)$ of $SL(3,R)$. They constitute the angular momentum part of $sl(3,R)$. The infinitesimal generators which belong to the $\mathbf{P}$ part of the Cartan

decomposition of $SL(3,R)$ are the symmetric traceless matrices T_{ij} $(i,j=1,\ 2,\ 3)$ whose matrix elements are given by $(T_{ij})_{mn}=\delta_{im}\delta_{jn}-\frac{1}{3}\delta_{ij}\delta_{mn}$. The Lie algebra of $SL(3,R)$ is now given by the following commutation relations:

$$[J_i,\ J_j]=i\epsilon_{ijk}J_k,$$
$$[J_i,\ T_{jk}]=i\epsilon_{ijm}T_{mk}+i\epsilon_{ikn}T_{jn},$$
$$[T_{ij},\ T_{kl}]=-i(\delta_{ik}\epsilon_{jlm}+\delta_{il}\epsilon_{jkm}+\delta_{jk}\epsilon_{ilm}+\delta_{jl}\epsilon_{ikm})J_m.$$

It is convenient to write the $SL(3,R)$ generators in a sperical basis. Instead of J_i and T_{ij} we consider the generators whose matrix representations are given by

$$J_0=J_z=\begin{pmatrix}0 & -i & 0\\ i & 0 & 0\\ 0 & 0 & 0\end{pmatrix},\quad J_\pm=J_x\pm iJ_y=\begin{pmatrix}0 & 0 & \mp 1\\ 0 & 0 & -i\\ \pm 1 & i & 0\end{pmatrix},$$

$$T_0=-i\sqrt{\tfrac{2}{3}}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -2\end{pmatrix},\quad T_{\pm 1}=\begin{pmatrix}0 & 0 & \mp i\\ 0 & 0 & 1\\ \mp i & 1 & 0\end{pmatrix},$$

$$T_{\pm 2}=\begin{pmatrix}i & \mp 1 & 0\\ \mp 1 & -i & 0\\ 0 & 0 & 0\end{pmatrix}. \tag{2.2}$$

The minimal set of the commutation relations in the spherical basis is

$$\begin{aligned}&[J_0,\ J_\pm]=\pm J_\pm,\\ &[J_+,\ J_-]=2J_0,\\ &[J_0,\ T_\mu]=\mu T_\mu\quad(\mu=0,\ \pm 1,\ \pm 2),\\ &[J_\pm,\ T_\mu]=[6-\mu(\mu\pm 1)]^{1/2}T_{\mu\pm 1}\end{aligned} \tag{2.3}$$

and

$$[T_2,\ T_{-2}]=-4J_0. \tag{2.4}$$

The last commutation relation is known as the $SL(3,R)$ condition. The remaining commutation relations can be obtained by means of the Jacobi identity. From the aforementioned relations between the J's and T's it is clear that T_μ $(\mu=0,\ \pm 1,\ \pm 2)$ is a second rank irreducible tensor operator with respect to the $SO(3)$ subgroup of $SL(3,R)$. The commutation relations of $SL(3,R)$ in the spherical basis are invariant under substitution of $-T_\mu$ for T_μ.

3. REPRESENTATIONS

A covenient way to parametrize any noncompact semisimple Lie group is given by means of the Iwasawa decomposition.[9] The Iwasawa decomposition tells us that any semisimple noncompact group G can be written as a product

$$G=NAK$$

where N, A, and K are subgroups of G, i.e., every element $g\in G$ can be written as a product $g=nak$, where $n\in N$, $a\in A$, and $k\in K$. N is a nilpotent subgroup of G and its elements are upper triangular matrices. A is an Abelian subgroup of G and can be given as the group of diagonal matrices (with positive elements) and K is the maximal compact subgroup of G. For $G=SL(3,R)$ the elements of nilpotent subgroup N are

$$n=\begin{pmatrix}1 & \nu_3 & -\nu_2\\ 0 & 1 & \nu_1\\ 0 & 0 & 1\end{pmatrix},\quad \nu_1,\ \nu_2,\ \nu_3\in R,$$

and the elements of the Abelian subgroup A are

$$a=\begin{pmatrix}e^{\lambda} & 0 & 0\\ 0 & e^{\mu} & 0\\ 0 & 0 & e^{-\lambda-\mu}\end{pmatrix},\quad \lambda,\ \mu\in R.$$

The maximal compact subgroup K of $SL(3,R)$, i.e., $K=SO(3)$, will be parametrized in terms of three Euler angles α, β, and γ. Every element $k\in K$ is then of the form

$$k=R_z(\alpha)R_y(\beta)R_z(\gamma),$$

where, for instance, $R_z(\alpha)$ is a rotation in the three-dimensional vector space for the angle α around the z axis. In the following we will use two properties of the Iwasawa decomposition, namely, (1) the Iwasawa decomposition $G=NAK$, i.e., $g=nak$, $g\in G$, $n\in N$, $a\in A$, and $k\in K$ is unique[9] and (2) the product of some element $k\in K$ and an arbitrary element $g\in G$ is in general an arbitrary element of G which can be uniquely written as

$$kg=na(k,g)k\cdot g, \tag{3.1}$$

where $n\in N$, $a(k,g)\in A$, and $k\cdot g\in K$. The elements $k\cdot g$ and $a(k,g)$ satisfy

$$k\cdot(g_1g_2)=(k\cdot g_1)\cdot g_2 \tag{3.2}$$

and

$$a(k,g_1g_2)=a(k,g_1)\,a(k\cdot g_1,g_2) \tag{3.3}$$

for every g_1, $g_2\in G$. These relations follow immediately from the associativity of the group multiplication law, i.e., $k(g_1g_2)=(kg_1)g_2$. Namely,

$$k(g_1g_2)=na(k,g_1g_2)k\cdot(g_1g_2).$$

and

$$\begin{aligned}(kg_1)g_2&=(n_1a(k,g_1)k\cdot g_1)g_2\\ &=n_1a(k,g_1)n_2a(k\cdot g_1,g_2)\,(k\cdot g_1)\cdot g_2.\end{aligned}$$

Since N is an invariant subgroup in NA one has $a(k,g_1)n_2=n_2'a(k,g_1)$, $n_2'\in N$, and writing $n_1n_2'=n$, one obtains

$$(kg_1)g_2=na(k,g_1)\,a(k\cdot g_1,g_2)\,(k\cdot g_1)\cdot g_2$$

The uniqueness of decomposition provides us with the above-stated relations. When $a(k,g)$ is written as $\exp[h(k,g)]$ the second relation reads

$$h(k,g_1g_2)=h(k,g_1)+h(k\circ g_1,\ g_2). \tag{3.4}$$

It is well known that all unitary irreducible representations of compact groups are finite-dimensional. For noncompact groups, however, every nontrivial unitary representation is necessarily infinite-dimensional and this partly accounts for the additional complexity which occurs when one deals with their unitary representations. The class of real semisimple Lie groups is especially complex, and most of the progress in this direction has been made by Harish-Chandra.[10]

From the point of view of quantum mechanics one is

interested not only in single-valued group representations but also in projective or ray representations. This is due to the fact that central extensions of symmetry groups are physically significant. In this respect, it is sufficient to consider only the single-valued representations of the corresponding covering group, which in our case is $\overline{SL}(3,R)$. From the work of Bargmann[11] it follows that there is one-to-one correspondence between ray representations of a group and the single-valued representations of its universal covering group.

Owing to the Iwasawa decomposition every element $g \in SL(3,R)$ can be uniquely written as

$$g = ne^h k. \tag{3.5}$$

The Abelian subgroup of $\overline{SL}(3,R)$ has two generators A_1 and A_2 and if λ and μ are the corresponding group parameters, respectively, one has

$$h = \lambda A_1 + \mu A_2.$$

Let α be a linear, in general complex, function such that

$$\alpha(h) = \lambda\alpha(A_1) + \mu\alpha(A_2),$$

and let us denote $\alpha(A_1)$ and $\alpha(A_2)$ by a and b, respectively. Existence of the mapping α is guaranteed by the one-dimensionality of the irreducible representations of the Abelian subgroup. One usually takes for $\exp[\alpha(h)]$ the characters of the corresponding irreducible representations of the Abelian subgroup A. The mapping α can be extended in a natural way to a mapping from the group NA into the complex numbers since N is an invariant subgroup in NA.

The set of cosets of $\overline{SL}(3,R)$ with respect to the subgroup NA, i.e., $\overline{SL}(3.R)/NA$ is in one-to-one correspondence with the group $K = SU(2)$ and can be parametrized by the elements of K. In the coset space $\overline{SL}(3,R)/NA$ one has as well a measure, which we choose to be the invariant measure dk on K. Let $H = L^2(K)$ be the (separable) Hilbert space of functions on K which are square integrable with respect to the invariant measure on K, i.e.,

$H = \{f(k) \mid k \in K\}$, such that $\int dk f^*(k) f(k) < \infty$, and let $\int dk = 1$.

Harish-Chandra now defines a representation $U(g)$ of $G = \overline{SL}(3,R)$ on H in the following way.

$U(g)$ is a homomorphic continuous mapping from G into the set of linear transformations on H given by

$$U(g)f(k) = \exp[\alpha(h(k,g))]f(k \cdot g), \tag{3.6}$$

where $g \in G$, $f \in H$, $k \in K$, $e^h \in A$ and where $U(g)f(k)$ denotes the value of $U(g)f$ at the point k.

That $U(g)$ is a representation of G is straightforward:

$$U(g_1)U(g_2)f(k) = \exp[\alpha(h(k,g_1))]\, U(g_2)f(k \cdot g_1)$$
$$= \exp[\alpha(h(k,g_1))]\exp[\alpha(h(k \cdot g_1, g_2))]\, f((k \cdot g_1) \cdot g_2)$$
$$= \exp[\alpha(h(k, g_1 g_2))]\, f(k \cdot (g_1 g_2)).$$

The last step is due to the properties of the Iwasawa decomposition. Thus, $U(g_1)\, U(g_2) = U(g_1 g_2)$.

In Dirac bra-ket notation $f(k) = \langle k \mid f\rangle$, $U(g)f(k) = \langle k \mid U(g) \mid f\rangle$ and the representation defined above reads

$$\langle k \mid U(g) \mid f\rangle = \langle k \mid f'\rangle = \exp[\alpha(h(k,g))]\, \langle k \cdot g \mid f\rangle.$$

This is one reason why we have chosen the Iwasawa decomposition in the form NAK, and we multiply an element $k \in K$ by an element $g \in G$ on the right. Owing to the fact that for the group $\overline{SL}(3,R)$

$$\alpha(h) = \lambda a + \mu b, \tag{3.7}$$

we write $U(g)$ more explicitly as

$$U^{a,b}(g)f(k) = \exp[\lambda(k,g)a + \mu(k,g)b]\, f(k \cdot g). \tag{3.8}$$

Harish-Chandra[12] now defines the concept of infinitesimal equivalence of two representations in the following manner: Two representations are infinitesimally equivalent if there exists a similarity transformation of one representation into the other, with a nonsingular, not necessarily unitary, operator. In the case of equivalence there exists a unitary operator by means of which the transformation between the two representations is carried out. If both of two infinitesimally equivalent representations are unitary, then they are equivalent. Unitarity of either of these infinitesimally equivalent representations does not necessarily mean unitarity of the other. Suppose now that $U(g)$ is a representation of a group G on a Hilbert space H. Suppose further that H_1 and H_2 are the two closed invariant subspaces of H, such that $H_2 \subset H_1 \subset H$, and $H_1 \neq H_2$. Then $U(g)$ induces a representation $U'(g)$ on the quotient Hilbert space H_1/H_2 in a natural way. The representation $U'(g)$ is said to be deducible from the representation $U(g)$. If $U(g)$ is unitary, this means that $U'(g)$ is obtainable from $U(g)$ by decomposition.

Harish-Chandra[12] has proved that every unitary irreducible representation is infinitesimally equivalent to some irreducible representation deducible from some representation $U(g)$ of the form (3.6). Thus it is always possible to construct a bilinear form $(\tilde{f}, \tilde{g})$ in some quotient space H_1/H_2, where $\tilde{f}$ and $\tilde{g} \in H_1/H_2$. One can extend the domain of this bilinear form to all H_1 uniquely by defining $(\ ,\ )$ to vanish on H_2, i.e.,

$$(f,f) = 0, \quad f \in H_2,$$
$$(f,f) \neq 0, \quad f \in H_1 \ominus H_2.$$

Unitarity now means that

$$(U(g)f,\ U(g)f) = (f,f), \quad f \in H_1,\ g \in G, \tag{3.9}$$

and the additional conditions that the bilinear form is a scalar product are hermiticity and positive definiteness

$$(f,g) = (g,f)^*, \quad f,g \in H_1, \tag{3.10}$$

$$(f,f) \geq 0, \quad \forall f \in H_1. \tag{3.11}$$

It is convenient to extend the domain of the scalar product to the whole space H.

In the general case the scalar product (f,g) can be written in the form[13]

$$(f,g) = \int\int dk_1 dk_2 f^*(k_1)\, \rho(k_1,k_2)\, g(k_2), \quad f,g \in H, \tag{3.12}$$

where $\rho(k_1,k_2)$ is a kernel and the integration is over K

(the covering group of the maximal compact subgroup of G) and dk is an invariant measure.

In order to find the infinitesimal operators for the representation $U^{a,b}(g)$, we consider the group elements $\{g(\epsilon)\}$. Here the set of elements $\{g(\epsilon)\}$ form a one-dimensional subgroup of G with ϵ as a parameter. Let us denote by $X_\epsilon{}^{a,b}$ the infinitesimal operator for Harish-Chandra's representation $U^{a,b}(g)$ which corresponds to an arbitrary group parameter ϵ of $\overline{SL}(3,R)$. We define $X_\epsilon{}^{a,b}$ by

$$X_\epsilon{}^{a,b}f(k)=\lim_{\epsilon\to 0}\frac{i}{\epsilon}(U^{a,b}(g(\epsilon))-1)f(k), \tag{3.13}$$

with $g(0)$ being the unit element of $\overline{SL}(3,R)$. Thus,

$$X_\epsilon{}^{a,b}f(k)=i\frac{d}{d\epsilon}f(k\cdot g(\epsilon))\Big|_{\epsilon=0} + i\frac{d}{d\epsilon}(\lambda(k,g(\epsilon))a+\mu(k,g(\epsilon)b))\Big|_{\epsilon=0}f(k).$$

Now for the infinitesimal value of parameter ϵ we have

$$f(k\cdot g(\epsilon))\simeq f(k)+\epsilon\left[\left(\frac{d\alpha}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\alpha}+\left(\frac{d\beta}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\beta}+\left(\frac{d\gamma}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\gamma}\right]f(k).$$

The reason that only parameters from K occur on the right-hand side of this equation is that $k\cdot g(\epsilon)$ is an element from K [K being $SU(2)$ when G is $\overline{SL}(3,R)$] and therefore depends only on group parameters from K, i.e., α, β, and γ. Substituting this expression for $f(k\cdot g(\epsilon))$ into the previous expression for $X_\epsilon{}^{a,b}f(k)$, we arrive at the expression

$$X_\epsilon{}^{a,b}f(k)=i\left[\left(\frac{d\alpha}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\alpha}+\left(\frac{d\beta}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\beta}+\left(\frac{d\gamma}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\gamma}+\left(\frac{d\lambda}{d\epsilon}\right)_{\epsilon=0}a+\left(\frac{d\mu}{d\epsilon}\right)_{\epsilon=0}b\right]f(k). \tag{3.14}$$

This expression was obtained by Kihlberg[13] using an alternative method.

If we extend our underlying space from $\overline{SL}(3,R)/NA$ to $\overline{SL}(3,R)/N$, i.e., to the space of functions $\exp(\lambda a+\mu b)f(k)$, then $X_\epsilon{}^{a,b}$, is a differential operator in the space of functions belonging to $H=L^2(\overline{SL}(3,R)/N)$, and has the form

$$X^{a,b}=i\left[\left(\frac{d\alpha}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\alpha}+\left(\frac{d\beta}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\beta}+\left(\frac{d\gamma}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\gamma}+\left(\frac{d\lambda}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\lambda}+\left(\frac{d\mu}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\mu}\right]. \tag{3.15}$$

It is interesting to point out, at this stage, that $X_\epsilon{}^{a,b}$ is not a differential operator in the group parameters corresponding to the nilpotent subgroup N. This holds for all generators of the $\overline{SL}(3,R)$ algebra since ϵ could be the parameter of any one-dimensional subgroup in $\overline{SL}(3,R)$.

In Appendix A we explicitly evaluate the $\overline{SL}(3,R)$ generators. They are

$$J_0=i\frac{\partial}{\partial\gamma},$$

$$J_\pm=\exp(\mp i\gamma)\left(-\frac{i}{s\beta}\frac{\partial}{\partial\alpha}\mp\frac{\partial}{\partial\beta}+ictg\frac{\partial}{\partial\gamma}\right),$$

$$T_0=i\Big(\left(\tfrac{2}{3}\right)^{1/2}(3c^2\alpha s^2\beta-1)\frac{\partial}{\partial\lambda}+\left(\tfrac{2}{3}\right)^{1/2}(3s^2\alpha s^2\beta-1)\frac{\partial}{\partial\mu}+\sqrt{6}\,s\alpha c\alpha s^2\beta\frac{\partial}{\partial\alpha}-\sqrt{6}\,s\beta c\beta\frac{\partial}{\partial\beta}\Big),$$

$$T_{\pm1}=\exp(\mp i\gamma)\Big((-2s\alpha c\alpha s\beta\mp 2ic^2\alpha s\beta c\beta)\frac{\partial}{\partial\lambda}+(2s\alpha c\alpha s\beta\mp 2is^2\alpha s\beta c\beta)\frac{\partial}{\partial\mu}+((c^2\alpha-s^2\alpha)s\beta-c\beta ctg\beta\mp 2is\alpha c\alpha s\beta c\beta)\frac{\partial}{\partial\alpha}\pm i(c^2\beta-s^2\beta)\frac{\partial}{\partial\beta}+ctg\beta\frac{\partial}{\partial\gamma}\Big), \tag{3.16}$$

$$T_{\pm2}=\exp(\mp 2i\gamma)\Big((\pm 2s\alpha c\alpha c\beta+i(c^2\alpha c^2\beta-s^2\alpha))\frac{\partial}{\partial\lambda}+(\mp 2s\alpha c\alpha c\beta+i(s^2\alpha c^2\beta-c^2\alpha))\frac{\partial}{\partial\mu}+(\mp 2c^2\alpha c\beta+is\alpha c\alpha(1+c^2\beta))\frac{\partial}{\partial\alpha}+is\beta c\beta\frac{\partial}{\partial\beta}\pm\frac{\partial}{\partial\gamma}\Big).$$

Besides the generators J_0 and $J_\pm$ of the maximal compact subgroup of $\overline{SL}(3,R)$, it is possible to introduce generators K_0 and $K_\pm$ of another $SU(2)$ group by exchanging the parameters α and γ in the expressions for J_0 and $J_\pm$, respectively. Explicitly,

$$K_0=i\frac{\partial}{\partial\alpha},$$

$$K_\pm=\exp(\mp i\alpha)\left(ictg\beta\frac{\partial}{\partial\alpha}\mp\frac{\partial}{\partial\beta}-\frac{i}{s\beta}\frac{\partial}{\partial\gamma}\right). \tag{3.17}$$

The operators **J** and **K** mutually commute. Also, the corresponding invariant operators are equal, i.e., $J^2=K^2$. The significance of the operator K_0, as it will be seen later, is in characterizing the degeneracy of the $SU(2)$ multiplets in the unitary irreducible representations of $\overline{SL}(3,R)$.

4. MATRIX ELEMENTS

In order to analyse the representations of $\overline{SL}(3,R)$ it is convenient to have the matrix elements of the group generators. Also, in this case the task of determining the scalar products of the unitary representations is considerably simplified. The Hilbert space H in which some representation of $\overline{SL}(3,R)$ is defined, is a symmetric homogeneous space over the group $SU(2)$, i.e., the set $L^2(SU(2))$ of square-integrable functions over $SU(2)$ with respect to the invariant measure over $SU(2)$. An Arbitrary function $f(k)$, $k\in SU(2)$ can be written as

$$f(k)=\sum_{JKM}f^J_{KM}D^J_{KM}(k),\quad k\in SU(2), \tag{4.1}$$

where $f^J_{KM}\in C$, and $D^J_{KM}(k)$ are the matrix elements of the unitary irreducible representations of $SU(2)$. Thus, it is sufficient to know the matrix elements of the group generators in the $|{}^J_{KM}\rangle$ basis, where

$$\left\langle k\left|{}^J_{KM}\right.\right\rangle\equiv(2J+1)^{1/2}D^J_{KM}(k)$$

and

$$\left\langle \begin{matrix} J' \\ K'M' \end{matrix} \middle| \begin{matrix} J \\ KM \end{matrix} \right\rangle = \delta_{J'J}\delta_{K'K}\delta_{M'M}.$$

Since $SU(2)$ is the maximal compact subgroup of $\overline{SL}(3,R)$, the label J can take both integer and half-integer values. Confining J to integer values only, one obtains results valid for the $SL(3,R)$ group.

The differential operators

$$J_0 = i\frac{\partial}{\partial\gamma}$$

and

$$J_\pm = \exp(\mp i\gamma)\left(\frac{-i}{s\beta}\frac{\partial}{\partial\alpha} \mp \frac{\partial}{\partial\beta} + ictg\beta\frac{\partial}{\partial\gamma}\right)$$

of the $SU(2)$ subgroup are well-known ones, and their matrix elements are

$$\begin{aligned} &\left\langle \begin{matrix} J' \\ K'M' \end{matrix} \middle| J_0 \middle| \begin{matrix} J \\ KM \end{matrix} \right\rangle = M\delta_{J'J}\delta_{K'K}\delta_{M'M}, \\ &\left\langle \begin{matrix} J' \\ K'M' \end{matrix} \middle| J_\pm \middle| \begin{matrix} J \\ KM \end{matrix} \right\rangle = [J(J+1) - M(M\pm1)]^{1/2}\delta_{J'J}\delta_{K'K}\delta_{M'M\pm1}. \end{aligned} \tag{4.2}$$

Furthermore, the generators of the compact subgroup are Hermitian, and therefore the representations of $\overline{SL}(3,R)$ when restricted to the maximal compact subgroup are unitary.

The coefficients in the differential form for T_0 can be expressed in terms of D^J_{KM} functions (cf. Appendix B) in the following form

$$\begin{aligned} T_0 = &-i\sqrt{\tfrac{2}{3}}(a+b)D^2_{00} + i(D^2_{0-1}J_+ + D^2_{01}J_-) \\ &+ i(2+a-b)(D^2_{20}+D^2_{-20}) + i(D^2_{20}-D^2_{-20})K_0, \end{aligned} \tag{4.3}$$

where a and b are the eigenvalues of the operators $\partial/\partial\lambda$ and $\partial/\partial\mu$, respectively.

Making use of the orthogonality relations for D^J_{KM} functions, we obtain (cf. Appendix B)

$$\begin{aligned} &\left\langle \begin{matrix} J' \\ K'M' \end{matrix} \middle| T_0 \middle| \begin{matrix} J \\ KM \end{matrix} \right\rangle \\ &= (-)^{J'-K'}(-)^{J'-M'}[(2J'+1)(2J+1)]^{1/2}\left\{ \begin{pmatrix} J' & 2 & J \\ -M' & 0 & M \end{pmatrix} \right. \\ &\times\left[-i\sqrt{\tfrac{2}{3}}(a+b)\begin{pmatrix} J' & 2 & J \\ -K' & 0 & K \end{pmatrix} + i(2+K+a-b)\begin{pmatrix} J' & 2 & J \\ -K' & 2 & K \end{pmatrix} \right. \\ &\left. + i(2-K+a-b)\begin{pmatrix} J' & 2 & J \\ -K' & -2 & K \end{pmatrix} \right] + i\begin{pmatrix} J' & 2 & J \\ -K' & 0 & K \end{pmatrix} \\ &\times\left[[J(J+1)-M(M+1)]^{1/2}\begin{pmatrix} J' & 2 & J \\ -M' & -1 & M+1 \end{pmatrix} \right. \\ &\left. \left. + [J(J+1)-M(M-1)^{1/2}\begin{pmatrix} J' & 2 & J \\ -M' & 1 & M-1 \end{pmatrix} \right] \right\}. \end{aligned}$$

The recurrence relation[14]

$$\begin{aligned} &[2m_1m_2 + j_1(j_1+1) - j_2(j_2+1) - j_3(j_3+1)]\begin{pmatrix} j_2 & j_3 & j_1 \\ m_2 & -m_1 & m_1-m_2 \end{pmatrix} \\ &= [j_3(j_3+1)-m_1(m_1+1)]^{1/2}[j_2(j_2+1)-m_2(m_2+1)]^{1/2} \\ &\times\begin{pmatrix} j_2 & j_3 & j_1 \\ m_2+1 & -m_1-1 & m_1-m_2 \end{pmatrix} \\ &+ [j_3(j_3+1)-m_1(m_1-1)]^{1/2}[j_2(j_2+1)-m_2(m_2-1)]^{1/2} \\ &\times\begin{pmatrix} j_2 & j_3 & j_1 \\ m_2-1 & -m_1+1 & m_1-m_2 \end{pmatrix}. \end{aligned}$$

for $j_1=J'$, $j_2=2$, $j_3=J$, $m_1=-K$, $m_2=0$, and $m_1-m_2=-K'$ now becomes

$$\begin{aligned} &[J(J+1)-M(M+1)]^{1/2}\begin{pmatrix} J' & 2 & J \\ -M' & -1 & M+1 \end{pmatrix} \\ &+ [J(J+1)-M(M-1)]^{1/2}\begin{pmatrix} J' & 2 & J \\ -M' & 1 & M-1 \end{pmatrix} \\ &= \left(\frac{J'(J'+1)-J(J+1)}{\sqrt{6}} - \sqrt{6}\right)\begin{pmatrix} J' & 2 & J \\ -M' & 0 & M \end{pmatrix}. \end{aligned}$$

Substituting this relation into the form for the matrix elements of T_0, we obtain

$$\left\langle \begin{matrix} J' \\ K'M' \end{matrix} \middle| T_0 \middle| \begin{matrix} J \\ KM \end{matrix} \right\rangle = (-)^{J'-M'}\begin{pmatrix} J' & 2 & J \\ -M' & 0 & M \end{pmatrix}\left\langle \begin{matrix} J' \\ K' \end{matrix} \middle\| T \middle\| \begin{matrix} J \\ K \end{matrix} \right\rangle \tag{4.4}$$

the reduced matrix elements $\langle {}^{J'}_{K'}\|T\|^{J}_{K}\rangle$ being given by

$$\begin{aligned} &\left\langle \begin{matrix} J' \\ K' \end{matrix} \middle\| T \middle\| \begin{matrix} J \\ K \end{matrix} \right\rangle = (-)^{J'-K'}[(2J'+1)(2J+1)]^{1/2} \\ &\times\left[-i\left(\sqrt{\tfrac{2}{3}}\sigma + \sqrt{6} - \frac{J'(J'+1)-J(J+1)}{\sqrt{6}}\right)\begin{pmatrix} J' & 2 & J \\ -K' & 0 & K \end{pmatrix} \right. \\ &+ i(2+K+\delta)\begin{pmatrix} J' & 2 & J \\ -K' & 2 & K \end{pmatrix} \\ &\left. + i(2-K+\delta)\begin{pmatrix} J' & 2 & J \\ -K' & -2 & K \end{pmatrix} \right]. \end{aligned} \tag{4.5}$$

In this expression we have introduced $\sigma = a+b$ and $\delta = a-b$. Note the selection rule: $\Delta K = 0, \pm2$.

Having this expression we can immediately write the matrix elements for an arbitrary component of the quadrupole tensor operator T_μ, $\mu = 0, \pm1, \pm2$ as

$$\left\langle \begin{matrix} J' \\ K'M' \end{matrix} \middle| T_\mu \middle| \begin{matrix} J \\ KM \end{matrix} \right\rangle = (-)^{J'-M'}\begin{pmatrix} J' & 2 & J \\ -M' & \mu & M \end{pmatrix}\left\langle \begin{matrix} J' \\ K' \end{matrix} \middle\| T \middle\| \begin{matrix} J \\ K \end{matrix} \right\rangle, \tag{4.6}$$

the reduced matrix elements being given by (4.5).

5. UNITARITY

In the previous part of this work we have evaluated

expressions for the group generators as differential operators and we have obtained the explicit form of their matrix elements in the basis $\{|{}^{J}_{KM}\rangle\}$ of vectors which span the space $L^2(SU(2))$. Both the differential forms and the matrix elements of the quadrupole operator are functions of two parameters $\sigma = a + b$ and $\delta = a - b$. In general the quadrupole operator T_μ is not Hermitian and therefore the corresponding representations of $\overline{SL}(3,R)$ are not unitary. From the requirement that the matrix elements of T_μ obey the hermiticity condition we will obtain the constraints on σ and δ, which are the unitary constraints of the representations of the $\overline{SL}(3,R)$ group.

Let us first consider the Hilbert space $L^2(SU(2))$ with the basis $\{|{}^{J}_{KM}\rangle\}$. The hermiticity condition on the T_μ operator, i.e.,

$$T^{\dagger}_{\mu} = (-)^{\mu} T_{-\mu}, \tag{5.1}$$

yields

$$(-)^{J} \left\langle {J' \atop K'} \middle\| T \middle\| {J \atop K} \right\rangle = (-)^{J'} \left\langle {J \atop K} \middle\| T \middle\| {J' \atop K'} \right\rangle^{*}. \tag{5.2}$$

Using the explicit form for $\langle {}^{J'}_{K'} \| T \| {}^{J}_{K} \rangle$ and the symmetry properties of the 3-j symbols, we extract, from the hermiticity condition, the equations

$$\sigma + \sigma^* = -6 \quad \text{for } K' = K$$

and

$$\delta + \delta^* = -2 \quad \text{for } K' = K \pm 2.$$

For $\sigma = \sigma_0 + i\sigma_2$ and $\delta = \delta_0 + i\delta_2$ it follows immediately $\sigma_0 = -3$, $\delta_0 = -1$, $\sigma_2, \delta_2 \in R$.

We now proceed to the general case. The unitarity of the representations (i.e., the hermiticity of the generators) is a property of the representations which depends on the scalar product of the Hilbert space in which each representation is defined. If $f(k)$ and $g(k)$ are any two functions, the most general scalar product is given by

$$(f,g) = \int\int dk_1\, dk_2 f^*(k_1)\, \rho(k_1, k_2)\, g(k_2), \tag{5.3}$$

$\rho(k_1, k_2)$ being the kernel of the scalar product while dk_1 and dk_2 are invariant measures over $SU(2)$. The Hilbert space with this scalar product we denote by $H = L^2(SU(2), \rho)$. Obviously, when $\rho(k_1, k_2) = \delta(k_1 k_2^{-1})$, the Dirac δ-function, we recover the usual product $(f,g) = \int dk f^*(k)\, g(k)$. The problem of finding the unitary representations is now reduced to that of finding all possible scalar products, or, in other words, to the problem of finding all possible kernels.

For an arbitrary element $k \in SU(2)$ the unitarity of the representation means that

$$(U(k)f, U(k)g) = (f,g),$$

and from the definition of the representations (3.8) we have that

$$(U(k)f, U(k)g) = \int\int dk_1\, dk_2 f^*(k_1 \cdot k)\, \rho(k_1, k_2)\, g(k_2 \cdot k).$$

Since dk is an invariant measure, we obtain the following constraint on the form of the kernel:

$$\rho(k_1, k_2) = \rho(k_1 k_2^{-1}). \tag{5.4}$$

In terms of the complete set of functions $\{(2J+1)^{1/2} D^{J}_{KM}\}$ over $SU(2)$ one writes

$$f(k) = \sum_{JKM} (2J+1)^{1/2} f^{J}{}_{KM} D^{J}{}_{KM}(k).$$

obtaining

$$(f,g) = \sum_{\substack{JKM \\ J'K'M'}} f^{J'}_{K'M'} {}^{*} g^{J}_{KM} \left[(2J'+1)(2J+1)\right]^{1/2} \times \int\int dk_1\, dk_2 D^{J'}_{K'M'} {}^{*}(k_1)\, \rho(k_1 k_2^{-1})\, D^{J}_{KM}(k_2).$$

Since dk is an invariant measure and

$$D^{J'}_{K'M'}(kk_2) = \sum_{N} D^{J'}_{K'N}(k)\, D^{J}_{NM'}(k_2),$$

we find that

$$(f,g) = \sum_{JK'K} \frac{1}{(2J+1)^{1/2}} \left(\sum_{M} f^{J*}_{K'M} g^{J}_{KM} \right) \left\langle {J \atop K'K} \middle| \rho \middle| {0 \atop 00} \right\rangle, \tag{5.5}$$

where

$$\left\langle {J \atop K'K} \middle| \rho \middle| {0 \atop 00} \right\rangle = (2J+1)^{1/2} \int dk\, D^{J*}_{K'K}(k)\, \rho(k). \tag{5.6}$$

In general, an arbitrary generator X of $SL(3,R)$ has matrix elements (f, Xg) given by

$$(f, Xg) = \sum_{\substack{J'K'M' \\ JKM}} f^{J'*}_{K'M'} g^{J}_{KM} \left[(2J'+1)(2J+1)\right]^{1/2} \times \int\int dk_1\, dk_2 D^{J'*}_{K'M'}(k_1)\, \rho(k_1 k_2^{-1}) X(k_2)\, D^{J}_{KM}(k_2).$$

Making use of the additivity relations for the D^{J}_{KM} functions and the fact that dk is an invariant measure, we obtain

$$(f, Xg) = \sum f^{J'*}_{K'M'} g^{J}_{KM} \frac{1}{(2J'+1)^{1/2}} \left\langle {J' \atop K'N} \middle| \rho \middle| {0 \atop 00} \right\rangle \left\langle {J' \atop NM'} \middle| X \middle| {J \atop KM} \right\rangle. \tag{5.7}$$

The hermiticity condition (i.e., the unitarity of the representation)

$$(f, Xg) = (g, Xf)^* \tag{5.8}$$

now reads

$$(2J+1)^{1/2} \sum_{N_1} \left\langle {J' \atop K'N_1} \middle| \rho \middle| {0 \atop 00} \right\rangle \left\langle {J' \atop N_1 M'} \middle| X \middle| {J \atop KM} \right\rangle = (2J'+1)^{1/2} \sum_{N_2} \left\langle {J \atop KN_2} \middle| \rho \middle| {0 \atop 00} \right\rangle^{*} \left\langle {J \atop N_2 M} \middle| X \middle| {J \atop K'M'} \right\rangle^{*}. \tag{5.9}$$

At this stage we confine ourselves to solutions of the form

$$\left\langle {J \atop KK'} \middle| \rho \middle| {0 \atop 00} \right\rangle = \rho(J;K)\, \delta_{KK'}. \tag{5.10}$$

If X is any of the compact generators (J_0, J_+, J_-), the hermiticity condition becomes

$$\left\langle {J \atop KK'} \middle| \rho \middle| {0 \atop 00} \right\rangle^{*} = \left\langle {J \atop K'K} \middle| \rho \middle| {0 \atop 00} \right\rangle.$$

If ρ were a direct sum of the tensor operators ρ^{l}_{m} with respect to the $SU(2)$ subgroup of $\overline{SL}(3,R)$ and if complex conjugation at most changed the absolute value of the magnetic quantum number m (i.e., $m \to \pm m$), then using the Wigner–Eckart theorem for each irreducible subspace we would obtain

$$(-)^{J-K'}\begin{pmatrix} J & l & 0 \\ -K' & \pm m & 0 \end{pmatrix}\left\langle {J \atop K}\middle\| \rho^{l} \middle\| {0 \atop 0}\right\rangle = (-)^{J-K}$$

$$\times\begin{pmatrix} J & l & 0 \\ -K & m & 0 \end{pmatrix}\left\langle {J \atop K'}\middle\| \rho^{l} \middle\| {0 \atop 0}\right\rangle.$$

The result $K' = \pm K$ now follows from the properties of the 3-j symbols and we can write

$$\left\langle {J \atop KK'}\middle| \rho \middle| {0 \atop 00}\right\rangle = \rho(J;K)\delta_{KK'} + \bar{\rho}(J;K)\delta_{K,-K'}. \tag{5.11}$$

Although the second term in this expression is of importance in the analysis of the irreducibility of the representations, it is sufficient here to take into account only the first term. The reason for this is that we rely solely on the little group technique in determining the irreducibility of the representations. From (5.11) it follows that $\rho(J;K)$ is a real function.

In the basis $\{|{}^{J}_{KM}\rangle\}$ of the Hilbert space $L^2(SU(2))$, unitarity of the representations requires the parameters σ and δ to be $-3+i\sigma_2$ and $-1+i\delta_2$, respectively. We are now dealing with the Hilbert space $L^2(SU(2),\rho)$ with a more complex scalar product and that allows us to analytically continue the parameters σ and δ to arbitrary complex numbers, i.e.,

$$\begin{aligned} \sigma &= -3+\sigma_1+i\sigma_2, \\ \delta &= -1+\delta_1+i\delta_2, \end{aligned} \tag{5.12}$$

where σ_1, σ_2, δ_1, $\delta_2 \in R$. Thus the reduced matrix elements (4.5), of the quadrupole operator, are now given by

$$\left\langle {J' \atop K'}\middle\| T \middle\| {J \atop K}\right\rangle = (-)^{J'-K'}[(2J'+1)(2J+1)]^{1/2}\Bigg[\sqrt{\tfrac{2}{3}}(\sigma_2 - i\sigma_1)$$

$$+\frac{i}{\sqrt{6}}[J'(J'+1)-J(J+1)]\begin{pmatrix} J' & 2 & J \\ -K' & 0 & K \end{pmatrix} + i(1+K+\delta_1+i\delta_2)$$

$$\begin{pmatrix} J' & 2 & J \\ -K' & 2 & K \end{pmatrix} + i(1-K+\delta_1+i\delta_2)\begin{pmatrix} J' & 2 & J \\ -K' & -2 & K \end{pmatrix}\Bigg]. \tag{5.13}$$

Since T_0 does not change the M quantum number and since $\langle {}^{J}_{KK'}|\rho|{}^{0}_{00}\rangle = \rho(J;K)\delta_{K'K}$, the hermiticity condition (5.9) reads

$$(-)^{J'}(2J+1)^{1/2}\rho(J';K')\left\langle {J' \atop K'}\middle\| T \middle\| {J \atop K}\right\rangle$$

$$= (-)^{J}(2J'+1)^{1/2}\rho(J;K)\left\langle {J \atop K}\middle\| T \middle\| {J' \atop K'}\right\rangle^{*}, \tag{5.14}$$

$\rho(J;K)$ being an arbitrary real function of J and K. Substituting into this equation the expression for the reduced matrix elements with the analytically continued parameters σ and δ, we obtain, for $K'=K$,

$$(2J+1)^{1/2}\{2\sigma_2 - 2i\sigma_1 + i[J'(J'+1)-J(J+1)]\}\rho(J';K)$$

$$= (2J'+1)^{1/2}\{2\sigma_2 + 2i\sigma_1 + i[J'(J'+1)-J(J+1)]\}\rho(J;K), \tag{5.15}$$

while, for $K'=K+2$, we obtain

$$(K+1+\delta_1+i\delta_2)\rho(J;K+2) = (K+1-\delta_1+i\delta_2)\rho(J;K). \tag{5.16}$$

Since $\rho(J;K)$ is a real function from the first of the above equations it follows that $\sigma_1=0$, σ_2 is an arbitrary real number and $(2J+1)^{1/2}\rho(J';K) = (2J'+1)^{1/2}\rho(J;K)$. From the second equation it follows either that $\delta_1=0$, δ_2 is an arbitrary real number and $\rho(J;K+2)=\rho(J;K)$, or that δ_1 is an arbitrary real number, $\delta_2=0$ and $(K+1+\delta_1)\rho(J;K+2) = (K+1-\delta_1)\rho(J;K)$. Thus there are two general cases allowed by unitarity:

Case I. $\sigma = -3+i\sigma_2$, $\delta = -1+i\delta_2$, $\sigma_2, \delta_2 \in R$,

$$(2J+1)^{1/2}\rho(J';K) = (2J'+1)^{1/2}\rho(J;K), \quad K \geq 0,$$

$$\rho(J;K+2) = \rho(J;K). \tag{5.17}$$

Case II. $\sigma = -3+i\sigma_2$, $\delta = -1+\delta_1$, $\sigma_2, \delta_1 \in R$, $K \geq 0$,

$$(2J+1)^{1/2}\rho(J';K) = (2J'+1)^{1/2}\rho(J;K),$$

$$(K+1+\delta_1)\rho(J;K+2) = (K+1-\delta_1)\rho(J;K). \tag{5.18}$$

Case I in our analysis (which in terms of the parameters a and b is given by $a=-2+ia_2$ and $b=-1+ib_2$) has been treated using the theory of the induced representations.[6]

Due to the constraint $K^2 = J^2$ only those values of K and J which satisfy the constraint $|K| \leq J$ are allowed so that, for fixed K, only those components of D^{J}_{KM} for which $J \geq |K|$ contribute. From now on we shall write K for $|K|$.

Now Case II provides us with scalar products which are characterized by an arbitrary minimal K value, i.e., K can be $K_{\min}$, $K_{\min}+2$, $K_{\min}+4$, $\cdots$. These solutions are characterized by $\rho(J;K_{\min}) \neq 0$ and $\rho(J;K_{\min}-2) = 0$. Thus from

$$(K-1+\delta_1)\rho(J;K) = (K-1-\delta_1)\rho(J;K-2), \tag{5.19}$$

we obtain

$$\delta_1 = 1 - K_{\min}, \tag{5.20}$$

and for the matrix elements of kernel

$$\rho(J;K) = \left(\frac{2J+1}{2J_{\min}+1}\right)^{1/2}$$

$$\times \frac{\Gamma(\frac{1}{2}(K+K_{\min}))}{\Gamma(\frac{1}{2}(K-K_{\min})+1)\,\Gamma(K_{\min})}\,\rho(J_{\min};K_{\min}) \tag{5.21}$$

where $K \equiv K_{\min} \pmod 2$ and $K \geq K_{\min}$.

In the special case $K_{\min}=0$, from $\rho(J;0) \neq 0$ it follows that, for every $K \neq 0$, $\rho(J;K)=0$, and therefore for a given J there is not any degeneracy besides the $2J+1$ degeneracy in the quantum number M. These representations are so-called "multiplicity free," "most degenerate," or "ladder representations."

The representations in Case II obtained above are characterized by a continuous parameter σ_2 and a discrete valued parameter δ_1.

Up to now we have not used the constraint due to the positive definiteness (3.11) of the scalar product (5.5). For $f=g$ and $\langle {}^{J}_{K'K}|\rho|{}^{0}_{00}\rangle = \rho(J;K)\delta_{K'K}$ it follows that

$$(f,f)=\sum_{JKM}\frac{1}{(2J+1)^{1/2}}\left|f^{J}_{KM}\right|^{2}\rho(J;K). \tag{5.22}$$

The positive definiteness of the scalar product yields

$$\rho(J;K)\geq 0, \tag{5.23}$$

for every J and K. By inspection, one sees immediately that the previously obtained representations in Case I as well as in Case II satisfy the condition $\rho(J;K)\geq 0$. In case II the positive definiteness condition, i.e., $\rho(J;K)>0$ and $\rho(J;K+2)>0$, implies

$$(K+1-\delta_1)/(K+1+\delta_1)>0. \tag{5.24}$$

Therefore, when $K+1-\delta_1>0$ and $K+1+\delta_1>0$, it follows that $|\delta_1|<|K+1|$ for every K. Thus we obtain the representations characterized by two continuous parameters:

$$\sigma=-3+i\sigma_2,\quad \delta=-1+\delta_1,\quad \sigma_2,\delta_1\in R \tag{5.25}$$

with $|\delta_1|<1$ for K integer and $|\delta_1|<\frac{1}{2}$ for K half-integer. When $K+1-\delta_1<0$ and $K+1+\delta_1<0$, it follows $|\delta_1|>|K+1|$, for every K. This cannot be satisfied and there are no additional solutions. For this case (δ_1 being continuous) the matrix elements $\rho(J;K)$ are given by

$$\rho(J;K)=\left(\frac{2J+1}{2J_{\min}+1}\right)^{1/2}\times\frac{\Gamma(\frac{1}{2}(K+1)-\frac{1}{2}\delta_1)\,\Gamma(\frac{1}{2}(K_{\min}+1)+\frac{1}{2}\delta_1)}{\Gamma(\frac{1}{2}(K+1)+\frac{1}{2}\delta_1)\,\Gamma(\frac{1}{2}(K_{\min}+1)-\frac{1}{2}\delta_1)}\times\rho(J_{\min};K_{\min}), \tag{5.26}$$

where $K\equiv K_{\min}\pmod 2$, and $J_{\min}$ and $K_{\min}$ are either 0 or $\frac{1}{2}$.

The classes of the representations obtained in this analysis are in agreement with those obtained by Gel'fand and Graev[15] and by Hulthen.[16]

6. INVARIANTS

Since $\overline{SL}(3,R)$ is a semisimple Lie group of rank two, there are two generators in the center of the universal enveloping algebra. These are the group invariants I_2 and I_3. The expressions for these invariants may be obtained from those of $SU(3)$ by a simple modification of the structure constants since $\overline{SL}(3,R)$ and $SU(3)$ both are real forms of the same complex Lie group A_2.

Convenient expressions for the invariants of $SU(3)$ in the spherical basis have been given by Racah.[17] With the necessary modifications we can obtain the invariants of $\overline{SL}(3,R)$ in the following form

$$I_2=\frac{1}{12}(J^2-\tfrac{1}{2}T^2)$$

and

$$I_3=\frac{i}{24}(JTJ+\tfrac{1}{3}T^3). \tag{6.1}$$

Substituting in these expressions the explicit differential forms for the $\overline{SL}(3,R)$ generators, we arrive at

$$I_2=\tfrac{1}{9}[(a+2)^2+(b+1)^2-(a+2)(b+1)-3]$$

and

$$I_3=\frac{1}{162}\times[2(a+2)^3+2(b+1)^3-3(a+2)^2(b+1)-3(a+2)(b+1)^2]. \tag{6.2}$$

These expressions are, up to a multiplicative factor, identical with those obtained by Hulthen.[16,18]

From the form of the $SU(3)$ invariants, in terms of the labels $[p,q,0]$,

$$I_2^{SU(3)}=\tfrac{1}{9}(p^2+q^2-pq+3p)$$

and

$$I_3^{SU(3)}=\frac{1}{162}(p-2q)(2p-q+3)(p+q+3),$$

it is easy to see that they are identical to those of $\overline{SL}(3,R)$ by means of the substitution $a=p$ and $b=q$. Therefore, we can characterize the unitary irreducible representations of the $\overline{SL}(3,R)$ group by the (generalized) Young pattern labels $[a,b,0]$, $a=-2+\frac{1}{2}(\sigma_1+\delta_1)+\frac{1}{2}i(\sigma_2-\delta_2)$ and $b=-1+\frac{1}{2}(\sigma_1-\delta_1)+\frac{1}{2}i(\sigma_2-\delta_2)$.

For Case I we have

$$\sigma=-3+i\sigma_2,\quad \delta=-1+i\delta_2,$$

$$I_2=-\frac{1}{36}(\sigma_2^2+3\delta_2^2+12),$$

and

$$I_3=\frac{i}{4\times 162}\sigma_2(\sigma_2+\delta_2)(\sigma_2-3\delta_2). \tag{6.3}$$

For Case II we have

$$\sigma=-3+i\sigma_2,\quad \delta=-1+\delta_1,$$

$$I_2=-\frac{1}{36}(\sigma_2^2-3\delta_1^2+12),$$

and

$$I_3=\frac{1}{4\times 162}\sigma_2(\sigma_2^2+9\delta_1^2). \tag{6.4}$$

The invariants I_2 and I_3 (the generators of the center of the universal enveloping algebra) are constants for the Harish-Chandra representations $U^{a,b}$ analysed previously so that they give no decomposition of the Hilbert space $L^2(SU(2),\rho)$.

We also, for the representations of Cases I and II, point out that the values of the second- and the third-rank invariant, i.e., I_2 and I_3, are real and pure imaginary respectively. This is an alternative proof for the unitarity of the representations.

The Nelson operator for $\overline{SL}(3,R)$ is

$$\Delta=2J^2-12I_2. \tag{6.5}$$

Since both J^2 and I_2 are Hermitian operators the above representations of the Lie algebra $sl(3,R)$ can be exponentiated to give continuous unitary irreducible representations[8] of the simply connected group $\overline{SL}(3,R)$. Also, since $\overline{SL}(3,R)$ is a semisimple Lie group, the analytic vectors of the maximal compact subgroup, which in this case is $SU(2)$, are also analytic vectors for the whole group.[12]

7. IRREDUCIBILITY

As we have already seen, the invariant operators I_2 and I_3 have constant values for given values of the

parameters σ and δ so that they do not require any decomposition of the Hilbert space $L^2(SU(2),\rho)$ in which the representation $U^{a,b}(g)$, $g\in\overline{SL}(3,R)$ is defined.

Although it is necessary for the irreducibility, in the case of noncompact groups, that the invariants take on the constant values on the corresponding Hilbert space, it is not in general sufficient. Besides the invariant Casimir operators it is usually necessary to find additional labels (quantum numbers) which account for the further decomposition of the space in which the representation is defined.

On the one hand, a unitary representation of a semisimple Lie group is always completely reducible so that the unitary representations of $\overline{SL}(3,R)$ are completely reducible. On the other hand, not much is known concerning the concrete methods for determining the irreducible unitary representations of these groups.

The most effective technique for obtaining the additional labels for the representations, besides invariant operators, as well as for imposing the constraints on the vectors belonging to an irreducible subspace, is the method of using little groups.[19]

For $\overline{SL}(3,R)$ we proceed in the following way. First we note that it is sufficient to take into account only the operator T_0. The reason for this is that the operator J_0 is diagonal, the operators J_+ and J_- change only the M value and they are responsible for the usual $2J+1$ degeneracy in M. Since $T_{\pm1}$ and $T_{\pm2}$ are related to T_0 through the commutation relations of T_0 and $J_\pm$, they do not contribute to any additional degeneracy in M.

Let ω be the group parameter corresponding to the one-dimensional subgroup of $\overline{SL}(3,R)$ generated by T_0. The elements of this subgroup are diagonal matrices of the form

$$g(\omega)=\begin{pmatrix} e^{\sqrt{2/3}a\omega} & 0 & 0\\ 0 & e^{\sqrt{2/3}b\omega} & 0\\ 0 & 0 & e^{-\sqrt{2/3}(a+b)\omega}\end{pmatrix},\quad \omega\in R.$$

Now the little group is the set of elements from $SU(2)$ subgroup which leaves the one parameter subgroup $\{g(\omega)\}$ unchanged. Obviously the form of the little group depends on whether $a\neq b$ or $a=b$. For the Case I representations, $a=-2+\frac{1}{2}i(\sigma_2+\delta_2)$ and $b=-1+\frac{1}{2}i(\sigma_2-\delta_2)$ so that $a\neq b$, while for the Case II representations, $a=-2+\frac{1}{2}\delta_1+\frac{1}{2}i\sigma_2$ and $b=-1-\frac{1}{2}\delta_1+\frac{1}{2}i\sigma_2$ giving $a\neq b$ unless $\delta_1=1$ (corresponding to $K_{\min}=0$). We will concentrate now on the case $a\neq b$. The special case $a=b$ ($\delta_1=1$) will be treated separately.

When $a\neq b$ it is easy to see that elements from the $SU(2)$ subgroup which leave the one-dimensional subgroup $\{g(\omega)\}$ unchanged belong to the centralizer of the subgroup $\{g(\omega)\}$ in $SU(2)$. The elements in the centralizer consist of the identity, a rotation about any axis through 2π [minus the identity in $SU(2)$] and rotations $\pm\pi$ about each coordinate axis; these transformations form the quaternion group Q_2, with elements given by

$$Q_2=\left\{\pm\begin{pmatrix}1&0\\0&1\end{pmatrix},\ \pm\begin{pmatrix}0&i\\i&0\end{pmatrix},\ \pm\begin{pmatrix}0&-1\\1&0\end{pmatrix},\ \pm\begin{pmatrix}i&0\\0&-i\end{pmatrix}\right\}. \tag{7.1}$$

The group Q_2 is a subgroup of $SU(2)$ and its elements, as listed above, can be obtained from the $D^{1/2}$ matrix upon substitution of the appropriate Euler angles.

The next step is to find the irreducible representations of the Q_2 group in order to obtain additional labels for the unitary irreducible representations of $\overline{SL}(3,R)$. The irreducible representations of Q_2 are well known. There are five of them and four of them are one-dimensional being labeled by $V^{\epsilon,\epsilon'}$, $\epsilon,\epsilon'=\pm1$. These four representations have the property that in an irreducible representation of $SU(2)$ labeled by J, $(-)^{2J}=1$. The fifth one is two-dimensional and we call it V^5. Using the representations of the Q_2 group one now can obtain the constraints on the basis vectors due to the little group.[20] From the equation

$$f(kk')=V(k')f(k),\quad k\in SU(2),\quad k'\in Q_2,$$

where $V(k')$ is one of the irreducible representations of the quaternion group Q_2, one obtains the constraints on the coefficients in the expansion for $f(k)$ in the D^J_{KM} basis. For the four one-dimensional irreducible representations one has

$$f(k)=\sum_{JKM} f^J_{KM}\left(D^J_{KM}(k)+\epsilon'(-)^{J-K}D^J_{-KM}(k)\right),\quad K\geq0,$$

Here ϵ' is the eigenvalue under $R_y(\pi)$, ϵ corresponds to the eigenvalue of f under $R_z(\pi)$. For $\epsilon=+1$ K is an even number and for $\epsilon=-1$ K is an odd number.

We now introduce a new basis

$$\left|\begin{matrix}J\\KM\end{matrix}\right\rangle\equiv\frac{1}{\sqrt{2}}\left(\left|\begin{matrix}J\\KM\end{matrix}\right\rangle+\epsilon'(-)^{J-K}\left|\begin{matrix}J\\-KM\end{matrix}\right\rangle\right),\quad K\geq0, \tag{7.2}$$

and the corresponding Hilbert space we denote by $H^{(\epsilon,\epsilon')}_{K_{\min}}=L^2(SU(2),\rho;\epsilon,\epsilon')$. It is not difficult to show that the new basis vectors are orthonormal

$$\left(\begin{matrix}J'\\K'M'\end{matrix}\middle|\begin{matrix}J\\KM\end{matrix}\right)=\delta_{J'J}\delta_{K'K}\delta_{M'M} \tag{7.3}$$

and that

$$T_\mu\left|\begin{matrix}J\\KM\end{matrix}\right\rangle=\sum_{J'K'M'}\left\langle\begin{matrix}J'\\K'M'\end{matrix}\middle|T_\mu\middle|\begin{matrix}J\\KM\end{matrix}\right\rangle\left|\begin{matrix}J'\\K'M'\end{matrix}\right\rangle, \tag{7.4}$$

i.e., the Hilbert space $H^{(\epsilon,\epsilon')}$ with the basis $\{|\,{}^J_{KM}\rangle\}$ is invariant under the action of the $\overline{SL}(3,R)$ generators. The hermiticity condition in this basis turns out to be

$$(2J+1)^{1/2}\left(\begin{matrix}J'\\K'N\end{matrix}\middle|\rho\middle|\begin{matrix}0\\00\end{matrix}\right)\left(\begin{matrix}J\\NM\end{matrix}\middle|T_0\middle|\begin{matrix}J'\\K'M'\end{matrix}\right)$$

$$=(2J'+1)^{1/2}\left(\begin{matrix}J\\KN\end{matrix}\middle|\rho\middle|\begin{matrix}0\\00\end{matrix}\right)^*\left(\begin{matrix}J\\NM\end{matrix}\middle|T_0\middle|\begin{matrix}J'\\K'M'\end{matrix}\right)^*. \tag{7.5}$$

Thus, since the equations which follow from this hermiticity condition are the same as (5.15) and (5.16) all the results we have obtained previously hold for the Hilbert space $H^{(\epsilon,\epsilon')}_{K_{\min}}$.

The fifth representation V^5 of the quaternion group Q_2 is two-dimensional and leads to half-integral J values. The group generators are now

$$J_\nu\otimes I_2,\quad \nu=0,\pm,$$

and

$$T_\mu\otimes I_2,\quad \mu=0,\ \pm1,\ \pm2, \tag{7.6}$$

I_2 being the two-dimensional unit matrix. We denote by $H^{(5)}_{K_{\min}} = L^2(SU(2), \rho; 5)$ the irreducible subspace of the whole Hilbert space $L^2(SU(2), \rho)$, which corresponds to the representation V^5 of the little group Q_2. The arbitrary function in this space is now of the form[20]

$$f(k) = \sum_{JNM} f^J_{NM} \begin{pmatrix} D^J_{1/2-2N,M}(k) \\ e^{i\pi(J-1/2)} D^J_{-1/2+2N,M}(k) \end{pmatrix}, \quad k \in SU(2).$$

Due to the label K all the irreducible representations discussed here possess, besides the $2J+1$ degeneracy in M, an additional degeneracy. In other words, if we decompose some irreducible representation of the $\overline{SL}(3,R)$ group with respect to the maximal compact subgroup, the irreducible representations with the same J value will, in general, appear several times. The simplest way to determine the additional degeneracy for a given J is to denumerate the possible K values of the vectors belonging to some invariant subspace. Now, excluding the M degeneracy we obtain the following result in terms of $(\epsilon, \epsilon') = (\pm, \pm)$.

Case I: When the Hilbert space is $H^{(\epsilon,\epsilon')} = L^2(SU(2); \epsilon, \epsilon')$, the J content is given by

$$\{J^n\} = \{0,\ 2^2,\ 3,\ 4^3,\ 5^2,\ 6^4, \cdots\}, \tag{7.7}$$

for $(\epsilon, \epsilon') = (+, +)$, and by

$$\{J^n\} = \{1,\ 2,\ 3^2,\ 4^2,\ 5^3,\ 6^3, \cdots\}, \tag{7.8}$$

for $(\epsilon, \epsilon') = (+, -)$, $(-, +)$, and $(-, -)$. In the Hilbert space $H^{(5)} = L^2(SU(2); 5)$, corresponding to the fifth irreducible representation V^5, the J content is given by

$$\{J^n\} = \{\tfrac{1}{2},\ (\tfrac{3}{2})^2,\ (\tfrac{5}{2})^3,\ (\tfrac{7}{2})^4, \cdots\}. \tag{7.9}$$

Case II: In the Hilbert space $H^{(\epsilon,\epsilon')}_{K_{\min}} = L^2(SU(2), \rho; \epsilon, \epsilon')$ the angular momentum content is

$$\{J^n\} = \{K_{\min},\ K_{\min}+1,\ (K_{\min}+2)^2,\ (K_{\min}+3)^2,\ (K_{\min}+4)^3, \cdots\}, \tag{7.10}$$

where $K_{\min} \neq 0$, whereas in the Hilbert space $H^{(5)}_{K_{\min}} = L^2(SU(2), \rho; 5)$, the angular momentum content is

$$\{J^n\} = \{K_{\min}, (K_{\min}+1)^2,\ (K_{\min}+2)^3,\ (K_{\min}+3)^4, \cdots\}. \tag{7.11}$$

8. MULTIPLICITY FREE REPRESENTATIONS

We will discuss now in more detail Case II when $\delta = 0$. The parameters of the unitary representations are $\sigma = -3 + i\sigma_2$, $\delta = -1 + \delta_1$, and the parameters a and b are given in terms of σ and δ by $a = -2 + \frac{1}{2}\delta_1 + \frac{1}{2}i\sigma_2$, $b = -1 - \frac{1}{2}\delta_1 + \frac{1}{2}i\sigma_2$. Furthermore, from (5.19) it follows when $K = 0$ and $\delta_1 = 1$ that $\rho(J;0) \neq 0$ implies $\rho(J;2) = \rho(J;-2) = 0$. Thus we obtain the multiplicity free representations, with the matrix elements of the group generators given by

$$\left\langle \begin{matrix} J' \\ OM' \end{matrix} \middle| J_0 \middle| \begin{matrix} J \\ OM \end{matrix} \right\rangle = M\delta_{J'J}\delta_{M'M},$$

$$\left\langle \begin{matrix} J' \\ OM' \end{matrix} \middle| J_\pm \middle| \begin{matrix} J \\ OM \end{matrix} \right\rangle = [J(J+1) - M(M\pm 1)]^{1/2}\delta_{J'J}\delta_{M',M\pm 1},$$

and

$$\left\langle \begin{matrix} J' \\ OM' \end{matrix} \middle| T_\mu \middle| \begin{matrix} J \\ OM \end{matrix} \right\rangle = (-)^{J'-M'} \begin{pmatrix} J' & 2 & J \\ -M' & \mu & M \end{pmatrix} (-)^{J'} [(2J'+1)(2J+1)]^{1/2}$$

$$\times \left(\sqrt{\tfrac{2}{3}}\sigma_2 + \frac{i}{\sqrt{6}}[J'(J'+1) - J(J+1)]\right) \begin{pmatrix} J' & 2 & J \\ 0 & 0 & 0 \end{pmatrix}, \tag{8.1}$$

σ_2 being an arbitrary real number. The 3-j symbol $\left(\begin{smallmatrix} J' & 2 & J' \\ 0 & 0 & 0 \end{smallmatrix}\right)$ is nonzero only if $J + J'$ is an even integer and this provides us with the $\Delta J = 2$ rule, i.e., the J content for this representation is[4,21]

$$\{J\} = \{0, 2, 4, \cdots\}, \tag{8.2}$$

or

$$\{J\} = \{1, 3, 5, \cdots\}. \tag{8.3}$$

The explicit forms for the reduced matrix elements of T_μ are

$$\left\langle \begin{matrix} J \\ O \end{matrix} \middle\| T \middle\| \begin{matrix} J \\ O \end{matrix} \right\rangle = -2\sqrt{\tfrac{2}{3}}\sigma_2 \left(\frac{J(J+1)(2J+1)}{(2J-1)(2J+3)}\right)^{1/2},$$

and

$$\left\langle \begin{matrix} J+2 \\ 0 \end{matrix} \middle\| T \middle\| \begin{matrix} J \\ O \end{matrix} \right\rangle = [J(J-1)(2J-1)]^{1/2}[i + \sigma_2/(2J-1)]. \tag{8.4}$$

They were first evaluated by the means of analytical continuation in the $SU(3)$ labels.[22] It is easy to see from the general expressions for the $\overline{SL}(3,R)$ generators as differential operators that, in this case, they are simply given by[23]

$$J_0 = i\frac{\partial}{\partial\gamma}, \quad J_\pm = \exp(\mp i\gamma)\left(\mp\frac{\partial}{\partial\beta} + ictg\beta\frac{\partial}{\partial\gamma}\right),$$

and

$$T_0 = \frac{i}{\sqrt{6}}(3c^2\beta - 1)\sigma_2 - \sqrt{6}\ s\beta\ c\beta\frac{\partial}{\partial\beta}. \tag{8.5}$$

The above explicit forms (8.4) for the reduced matrix elements of T_0 are obtained under the assumption that J is an integer. If one continues these forms in J to half-integer values, then they satisfy the $\overline{SL}(3,R)$ condition (2.4) only if $J_{\min} = \frac{1}{2}$ and $\sigma_2 = 0$. Thus there is a multiplicity free representation of the $\overline{SL}(3,R)$ group with $\Delta J = 2$, $J_{\min} = \frac{1}{2}$, and $\sigma_2 = 0$ (Refs. 24, 25), i.e.,

$$\{J\} = \{\tfrac{1}{2},\ \tfrac{5}{2},\ \tfrac{9}{2}, \cdots\}. \tag{8.6}$$

Using the general expressions for the $\overline{SL}(3,R)$ invariants, Case II with $\delta_1 = 1$ yields

$$I_2 = -\tfrac{1}{4} - \tfrac{1}{36}\sigma_2{}^2$$

and

$$I_3 = \frac{i}{4\times 162}\sigma_2(9 + \sigma_2{}^2). \tag{8.7}$$

Note that I_3 is proportional to I_2, i.e., $I_3 = -(i/18)\sigma_2 I_2$.

9. SUMMARY

In this work we have parameterized the unitary irreducible representations of the $\overline{SL}(3,R)$ group in

terms of two parameters $\sigma=-3+\sigma_1+i\sigma_2$ and $\delta=-1+\delta_1+i\delta_2$ where $\sigma=a+b$ and $\delta=a-b$. The main results for the unitary irreducible representations are: The matrix elements for the group generators are given by Eqs. (4.2) and (4.6). The general scalar product is given by Eq. (5.7), and the basis vectors in the irreducible subspaces are given by (7.2). There are three series of representations.

Principal series: $\sigma=-3+i\sigma_2$, $\delta=-1+i\delta_2$, $\sigma_2,\delta_2\in R$. The matrix elements of the kernel are $\rho(J;K)=(2J+1)^{1/2}$, the invariants are given by (6.3), and the irreducible Hilbert spaces are $H^{(\epsilon,\epsilon')}$ and $H^{(5)}$ with the J contents given by (7.8) and (7.9), respectively.

Supplementary series: (a) $\sigma=-3+i\sigma_2$, $\delta=-1+\delta_1$, $\sigma_2\in R$, $|\delta_1|<1$ for J integer and $|\delta_1|<\frac{1}{2}$ for J half-integer. The invariants are given by (6.4) and the matrix elements of the kernel ρ are given by (5.26). The irreducible Hilbert spaces are $H^{(\epsilon,\epsilon')}$ and $H^{(5)}$ with the J contents given by (7.8) and (7.9), respectively.

(b) $\sigma=-3+i\sigma_2$, $\delta=-1+\delta_1$, $\sigma_2\in R$, $\delta_1=1-K_{\min}$ ($K_{\min}=0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\cdots$). The matrix elements of the Kernel ρ are given by (5.21) and the invariants by (6.4). The irreducible Hilbert spaces are $H^{(\epsilon,\epsilon')}_{K_{\min}}$ and $H^{(5)}_{K_{\min}}$ with J contents given by (7.10) and (7.11), respectively.

Multiplicity free representations: $\sigma=-3+i\sigma_2$, $\delta=0$, $\sigma_2\in R$, $\delta_1=1$. The matrix elements of the quadrupole operator are given by (8.1) and the invariants are given by (8.7). The matrix elements of the kernel are $\rho(J;K)=(2J+1)^{1/2}$, $K=0$ and the J content for these representations is given by (8.2) and (8.3). For $\sigma_2=0$, J can also be half-integer and the J content is given by (8.6).

We have argued that the matrix elements of the kernel ρ can be taken in the form $\rho(J;K)\delta_{KK'}$ and the whole analysis is done under this assumption. This indicates that a complete set of unitary irreducible representations is obtained, but we are not aware of a rigorous proof.

10. CONTRACTION TO $T_5 \circledS SU(2)$

Let G be a noncompact, connected semisimple Lie group with finite center and let K be its maximal compact subgroup. The Cartan decomposition of the algebra $\mathbf{G}$ of the group G is $\mathbf{G}=\mathbf{K}\oplus\mathbf{P}$, where $\mathbf{K}$ is the Lie algebra of the maximal compact subgroup K of the group G and $[\mathbf{K},\mathbf{P}]\subset\mathbf{P}$, $[\mathbf{P},\mathbf{P}]\subset\mathbf{K}$. Contraction in the sense of Inönü and Wigner[26] is a limiting process. $\mathbf{K}\to\mathbf{K}$, $\mathbf{P}\to\epsilon\mathbf{P}$, where $\epsilon\to 0$. Thus, in this manner one arrives at the new algebra $\mathbf{G}'=\mathbf{K}\oplus\mathbf{P}'$, with $[\mathbf{K},\mathbf{P}']\subset\mathbf{P}'$, $[\mathbf{P}',\mathbf{P}']=0$. The corresponding group G' is a semidirect product of its subgroups K and P', i.e., $G=P'\circledS K$, P' being the invariant subgroup in G.

In our case G is $\overline{SL}(3,R)$, and it is obvious that the corresponding contracted group, with respect to $SU(2)$, is $T_5\circledS SU(2)$. If we denote the generators of the contracted group by Q_μ, i.e.,

$$Q_\mu=\lim_{\epsilon\to 0}\epsilon T_\mu, \tag{10.1}$$

then the new commutation relations are

$$[J_0,J_\pm]=\pm J_\pm,$$

$$[J_+,J_-]=2J_0,$$

$$[J_0,Q_\mu]=\mu Q_\mu,$$

$$[J_\pm,Q_\mu]=[2-\mu(\mu\pm1)]^{1/2}Q_{\mu\pm1},$$

and

$$[Q_\mu,Q_\nu]=0\quad(\mu,\nu=0,\pm1,\pm2). \tag{10.2}$$

The invariants I'_2 and I'_3 of the $T_5\circledS SU(2)$ group are related to those of $\overline{SL}(3,R)$ in the following way:

$$I'_2=\lim_{\epsilon\to0}\epsilon^2 I_2,\qquad I'_3=\lim_{\epsilon\to0}\epsilon^3 I_3. \tag{10.3}$$

In order to relate the unitary irreducible representation of $\overline{SL}(3,R)$ to those of $T_5\circledS SU(2)$ (Refs. 20, 27) we present the summary of the unitary irreducible representations of $T_5\circledS SU(2)$. In terms of the parameters a and b in this work, the so-called α, β, and γ irreducible representations[20] of the group $T_5\circledS SU(2)$ are given by

$$\alpha_J\ bands:\left\langle\begin{matrix}J'\\K'\end{matrix}\middle\|Q\middle\|\begin{matrix}J\\K\end{matrix}\right\rangle=0, \tag{10.4}$$

with invariants

$$I_2=I_3=0, \tag{10.5}$$

$$\beta_K\ bands:\left\langle\begin{matrix}J'\\K'\end{matrix}\middle\|Q\middle\|\begin{matrix}J\\K\end{matrix}\right\rangle=(-)^{J'-K'}[(2J'+1)(2J+1)]^{1/2}\times\sqrt{\tfrac{2}{3}}\sigma_2\begin{pmatrix}J'&2&J\\-K'&0&K\end{pmatrix}. \tag{10.6}$$

with invariants

$$I_2=\tfrac{1}{9}a_2^2,\qquad I_3=-\frac{1}{3^4}a_2^3, \tag{10.7}$$

and

$$\gamma\ bands:\left\langle\begin{matrix}J'\\K'\end{matrix}\middle\|Q\middle\|\begin{matrix}J\\K\end{matrix}\right\rangle=(-)^{J'-K'}[(2J'+1)(2J+1)]^{1/2}\times\left\{\sqrt{\tfrac{2}{3}}\sigma_2\begin{pmatrix}J'&2&J\\-K'&0&K\end{pmatrix}-\delta_2\left[\begin{pmatrix}J'&2&J\\-K'&2&K\end{pmatrix}+\begin{pmatrix}J'&2&J\\-K'&-2&K\end{pmatrix}\right]\right\}, \tag{10.8}$$

with invariants

$$I_2=\tfrac{1}{9}(a_2^2+b_2^2-a_2b_2),\qquad I_3=\frac{1}{162}(2a_2^3+2b_2^3-3a_2^2b_2-3a_2b_2^2). \tag{10.9}$$

Now, it is straightforward to see that in the Inönü–Wigner contraction limit of the unitary irreducible representations of the $\overline{SL}(3,R)$ group to those of the $T_5\circledS SU(2)$ group one has the following situation.

Case I: For σ_2 and δ_2 kept fixed we arrive at the α_J bands ($J=0,\ \frac{1}{2},\ 1,\cdots$), for $\epsilon\sigma_2$ and δ_2 kept fixed we arrive at the β_K bands ($K=0,\ \frac{1}{2},\ 1,\cdots$), and finally for $\epsilon\sigma_2$ and $\epsilon\delta_2$ kept fixed we obtain the corresponding γ bands characterized by $(\epsilon,\epsilon')=(\pm,\pm)$.

Case II: The irreducible representations of the

$\overline{SL}(3,R)$ group are defined on the Hilbert spaces $H^{(\epsilon,\epsilon')}_{K_{\min}} = L^2(SU(2),\rho;\epsilon,\epsilon')$ and $H^{(5)}_{K_{\min}} = L^2(SU(2),\rho;5)$. Now, since δ_1 is a finite discrete number and $\epsilon\delta_1 \to 0$, we obtain for σ_2 kept fixed the α_J bands ($J = K_{\min}, K_{\min}+1, K_{\min}+2, \cdots$) and for $\epsilon\sigma_2$ kept fixed the β_K bands ($J_{\min} = K_{\min}, K_{\min}+1, K_{\min}+2, \cdots$). The multiplicity free representation ($\delta_1 = 1$ and $\Delta J = 2$) characterized by $J_{\min} = \frac{1}{2}$ in the contraction process goes into the α_J bands ($J = \frac{1}{2}, \frac{5}{2}, \cdots$). The other two classes of the multiplicity free representations, characterized by $J_{\min} = 0$ or 1 and σ_2 arbitrary real number, in the contraction process with σ_2 kept fixed go into the α_J bands ($J = 0, 2, 4, \cdots$, or $J = 1, 3, 5, \cdots$) and for $\epsilon\sigma_2$ kept fixed into the β_0 bands ($J_{\min} = 0$ or $J_{\min} = 1$).

ACKNOWLEDGMENTS

I wish to thank Professor L. C. Biedenharn for suggesting the problem and numerous discussions. Thanks are also due to Dr. S. Browne for discussions. A research grant from RZNR SRS (Belgrade) is gratefully acknowledged.

APPENDIX A

In the Iwasawa decomposition of $SL(3,R)$ the nilpotent subgroup N is the set of 3×3 upper triangular matrices with the group parameters ν_1, ν_2, and ν_3, the Abelian subgroup A is the set of diagonal matrices $\mathrm{diag}(e^{\lambda}, e^{\mu}, e^{-\lambda-\mu})$, and the maximal compact subgroup $SO(3)$ can be given in terms of Euler angles α, β, and γ through the relation $k(\alpha,\beta,\gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma)$, where for example $R_z(\alpha)$ is a rotation through an angle α about the Z axis. Explicitly,

$$k(\alpha,\beta,\gamma) = \begin{pmatrix} c\alpha c\beta c\gamma - s\alpha s\gamma & -c\alpha c\beta s\gamma - s\alpha c\gamma & c\alpha s\beta \\ s\alpha c\beta c\gamma + c\alpha s\gamma & -s\alpha c\beta s\gamma + c\alpha c\gamma & s\alpha s\beta \\ -s\beta c\gamma & c\beta s\gamma & c\beta \end{pmatrix},$$

where, for example, $s\alpha$ and $c\alpha$ stand for $\sin\alpha$ and $\cos\alpha$, respectively. An arbitrary group element of $SL(3,R)$ is given by $g = nak = \{g_{ij}\}$, $i,j = 1, 2, 3$, where

$$\begin{aligned}
g_{11} &= e^{\lambda}(c\alpha c\beta c\gamma - s\alpha s\gamma) + \nu_3 e^{\mu}(s\alpha c\beta c\gamma + c\alpha s\gamma) \\
&\quad + \nu_2 e^{-\lambda-\mu} s\beta c\gamma, \\
g_{12} &= e^{\lambda}(-c\alpha c\beta s\gamma - s\alpha c\gamma) + \nu_3 e^{\mu}(-s\alpha c\beta s\gamma + c\alpha c\gamma) \\
&\quad - \nu_2 e^{-\lambda-\mu}(s\beta s\gamma), \\
g_{13} &= e^{\lambda}(c\alpha s\beta) + \nu_3 e^{\mu}(s\alpha s\beta) - \nu_2 e^{-\lambda-\mu} c\beta, \\
g_{21} &= e^{\mu}(s\alpha c\beta c\gamma + c\alpha s\gamma) + \nu_1 e^{-\lambda-\mu}(-s\beta c\gamma), \\
g_{22} &= e^{\mu}(-s\alpha c\beta s\gamma + c\alpha c\gamma) + \nu_1 e^{-\lambda-\mu} s\beta s\gamma, \\
g_{23} &= e^{\mu}(s\alpha s\beta) + \nu_1 e^{-\lambda-\mu} c\beta, \\
g_{31} &= e^{-\lambda-\mu}(-s\beta c\gamma), \\
g_{32} &= e^{-\lambda-\mu} s\beta s\gamma,
\end{aligned}$$

and

$$g_{33} = e^{-\lambda-\mu} c\beta.$$

The differential operator $X^{a,b}_{\epsilon}$ corresponding to an arbitrary group parameter ϵ is given by (3.15), i.e.,

$$X^{a,b}_{\epsilon} = \left(\frac{d\lambda}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\lambda} + \left(\frac{d\mu}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\mu} + \left(\frac{d\alpha}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\alpha} + \left(\frac{d\beta}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\beta} + \left(\frac{d\gamma}{d\epsilon}\right)_{\epsilon=0}\frac{\partial}{\partial\gamma}.$$

In order to evaluate explicitly the coefficients of this differential operator we proceed in the following manner: Let X_ϵ be the 3×3 matrix form for the generator corresponding to the parameter ϵ and consider an infinitesimal $SL(3,R)$ transformation $g(\epsilon) = 1 + \epsilon X_\epsilon$. For an arbitrary element $g(\nu_1,\nu_2,\nu_3,\lambda,\mu,\alpha,\beta,\gamma) \in SL(3,R)$ one has

$$gg(\epsilon) = g', \tag{A1}$$

where $g' \in SL(3,R)$ has new values of the group parameters; i.e., $g' = g(\nu_1',\nu_2',\nu_3',\lambda',\mu',\alpha',\beta',\gamma')$, where the old and new values for the group parameters differ infinitesimally. The matrix equation (A1) provides us with nine equations, which if we differentiate with respect to ϵ and set $\epsilon = 0$, lead to a set of algebraic equations in $(d\lambda/d\epsilon)_{\epsilon=0}, \ldots, (d\gamma/d\epsilon)_{\epsilon=0}$.

For the case when ϵ corresponds to any of the $SO(3)$ subgroup parameters we obtain the familiar differential forms for the generators J_i. In the spherical basis the explicit forms for these operators are

$$\begin{aligned}
J_0 &= i\frac{\partial}{\partial\gamma}, \\
J_{\pm} &= \exp(\mp i\gamma)\left(\frac{-i}{s\beta}\frac{\partial}{\partial\alpha} \mp \frac{\partial}{\partial\beta} + ictg\beta\frac{\partial}{\partial\gamma}\right).
\end{aligned} \tag{A2}$$

For the infinitesimal transformations of the one-dimensional subgroup of $SL(3,R)$ generated by $A_1 + A_2$, A_1 and A_2 being the infinitesimal generators corresponding to the parameters λ and μ, respectively, $g(\epsilon)$ is given by $\mathrm{diag}(1+\epsilon, 1+\epsilon, 1-2\epsilon)$. Now equation (A1) provides us with[5]

$$\left(\frac{d\lambda}{d\epsilon}\right)_{\epsilon=0} = 1 - 3c^2\alpha s^2\beta, \quad \left(\frac{d\mu}{d\epsilon}\right)_{\epsilon=0} = 1 - 3s^2\alpha s^2\beta,$$

$$\left(\frac{d\alpha}{d\epsilon}\right)_{\epsilon=0} = -3s\alpha c\alpha s^2\beta, \quad \left(\frac{d\beta}{d\epsilon}\right)_{\epsilon=0} = 3s\beta c\beta,$$

and

$$\left(\frac{d\gamma}{d\epsilon}\right)_{\epsilon=0} = 0.$$

Since the zeroth component T_0 of the quadrupole operator T_μ is related to generator $A_1 + A_2$ by $T_0 = -i\sqrt{\frac{2}{3}}(A_1 + A_2)$, we obtain

$$T_0 = i\left(\sqrt{\tfrac{2}{3}}(3c^2\alpha s^2\beta - 1)\frac{\partial}{\partial\lambda} + \sqrt{\tfrac{2}{3}}(3s^2\alpha s^2\beta - 1)\frac{\partial}{\partial\mu} + \sqrt{6}s\alpha c\alpha s^2\beta\frac{\partial}{\partial\alpha} - \sqrt{6}s\beta c\beta\frac{\partial}{\partial\beta}\right). \tag{A3}$$

The simplest way to obtain the remaining components $T_{\pm1}$ and $T_{\pm2}$ of the quadrupole operator is by means of the commutation relations

$$T_{\pm1} = \frac{1}{\sqrt{6}}[J_{\pm}, T_0]$$

and

$$T_{\pm2} = \tfrac{1}{2}[J_{\pm}, T_{\pm1}] = \frac{1}{2\sqrt{6}}[J_{\pm}, [J_{\pm}, T_0]].$$

The result is[5]

$$T_{\pm1}=\exp(\mp i\gamma)\Big(\Big(-2s\alpha c\alpha s\beta\mp 2ic^2\alpha s\beta c\beta\Big)\frac{\partial}{\partial\lambda}$$
$$+(2s\alpha c\alpha s\beta\mp 2is^2\alpha s\beta c\beta)\frac{\partial}{\partial\mu}+[(c^2\alpha-s^2\alpha)s\beta-c\beta ctg\beta$$
$$\mp 2is\alpha c\alpha s\beta c\beta]\frac{\partial}{\partial\alpha}\pm i(c^2\beta-s^2\beta)\frac{\partial}{\partial\beta}+ctg\beta\frac{\partial}{\partial\gamma}\Big)$$

and

$$T_{\pm2}=\exp(\mp 2i\gamma)\Big(\Big[\pm 2s\alpha c\alpha c\beta+i(c^2\alpha c^2\beta-s^2\alpha)\Big]\frac{\partial}{\partial\lambda}$$
$$+\Big[\mp 2s\alpha c\alpha c\beta+i(s^2\alpha c^2\beta-c^2\alpha)\Big]\frac{\partial}{\partial\mu}$$
$$+\Big[\mp 2c^2\alpha c\beta+is\alpha c\alpha(1+c^2\beta)\Big]\frac{\partial}{\partial\alpha}+is\beta c\beta\frac{\partial}{\partial\beta}\pm\frac{\partial}{\partial\gamma}\Big). \tag{A4}$$

From the explicit form of the operators J_0, T_0, $T_{\pm1}$, and $T_{\pm2}$ it is obvious that $[J_0, T_\mu]=\mu T_\mu$ ($\mu=0, \pm1, \pm2$). In order to make sure that the generators we have obtained actually generate the $SL(3,R)$ algebra we must verify that the $SL(3,R)$ condition, $[T_2, T_{-2}]=-4J_0$, is satisfied. An explicit calculation shows that this is indeed the case.

We know that the differential forms given for the $SL(3,R)$ group generators are also valid for $\overline{SL}(3,R)$ since, in so far as these differential operators are concerned, the only difference between $SL(3,R)$ and $\overline{SL}(3,R)$ lies in the range of the maximal compact subgroup parameters.

APPENDIX B

The elements of the $SU(2)$ group can be parametrized in terms of the three Auler angles α, β, and γ by

$$k(\alpha\beta\gamma)=\exp(-i\alpha J_Z)\exp(-i\beta J_Y)\exp(-i\gamma J_Z).$$

The unitary irreducible representations of this group are labeled by the value J of the Casimir operator J^2, and their matrix elements are D^J_{KM}-functions

$$D^J_{KM}(\alpha\beta\gamma)=\langle JK|\,k(\alpha\beta\gamma)\,|JM\rangle,$$

where $J=0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$, $-J\le K$, $M\le J$, and $\{|JM\rangle\}$ span the basis of the representation. Since J_Z is diagonal in the $\{|JM\rangle\}$ basis one has

$$D^J_{KM}(\alpha\beta\gamma)=\exp(-i\alpha K)\,d^J_{KM}(\beta)\exp(-i\gamma M), \tag{B1}$$

where

$$d^J_{KM}(\beta)=\langle JK|\,\exp(-iJ_Y\beta)\,|JM\rangle. \tag{B2}$$

The latter functions are given explicitly by

$$d^J_{KM}(\beta)=\left[\frac{(J+K)!\,(J-K)!}{(J+M)!\,(J-M)!}\right]^{1/2}\sum_N\binom{J+M}{J-K-N}\binom{J-M}{N}$$
$$\times(-)^{J+M+N}\left(c\frac{\beta}{2}\right)^{2N+K+M}\left(-s\frac{\beta}{2}\right)^{2J-2N-K-M},$$

and they are subject to the following symmetry properties

$$\begin{aligned} &d^J_{KM}(\beta)^*=d^J_{KM}(\beta),\\ &d^J_{KM}(\beta)=(-)^{K-M}d^J_{-K-M}(\beta),\\ &d^J_{KM}(\beta)=(-)^{K-M}d^J_{MK}(\beta). \end{aligned} \tag{B3}$$

The coefficients of the operator

$$T_0=i\Big(\sqrt{\tfrac{2}{3}}(3c^2\alpha s^2\beta-1)a+\sqrt{\tfrac{2}{3}}(3s^2\alpha s^2\beta-1)b$$
$$+\sqrt{6}\,s\alpha c\alpha s^2\beta\frac{\partial}{\partial\alpha}-\sqrt{6}\,s\beta c\beta\frac{\partial}{\partial\beta}\Big),$$

can be expressed in terms of the D^2_{KM} functions. Making use of

$$\begin{aligned} &d^2_{00}(\beta)=\tfrac{1}{2}(3c^2\beta-1),\\ &d^2_{10}(\beta)=\sqrt{\tfrac{2}{3}}\,s\beta c\beta,\\ &d^2_{20}(\beta)=\frac{\sqrt{6}}{4}s^2\beta, \end{aligned} \tag{B4}$$

and the symmetry properties of d^J_{KM} functions, one obtains

$$\sqrt{\tfrac{2}{3}}(3c^2\alpha s^2\beta-1)=-\sqrt{\tfrac{2}{3}}D^2_{00}+D^2_{20}+D^2_{-20},$$
$$\sqrt{\tfrac{2}{3}}(3s^2\alpha s^2\beta-1)=-\sqrt{\tfrac{2}{3}}D^2_{00}-D^2_{20}-D^2_{-20},$$

and

$$\sqrt{6}\,s\alpha c\alpha s^2\beta=i(D^2_{20}-D^2_{-20}). \tag{B5}$$

From the expressions

$$J_\pm=\exp(\mp i\gamma)\Big(-i\frac{1}{s\beta}\frac{\partial}{\partial\alpha}\mp\frac{\partial}{\partial\beta}+ictg\beta\frac{\partial}{\partial\gamma}\Big),$$

one obtains

$$-\sqrt{6}\,s\beta c\beta\frac{\partial}{\partial\beta}=\sqrt{\tfrac{2}{3}}\,s\beta c\beta[\exp(i\gamma)J_+-\exp(-i\gamma)J_-],$$

and this can be written as

$$-\sqrt{6}\,s\alpha c\beta\frac{\partial}{\partial\beta}=D^2_{01}J_-+D^2_{0-1}J_+. \tag{B6}$$

In the expression for T_0 the operator $K_0=i\partial/\partial\alpha$ actually acts on the left. However, making use of the relation

$$[K_0, D^J_{KM}]=KD^J_{KM},$$

one has

$$K_0(D^2_{20}-D^2_{-20})=2(D^2_{20}+D^2_{-20})+(D^2_{20}-D^2_{-20})K_0, \tag{B7}$$

so that we finally obtain

$$T_0=-i\sqrt{\tfrac{2}{3}}(a+b)D^2_{00}+i(D^2_{0-1}J_++D^2_{01}J_-)$$
$$+i(2+a-b)(D^2_{20}+D^2_{-20})+i(D^2_{20}-D^2_{-20})K_0. \tag{B8}$$

We work in the basis of the orthonormal vectors $|{}^J_{KM}\rangle$, i.e.,

$$\left\langle\begin{matrix}J'\\K'M'\end{matrix}\middle|\begin{matrix}J\\KM\end{matrix}\right\rangle=\delta_{J'J}\delta_{K'K}\delta_{M'M},$$

which are related to the D^J_{KM} functions by

$$\left\langle k\middle|\begin{matrix}J\\KM\end{matrix}\right\rangle=(2J+1)^{1/2}D^J_{KM}(k),\quad k\in SU(2).$$

By making use of the relation

$$\int dk\,D^{J'*}_{K'M'}(k)D^{J''}_{K''M''}(k)D^J_{KM}(k)=(-)^{J'-K'}(-)^{J'-M'}$$

$$\times\begin{pmatrix} J' & J'' & J \\ -K' & K'' & K \end{pmatrix}\begin{pmatrix} J' & J'' & J \\ -M' & M'' & M \end{pmatrix}, \tag{B9}$$

it is now straightforward to evaluate the matrix elements of T_0 in the $\{|{}^{J}_{KM}\rangle\}$ basis; we obtain

$$\left\langle {}^{J'}_{K'M'}\middle| T_0 \middle|{}^{J}_{KM}\right\rangle = (-)^{J'-K'}(-)^{J'-M'}[(2J'+1)(2J+1)]^{1/2}$$
$$\left\{\begin{pmatrix} J' & 2 & J \\ -M' & 0 & M \end{pmatrix}\times\left[-i\sqrt{\tfrac{2}{3}}(a+b)\begin{pmatrix} J' & 2 & J \\ -K' & 0 & K \end{pmatrix}\right.\right.$$
$$+i(2+K+a-b)\begin{pmatrix} J' & 2 & J \\ -K' & 2 & K \end{pmatrix}$$
$$\left.+i(2-K+a-b)\begin{pmatrix} J' & 2 & J \\ -K' & -2 & K \end{pmatrix}\right]$$
$$+i\begin{pmatrix} J' & 2 & J \\ -K' & 0 & K \end{pmatrix}\times\left[[J(J+1)-M(M+1)]^{1/2}\right.$$
$$\times\begin{pmatrix} J' & 2 & J \\ -M' & -1 & M+1 \end{pmatrix}$$
$$\left.\left.+[J(J+1)-M(M-1)]\begin{pmatrix} J' & 2 & J \\ -M' & 1 & M-1 \end{pmatrix}\right]\right\}. \tag{B10}$$

*Research supported in part by NSF and ARO (D).
†Present address: "B. Kidrich" Institute, Vincha, Belgrade, Yugoslavia.

[1]Y. Ne'eman, *Algebraic Theory of Particle Physics* (Benjamin, New York, 1967).
[2]L. O'Raifeartaigh, in *Symmetry Principles at High Energy*, Fifth Coral Gables Conference, edited by A. Perlmutter, C. A. Hurst, and B. Kursunoglu (Benjamin, New York, 1968).
[3]M. Gell-Mann, Phys. Rev. Lett. **14**, 77 (1965).
[4]Y. Dothan, M. Gell-Mann, and Y. Ne'eman, Phys. Lett. **17**, 148 (1965).
[5]Dj. Šijači, Ph.D. thesis, Duke University, 1974.
[6]L. Weaver and L.C. Biedenharn, Nucl. Phys. A **185**, 1 (1972).
[7]G. Rosen, Nuovo Cimento **42**, 797 (1966).
[8]E. Nelson, Ann. Math. **70**, 572 (1959).
[9]R. Hermann, *Lie Groups for Physicists* (Benjamin, New York, 1966).
[10]E.M. Stein, A Survey of Representations of Non Compact Groups, IAEA, Trieste, 1965.
[11]V. Bargmann, Ann. Math. **59**, 1 (1954).
[12]Harish-Chandra, Proc. Natl. Acad. Sci. USA **37**, 170 362, 366, 691 (1951).
[13]A. Kihlberg, Ark. Fys. **30**, 121 (1965).
[14]M. Rotenberg, R. Bivins, N. Metropolis, and J.K. Wooten, *The 3-j and 6-j Symbols* (MIT Press, Cambridge, 1959).
[15]I.M. Gel'fand and M.I. Graev, Am. Math. Soc. Trans. **2**, (2) 147 (1956).
[16]I. Hulthen, Ark. Fys. **38**, 175 (1968).
[17]G. Racah, "Lectures on Lie Groups," in *Group Theoretical Concepts and Methods in Elementary Particle Physics*, edited by F. Gürsey (Gordon and Breach, New York, 1964).
[18]If we denote by a^H and b^H the parameters in Hulthen's work, then our parameters a and b are related to his through $a=a^H$, $b=a^H-b^H$, and $\sigma=2a^H-b^H$, $\delta=b^H$. These substitutions, i.e., $a\to a$, $b\to a-b$, send I_2 into itself and I_3 into $-I_3$. Two representations with invariants I_2, I_3 and I_2, $-I_3$ are mutually contragradient.
[19]G.W. Mackey, *Induced Representations of Groups and Quantum Mechanics* (Benjamin, New York, 1968).
[20]L. Weaver, L.C. Biedenharn, and R.Y. Cusson, Ann. Phys. **77**, 250 (1973).
[21]Y. Dothan and Y. Ne'eman, "Band Spectra Generated by Non-Compact Algebra", in *Symmetry Groups in Nuclear and Particle Physics*, edited by F.J. Dyson (W.A. Benjamin, Inc., New York, 1966).
[22]N. Mukunda, L. O'Raifeartaigh, and E.C.G. Sudarshan, Phys. Rev. Lett. **15**, 1041 (1965).
[23]G. Rosen, J. Math. Phys. **7**, 1284 (1966).
[24]D.W. Joseph, University of Nebraska preprint (1970).
[25]L.C. Biedenharn, R.Y. Cusson, M.Y. Han, and O.L. Weaver, Phys. Lett. B **42**, 257 (1972), and Errata in which it is stated that the $SL(3,R)$ condition is valid for every J but $J=\frac{3}{2}$.
[26]E. Inönü and E.P. Wigner, Proc. Natl. Acad. Sci. USA **39**, 510 (1953).
[27]C. Goebel, "Representations of Strong Coupling Groups" in *Non-Compact Groups in Particle Physics*, edited by Y. Chow (Benjamin, New York, 1966).

Representation theory of the universal covering of the Euclidean conformal group and conformal invariant Green's functions

G. Grensing*

Deutsches Elekronen-Synchrotron DESY, Hamburg, Germany
(Received 13 August 1974)

We present a special realization for the universal covering of the Euclidean conformal group. This group can be defined as a transformation group on the Euclidean version of compactified Minkowski space such that the action on this space coincides with the usual one of the Euclidean conformal group. We construct the representations of the principal and complementary series and derive the intertwining kernels for the equivalent representations. The connection between representation theory and conformal invariant quantum field theory is studied. To this end we also give the reduction of the tensor product of two representations of the supplementary series.

The group $SU^*(4)$ (Ref. 1) is the universal covering of the Euclidean conformal group $SO_0(5,1)$. The unitary irreducible representations of this group are completely known on the pure Lie algebra level of representation theory.[2-5] It is the main goal of this paper to develop the global representation theory.

The investigation is motivated by the recent interest in conformal invariant Euclidean quantum field theory,[6-8] the underlying symmetry group being the twofold covering of $SO_0(5,1)$. It is generally accepted that such a theory becomes relevant if interpreted as a Gell-Mann Low limit[9] of a renormalizable field theory.[10]

As is well known, the two- and three-point functions of a conformal invariant quantum field theory are, up to a multiplicative constant, uniquely determined.[11,12] For the four-point function a conformal invariant partial wave decomposition has been given, which contains only products of two- and three-point functions.[13] With this knowledge the system of coupled integral equations the n-point functions are known to obey[14] becomes tractable. This has been shown by Mack.[6] Furthermore, he has drawn attention to the fact that the two- and three-point functions have a pure group theoretical meaning. A thorough investigation of this point will be given in this paper.

We present a special realization G of $SU^*(4)$ (Sec. 1), which will be useful in exhibiting the relation between representation theory and conformal invariant quantum field theory. A factorization of this group can be derived (Sec. 2) such that the Euclidean version of compactified Minkowski space is obtained as the factor space G/G', where G' is the inducing subgroup for the unitary irreducible representations of G. Our choice of the universal covering group proves to be significant because the action of G on G/G' coincides with the usual one of the conformal group (Sec. 3). Thus, the representations of the principal (Sec. 4) and supplementary series (Sec. 5) of G act on fields over Euclidean space and they yield a transformation law, which is adapted to the conformal group. The two-point functions appear in this context as kernels for the scalar product of the supplementary series or as intertwining kernels for equivalent representations (Sec. 6), and the three-point functions as Clebsch–Gordan kernels for the tensor product of two representations of the supplementary series (Sec. 7).

After completion of this work we have received a preprint by Koller,[15] in which a similar program is carried out for the groups $SO_0(n,1)$. He gives an independent derivation of the intertwining kernels by means of the nontrivial element of the Weyl group.

1. THE UNIVERSAL CONVERING GROUP OF SO_0 (5,1)

We use the following realization of the universal covering group of $SO_0(5,1)$: G consists of the elements g of $SL(4,\mathbb{C})$ which obey the condition

$$gE = Eg^* \tag{1.1}$$

where

$$E = \begin{pmatrix} -\epsilon & 0 \\ 0 & +\epsilon \end{pmatrix} \quad \text{with} \quad \epsilon = \begin{pmatrix} 0 & +1 \\ -1 & 0 \end{pmatrix} \tag{1.2}$$

and $*$ means complex conjugation.

If g is split into 2×2 matrices such that

$$g = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}, \tag{1.3}$$

the condition (1.1) reads

$$\begin{aligned} g_{11}\epsilon &= +\epsilon g_{11}^*, \quad g_{12}\epsilon = -\epsilon g_{12}^*, \\ g_{21}\epsilon &= -\epsilon g_{21}^*, \quad g_{22}\epsilon = +\epsilon g_{22}^*. \end{aligned} \tag{1.4}$$

It is readily verified that G is isomorphic to $SU^*(4)$ (Ref. 1) by virtue of a suitable real, orthogonal transformation. Furthermore, it can be shown that there exists a homomorphism of G onto $SO_0(5,1)$ which has kernel $Z_2 = \{+e, -e\}$, this yielding the well-known isomorphism

$$SU^*(4)/Z_2 \hat{=} SO_0(5,1). \tag{1.5}$$

2. FACTORIZATION OF G

The set of conditions (1.4) on g is now used to obtain a parametrization for G, which will be valid for all elements of G with the exception of a lower dimensional manifold.

The elements g of G with subdeterminant $|g_{11}| \neq 0$ can uniquely be factorized into

$$g = g_1 g_2 g_3 g_4, \tag{2.1}$$

each factor constituting a subgroup $G_a (a = 1,2,3,4)$ of G. They are explicitly given by:

$$(1)\quad g_1 = \begin{pmatrix} 1 & 0 \\ A & 1 \end{pmatrix}. \tag{2.2}$$

The 2×2 matrix A is built up by a 4-vector $a = (a^1, a^2, a^3, a^4)$ according to

$$A = \begin{pmatrix} -a^1 - ia^2 & ia^4 + a^3 \\ -ia^4 + a^3 & a^1 - ia^2 \end{pmatrix}. \tag{2.3}$$

G_1 is an Abelian, four-dimensional subgroup of G.

$$(2)\quad g_2 = \begin{pmatrix} \mathcal{A}' & 0 \\ 0 & \mathcal{A}''^* \end{pmatrix}, \tag{2.4}$$

with $\mathcal{A}', \mathcal{A}'' \in SU(2)$.

Because G_2 is simply $SU(2)\otimes SU(2)$, we use the notation $\mathcal{A} = (\mathcal{A}', \mathcal{A}'')$.

$$(3)\quad g_3 = \begin{pmatrix} e^{-\lambda/2}\,1 & 0 \\ 0 & e^{+\lambda/2}\,1 \end{pmatrix}, \tag{2.5}$$

with $\lambda \in (\mathbb{R})$.

This subgroup is isomorphic to the multiplicative group of the positive real numbers.

$$(4)\quad g_4 = \begin{pmatrix} 1 & -C^\dagger \\ 0 & 1 \end{pmatrix}. \tag{2.6}$$

The matrix C has the same form as has been given in (2.3), the 4-vector a replaced by c.

3. THE CONNECTION WITH THE EUCLIDEAN CONFORMAL GROUP

Obviously, the elements of $G_1 \cdot G_2$ form a subgroup of G isomorphic to the twofold covering group of the Euclidean Poincaré group. This suggests examining the action of G on the translational part G_1 of G, the elements of which we now write as

$$g_x = \begin{pmatrix} 1 & 0 \\ X & 1 \end{pmatrix}. \tag{3.1}$$

According to the results of Sec. 2 we can uniquely factorize $g g_x$ for $|(g g_x)_{11}| \neq 0$ into

$$g g_x = g_{x'} g' \tag{3.2}$$

where $g' \in G' = G_2 G_3 G_4$. Hence, we obtain

$$X' = g_{21} + g_{22} X / g_{11} + g_{12} X. \tag{3.3}$$

The computation of (3.3) for the various subgroups G_a $(a = 1, 2, 3, 4)$ is facilitated by means of the identity

$$\tfrac{1}{2}(X_1 X_2^\dagger + X_2 X_1^\dagger) = x_1 \cdot x_2 \, 1, \tag{3.4}$$

the right-hand side being the usual scalar product over Euclidean space M; the result is:

$$\begin{aligned} &(1)\ x' = x + a, \\ &(2)\ x' = \Lambda x, \\ &(3)\ x' = e^{\lambda} x, \\ &(4)\ x' = x - x^2 c / 1 - 2x \cdot c + x^2 c^2. \end{aligned} \tag{3.5}$$

In deriving Eq. (2) of (3.5) we have used the covering map π of $SU(2)\otimes SU(2)$ onto the Euclidean Lorentz group $L_0 = SO(4)$ given by

$$\Lambda^\mu{}_\nu = \tfrac{1}{2} \operatorname{tr}\{\sigma^\mu \mathcal{A}' \tilde{\sigma}_\nu \mathcal{A}''^\dagger\}, \tag{3.6}$$

where $\sigma = (\sigma^k, -i1)_{k=1,2,3}$ and $\tilde{\sigma} = (\sigma^k, +i1)_{k=1,2,3}$.

From (3.5) we see that the elements of G with $|g_{11}| \neq 0$ act on $g_x G'$ in exactly the same way as the Euclidean conformal group on elements x of M. Thus we can identify the factors occurring in (2.1) as translations, Lorentz transformations, dilatations, and special conformal transformations. For that reason we shall use the notation

$$g = (a, \mathcal{A} | d, c) \tag{3.7}$$

for the elements with $|g_{11}| \neq 0$.

4. THE PRINCIPAL SERIES OF REPRESENTATIONS OF G

In this section we will construct the representations of the principal series. To do this we must know the Iwasawa decomposition $G = K\,A\,N$ of G with K being the maximal compact, A an Abelian and N a nilpotent subgroup.[1] In our specific case, K is the unitary subgroup of G

$$K = \{g \in G \,|\, g^\dagger g = e\}. \tag{4.1}$$

It is easy to prove that K is isomorphic to $Sp(2)$. The remaining factors occurred already, $A = G_3$ and $N = G_4$.

The representations of the principal series are obtained as induced representations on the homogeneous space G/G', where $G' = K' A\, N$ and K' being the centralizer of A in K.[16] We have $K' = G_2$, so that

$$G' = G_2 G_3 G_4. \tag{4.2}$$

The factor space G/G' is diffeomorphic to K/K'. However, we want to use the compactified M as homogeneous space, this leading to the decomposition (2.1).

According to the general theory, the representations are induced by the unitary and irreducible representations of G' which are trivial on N. The complete system of representations of $G_2 G_3$, which is just $SU(2)\otimes SU(2) \otimes \mathbb{R}_+$, can immediately be written down

$$D(g') = (e^{\lambda})^{-i\rho} D^{(l_1, l_2)}(\mathcal{A}), \tag{4.3}$$

where ρ is real and $D^{(l_1,l_2)}(\mathcal{A}) = D^{(l_1)}(\mathcal{A}') \otimes D^{(l_2)}(\mathcal{A}'')$ is a unitary and irreducible representation of $SU(2)\otimes SU(2)$ on the tensor product $\mathbb{C}^{2l_1+1} \otimes \mathbb{C}^{2l_2+1}$ with $l_1, l_2 = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \cdots$.

The representation space H consists of the functions

$$\phi : M \to \mathbb{C}^{2l_1+1} \otimes \mathbb{C}^{2l_2+1} \tag{4.4}$$

infinitely differentiable on M including infinity. Finally, we may define the representations of the principal series

$$U(g)\,\phi(x) = \left| \frac{\partial x'}{\partial x} \right|^{1/2} D(g')\,\phi(x') \tag{4.5}$$

with

$$g^{-1}g_x = g_{x'}g'^{-1}, \quad x' = g^{-1}\cdot x. \tag{4.6}$$

and where we have to exclude the point at infinity. The representation is unitary relative to the scalar product

$$(\phi_1, \phi_2) = \int \phi_1^\dagger(x)\,\phi_2(x)\,d^4x \tag{4.7}$$

and known to be irreducible, too.[16,17]

To get a more explicit form of the transformation law (4.5), we use the decomposition (2.1); the factors g_a $(a=1, 2, 3, 4)$ yield:

$$\begin{aligned}
&(1)\ U^\chi(a)\,\phi(x) = \phi(x-a),\\
&(2)\ U^\chi(\mathcal{A})\,\phi(x) = D^{(l_1,l_2)}(\mathcal{A})\,\phi(\Lambda^{-1}x),\\
&(3)\ U^\chi(d)\,\phi(x) = (e^\lambda)^{-\Delta}\phi(e^{-\lambda}x),\\
&(4)\ U^\chi(c)\,\phi(x) = (1+2x\cdot c+x^2c^2)^{-\Delta}D^{(l_1,l_2)}(\mathcal{A}(x,c))\\
&\qquad\qquad \times\phi(x+x^2c/1+2x\cdot c+x^2c^2).
\end{aligned} \tag{4.8}$$

In the last equation the $SU(2)\otimes SU(2)$ element

$$\mathcal{A}(x,c) = (\mathcal{A}'(x,c),\ \mathcal{A}''(x,c)) \tag{4.9}$$

is given by

$$\begin{aligned}
\mathcal{A}'(x,c) &= (1+X^\dagger C)/(1+2x\cdot c+x^2c^2)^{1/2}\\
\mathcal{A}''(x,c) &= (1+XC^\dagger)^*/(1+2x\cdot c+x^2c^2)^{1/2}.
\end{aligned} \tag{4.10}$$

Furthermore, we have defined $\Delta = 2+i\rho$ and more accurately written U^χ with χ being an abbreviation for the labels of the representation of G'

$$\chi = (l_1, l_2;\ \Delta = 2+i\rho). \tag{4.11}$$

After all, there remains to be shown that the integral (4.7) is actually convergent for $\phi \in H^\chi$. This will be proved by the aid of the element

$$g_\infty = \begin{pmatrix} 0 & -i\epsilon \\ +i\epsilon & 0 \end{pmatrix} \tag{4.12}$$

of the maximal compact subgroup of G, which maps x into

$$g_\infty \cdot x = \Lambda_T\, x/x^2 \tag{4.13}$$

where Λ_T is the time inversion. Obviously, this element is appropriate for studying the behavior of ϕ for large x. To compute the action of $U^\chi(g_\infty)$ on ϕ, we observe that $\mathcal{A}(x,c)$ can be cast into the form

$$\mathcal{A}(x,c) = \mathcal{A}'_x\,\mathcal{A}^\dagger_{x'}, \tag{4.14}$$

where $x' = x+x^2c/1+2x\cdot c+x^2c^2$ and $\mathcal{A}_x = (\mathcal{A}'_x,\ \mathcal{A}''_x)$ with

$$\mathcal{A}'^\dagger_x = \frac{i}{(x^2)^{1/2}}\,X = \mathcal{A}''^*_x. \tag{4.15}$$

Then the final result is

$$U^\chi(g_\infty)\,\phi(x) = (x^2)^{-\Delta}D^{(l_1,l_2)}(\mathcal{A}_x)D^{(l_1,l_2)}(\mathcal{A}_\epsilon)\,\phi(\Lambda_T\,x/x^2) \tag{4.16}$$

with $\mathcal{A}_\epsilon = (-\epsilon, +\epsilon)$. From (4.16) we may derive the asymptotic behavior

$$|\phi(x)| \sim c\,(x^2)^{-(\Delta+\Delta^*)/2} \quad \text{for } x^2 \to \infty, \tag{4.17}$$

which yields the convergence of the integral (4.7) for arbitrary ρ in R.

With (4.8) we come into contact with conformal quantum field theory: If we continue x^4 to imaginary values, that is, define $ix^4 = x^0$ such that $SO(4)$ gets $SO_0(3,1)$, we obtain the integrated form of the representation of the conformal Lie algebra acting on fields over Monkowski space.[18]

5. THE SUPPLEMENTARY SERIES OF REPRESENTATIONS OF G

We proceed in analogy to the analysis for the universal covering of $SO_0(3,1)$ (Refs. 19, 20) and try to generalize Δ in (4.8) to arbitrary complex values δ and the scalar product (4.7) to

$$(\phi_1, \phi_2) = \int \phi_1^\dagger(x_1)K(x_1,x_2)\,\phi_2(x_2)\,d^4x_1\,d^4x_2 \tag{5.1}$$

such that the resulting representation is unitary and irreducible. Thus, U^χ must leave invariant the bilinear form (5.1), which will serve to fix the dependence of the kernel $K(x_1,x_2)$ on (x_1,x_2). Furthermore, the integral (5.1) must exist and be positive definite, which will yield the admissible range of δ.

At first, we analyze the invariance condition on the kernel:

$$\begin{aligned}
&(1)\ K(x_1-a,\ x_2-a) = K(x_1,\ x_2),\\
&(2)\ D^{(l_1,l_2)}(\mathcal{A})K(\Lambda^{-1}x_1,\ \Lambda^{-1}x_2)D^{(l_1,l_2)}(\mathcal{A})^\dagger = K(x_1,\ x_2),\\
&(3)\ (e^\lambda)^{-(4-\delta^*)}K(e^{-\lambda}x_1,\ e^{-\lambda}x_2)(e^\lambda)^{-(4-\delta)} = K(x_1,\ x_2),\\
&(4)\ (x_1'^2/x_1^2)^{-(4-\delta^*)}D^{(l_1,l_2)}(\mathcal{A}_{x_1}\mathcal{A}^\dagger_{x_1'})K(x_1',\ x_2')D^{(l_1,l_2)}\\
&\qquad\times(\mathcal{A}_{x_2}\mathcal{A}^\dagger_{x_2'})^\dagger(x_2'^2/x_2^2)^{-(4-\delta)} = K(x_1,\ x_2).
\end{aligned} \tag{5.2}$$

Choosing $a = -x_2$ in Eq. (1) of (5.2), we see the $K(x_1,\ x_2)$ is a function of $x_1 - x_2$ only:

$$K(x_1,\ x_2) = K(x_1-x_2,\ 0) =: K(x_1-x_2). \tag{5.3}$$

We take $x_2 = 0$, $x_1 = x$ and define

$$K(x) = [1/(x^2)^{4-\delta^*}]\hat{K}(x) \tag{5.4}$$

to simplify Eq. (4) of (5.2), this yielding

$$\begin{aligned}
&(2')\ D^{(l_1,l_2)}(\mathcal{A})\hat{K}(\Lambda^{-1}x)D^{(l_1,l_2)}(\mathcal{A})^\dagger = \hat{K}(x)\\
&(3')\ (e^\lambda)^{\delta-\delta^*}\hat{K}(e^{-\lambda}x) = \hat{K}(x)\\
&(4')\ D^{(l_1,l_2)}(\mathcal{A}_{x'})^\dagger\hat{K}(x') = D^{(l_1,l_2)}(\mathcal{A}_x)^\dagger\hat{K}(x)
\end{aligned} \tag{5.5}$$

where $x' = x+x^2c/1+2x\cdot c+x^2c^2$.

Equation (4') of (5.5) requires

$$\hat{K}(x) = D^{(l_1,l_2)}(\mathcal{A}_x)\hat{K}_0 \tag{5.6}$$

with $\hat{K}_0$ being a constant, invertible matrix. To see whether Eq. (2') of (5.5) is satisfied by (5.6), we use

$$(\mathcal{A}'_{\Lambda x}, \mathcal{A}''_{\Lambda x}) = (\mathcal{A}', \mathcal{A}'')(\mathcal{A}'_x, \mathcal{A}''_x)(\mathcal{A}''^*, \mathcal{A}'^*) \tag{5.7}$$

to transform (2') into

$$K_0D^{(l_1)}(\mathcal{A}')\otimes D^{(l_2)}(\mathcal{A}'') = D^{(l_1)}(\mathcal{A}'')\otimes D^{(l_2)}(\mathcal{A}')K_0 \tag{5.8}$$

where we have set

$$\hat{K}_0 = D^{(l_1,l_2)}(\mathcal{A}_\epsilon)K_0. \tag{5.9}$$

The condition (5.8) requires $l_1 = l_2$ and K_0 is determined to be

$$K_0{}^{i_1 i_2}{}_{k_1 k_2} = k\delta^{i_1}{}_{k_2}\,\delta^{i_2}{}_{k_1}. \tag{5.10}$$

Finally, Eq. (3′) of (5.5) shows that δ must be real.

Collecting the results, we get

$$K(x) = \frac{1}{(x^2)^{4-\delta}} D^{(l_1,l_2)}(\mathcal{A}_x) D^{(l_1,l_2)}(\mathcal{A}_e) K_0,$$

$$K_0{}^{i_1 i_2}{}_{k_1 k_2} = k\delta^{i_1}{}_{k_2}\,\delta^{i_2}{}_{k_1}, \quad \begin{matrix} i_1, k_1 = -l_1, \ldots, +l_1, \\ i_2, k_2 = -l_2, \ldots, +l_2 \end{matrix} \tag{5.11}$$

with $\delta = \delta^*$ and $l_1 = l_2$.

At this stage, it is convenient to use another realization of $D^{(l_1,l_2)}(\mathcal{A})$ with $l_1 = l_2$. This representation is equivalent to the l-fold tensor product $\overset{l}{\otimes}\Lambda$ with $l_1 = l_2 = l/2$ acting on the completely symmetric, traceless tensors of rank l over M.[21] We denote these representations by $D^{(l)}(\Lambda)$.

To determine the analogue of (5.11) in this realization, we must know

$$\pi(\mathcal{A}(x,c)) = \Lambda(x,c). \tag{5.12}$$

This is done with the help of (3.6); the result is

$$\Lambda(x,c) = g(x)\, g(x + x^2 c/1 + 2x\cdot c + x^2 c^2), \tag{5.13}$$

where

$$g_{\mu\nu}(x) = g_{\mu\nu} - 2(x_\mu x_\nu / x^2). \tag{5.14}$$

We remark that $g(x)$ is not contained in L_0, $\pi(\mathcal{A}_x)$ is the product of $g(x)$ and the reflection with respect to the second axis. However, the representation $D^{(l)}(\Lambda)$ of L_0 can be extended to a representation of the complete Lorentz group. This property of $D^{(l)}(\Lambda)$ will be used in the sequel.

Repeating the arguments used to derive (5.11), we get

$$K^{\bar\chi}(x)^{\mu_1\cdots\mu_l}{}_{\nu_1\cdots\nu_l} = k(\bar\chi)\frac{1}{(x^2)^{4-\delta}}\frac{1}{l!}\sum_{\pi\in S_l}\{g(x)^{\mu_1\nu_{\pi(1)}}\cdots \times g(x)^{\mu_2\nu_{\pi(1)}} - \text{traces}\}, \tag{5.15}$$

where S_l with elements π is the permutation group of l objects. We slightly changed the notation in (5.15) in writing more specifically $K^{\bar\chi}$ for K because of the invariance property

$$U^{\bar\chi}(g) \otimes U^{\bar\chi}(g)\, K^{\bar\chi}(x_1, x_2) = K^{\bar\chi}(x_1, x_2) \tag{5.16}$$

with $\bar\chi = (\bar l; \bar\delta)$ being defined by

$$\bar l = l, \quad \bar\delta = 4 - \delta. \tag{5.17}$$

The last requirement we have to fulfill is that the invariant bilinear form

$$(\phi_1, \phi_2)_\chi = \int \phi_1^\dagger(x_1) K^{\bar\chi}(x_1, x_2)\, \phi_2(x_2)\, d^4x_1 d^4x_2 \tag{5.18}$$

be finite and positive. For this integral to be convergent, it is obviously necessary that $\delta > 2$. To give a sufficient answer, we use Fourier transformation because $K^{\bar\chi}(x_1, x_2)$ is a function of $x_1 - x_2$ only. The Fourier transform

$$\hat\phi(p) = \int d^4x \exp(-ipx)\, \phi(x) \tag{5.19}$$

of $\phi(x)$ exists owing to $\delta > 2$ and the asymptotic behavior (4.17). The computation of

$$\hat K^{\bar\chi}(p) = \int d^4x \exp(-ipx) K^{\bar\chi}(x) \tag{5.20}$$

will be reduced to[22]

$$(x^2)^{-\delta} = \frac{(4\pi)^2}{2^{2\delta}}\frac{\Gamma(2-\delta)}{\Gamma(\delta)}\int \frac{d^4p}{(2\pi)^4}\exp(ipx)(p^2)^{-2+\delta}, \tag{5.21}$$

which in turn may be used to show that

$$\hat K^{\bar\chi}(p) \sim C(p^2)^{2-\delta} \quad \text{for } p^2 \to 0 \tag{5.22}$$

and leads to the restriction $\delta < 4$. Hence, δ must lie in the interval $2 < \delta < 4$.

If (5.18) is expressed in terms of the Fourier transforms, we get

$$(\phi, \phi)_\chi = \int \frac{d^4p}{(2\pi)^4}\,\hat\phi(p)\hat K^{\bar\chi}(p)\,\hat\phi(p). \tag{5.23}$$

Thus, the positivity condition amounts to the requirement that

$$z_{\mu_1\cdots\mu_l}\hat K^{\bar\chi}(p)^{\mu_1\cdots\mu_l\nu_1\cdots\nu_l} z_{\nu_1\cdots\nu_l} > 0 \tag{5.24}$$

for a nonzero, completely symmetric, and traceless tensor z of rank l. This condition is trivially fulfilled for $l = 0$, if the constant $k(0, 4-\delta)$ appearing in (5.15) is chosen to be real and positive. For $l = 1$ it yields because of

$$\hat K^{(1,4-\delta)}(p) = k(1, 4-\delta)\frac{(4\pi)^2}{2^{2(4-\delta)}}\frac{\Gamma(\delta-2)}{\Gamma(5-\delta)}(p^2)^{2-\delta} \times\left((3-\delta)g_{\mu\nu} + (\delta-2)2\frac{p_\mu p_\nu}{p^2}\right) \tag{5.25}$$

that $\delta \leq 3$ and $k(1, 4-\delta) > 0$. This condition is equally valid for $l = 2$, as can be shown by a somewhat lengthy computation, and is indeed known to hold true for all $l \neq 0$.[4]

We state the final results: If δ is restricted to the interval

$$\begin{aligned} &2 < \delta < 4 \quad \text{for } l = 0, \\ &2 < \delta < 3 \quad \text{for } l \neq 0, \end{aligned} \tag{5.26}$$

the representation (4.8) with $\chi = (l;\delta)$ and $l_1 = l_2 = l/2$ is unitary with respect to the scalar product (5.18), where the kernel is given by (5.15). These representations are said to belong to the supplementary series. We assert, but do not prove that they are irreducible. At the integer points $\chi = (l; \delta = 3)$ the exceptional series occurs.

The kernel $K^{\bar\chi}(x_1, x_2)$ is the inverse dressed propagator of conformal invariant Euclidean quantum field theory. The two-point function for a field with half-integer spin can be obtained by taking an appropriate direct sum of representations (4.3). But these representations are not irreducible.

6. EQUIVALENCE OF REPRESENTATIONS OF G

The series of representations derived in Secs. 4 and 5 are known to exhaust the irreducible and unitary representations of G.[4] These representations are, however, not all inequivalent.

To study this question, we look for a bounded operator K of H^χ into $H^{\chi'}$ such that K intertwines the action of $U^\chi(g)$ and $U^{\chi'}(g)$ for arbitrary g in G:

$$U^{\chi'}(g)K = KU^{\chi}(g). \tag{6.1}$$

The kernel $K(x, x')$ of K with

$$\phi'(x') = (K\phi)(x) = \int K(x', x)\,\phi(x)\,d^4x \tag{6.2}$$

thus has to satisfy

$$U^{\chi'}(g)\,\phi'(x') = \int K(x', x)\,U^{\chi}(g)\,\phi(x)\,d^4x. \tag{6.3}$$

The computation proceeds along the lines of Sec. 5. We only state the results: An intertwining operator for the representations U^{χ} and $U^{\chi'}$ exists if $\chi' = \bar{\chi}$ with

$$\overline{(l_1, l_2)} = (l_2, l_1), \quad \bar{\Delta} = 4 - \Delta. \tag{6.4}$$

The kernel $K(x', x) = K(x' - x)$ has the form

$$K(x) = \frac{1}{(x^2)^{4-\Delta}}\,D^{(l_1,l_2)}(\mathcal{A}_x)\,D^{(l_1,l_2)}(\mathcal{A}_\epsilon)\,K_0, \tag{6.5}$$

where

$$K_0{}^{i_1 i_2}{}_{k_1 k_2} = k\delta^{i_1}{}_{k_2}\,\delta^{i_2}{}_{k_1}, \quad i_1, k_2 = -l_1, \ldots, +l_1, \quad i_2, k_1 = -l_2, \ldots, +l_2.$$

We note that, if necessary, the kernel has to be appropriately regularized.[22]

The operator K is invertible, thus yielding that the following representations are equivalent:

$$U^{\chi}(g) \hat{=} U^{\bar{\chi}}(g), \qquad \chi = (l_1, l_2; \Delta), \quad \bar{\chi} = (l_2, l_1; 4 - \Delta). \tag{6.6}$$

However, we do not give the proof that K is invertible and will treat in some detail only the case of the principal and supplementary series with $l_1 = l_2 = l/2$. As has been explained in Sec. 5, we may change to an equivalent representation acting on completely symmetric, traceless tensors over M such that the kernel takes the form

$$K^{\bar{\chi}}(x)^{\mu_1\cdots\mu_l}{}_{\nu_1\cdots\nu_l} = k(\bar{\chi})\frac{1}{(x^2)^{4-\Delta}}\frac{1}{l!}\sum_{\pi\in S_l} \times \{g(x)^{\mu_1\nu_{\pi(1)}}\cdots g(x)^{\mu_l\nu_{\pi(l)}} - \text{traces}\}. \tag{6.7}$$

Making the substitution $\delta \to \Delta$ in (5.15), we observe that the resulting expression is identical with (6.7). Thus the kernel (6.7) plays a twofold role for the representations of the supplementary series as intertwining kernel and as kernel for the scalar product (5.18). The kernel for the principal series of representations may be obtained from the kernel of the supplementary series by analytic continuation in δ.

As can immediately be shown, K^{χ} and $K^{\bar{\chi}}$ must obey the relation

$$\int K^{\chi}(x - x'')\,K^{\bar{\chi}}(x'' - x')\,d^4x'' = 1\,\delta^4(x - x'), \tag{6.8}$$

which imposes a restriction on the constant $k(\bar{\chi})$ in (6.7). The value of this constant may be guessed by treating the low-dimensional cases, e.g., for $l = 2$ one must use

$$K^{(2,\Delta)}(x)^{\mu_1\mu_2\nu_1\nu_2} = k(2, \Delta)\frac{1}{(x^2)^{\Delta}}\frac{1}{2}\{g(x)^{\mu_1\nu_1}g(x)^{\mu_2\nu_2} + g(x)^{\mu_1\nu_2}g(x)^{\mu_2\nu_1} - \tfrac{1}{2}g^{\mu_1\mu_2}g^{\nu_1\nu_2}\}. \tag{6.9}$$

A possible choice, which is compatible with the results of Sec. 5, is

$$k(\bar{\chi}) = \frac{1}{\pi^2}\frac{\Gamma(4 - \delta + l)}{\Gamma(-2 + \delta)}\{(\delta - 1)\,\delta(\delta + 1)\cdots(\delta + l - 2)\}^{-1}. \tag{6.10}$$

Thus, for χ in the principal series the representations with $\chi = (l; 2 + i\rho)$ and $\bar{\chi} = (l; 2 - i\rho)$ are equivalent. For χ in the supplementary series we get new representations, because $\bar{\delta}$ takes the values $0 < \bar{\delta} < 2$ for $l = 0$ and $1 < \bar{\delta} < 2$ for $l \neq 0$. Hence, the admissible range of δ for the supplementary series is

$$\left.\begin{array}{l} 0 < \delta < 4 \quad \text{for } l = 0 \\ 1 < \delta < 3 \quad \text{for } l \neq 0 \end{array}\right\} \quad (\delta \neq 2). \tag{6.11}$$

If $\delta < 2$ the kernel (6.7) must be regularized such that the integral (5.18) is convergent.

7. CLEBSCH-GORDAN COEFFICIENTS FOR THE TENSOR PRODUCT OF REPRESENTATIONS OF THE SUPPLEMENTARY SERIES

In this concluding section we make some qualitative remarks about Clebsch–Gordan kernels of two representations of the supplementary series, which play the role of three-point functions in conformal invariant Euclidean quantum field theory.

The problem we want to solve is the decomposition of the tensor product $U^1(g) \otimes U^2(g)$ acting on $\phi_{12}(x_1, x_2)$ into irreducible components. We will restrict ourselves to the spinless case, $\chi_1 = (0; \delta_1)$ and $\chi_2 = (0; \delta_2)$; the general case can be dealt with in an analogous way. The decomposition of the Kronecker product is implemented by a kernel $C^{\bar{\chi}}(x_1, x_2; x_3)$ such that

$$\phi_{12}(x_1, x_2) = \sum\!\!\!\!\!\!\int d\chi\, C^{\bar{\chi}}(x_1, x_2; x_3)\,\phi^{\chi}(x_3)\,d^4x_3 \tag{7.1}$$

where $d\chi$ means summation over (l_1, l_2) and integration over Δ.

We have to specify which series of representations of G contribute to the integral (7.1): From the analogous problem for the universal covering group of $SO_0(3,1)$ we know that only representations of the principal series occur.[23] We shall assume this to hold true in our case.

The decomposition (7.1) must accomplish

$$U_1(g) \otimes U_2(g)\,\phi_{12}(x_1, x_2) = \sum\!\!\!\!\!\!\int d\chi \int C^{\bar{\chi}}(x_1, x_2; x_3)U^{\chi}(g)\,\phi^{\chi}(x_3)\,d^4x_3, \tag{7.2}$$

which serves to determine the explicit form of the kernel:

$$\begin{aligned} &(1)\quad C^{\bar{\chi}}(x_1 - a, x_2 - a; x_3 - a) = C^{\bar{\chi}}(x_1, x_2; x_3), \\ &(2)\quad C^{\bar{\chi}}(\Lambda^{-1}x_1, \Lambda^{-1}x_2; \Lambda^{-1}x_3) = C^{\bar{\chi}}(x_1, x_2; x_3)\,D^{(l_1,l_2)}(\mathcal{A}), \\ &(3)\quad (e^{\lambda})^{-\delta_1}(e^{\lambda})^{-\delta_2}\,C^{\bar{\chi}}(e^{-\lambda}x_1, e^{-\lambda}x_2; e^{-\lambda}x_3) \\ &\qquad = C^{\bar{\chi}}(x_1, x_2; x_3)(e^{\lambda})^{4-\Delta}, \\ &(4)\quad (x_1^2/x_1'^2)^{-\delta_1}(x_2^2/x_2'^2)^{-\delta_2}\,C^{\bar{\chi}}(x_1', x_2'; x_3') = C^{\bar{\chi}}(x_1, x_2; x_3) \\ &\qquad \times D^{(l_1,l_2)}(\mathcal{A}(x_3, c))\,(x_3^2/x_3'^2)^{4-\Delta} \end{aligned} \tag{7.3}$$

where $x_i' = x_i + x_i^2 c/1 + 2x_i \cdot c + x_i^2 c^2$, $i = 1, 2, 3$.

We use Eq. (1) to define

$$C^{\bar{\chi}}(x_1 - x_2, x_2 - x_3) := C^{\bar{\chi}}(x_1 - x_3, x_2 - x_3; 0) \tag{7.4}$$

and, furthermore, set

$$C^{\bar{\chi}}(x_1, x_2) = (x_1^2)^{-\delta_1}(x_2^2)^{-\delta_2}\,\hat{C}^{\bar{\chi}}(x_1, x_2), \tag{7.5}$$

so that the remaining Equations of (7.3) take the simple form

$$\begin{aligned}
&(2')\quad \hat{C}^{\bar{\chi}}(\Lambda^{-1}x_1, \Lambda^{-1}x_2) = \hat{C}^{\bar{\chi}}(x_1, x_2)\,D^{(l_1, l_2)}(\mathcal{A}),\\
&(3')\quad (e^{\lambda})^{-4+\Delta+\delta_1+\delta_2}\,\hat{C}^{\bar{\chi}}(e^{-\lambda}x_1, e^{-\lambda}x_2) = \hat{C}^{\bar{\chi}}(x_1, x_2),\\
&(4')\quad \hat{C}^{\bar{\chi}}(x_1', x_2') = \hat{C}^{\bar{\chi}}(x_1, x_2).
\end{aligned} \tag{7.6}$$

Equation (4') requires $\hat{C}^{\bar{\chi}}(x_1, x_2)$ to be invariant with respect to special conformal transformations. Because

$$x_1/x_1^2 - x_2/x_2^2 \tag{7.7}$$

and

$$(x_1/x_1^2 - x_2/x_2^2)^2 = (x_1 - x_2)^2/x_1^2\,x_2^2 \tag{7.8}$$

are the only invariant expressions which can be formed out of x_1 and x_2, we get

$$\hat{C}^{\bar{\chi}}(x_1, x_2) = \hat{C}^{\bar{\chi}}\left((x_1 - x_2)^2/x_1^2\,x_2^2;\ \frac{x_1}{x_1^2} - \frac{x_2}{x_2^2}\right). \tag{7.9}$$

The dependence of $\hat{C}^{\bar{\chi}}$ on $x_1/x_1^2 - x_2/x_2^2$ must be used to build up the tensor character of the kernel.

Now we investigate Eq. (2') of (7.6): Obviously, the sum $l_1 + l_2$ must be an integer. For these representations it is known how to construct the representation space with the help of tensor products of M.[21] Thus, only those representations of the principal series contribute to (7.1) which satisfy $l_1 = l_2$, because on l-fold tensor product of $x_1/x_1^2 - x_2/x_2^2$ is already symmetric so that the antisymmetrization with respect to two indices yields zero.

Hence, we obtain that $\hat{C}^{\bar{\chi}}(x_1, x_2)$ looks as follows:

$$\hat{C}^{\bar{\chi}}(x_1, x_2)^{\mu_1\cdots\mu_l} = c^{\bar{\chi}}((x_1 - x_2)^2/x_1^2\,x_2^2) \times\left\{\left(\frac{x_1}{x_1^2} - \frac{x_2}{x_2^2}\right)^{\mu_1}\cdots\left(\frac{x_1}{x_1^2} - \frac{x_2}{x_2^2}\right)^{\mu_l} - \text{traces}\right\}. \tag{7.10}$$

The determination of the scalar factor in (7.10) is easily done with the remaining Eq. (3') of (7.6):

$$c^{\bar{\chi}}((x_1 - x_2)^2/x_1^2\,x_2^2) = c(\bar{\chi})\,[x_1^2 x_2^2/(x_1 - x_2)^2]^{(-2+l+ip+\delta_1+\delta_2)/2}. \tag{7.11}$$

To write down the final result it is convenient to introduce the abbreviation

$$x_{ik} = x_i - x_k\ \ (i, k = 1, 2, 3),\quad \hat{x} = x_{13}/x_{13}^2 - x_{23}/x_{23}^2, \tag{7.12}$$

so that

$$\begin{aligned}
&C^{\bar{\chi}}(x_1, x_2; x_3)^{\mu_1\cdots\mu_l}\\
&= c(\bar{\chi})\,(x_{12}^2)^{(-2+l+ip+\delta_1+\delta_2)/2}(x_{13}^2)^{(-2+l+ip-\delta_1+\delta_2)/2}\\
&\quad (x_{23}^2)^{(-2+l+ip+\delta_1-\delta_2)/2} \times\{\hat{x}^{\mu_1}\cdots\hat{x}^{\mu_l} - \text{traces}\}.
\end{aligned} \tag{7.13}$$

This kernel has the following invariance property:

$$U_1(g)\otimes U_2(g)\otimes U^{\bar{\chi}}(g)\,C^{\bar{\chi}}(x_1, x_2; x_3) = C^{\bar{\chi}}(x_1, x_2; x_3). \tag{7.14}$$

ACKNOWLEDGMENTS

The author would like to thank H. Joos and K. Koller for interesting discussions.

*Present address: Department of Physics, University of Kiel, Germany.

[1] S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962).
[2] T. Hirai, Proc. Japan Acad. 38, 83 (1962).
[3] R. Takahashi, Japan J. Math. 31, 55 (1963).
[4] U. Ottoson, Commun. Math. Phys. 8, 228 (1968).
[5] J. G. Kuriyan, N. Mukunda, and E. C. C. Sudarshan, Commun. Math. Phys. 8, 204 (1968).
[6] G. Mack, "Group theoretical approach to conformal invariant quantum field theory", preprint, Universität Bern, 1973.
[7] T. T. Todorov, in *Recent Developments in Mathematical Physics*, edited by P. Urban (Springer-Verlag, Vienna, 1973).
[8] S. Ferrara, R. Gatto, and A. F. Grillo, *Springer Tracts in Modern Physics* (Springer-Verlag, Berlin, 1973). Vol. 67.
[9] M. Gell-Mann and F. E. Low, Phys. Rev. 95, 1300 (1954).
[10] G. Mack, in *Lecture Notes in Physics*, edited by W. Rühl and A. Vancura (Springer-Verlag, Heidelberg, 1973).
[11] A. M. Polyakov, JETP Lett. 12, 381 (1970).
[12] E. Schreier, Phys. Rev. D 3, 980 (1971).
[13] A. A. Migdal, "4-dimensional soluble models of conformal field theory," preprint, Landau Institute, Chernogolovka, 1972.
[14] K. Symanzik, in *Lectures on High Energy Physics*, edited by B. Jaksic (Gordon and Breach, New York, 1965).
[15] K. Koller, "The significance of conformal inversion in quantum field theory," DESY Preprint 74/8, Hamburg, 1974.
[16] G. Warner, *Harmonic Analysis on Semi-Simple Lie Groups* (Springer-Verlag, Berlin, 1972), Vol. 1.
[17] T. Hirai, Proc. Japan Acad. 38, 258 (1962).
[18] G. Mack and A. Salam, Ann. Phys. (N.Y.) 53, 174 (1969).
[19] M. A. Naimark, *Lineare Darstellungen der Lorentz Gruppe* (VEB, Berlin, 1963).
[20] I. M. Gelfand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5.
[21] H. Weyl, *The Classical Groups* (Princeton U. P., Princeton, 1939).
[22] I. M. Gelfand and G. E. Shilov, *Generalized Functions* (Academic, New York, 1964), Vol. 1.
[23] M. A. Naimark, Am. Math. Soc. Transl. 36, 189 (1964).

Novel identities for simple n-j symbols*

B. I. Dunlap and B. R. Judd

Physics Department, The Johns Hopkins University, Baltimore, Maryland 21218
(Received 11 September 1974)

A number of identities involving 3-j, 6-j, and 9-j symbols in combinations not commonly encountered in angular-momentum theory are established and briefly discussed.

In the course of extending Shibuya and Wulfman's analysis of the hydrogen molecular ion,[1] we have found that certain sums over products of n-j symbols yield surprisingly simple results. By pursuing the implications of the analysis, we have been able to uncover a number of identities that the n-j symbols satisfy and that lie outside the normal scope of angular-momentum theory.[2] The use of the n-j symbols is so widespread in physics that we cannot be sure that some of the identities (particularly special cases of them) are not already known. However, they are sufficiently novel and striking to warrant, in our opinion, a brief listing. A common feature is the absence of the characteristic weighting factor $2x+1$ in sums over x.

We begin with

$$\sum_m (-1)^{a+c+k-M}\begin{pmatrix} a & k & a \\ -m & 0 & m \end{pmatrix}\begin{pmatrix} c & k & c \\ m-M & 0 & M-m \end{pmatrix}$$

$$=\sum_L (-1)^{L+a+c+k}\begin{Bmatrix} a & c & L \\ c & a & k \end{Bmatrix} \tag{1}$$

$$=\frac{1}{2k+1}\left(\frac{(2a-k)!(2c+k+1)!}{(2c-k)!(2a+k+1)!}\right)^{1/2}, \tag{2}$$

where $a-c\geq|M|$. Equation (1) can be proved by summing the diagonal matrix elements of the scalar product of two spherical tensors of rank k, these matrix elements being calculated in two alternative bases: that formed from the coupled states $|(ac)LM\rangle$, and that formed from the uncoupled states $|am,\ c\ M-m\rangle$. Equation (2) is not so easy to derive. Although it has a very simple structure, we had to rely on a detailed combinatorial analysis similar to that described by Racah[3] in order to obtain a proof. The only complication in a straightforward though rather long analysis lies in the occasional need to restrict sums over running indices to smaller ranges than those implied in the various factorial functions. A similar approach can be used to obtain

$$\sum_L (-1)^{L+k+2c}\begin{Bmatrix} a & c & L \\ a & c & k \end{Bmatrix}=\frac{1}{2k+1}\quad (c\leq a). \tag{3}$$

An analog to Eq. (1) is provided by

$$\sum_L (-1)^{2k}\begin{Bmatrix} a & c & L \\ a & c & k \end{Bmatrix}$$

$$=\sum_{m,q}\begin{pmatrix} a & k & c \\ -m & q & M-m \end{pmatrix}^2 \quad (a-c\geq|M|), \tag{4}$$

which can be proved by considering exchange rather than direct matrix elements of the scalar product of two spherical tensors. The sums of Eq. (4) do not appear to have a simple analytic form. When $a=c$, Eqs. (2) and (3) collapse to a common form for which an elementary proof can be given,[4] since the sum over the 3-j symbols in Eq. (1) can now be readily carried out.

The identities above can be used in conjunction with well-established ones to derive further results. Multiplication of both sides of Eq. (3) by

$$(2k+1)\begin{Bmatrix} a & c & k \\ c & a & S \end{Bmatrix}$$

and summation over k yields

$$\sum_L (-1)^{2c-S}\begin{Bmatrix} c & a & L \\ a & c & S \end{Bmatrix}=\sum_k \begin{Bmatrix} c & a & k \\ a & c & S \end{Bmatrix},$$

from which we can deduce

$$\sum_k \begin{Bmatrix} c & a & k \\ a & c & S \end{Bmatrix}=0\quad (c\leq a,\ 2c-S \text{ odd}). \tag{5}$$

In a similar way, products of n-j symbols can be used with Eq. (3) to obtain (for $c\leq a$)

$$\sum_L \begin{Bmatrix} a & c & L \\ b & d & e \end{Bmatrix}\begin{Bmatrix} b & d & L \\ c & a & f \end{Bmatrix}$$

$$=\sum_k (-1)^{b-d+e-f}\begin{Bmatrix} a & c & k \\ e & f & b \end{Bmatrix}\begin{Bmatrix} e & f & k \\ c & a & d \end{Bmatrix} \tag{6}$$

and

$$\sum_L (-1)^L\begin{Bmatrix} a & c & L \\ b & d & e \end{Bmatrix}\begin{Bmatrix} b & d & L \\ a & c & f \end{Bmatrix}$$

$$=\sum_k (-1)^{k+2c}\begin{Bmatrix} k & a & c \\ a & d & f \\ c & e & b \end{Bmatrix}. \tag{7}$$

These equations can be extended by making the interchanges $a\longleftrightarrow b$, $c\longleftrightarrow d$ (for $b\geq d$) or $a\longleftrightarrow d$, $c\longleftrightarrow b$ (for $d\geq b$). All the results of this paragraph can be paralleled by taking Eq. (2) rather than Eq. (3) as the starting point.

Recursion relations and inductive methods can be used to introduce a variety of weighting factors in the sums. Thus, Eq. (22) of Biedenharn, Blatt, and Rose[5] can be combined with Eq. (2) to derive

$$\sum_L (-1)^{L+a+c+k}\frac{1}{L(L+1)}\begin{Bmatrix} a & c & L \\ c & a & k \end{Bmatrix}$$

$$= \frac{1}{a(a+1)-c(c+1)} \left(\frac{(2a-k)!(2c+k+1)!}{(2c-k)!(2a+k+1)!} \right)^{1/2}$$

$$(a > c). \tag{8}$$

It is interesting to notice that such equations as this are invariant with respect to the transformation $k \to -k-1$ provided the interchange $a \longleftrightarrow c$ is also made. [See Eq. (22.19) of Jucys and Bandzaitis.[6]]

It is well known that 9-j symbols for which one of the arguments is unity can be frequently reduced to a simpler form. If we use Eq. (25.37) of Jucys and Bandzaitis,[6] together with Eq. (8), we find

$$\sum_L \left(\frac{2L+1}{L(L+1)} \right)^{1/2} \begin{Bmatrix} a & a & k+1 \\ c & c & k \\ L & L & 1 \end{Bmatrix}$$

$$= \left(\frac{(2a-k-1)!(2c+k+1)!(4k+4)}{(2c-k)!(2a+k+2)!(2k+1)(2k+3)} \right)^{1/2}$$

$$(a > c) \tag{9}$$

$$= 0 \quad (a < c). \tag{10}$$

It was the empirical discovery of Eq. (10) that led us to uncover the existence of the other unusual identities described in this article.

*Work supported in part by the National Science Foundation.

[1]T. Shibuya and C. E. Wulfman, Proc. Roy. Soc. (London) A **286**, 376 (1965).
[2]A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U. P., Princeton, N.J., 1960).
[3]G. Racah, Phys. Rev. **62**, 438 (1942), Appendix B.
[4]B. R. Judd, *Angular Momentum Theory for Diatomic Molecules* (Academic, New York, to be published).
[5]L. C. Biedenharn, J. M. Blatt, and M. E. Rose, Rev. Mod. Phys. **24**, 249 (1952).
[6]A. Jucys and A. Bandzaitis, *Teoriya Momenta Kolichestva Dvizheniya v Kvantovoi Mekhanike* (Mintis, Vilnius, Lithuania, 1965).

A correction to the Sen and Dunn gravitational field equations

J. S. Jeavons

Department of Mathematics, Caulfield Institute of Technology, Caulfield East, Victoria, 3145, Australia

C. B. G. McIntosh

Department of Mathematics, Monash University, Clayton, Victoria, 3168, Australia

D. K. Sen

Department of Mathematics, University of Toronto, Canada, M5S 1A1
(Received 9 October 1974)

It is shown that the gravitational field equations proposed by Sen and Dunn do not follow from their variational principle and that in the vacuum case the correct field equations are related to Einstein's field equations by a conformal mapping.

Sen and Dunn[1] proposed a new scalar—tensor theory of gravitation based on a Lyra manifold rather than a Riemannian manifold. We show in this note that this new theory is related to Einstein's theory of gravitation in a manner similar to the way in which the conformally mapped set of Einstein's equations are related to the Einstein equations.

Since the publication of Einstein's theory of gravitation, attempts have been made at unified field theories. It was clear that such a theory would require a generalization of the usual Riemannian space time. The first attempt was by Weyl[2]; however, the theory was unacceptable due to the feature of nonintegrability of length transfer contained in the theory.

Lyra[3] put forward another modification of the Riemannian manifold and a summary of field theories based on the Lyra manifold has been given by Sen.[4] In this paper we are only interested in the gravitational theory based on the Lyra manifold.

Details of the basis of the Lyra manifold can be found in the paper by Sen and Dunn,[1] in the references given there and in a later paper by Sen and Vanstone.[5] We give here only an outline of the main features of a Lyra manifold. The displacement vector PP' between two neighboring points $P(x^\mu)$ and $P'(x^\mu + dx^\mu)$ has components $\xi^\mu = x^0 dx^\mu$, where $x^0(x^\mu)$ is a nonzero gauge function. The coordinate system x^μ together with x^0 form a reference system $(x^0; x^\mu)$. Tensors are characterized by the way components transform under a general transformation of reference system. The metric is the absolute invariant (that is, invariant under change of reference system)

$$ds^2 = g_{\mu\nu} x^0 dx^\mu x^0 dx^\nu.$$

The components of the affine connection are no longer symmetric in the lower indices and cannot be identified with the Christoffel symbols as is the case in Riemannian geometry. In Lyra geometry the components of the affine connection contain not only the Christoffel symbols, but also a function ϕ_α, which arises as a consequence of the introduction of a gauge function. The Lyra curvature tensor $K^\mu_{\kappa\alpha\beta}$ can be obtained in a similar manner to the Riemannian case together with the contracted curvature tensor $K_{\kappa\alpha}$, and the curvature scalar K.

Adopting the conventions in the Sen and Dunn[1] paper, the field equations are derived from the variational principle

$$\delta \int (K\sqrt{-g} + L_m \sqrt{-g})\, x^0\, dx^1 \cdots x^0\, dx^4 = 0, \tag{1}$$

where the curvature scalar K is given by

$$K = R/(x^0)^2 + (3/x^0)\,\phi^\alpha_{;\alpha} + \tfrac{3}{2}\phi_\alpha \phi^\alpha + \tfrac{3}{2}\mathring{\phi}_\alpha \phi^\alpha, \tag{2}$$

where R is the Riemannian curvature scalar and $\mathring{\phi}_\alpha$ is defined by

$$\mathring{\phi}_\alpha = (x^0)^{-1}[\ln(x^0)^2]_{;\alpha}. \tag{3}$$

The matter Lagrangian L_m is chosen such that under change of gauge $x^0 \to x^{0\prime}$,

$$L_m \to L'_m. \tag{4}$$

The variation in (1) is to be carried out with respect to $g_{\mu\nu}$ and ϕ_α. Variation with respect to $g_{\mu\nu}$ gives

$$\begin{aligned}\int [&(x^0)^2(R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R) \\ &+ (x^0)^3(\mathring{\phi}_{\mu;\nu} - g_{\mu\nu}\Box\mathring{\phi}) + \tfrac{3}{2}(x^0)^4(\mathring{\phi}_\mu\mathring{\phi}_\nu - \tfrac{1}{2}g_{\mu\nu}\mathring{\phi}_\alpha\mathring{\phi}^\alpha) \\ &+ \tfrac{3}{2}(x^0)^4(\phi_\mu\phi_\nu - \tfrac{1}{2}g_{\mu\nu}\phi_\alpha\phi^\alpha) - 3(x^0)^4(\mathring{\phi}_\mu\phi_\nu - \tfrac{1}{2}g_{\mu\nu}\phi_\alpha\mathring{\phi}^\alpha) \\ &+ 8\pi G T_{\mu\nu}(x^0)^4]\sqrt{-g}\,\delta g^{\mu\nu}\, d^4x = 0.\end{aligned} \tag{5}$$

Variation with respect to ϕ_μ gives

$$\int 3(x^0)^4(\phi^\mu - \mathring{\phi}^\mu)\,\delta\phi_\mu \sqrt{-g}\, d^4x = 0, \tag{6}$$

which immediately gives

$$\phi^\mu = \mathring{\phi}^\mu. \tag{7}$$

Equation (7), together with (3), enables us to write the field equation resulting from (5) as

$$\begin{aligned}&R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R + x^0(\phi_{\mu;\nu} - g_{\mu\nu}\Box\phi) \\ &\quad - \tfrac{3}{4}g_{\mu\nu}(x^0)^2\phi_\alpha\phi^\alpha = -8\pi G(x^0)^2 T_{\mu\nu},\end{aligned} \tag{8a}$$

while Eqs. (7) and (3) give the relation

$$\phi_{\mu,\nu} = \phi_{\nu,\mu}. \tag{8b}$$

Equation (5) differs from the corresponding relation in the Sen and Dunn paper[1] due to the fact that Sen and Dunn have neglected the contribution from the term $\int (x^0)^3(\phi_{\mu;\nu} - g_{\mu\nu}\Box\phi)\sqrt{-g}\,\delta g^{\mu\nu}\, d^4x$. This terms arises from the presence of the gauge function in the term

$\int (x^0)^2\delta(R\sqrt{-g})\,d^4x$ in an analogous way to the corresponding term in the theory proposed by Brans and Dicke.[6] The relation (7) also differs from the corresponding result by Sen and Dunn because Sen and Dunn appear to have neglected the contribution from the term $3\int (x^0)^3\,\delta((\sqrt{-g}\,\phi_\mu g^{\mu\nu})_{,\nu})\,d^4x$ when carrying out the variation with respect to ϕ_μ. The condition (7) ensures that autoparallels in Lyra geometry can be identified with the geodesics of Lyra geometry. The field equations proposed by Sen and Dunn are not invariant under change of gauge $x^0 \to x^{0\prime}$, whereas Eqs. (8) proposed here are invariant under such gauge transformations.[7]

The similarity between the field equations proposed here and the conformally mapped Einstein equations can be seen by writing the equation obtained from (5) in the form

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R + (2/x^0)(x^0_{,\mu;\nu} - g_{\mu\nu}\Box x^0) - [4/(x^0)^2](x^0_{,\mu}x^0_{,\nu} - \tfrac{1}{4}g_{\mu\nu}x^0_{,\alpha}x^{0,\alpha}) = -8\pi G(x^0)^2T_{\mu\nu}, \tag{9}$$

where (3) has been used to rewrite terms in $\overset{\circ}{\phi}$ in terms of x^0, and then substituting $(x^0)^2 = \phi$ so that (9) gives

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R + \phi^{-1}(\phi_{\mu;\nu} - g_{\mu\nu}\Box\phi) - \tfrac{3}{2}\phi^{-2}(\phi_\mu\phi_\nu - \tfrac{1}{2}g_{\mu\nu}\phi_\alpha\phi^\alpha) = -8\pi G\phi T_{\mu\nu}. \tag{10}$$

In the vacuum case, Eq. (10) is algebraically identical to the conformally mapped Einstein equations, which can be obtained by applying the mapping

$$g_{\mu\nu} \to \phi\bar{g}_{\mu\nu}$$

to the Einstein equations. Thus the field equations (9) can be conformally mapped in the vacuum case to

$$R_{\mu\nu} = 0,$$

which are the Einstein vacuum field equations.

In the nonvacuum case Eq. (10) differs from the conformally mapped Einstein equations because in the conformally mapped Einstein equations the term involving the stress–energy tensor has the form $(-8\pi G/\phi)T_{\mu\nu}$, where in both cases the components $T_{\mu\nu}$ are independent of ϕ.

In conclusion we should point out that the original field equations proposed by Sen and Dunn may still prove to be heuristically useful even though they are not derivable from the usual variational principle.

[1]D. K. Sen and K. A. Dunn, J. Math. Phys. **12**, 578 (1971).
[2]H. Weyl, S.-B. Preuss, Akad. Wiss., Berlin, 465 (1918).
[3]G. Lyra, Math. Z. **54**, 52 (1951); E. Scheibe, Math. Z. **57**, 65 (1952).
[4]D. K. Sen, *Fields and/or Particles* (Academic, London, 1968).
[5]D. K. Sen and J. R. Vanstone, J. Math. Phys. **13**, 990 (1972).
[6]C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1961).

[7]The authors are grateful to Professor R. Kerr for pointing out that the field equations of Sen and Dunn are not invariant under gauge change.

Some properties of asymptotically flat, static, Einstein–Maxwell space–times

William Hallidy

California State College, California, Pennsylvania 15419
(Received 9 July 1974; revised manuscript received 30 September 1974)

We show that the generators of the Weyl tensor for asymptotically flat, static, Einstein–Maxwell space–times are electric type to asymptotic order r^{-7} and that for space–times which are stable under static perturbations the generators of the electromagnetic field tensor are electric type to order r^{-5}. We also argue that the Weyl tensor, the Ricci tensor, and the electromagnetic field tensor may fail to be electric type in general for asymptotically flat, static, Einstein–Maxwell space–times but that the generators of the Weyl tensor and the electromagnetic field tensor may still be electric type in general. The variable r is an affine parameter along null geodesics.

I. INTRODUCTION

Stationary space–times, space–times possessing a timelike Killing vector field, are certainly of current interest in general relativity. A subset of these stationary space–times, the static space–times, may be distinguished by the additional requirement that the Killing field be hypersurface orthogonal. These geometrical conditions can be used to pick a preferred coordinate system for static space–times in which the metric takes an especially simple form.

In general relativity the gravitational field is embodied in the Riemannian curvature of the space–time expressed in terms of the Riemann tensor. The Riemann tensor, which is in principle directly accessible to physical measurement, is canonically decomposable into two other tensors, the Weyl and Ricci tensors. The Ricci tensor is nonzero only in the presence of matter fields. In a formalism developed by Newman and Penrose[1] these two tensors play a preponderant role. Using an asymptotic form of this formalism which was developed by Newman and Unti,[2] Newman and Penrose were able to write the conditions for a vacuum space–time to be stationary as conditions on the Weyl tensor.[3] This was later extended to Einstein–Maxwell space–times in a paper by Exton, Newman, and Penrose in which conditions for a space–time to be stationary were elucidated for both the Weyl tensor and the electromagnetic field tensor.[4] Although it is clear that the region of space–time near the earth is not stationary, it is still satisfying to see that the conditions for stationarity can be interpreted as conditions on directly physically observable quantities.

The words stationary and static are regularly applied in the context of electromagnetism in flat space–time to denote certain conditions holding on the fields and their sources. To be stationary there must be a coordinate system in which the fields and their sources are time independent; to be static they must satisfy the additional requirement that there be an isometry under time reversal in this coordinate system. In particular, if we consider the fields to be generated only by the motion of structureless charge distributions (ignoring the spin and other higher intrinsic moments of the charged particles making up the distribution), then we find that to be stationary the fields must be time independent in some coordinate system and to be static there must be a coordinate system in which only time independent electric fields are present. The Newman–Penrose formalism may be applied in flat space–time to the electromagnetic field. It can be shown that using the differential operator edth (the symbol for which is $\eth$) defined by Newman and Penrose,[5] the electromagnetic field components can be characterized as electric type in a given coordinate system if and only if there is no magnetic field present in the coordinate system, with a similar statement holding for magnetic type fields. Therefore, in flat space–time the electromagnetic field for structureless charge distributions is static if and only if there is a coordinate system in which the Newman–Penrose components of the field are time independent and electric type.

As the analogy between electromagnetism and gravitation is made very clear by the use of the Newman–Penrose formalism, it was hoped that a distinction similar to that holding for electromagnetic fields would be found to hold for stationary and static gravitational fields. In an earlier paper by the author[6] this type of distinction was found to hold to a certain asymptotic order for vacuum space–times. This result was in agreement with the long standing assumption that the Weyl tensor must be electric type for any asymptotically flat, static, vacuum space–time. It is natural to ask whether a similar result will hold in the presence of matter fields, in particular, in the presence of electromagnetic fields. In this case we might ask whether, in addition to the Weyl tensor, the Ricci tensor and the electromagnetic field tensor are asymptotically electric type. It would certainly be nice if we could find such a characterization of static space–times since then both stationary and static space–times would be characterizable in terms of their physical fields rather than in terms of integrals of those fields. Furthermore, it would make the characterization of stationary and static space–times easier in the Newman–Penrose formalism since the metric tensor plays only a subsidiary role in this formalism. In this paper we attempt to answer the above questions.

As the results of this paper depend on the Newman–Penrose formalism, we review in Sec. II the conditions for an asymptotically flat Einstein–Maxwell space–time to be stationary, restricting the investigation to space–times having an asymptotically timelike Killing

vector.[7] In Sec. III we impose the conditions for the space–time to be static in the Newman–Penrose formalism and find that these conditions are interpretable as conditions on the Weyl tensor and the electromagnetic field tensor. In Sec. IV we present our conclusions. In an appendix we include some properties of spin-weighted functions which are useful in deriving the results of this paper.

II. STATIONARY CONDITIONS

In the Newman–Penrose formalism we can express the Weyl tensor components in terms of five complex functions ψ_A $(A=0,\ldots,4)$, the electromagnetic field components in terms of three complex functions ϕ_a $(a=0,1,2)$, and the Ricci tensor components as $\Phi_{ab}=\phi_a\bar{\phi}_b$, where we have set the coupling constant equal to one. In a Newman–Penrose null coordinate system the metric tensor components take the form[2]

$$g^{0\mu}=\delta^\mu_1,\quad g^{1k}=X^k-(\xi^k\bar{\omega}+\bar{\xi}^k\omega)$$
$$g^{11}=2(U-\omega\bar{\omega}),\quad g^{kl}=-(\xi^k\xi^l+\bar{\xi}^k\bar{\xi}^l), \tag{1}$$

where $k,l=(2,3)$ and $(x^0,x^1)=(u,r)$ in which u labels outgoing null hypersurfaces and r is an affine parameter along null geodesics on the hypersurfaces. For a stationary space–time in Newman–Penrose formalism we can choose our null coordinate system such that the Killing vector is given by $\partial/\partial u$. Then in this coordinate system all physical fields are independent of u (u is essentially the retarded time). We can further specialize the coordinate system[3,4] so that if we assume that[8]

$$\psi_0=\psi_0^0r^{-5}+\psi_0^1r^{-6}+O(r^{-7})$$

and

$$\phi_0=\phi_0^0r^{-3}+\phi_0^1r^{-4}+O(r^{-5}),$$

then the Weyl tensor, the metric tensor functions, the electromagnetic field tensor and certain components of the Ricci tensor can be written[5]

$$\psi_1=\psi_1^0r^{-4}+\psi_1^1r^{-5}+\psi_1^2r^{-6}+O(r^{-7}),$$
$$\psi_1^1=\bar{\eth}\psi_0^0+3\phi_0^0\bar{\phi}_1^0,\quad \psi_1^2=\tfrac{1}{2}(\bar{\eth}\psi_0^1+4\phi_0^1\bar{\phi}_1^0+3\phi_0^0\bar{\eth}\bar{\phi}_0^0), \tag{2a}$$
$$\psi_2=\psi_2^0r^{-3}+\psi_2^1r^{-4}+\psi_2^2r^{-5}+\psi_2^3r^{-6}+O(r^{-7}),$$
$$\psi_2^1=\bar{\eth}\psi_1^0+2\phi_1^0\bar{\phi}_1^0,$$
$$\psi_2^2=\tfrac{1}{2}[\bar{\eth}\psi_1^1+\bar{\phi}_1^0\bar{\eth}\phi_0^0+2(\phi_1^0\bar{\phi}_1^1+\bar{\phi}_1^0\phi_1^1)],$$
$$\psi_2^3=\tfrac{1}{3}[-2\psi_1^0\bar{\psi}_1^0+\bar{\eth}\psi_1^2+3\phi_0^0\bar{\phi}_0^0+\bar{\phi}_1^1\bar{\eth}\phi_0^0+\bar{\phi}_1^0\bar{\eth}\phi_0^1$$
$$+2(\bar{\phi}_1^0\phi_1^2+\bar{\phi}_1^2\phi_1^0+\phi_1^1\bar{\phi}_1^1)], \tag{2b}$$
$$\psi_3=\psi_3^2r^{-4}+\psi_3^3r^{-5}+\psi_3^4r^{-6}+O(r^{-7}),$$
$$\psi_3^2=\tfrac{1}{2}\bar{\eth}\psi_2^1,\quad \psi_3^3=\tfrac{1}{3}[3\bar{\omega}_0\psi_2^0+\bar{\eth}\psi_2^2+3\bar{\phi}_1^0\phi_2^2+2\bar{\phi}_0^0\phi_1^0],$$
$$\psi_3^4=\tfrac{1}{4}[4\bar{\omega}_0\psi_2^1+3\bar{\omega}_1\psi_2^0+\bar{\eth}\psi_2^3-\bar{\phi}_0^0\bar{\eth}\phi_2^2+3\bar{\phi}_1^1\phi_2^2$$
$$+4\bar{\phi}_1^0\phi_2^2+2\bar{\phi}_0^0\phi_1^1+2\bar{\phi}_0^1\phi_1^0+2\psi_2^0\bar{\phi}_0^0\phi_1^0], \tag{2c}$$
$$\psi_4=\psi_4^4r^{-5}+\psi_4^5r^{-6}+O(r^{-7}),$$
$$\psi_4^4=\tfrac{1}{4}(\bar{\eth}\psi_3^3+\bar{\phi}_1^0\bar{\eth}\phi_2^2),$$
$$\psi_4^5=\tfrac{1}{5}(\bar{\eth}\psi_3^4+4\bar{\omega}_0\psi_3^2-\tfrac{1}{2}\bar{\psi}_0^0\psi_2^0+3\bar{\phi}_0^0\phi_2^2+\bar{\phi}_1^0\bar{\eth}\phi_2^3+\bar{\phi}_1^1\bar{\eth}\phi_2^2), \tag{2d}$$
$$U=-1+U_1r^{-1}+U_2r^{-2}+U_3r^{-3}+U_4r^{-4}+O(r^{-5}),$$
$$U_1=-\psi_2^0,\quad U_2=-\tfrac{1}{6}(\psi_2^1+\bar{\psi}_2^1+2\phi_1^0\bar{\phi}_1^0),$$
$$U_3=-\tfrac{1}{12}[\psi_2^2+\bar{\psi}_2^2+2(\phi_1^0\bar{\phi}_1^1+\bar{\phi}_1^0\phi_1^1)],$$
$$U_4=-\tfrac{1}{20}[\psi_2^3+\bar{\psi}_2^3+12\omega_0\bar{\omega}_0+2(\phi_1^0\bar{\phi}_1^2+\bar{\phi}_1^0\phi_1^2+\phi_1^1\bar{\phi}_1^1)], \tag{3a}$$
$$\omega=\omega_0r^{-2}+\omega_1r^{-3}+\omega_2r^{-4}+O(r^{-5}),$$
$$\omega_0=-\tfrac{1}{2}\psi_1^0,\quad \omega_1=-\tfrac{1}{6}(\bar{\eth}\psi_0^0+4\phi_0^0\bar{\phi}_1^0),$$
$$\omega_2=-\tfrac{1}{24}(\bar{\eth}\psi_0^1+6\phi_0^1\bar{\phi}_1^0+5\phi_0^0\bar{\eth}\bar{\phi}_0^0), \tag{3b}$$
$$X^k=X_0^kr^{-3}+X_1^kr^{-4}+X_2^kr^{-5}+O(r^{-6}),$$
$$X_0^k=-\tfrac{1}{3}(\omega_0\bar{\xi}^{0k}+\bar{\omega}_0\xi^{0k}),\quad X_1^k=-\tfrac{1}{2}(\omega_1\bar{\xi}^{0k}+\bar{\omega}_1\xi^{0k}),$$
$$X_2^k=-\tfrac{3}{5}(\omega_2\bar{\xi}^{0k}+\bar{\omega}_2\xi^{0k}), \tag{3c}$$
$$\xi^k=\xi^{0k}r^{-1}+\tfrac{1}{6}\psi_0^0\bar{\xi}^{0k}r^{-4}+\tfrac{1}{12}(\psi_0^1\bar{\xi}^{0k}+\phi_0^0\bar{\phi}_0^0\xi^{0k})r^{-5}+O(r^{-6}), \tag{3d}$$
$$\phi_1=\phi_1^0r^{-2}+\phi_1^1r^{-3}+\phi_1^2r^{-4}+O(r^{-5}),$$
$$\phi_1^1=\bar{\eth}\phi_0^0,\quad \phi_1^2=\tfrac{1}{2}\bar{\eth}\phi_0^1, \tag{4a}$$
$$\phi_2=\phi_2^2r^{-3}+\phi_2^3r^{-4}+O(r^{-5}),$$
$$\phi_2^2=\tfrac{1}{2}\bar{\eth}\phi_1^1,\quad \phi_2^3=\tfrac{1}{3}(\bar{\eth}\phi_1^2-\phi_1^0\bar{\psi}_1^0), \tag{4b}$$
$$\Phi_{01}=\phi_0^0\bar{\phi}_1^0r^{-5}+(\phi_0^0\bar{\phi}_1^1+\bar{\phi}_1^0\phi_0^1)r^{-6}+O(r^{-7}), \tag{5a}$$
$$\Phi_{02}=\phi_0^0\bar{\phi}_2^2r^{-6}+(\phi_0^0\bar{\phi}_2^3+\phi_0^1\bar{\phi}_2^2)r^{-7}+O(r^{-8}), \tag{5b}$$
$$\Phi_{12}=\phi_1^0\bar{\phi}_2^2r^{-5}+(\phi_1^0\bar{\phi}_2^3+\phi_1^1\bar{\phi}_2^2)r^{-6}+O(r^{-7}), \tag{5c}$$

with $\xi^{02}=-i\xi^{03}=P$ and $P=\frac{1}{2}(1+\zeta\bar{\zeta})$, where $\zeta=-x^2+ix^3$ and $\lim_{r\to\infty}r^2g^{kl}=-2P^2\delta^{kl}$. We may then write the metric components as $g^{0\mu}=\delta^\mu_1$, $g_{1\mu}=\delta^0_\mu$,

$$g^{11}=2U-2\omega_0\bar{\omega}_0r^{-4}+O(r^{-5}), \tag{6a}$$
$$g^{1k}=X^k-r^{-1}(\xi^{0k}\bar{\omega}+\bar{\xi}^{0k}\omega)+O(r^{-6}), \tag{6b}$$
$$g^{kl}=-2P^2\delta^{kl}r^{-2}-\tfrac{1}{6}(\psi_0^0\bar{\xi}^{0k}\bar{\xi}^{0l}+\bar{\psi}_0^0\xi^{0k}\xi^{0l})r^{-5}$$
$$-\tfrac{1}{12}(\psi_0^1\bar{\xi}^{0k}\bar{\xi}^{0l}+\bar{\psi}_0^1\xi^{0k}\xi^{0l}+2P^2\phi_0^0\bar{\phi}_0^0\delta^{kl})r^{-6}+O(r^{-7}), \tag{6c}$$
$$g_{00}=2-2U_1r^{-1}-2U_2r^{-2}-2U_3r^{-3}$$
$$-(2U_4-2\omega_0\bar{\omega}_0+g_{0k}^{(0)}X^{0k})r^{-4}+O(r^{-5}), \tag{7a}$$
$$g_{0k}=-P^{-2}[\tfrac{2}{3}(\bar{\omega}_0\xi^{0k}+\omega_0\bar{\xi}^{0k})r^{-1}+\tfrac{3}{4}(\bar{\omega}_1\xi^{0k}+\omega_1\bar{\xi}^{0k})r^{-2}$$
$$+\tfrac{4}{5}(\bar{\omega}_2\xi^{0k}+\omega_2\bar{\xi}^{0k})r^{-3}]+O(r^{-4}), \tag{7b}$$
$$g_{kl}=-\tfrac{1}{2}P^{-2}\delta_{kl}r^2+\tfrac{1}{24}P^{-4}(\psi_0^0\bar{\xi}^{0k}\bar{\xi}^{0l}+\bar{\psi}_0^0\xi^{0k}\xi^{0l})r^{-1}$$
$$+\tfrac{1}{48}P^{-4}(\psi_0^1\bar{\xi}^{0k}\bar{\xi}^{0l}+\bar{\psi}_0^1\xi^{0k}\xi^{0l}+2P^2\phi_0^0\bar{\phi}_0^0\delta_{kl})r^{-2}+O(r^{-3}), \tag{7c}$$

where $g_{0k}^{(0)}$ is the coefficient of the r^{-1} term in g_{0k}. We now find that the remaining conditions on the Weyl tensor and the electromagnetic field components imposed by the field equations and the condition that the space–time be stationary are given by[3,4]

$$\psi_2^0=\bar{\psi}_2^0=\text{const},\quad \psi_1^0=\sum_{m=-1}^{1}a_{m1}Y_{1m},$$
$$\psi_0^0=\sum_{m=-2}^{2}b_{m2}Y_{2m},\quad \psi_0^1=\sum_{m=-3}^{3}c_{m2}Y_{3m}+\tfrac{5}{3}[(\psi_1^0)^2-\tfrac{3}{2}\psi_0^0\psi_2^0]$$
$$-\bar{\phi}_1^0\bar{\eth}\phi_1^0, \tag{8}$$
$$\phi_1^0=\text{const},\quad \phi_0^0=\sum_{m=-1}^{1}d_{m1}Y_{1m},$$
$$\phi_0^1=\sum_{m=-2}^{2}e_{m1}Y_{2m}+\{\phi_1^0\psi_1^0-\tfrac{3}{2}\psi_2^0\phi_0^0\}, \tag{9}$$

where a_m, b_m, c_m, d_m, and e_m are constants. Notice that

there is no requirement that $\phi_1^0 = \bar{\phi}_1^0$. The quantities in square brackets are the Weyl conserved quantities and those in curly brackets are the Maxwell conserved quantities.[4]

III. STATIC CONDITIONS

In our coordinate system we may write the condition for the metric to be static as[6]

$$g_{00,1}g_{02} - g_{02,1}g_{00} - g_{00,2} = 0, \tag{10a}$$

$$g_{00,1}g_{03} - g_{03,1}g_{00} - g_{00,3} = 0, \tag{10b}$$

$$g_{00,2}g_{03} - g_{00,3}g_{02} + (g_{02,3} - g_{03,2})g_{00} = 0, \tag{10c}$$

$$g_{03,1}g_{02} - g_{02,1}g_{03} + g_{02,3} - g_{03,2} = 0, \tag{10d}$$

where only three of these equations are actually independent. We use Eqs. (7) to expand Eqs. (10), asymptotically equating the coefficient of each power of r^{-1} separately to zero. Equation (10d) is the easiest to write out. The coefficients of the r^{-1}, r^{-2}, r^{-3} terms of this equation yield the following new information about the Weyl and electromagnetic fields:

$$\bar{\eth}\psi_1^0 = \eth\bar{\psi}_1^0, \quad \bar{\eth}^2\psi_0^0 = \eth^2\bar{\psi}_0^0, \quad \bar{\eth}^2\psi_0^1 = \eth^2\bar{\psi}_0^1, \tag{11}$$

$$\bar{\eth}(\phi_0^0\bar{\phi}_1^0) = \eth(\bar{\phi}_0^0\phi_1^0), \quad \bar{\eth}(\phi_0^1\bar{\phi}_1^0) = \eth(\bar{\phi}_0^1\phi_1^0). \tag{12}$$

For the Weyl field Eqs. (11) are just the conditions that ψ_1^0, ψ_0^0, and ψ_0^1 be electric type.[5] For the electromagnetic field, from Eqs. (12) we find that $\phi_0^0\bar{\phi}_1^0$ and $\phi_0^1\bar{\phi}_1^0$ are electric type. We may only infer from Eqs. (12) that ϕ_0^0 and ϕ_0^1 are electric type if $\phi_1^0 = \bar{\phi}_1^0 \neq 0$. Physically, $\phi_1^0 = \bar{\phi}_1^0 \neq 0$ implies that an electric monopole is present but that no magnetic monopole is present. (The conditions on the Weyl field hold independently of any condition on ϕ_1^0.) After a tedious calculation we find that the r^{-4} part of Eqs. (10a, b) yield the equation

$$2\bar{\phi}_0^0\bar{\eth}\phi_0^0 + (\eth\bar{\phi}_0^0)(\bar{\eth}^2\phi_0^0) = 0, \tag{13}$$

and its complex conjugate. If $\phi_1^0 \neq 0$, then we can multiply Eq. (13) by $\phi_1^0\bar{\phi}_1^0$ and apply Eqs. (9) and (12) to the resulting expression to yield an identity, but if $\phi_1^0 = 0$, then this is the only nontrivial static condition imposed on the electromagnetic field to the order we have investigated in r^{-1}. Using Eq. (9) for ϕ_0^0 we find that Eq. (13) imposes the following conditions on the coefficients in the expansion of ϕ_0^0:

$$d_{-1}\bar{d}_{-1} = d_1\bar{d}_1$$

and

$$d_0\bar{d}_{-1} = -d_1\bar{d}_0.$$

We may express the conditions that ϕ_0^0 be electric type as

$$d_1 = -\bar{d}_{-1} \quad \text{and} \quad d_0 = \bar{d}_0,$$

so that the conditions imposed on ϕ_0^0 by Eq. (13) are weaker than the condition that ϕ_0^0 be electric type.

If we apply the above conditions to the coefficient of each power of the asymptotic expansion of the Weyl tensor in Eqs. (2) and employ some of the results from the Appendix, we then find that the Weyl tensor is electric type to order r^{-7} in the asymptotic expansion.

We know that static space–times possess the discrete symmetry of time reversal invariance. We extend this symmetry to $\mathcal{I}$ (null infinity) by requiring that the Maxwell constants and the Weyl constants defined on $\mathcal{I}^-$ be equal to their counterparts defined on $\mathcal{I}^+$. The Maxwell constants and the Weyl constants are given on $\mathcal{I}^-$ respectively by:

$$-\tfrac{1}{2}\bar{\phi}_1^0\eth^2\bar{\psi}_1^0 + \tfrac{3}{4}\psi_2^0\eth^2\bar{\phi}_0^0$$

and

$$\tfrac{1}{4}(\eth^2\psi_1^0)^2 - \tfrac{1}{16}\psi_2^0\eth^4\bar{\psi}_0^0,$$

as can be seen to follow from Eqs. (2) and (4) with a little work. Using the conditions imposed above on ψ_1^0 and ψ_0^0, we can see immediately that the Weyl constants are equal on $\mathcal{I}^-$ and $\mathcal{I}^+$, but although for the Maxwell constants accidental equality may be possible for a particular space–time, this quality will be stable under static perturbations only if $\phi_1^0 = \bar{\phi}_1^0$ and $\eth^2\bar{\phi}_0^0 = -2\phi_0^0$ which implies here that $\eth\bar{\phi}_0^0 = \bar{\eth}\phi_0^0$. Therefore, we see that the static time reversal isometry together with the additional condition of static stability implies that ϕ_1^0 and ϕ_0^0 must be electric type. We may then return to the Eq. (12) for ϕ_0^1 and use these conditions to infer that ϕ_0^1 is electric. Applying these conditions to Eqs. (4) for the electromagnetic field, we find that the electromagnetic field will be electric type to order r^{-5} in the asymptotic expansion.

Finally, we apply our conditions, including the condition of static stability, to the Ricci tensor components. The components Φ_{00}, Φ_{11}, and Φ_{22} clearly must be electric type because they are spin weight zero, real functions. Thus we only need to look at Eqs. (5) for the components Φ_{01}, Φ_{02}, and Φ_{12}. With a little calculation we find that the coefficient of the r^{-7} terms in Φ_{02} will not in general be electric type unless further conditions on the variables arising at higher order in Eqs. (10) than we have investigated here suitably restrict them. Notice that each of these terms is essentially the product of two electric terms of spin weight one, but that $\eth\phi_0^1 \neq 0$ in general so that it does not satisfy the criteria of Statement A4. A situation for which these terms would not be electric can be found by considering for example ϕ_0^0 not to be axially symmetric and ϕ_0^1 to be axially symmetric but such that $\eth\phi_0^1 \neq 0$. The Ricci tensor will be electric type to the order investigated if ϕ_0^1 contains no quadrupole part, because then $\eth\phi_0^1 = 0$, or if both ϕ_0^0 and ϕ_0^1 are axially symmetric. Other conditions may also be possible for which Φ_{02} will be electric type to the order r^{-8}.

Although we have found that the Weyl tensor is electric type to the order r^{-7}, the result on the Ricci tensor tended to make us less optimistic about an exact result holding even for vacuum space–times. If we inspect the way in which the asymptotic terms entered the Weyl tensor, we find that terms of the type occurring in the equation for Φ_{02} can be present only at higher asymptotic orders in the Weyl tensor than we investigated here.

The author has recently seen a paper by T. W. J. Unti[9] in which a computer program was used to write out the Weyl tensor components for vacuum space–times to a higher order than was done here. In particular, the r^{-7}

coefficient for ψ_1 was written out explicitly. This term contains an expression of the form $\psi_1^0\psi_0^0$ which will not in general be electric type unless $\psi_1^0 = \overline{\psi}_1^0$ or $\psi_0^0 = 0$. Although it is possible to make $\psi_1^0 = \overline{\psi}_1^0$ by a translation of the spatial coordinates, this reduces the remaining coordinate freedom to the rotation group; therefore similar expressions occurring at higher asymptotic order could not be made electric type. Thus it is plausible to infer that the Weyl tensor will *not* in general be electric type for vacuum or Einstein–Maxwell space–times.

IV. CONCLUSIONS

To recapitulate the major results which were presented in the body of the paper, we have seen that the condition that an asymptotically flat, Einstein–Maxwell space–time be static imposes the asymptotic conditions in the Newman–Penrose formalism that ψ_2^0, ψ_1^0, ψ_0^0, ψ_0^1 and $\bar{\phi}_1^0\phi_0^0$, $\bar{\phi}_1^0\phi_0^1$ be time independent and electric type. If no further conditions are added to this, then we may not assert that ϕ_0^0 and ϕ_0^1 are electric type unless we happen to have $\phi_1^0 = \bar{\phi}_1^0 \neq 0$. If $\phi_1^0 = 0$, then we find that ϕ_0^0 is constrained only by a weaker condition than that it be electric type and that no static conditions are imposed on ϕ_0^1 to the asymptotic order investigated here. If, however, we add the condition that our space–time be stable under static perturbation, we find that the equality of the electromagnetic Newman–Penrose constants on $\mathcal{I}^+$ and $\mathcal{I}^-$ which is required by the time reversal isometry which is always defined for static space–times implies that we must have $\phi_1^0 = \bar{\phi}_1^0$ and that ϕ_0^0 and ϕ_0^1 must be electric type.

We have further found that, although the electromagnetic field tensor is electric type to the order investigated here, the Ricci tensor and the Weyl tensor are *not* expected to be electric type in general. In fact, the terms which cause the Ricci tensor and Weyl tensor to be nonelectric type arise from the nonlinearity of the theory, although at a relatively high asymptotic order; therefore, it might also be expected that a sufficiently large asymptotic expansion of the electromagnetic field tensor would expose terms which would make it, too, manifestly nonelectric type.

It thus appears that the physical fields will not in general be electric type in static, Einstein–Maxwell space–times. However, there is a weaker characterization which can still be conjectured to hold generally. Our physical fields are generated for asymptotically flat, stationary, Einstein–Maxwell space–times by the Newman–Penrose functions ψ_2^0, ψ_1^0, ψ_0 for the Weyl tensor and ϕ_1^0, ϕ_0 for the electromagnetic field tensor, and to the asymptotic order investigated here these are electric type. Therefore, we might hope that the following statement may be true: In any asymptotically flat, static, Einstein–Maxwell space–time there is a coordinate system in which the generators of the physical fields are time independent and electric type.

ACKNOWLEDGMENT

The author would like to thank Dr. Allen I. Janis for his helpful criticism of the manuscript.

APPENDIX

Statement A1: For any function $\alpha(\zeta, \bar{\zeta})$ of spin weight zero $\eth^s\bar{\eth}^s\alpha = \bar{\eth}^s\eth^s\alpha$ for all $s \geq 0$.

This statement is easily proved by induction.

Statement A2: If η is a spin weight $s \geq 0$ function of ζ and $\bar{\zeta}$, then η will be electric type if and only if $\bar{\eth}^s\eta = \eth^s\bar{\eta}$.

The proof of this statement uses only elementary properties of the operator edth and the result of Statement A1. The statement can be extended to negative spin weight functions by considering them to be the complex conjugates of positive spin weight functions. This shows in particular that if η is electric type so is $\bar{\eta}$ which may also be seen to follow directly from the definition.[5]

Statement A3: If η is electric, then $\eth\eta$ and $\bar{\eth}\eta$ are electric.

This is easily proved using the commutation properties of edth and its complex conjugate.

Statement A4: If η and ρ are two nonnegative spin weight functions which are electric type and if $\eth\eta = \eth\rho = 0$, then $\eta\rho$ is electric.

Proof: From the assumptions we may write $\eta = \eth^s\alpha$ and $\rho = \eth^t\beta$ where $s, t \geq 0$ and α and β are real, spin weight zero quantities. Using Statement A2 we need to show that $\bar{\eth}^{s+t}(\eth^s\alpha\eth^t\beta) = \eth^{s+t}(\bar{\eth}^s\alpha\bar{\eth}^t\beta)$. But we have

$$\bar{\eth}^{s+t}(\eth^s\alpha\eth^t\beta) = \left[1\Big/\binom{s+t}{s}\right]\bar{\eth}^{s+t}\eth^{s+t}(\alpha\beta), \quad \text{using } \eth\eta = \eth\rho = 0,$$

$$= \left[1\Big/\binom{s+t}{s}\right]\eth^{s+t}\bar{\eth}^{s+t}(\alpha\beta), \quad \text{using A1},$$

$$= \eth^{s+t}(\bar{\eth}^s\alpha\bar{\eth}^t\beta), \quad \text{because } \bar{\eth}^{s+1}\alpha = \bar{\eth}^{t+1}\beta = 0.$$

A similar result holds if η and ρ are both of negative spin weight. However, if their spin weights are of opposite sign, then $\eta\rho$ will not be electric type in general.

[1] E. T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962); Erratum, J. Math. Phys. **4**, 998 (1963).

[2] E. T. Newman and T. W. J. Unti, J. Math. Phys. **3**, 891 (1962).

[3] E. T. Newman and R. Penrose, Proc. Roy. Soc. (London) A **305**, 175 (1968).

[4] A. Exton, E. T. Newman, and R. Penrose, J. Math. Phys. **10**, 1566 (1969).

[5] We make extensive use of the operator edth in this paper. The ${}_sY_{lm}$ are spin weighted eigenfunctions of the operator $\eth\bar{\eth}$. For details see E. T. Newman and R. Penrose, J. Math. Phys. **7**, 863 (1966) and J. N. Goldberg *et al.*, J. Math. Phys. **8**, 2155 (1967).

[6] W. Hallidy, J. Math. Phys. **15**, 413 (1974).

[7] Some would say that this is no restriction at all but rather what is meant when one says that an asymptotically flat space–time is stationary. However, we would like to make it explicit here that we are not considering certain types of asymptotically flat space–times which might occur. For instance, the Kerr metric has a timelike Killing vector in the region between the ergosphere and the event horizon, but it is asymptotically spacelike and therefore is not the same as the asymptotically timelike Killing vector which also exists for that space–time. Space–times possessing only the former type of timelike Killing vector are not considered in this paper.

[8] We also assume that the functions satisfy certain differentiability conditions analogous to the asymptotic smoothness conditions of Ref. 2.

[9] T. W. J. Unti, General Relativity Gravitation **3**, 43 (1972).

Explicit representations of a single parabose operator

D. A. Gray and C. A. Hurst

Department of Mathematical Physics, University of Adelaide, Adelaide, South Australia
(Received 7 May 1974; revised manuscript received 26 August 1974)

Representations for a single parabose operator are found as operators over a Bargmann and over a Schrödinger or harmonic oscillator space. The equivalence of the two representations is proved by the construction of a unitary integral transform connecting them.

1. INTRODUCTION

Three well-known representations of the Bose commutation relations are (a) the matrix representation[1] of the operators, (b) the Bargmann representation[2] in terms of complex variables, and (c) the familiar quantum mechanical harmonic oscillator representation.[3] It is of interest, especially in view of their association with representations of the symplectic group,[4] to find the corresponding irreducible representations of parabose operators.

Irreducible representations of a single parabose oscillator were discussed by Jordan, Mukunda, and Pepper,[5] who classified them and found matrix representations. They showed that in this case the algebra generated by a single parabose operator and its adjoint contained the Lie algebra $sp(2,R)$ as a subalgebra, and a unitary representation of the parabose algebra was the direct sum of two irreducible representations of $sp(2,R)$ with Casimir labels g and $g+\frac{1}{2}$ only, where g is a positive real number. If g is a positive integer or half-integer, the representations of $sp(2,R)$ are integrable to unitary representations of the group $sp(2,R)$, whereas for the other values of g the representations are integrable to unitary representations of the covering group of $Sp(2,R)$.[6] The other unitary representations of $Sp(2,R)$ for which g is complex do not lead to representations of $Sp(2,R)$ which are contained in unitary representations of the parabose algebra.

In this paper we start by constructing representations of $sp(2,R)$ by using a pair of boson operators and then by integration obtain irreducible representations of the parabose algebra in a complex space which is a generalization of the Bargmann space. In this way representations for g integral or half-integral are found. Then, by changing the metric in the two-dimensional complex variable space and integrating, representations for the remaining values of g are found. An alternative way of obtaining the same representations more directly is also given. By a standard procedure the same representations are also constructed in a harmonic oscillator space. As these two representations are equivalent, there must exist a unitary mapping between them, and this mapping is given explicitly.

Some detailed calculations are given in the Appendix.

2. BARGMANN SPACE OF A SINGLE PARABOSE OPERATOR

As the three 2×2 matrices $-\frac{1}{2}i\sigma_2$, $\frac{1}{2}i\sigma_1$, $\frac{1}{2}\sigma_3$ from a two-dimensional irreducible nonunitary representation $sp(2,R)$ the operators

$$\begin{aligned} H_1 &= -\tfrac{1}{2}ia^\dagger\sigma_2\epsilon a = \tfrac{1}{2}(a_1^\dagger a_2^\dagger + a_1 a_2), \\ H_2 &= \tfrac{1}{2}ia^\dagger\sigma_1\epsilon a = \tfrac{1}{2}i(-a_1^\dagger a_2^\dagger + a_1 a_2), \\ H_3 &= \tfrac{1}{2}a^\dagger\sigma_3\epsilon a = \tfrac{1}{2}(a_1^\dagger a_1 + a_2 a_2^\dagger), \end{aligned} \tag{1}$$

where

$$a = \begin{pmatrix} a_1 \\ a_2^\dagger \end{pmatrix}, \quad \epsilon = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2}$$

are a set of hermitian operators which form a unitary representation of $sp(2,R)$ and hence satisfy the commutation relations

$$[H_1,H_2] = -iH_3, \quad [H_2,H_3] = iH_1, \quad [H_3,H_1] = iH_2. \tag{3}$$

a_1 and a_2 are boson operators which satisfy the usual commutation relations.[7] The standard representation of these operators is in a Bargmann space of entire functions $f(z_1,z_2)$, where z_1, z_2 are two complex variables. The scalar product is given by

$$(f,g) = \pi^{-2}\int\int dz_1 dz_2 \exp(-\bar{z}_1 z_1 - \bar{z}_2 z_2) \times \overline{f(z_1,z_2)}\; g(z_1,z_2), \tag{4}$$

and we choose as mutually adjoint operators

$$a_i = \frac{\partial}{\partial z_i}, \qquad a_i^\dagger = z_i. \tag{5}$$

In terms of the operators (5) the representations (1) become

$$\begin{aligned} H_1 &= \frac{1}{2}\left(\frac{\partial^2}{\partial z_1 \partial z_2} + z_1 z_2\right), \\ H_2 &= \frac{1}{2}i\left(\frac{\partial^2}{\partial z_1 \partial z_2} - z_1 z_2\right), \\ H_3 &= \frac{1}{2}\left(z_1\frac{\partial}{\partial z_1} + z_2\frac{\partial}{\partial z_2} + 1\right), \end{aligned} \tag{6}$$

and

$$H_+ = H_1 + iH_2 = z_1 z_2, \quad H_- = H_1 - iH_2 = \frac{\partial^2}{\partial z_1 \partial z_2}.$$

The operator

$$\begin{aligned} H &= \tfrac{1}{2}a^\dagger\epsilon a = \tfrac{1}{2}(a_1^\dagger a_1 - a_2 a_2^\dagger) \\ &= \frac{1}{2}\left(z_1\frac{\partial}{\partial z_1} - z_2\frac{\partial}{\partial z_2} - 1\right) \end{aligned} \tag{7}$$

commutes with all the operators in (6). The Casimir operator

$$C = H_1^2 + H_2^2 - H_3^2 = -H(H+1) \tag{8}$$

obviously also commutes with all the operators (6). The simultaneous normalized eigenstates of H_3 and H are

$$f_{g_1,h_1}(z_1,z_2)=\frac{z_1^{\,h_1-g_1}\,z_2^{\,h_1+g_1}}{[(h_1-g_1)!(h_1+g_1)!]^{1/2}} \tag{9}$$

where $-g=-(g_1+\frac{1}{2})$ is the eigenvalue of H and $h=h_1+\frac{1}{2}$ of H_3. The operators $H_\pm$ raise and lower h_1 by 1 respectively, and as H_3 can have only positive eigenvalues, there must be a lower cutoff value of h_1 which can be seen to be $|g_1|$. The two values $\pm|g_1|$ do not give different representations because the interchange $z_1 \longleftrightarrow z_2$ leaves H_i and hence C invariant while carrying $H \to -(H+1)$ and consequently taking $g_1 \to -g_1$. As $h_1 \pm g_1$ must both be integers because f_{g_1,h_1} is entire, this means that h_1 and g_1 are simultaneously integers or half-integers as already asserted. If we denote by B_n the closed space spanned by the states (9) with $g_1=n$, the same representation of $sp(2,R)$ is obtained with states (9) for which $g_1=-n$. Similarly the spaces $B_{n+1/2}$ for $g_1=n+\frac{1}{2}$ and $g_1=-(n+\frac{1}{2})$ give equivalent representations. The whole space B can then be broken into the direct sum

$$B=B_0\oplus 2\Big(\sum_{n=1}^{\infty}\oplus B_n\Big)\oplus 2\Big(\sum_{n=0}^{\infty}\oplus B_{n+1/2}\Big). \tag{10}$$

The integral and half-integral values of g (and h) can be distinguished by the eigenvalues of an operator P corresponding to the transformation $z_1 \to z_1\exp(i\pi)$, $z_2 \to z_2\exp(i\pi)$. Under this transformation $f_{g_1,h_1} \to \exp(2\pi i h_1) f_{g_1,h_1}$ and so $P=\exp[2\pi i(H_3-\frac{1}{2})]$ has the eigenvalues ± 1 for h_1 (and g_1) integral and half-integral respectively.

In order to construct parabose operators A and A^*, we require operators which change g_1 and h_1 by $\pm\frac{1}{2}$. The following expressions have the required property:

$$\begin{aligned} A&=\frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial z_2}(I\mp P)+\frac{\partial}{\partial z_1}(I\pm P)\right),\\ A^*&=\frac{1}{\sqrt{2}}\left[z_1(I\mp P)+z_2(I\pm P)\right]. \end{aligned} \tag{11}$$

The upper sign, which carries $B_n \rightleftarrows B_{n+1/2}$, is chosen for $|g_1|$ integral, while the lower sign, which carries $B_{n-1/2} \rightleftarrows B_n$, is chosen for $|g_1|$ half-integral. Since we are only interested in the variety of representations which are contained in B, we shall restrict ourselves from now on to the case $g_1 \geq 0$. It can be verified that both expressions (11) satisfy the parabose commutation relations:

$$\begin{aligned} &[\tfrac{1}{2}[A,A^*]_+,A]_- =2[H_3,A]_-=-A,\\ &[\tfrac{1}{2}[A,A^*]_+,A^*]_-=2[H,A^*]_-=A^*,\\ &[A^2,A^*]_-=2[H_-,A^*]_-=2A,\\ &[A^{*2},A]_-=2[H_+,A]_-=-2A^*, \end{aligned} \tag{12}$$

and the remaining trivial relations. Also A and A^* are mutually adjoint. It is interesting to note that (11) is a representation of the parabose algebra which uses *only two bose operators* together with the operator P. From the results of Jordan, Mukunda, and Pepper we know that (11) will give irreducible representations of the parabose algebra when confined to the subspaces $B_n \oplus B_{n+1/2}$ and $B_{n-1/2}\oplus B_n$ respectively, and these are all the irreducible unitary representations which contain integrable representations of $sp(2,R)$ with only a finite multiplicity. Although it is possible to construct unitary representations of $sp(2,R)$ using only a single bose operator, this, of course, will give only the representations of a bose operator. It is necessary to use a minimum of two bose operators, to construct parabose operators of arbitrary order, although more can be used as in, for example, the Green ansatz.

Also it can be verified that

$$AA^* f_{g_1,g_1}=(4g_1+2)f_{g_1,g_1} \tag{13}$$

with

$$A f_{g_1,g_1}=0 \quad \text{for} \quad g_1 \geq 0.$$

so that if we put $p=4g_1+2$, these are the representations of the parabose algebra of even order, excluding $p=0$.

Now if we make the change of variables

$$w=z_1/z_2, \quad z^2=2z_1z_2 \tag{14}$$

(6) becomes

$$\begin{aligned} H_1&=\left[-\frac{1}{z^2}\left(w\frac{\partial}{\partial w}\right)^2+\frac{1}{4z}\frac{\partial}{\partial z}+\frac{1}{4}\frac{\partial^2}{\partial z^2}+\frac{1}{4}z^2\right],\\ H_2&=i\left[-\frac{1}{z^2}\left(w\frac{\partial}{\partial w}\right)^2+\frac{1}{4z}\frac{\partial}{\partial z}+\frac{1}{4}\frac{\partial^2}{\partial z^2}-\frac{1}{4}z^2\right], \end{aligned} \tag{15}$$

$$H_3=\frac{1}{2}\left(z\frac{\partial}{\partial z}+1\right), \quad H=w\frac{\partial}{\partial w}-\frac{1}{2},$$

and the metric in (4) is now

$$\frac{1}{4}\left|\frac{z}{w}\right|^2\exp\left[-\frac{1}{2}|z|^2\left(|w|+\frac{1}{|w|}\right)\right]. \tag{16}$$

A state $\psi_{g_1}(z_1,z_2)$ for fixed g_1 is of the form $w^{-g_1}\psi_{g_1}(z)$ where $\psi_{g_1}(z)$ is a power series in z which contains only powers $\geq 2g_1$, of even order for g_1 integral and odd order for g_1 half-integral. As operators on the states $\psi_{g_1}(z)$, (15) becomes

$$\begin{aligned} H_1&=\frac{1}{4}\left(\frac{d^2}{dz^2}+\frac{1}{z}\frac{d}{dz}+z^2-\frac{4g_1^2}{z^2}\right),\\ H_2&=\frac{1}{4}i\left(\frac{d^2}{dz^2}+\frac{1}{z}\frac{d}{dz}-z^2-\frac{4g_1^2}{z^2}\right), \end{aligned} \tag{15'}$$

$$H_3=\frac{1}{2}\left(z\frac{d}{dz}+1\right), \quad H=-(g_1+\tfrac{1}{2}).$$

The metric in z-space is given by

$$\frac{1}{4\pi}\int\frac{dw\,|z|^2}{|w|^{2g_1+2}}\exp\left(\frac{-|z|^2(|w|+1/|w|)}{2}\right)$$

$$=|z|^2K_{2g_1}(|z|^2). \tag{17}$$

In order to construct representations of parabose operators, we must form the direct sum $B_{g_1}\oplus B_{g_1+1/2}$ with B_{g_1} containing even functions of z for g_1 integral and odd functions for g_1 half-integral. If P now denotes the operator corresponding to the transformation $z \to -z$, the operators (11) in the z representation can be written

$$A=\frac{d}{dz}\mp\frac{2g_1}{z}P+\frac{1}{2z}(I\mp P),\quad A^*=z. \tag{18}$$

These operators can be written in the matrix form:

$$A=\begin{pmatrix}0 & d/dz+(2g_1+1)/z\\ d/dz-2g_1/z & 0\end{pmatrix},\quad g_1 \text{ integral},$$

$$=\begin{pmatrix}0 & d/dz-2g_1/z\\ d/dz+(2g_1+1)/z & 0\end{pmatrix},\quad g_1 \text{ half-integral},$$

$$A^*=\begin{pmatrix}0 & z\\ z & 0\end{pmatrix}, \tag{19}$$

where the first row and column corresponds to even functions and the second row to odd functions. The operators (15′) will be diagonal in this matrix representation with the appropriate values of g_1 inserted. The inverse powers of z which appear in (15′) and (19) cause no difficulty because they are always multiplied with functions which contain sufficiently high powers of z.

As already pointed out, only the representations for even $p\geq 2$ have been found, and in order to obtain the remaining representations it is necessary to change the Bargmann space B and the operators H_i, H, A, and A^*. First of all the metric $\pi^{-2}\exp(-\bar{z}_1z_1-\bar{z}_2z_2)$ in (4) is replaced by the metric $\pi^{-2}|z_2|^{4g'}\exp(-\bar{z}_1z_1-\bar{z}_2z_2)$, with $-\frac{1}{2}<g'<0$, and the operators H_i and H are replaced by

$$H_1\to H_1+\frac{g'}{z_2}\frac{\partial}{\partial z_1},\quad H_2\to H_2+\frac{ig'}{z_2}\frac{\partial}{\partial z_1},$$

$$H_3\to H_3+g',\qquad H\to H-g'. \tag{6′}$$

These operators are formally symmetric in the new metric. The class of functions $f(z_1,z_2)$ to be considered are still entire functions, but now, with

$$f(z_1,z_2)=\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}a_{m_1m_2}{z_1}^{m_1}{z_2}^{m_2},$$

$$\|f\|^2=\sum_{m_1}\sum_{m_2}|a_{m_1m_2}|^2\,m_1!\,\Gamma(m_2+2g'+1),$$

and then, by following the same method as Bergmann, it can be easily shown that

$$|f(z_1,z_2)|^2\leq\frac{\exp(|z_1|^2)\,{}_1F_1(1,2g'+1;|z_2|^2)}{\Gamma(2g'+1)}\|f\|^2,$$

where ${}_1F_1$ is the confluent hypergeometric function. For $|z_2|$ large this gives:

$$|f(z_1,z_2)|^2\leq\frac{\exp(|z_1|^2+|z_2|^2)}{|z_2|^{4g'}}\left[1+O\left(\frac{1}{|z_2|^2}\right)\right]\|f\|^2,$$

which has the same exponential dominance as before. However, because of the additional factor $|z_2|^{-4g'}$ the class of functions $f(z_1,z_2)$ allowed will be different. It is clear that H_3 and H are now not only symmetric but also self-adjoint. The operators H_1 and H_2 will be defined if we restrict ourselves to the subspace of functions $f(z_1,z_2)$ for which $m_2\geq m_1$. This corresponds, as we shall see, to choosing $g_1\geq 0$, which gives no loss of generality as we are interested only in constructing new representations. In order to consider (if desired) the case $g_1<0$, the factor $|z_2|^{4g'}$ should be replaced by $|z_1|^{4g'}$ with consequent changes in (6′). With the restriction to the subspace corresponding to $g_1\geq 0$, H_1 and H_2 will be self-adjoint. Furthermore, because of the change in metric, although z_1 and $\partial/\partial z_1$ have the same properties as before, z_2 and $\partial/\partial z_2$ will no longer be mutually adjoint, although they are still closed operators with dense domains (for example, the set of polynomials in z_1 and z_2). It can be shown that now

$$z_2^*=\frac{\partial}{\partial z_2}+\frac{2g'}{z_2}$$

when $m_2\geq 1$ and $z_2^*=0$ when $m_2=0$.

If (11) is replaced by

$$A\to A+\frac{\sqrt{2}\,g'}{z_2}(I\mp P),\quad A^*\to A^*, \tag{11′}$$

it can be verified that they are mutually adjoint densely defined closed operators, satisfying the parabose commutation relations. The functions (9) are still eigenstates of H_3 and H {with a new normalization factor $[\Gamma(h_1+g_1+2g'+1)]^{-1/2}$ instead of $[(h_1+g_1)!]^{-1/2}$} and new eigenvalues $h+g'=h_1+g'+\frac{1}{2}$ and $-(g+g')=-(g_1+g'+\frac{1}{2})$ respectively, and the eigenvalue of AA^* when applied to f_{g_1,g_1} is now $4(g_1+g')+2$. If we choose $g_1=0$ and $g'=-\frac{1}{4}$, we should regain the ordinary bose operators, while for $g_1\neq 0$ we have parabose statistics for odd p. After making the change of variables (14) and integrating over w-space, the metric in z-space is given by

$$|z|^{4g'+2}K_{2(g+g')}(|z|^2) \tag{20}$$

and for the boson case this is

$$\sqrt{2/\pi}\;|z|K_{1/2}(|z|^2)=\exp(-|z|^2),$$

which is the usual metric. The functions $f(z)$ will be entire functions of z with the lowest power z^{2g_1} as before, and the operators (11′) have the matrix representation (for $g'=-\frac{1}{4}$)

$$A=\begin{pmatrix}0 & d/dz\pm 2g_1/z\\ d/dz\mp 2g_1/z & 0\end{pmatrix},\quad A^*=z, \tag{21}$$

with the upper and lower signs for g_1 integral and half-integral respectively. For $g_1\neq 0$ or $-\frac{1}{4}$ we obtain other representations of the parabose operators, but these do not correspond to integral values of p and so do not have a direct particle interpretation.

The above approach may appear rather indirect, and so it is interesting to see whether it is possible to construct parabose representation by generalizing the simple correspondence $A\to d/dz$, $A^*\to z$ of the boson calculus. The expressions:

$$A=\frac{d}{dz}+\frac{\tau}{z}P,\quad A^*=z, \tag{21′}$$

where τ is a parameter to be suitably chosen and P is the reflection operator, satisfy the parabose commutation relations formally. If these operators are to act on suitable entire functions, an appropriate metric must be found. In order to do this we decompose $f(z)$ according to the eigenvalues of P so that

$$f(z)=\begin{pmatrix} f_e(z) \\ f_0(z) \end{pmatrix}, \tag{22}$$

corresponding to the eigenvalues $P=\pm 1$ respectively. The condition on the ground state, $f^0(z)$, for a parabose system of order p, namely,

$$Af^0(z)=0=(AA^*-p)f^0(z), \tag{23}$$

leads to the expression

$$f^0(z)=C(\bar{z})\,z^{(p-1)/2} \tag{24}$$

where $C(\bar{z})$ satisfies

$$C(\bar{z})\,[(p-1)/2]\,z^{(p-1)/2}+\tau\,C(-\bar{z})(-z)^{(p-1)/2}=0. \tag{25}$$

The solution of (25) is

$$\begin{aligned} &\tau=(-1)^{\alpha+1}, \quad C(\bar{z})=1 \quad \text{for } p \text{ odd}, \quad p=2\alpha+1, \\ &=(-1)^{\alpha+1}(\alpha+\tfrac{1}{2}), \quad C(\bar{z})=(\bar{z})^{1/2} \quad \text{for } p \text{ even}, \quad p=2\alpha+2. \end{aligned} \tag{26}$$

The case p odd is quite straightforward, but p even is rather paradoxical because from (24) $f^0(z)=(\bar{z})^{1/2}z^{1/2}z^{\alpha}$, and it is neither entire nor analytic. The resolution of this undesirable conclusion will be given after completing the determination of the metric for p odd. In order to obtain this metric, we put

$$\rho(\bar{z},z)=\begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix}, \tag{27}$$

(we write 1 and 2 as suffixes instead of e and 0 for convenience) and require that

$$(f,Ag)=(A^*f,g), \tag{28}$$

where f and g are suitably convergent at infinity. If we substitute for (21′) and (27) in (28), we obtain the following system of differential equations for ρ_{ij}:

$$\begin{aligned} &\frac{d\rho_{12}}{dz}-\frac{\tau}{z}\,\rho_{12}=-\bar{z}\,\rho_{21}, \\ &\frac{d\rho_{12}}{dz}+\frac{\tau}{z}\,\rho_{21}=-\bar{z}\,\rho_{12}, \\ &\frac{d\rho_{11}}{dz}+\frac{\tau}{z}\,\rho_{11}=-\bar{z}\,\rho_{22}, \\ &\frac{d\rho_{22}}{dz}-\frac{\tau}{z}\,\rho_{22}=-\bar{z}\,\rho_{11}. \end{aligned} \tag{29}$$

The solution of these equations subject to the requirement that $\rho\to 0$ as $|z|\to\infty$ is

$$\begin{aligned} \rho_{11}(|z|^2)&=|z|\,K_{\tau+1/2}(|z|^2), \\ \rho_{22}(|z|^2)&=|z|\,K_{\tau-1/2}(|z|^2), \end{aligned} \tag{30}$$

which is the same as (20) for $g'=-\frac{1}{4}$. The off-diagonal elements are zero if the unitarity condition

$$(Pf,\ Pg)=(f,g) \tag{31}$$

is imposed. So we have recovered the results already obtained for p odd, as would be expected from a comparison of the assumption (21′) with the result (21).

The complications for p even arise from the discrepancy between (21′) and (19) and the second equation in (26) has to be interpreted in that light. If we nevertheless assume the expressions (27) for $\rho(z,\bar{z})$ and substitute in (28), making use of (21′) and (26), equations similar to (29) are obtained, but their solution is now

$$\begin{aligned} \rho_{11}(|z|^2)&=K_{\tau+1/2}(|z|^2), \\ \rho_{22}(|z|^2)&=K_{\tau-1/2}(|z|^2). \end{aligned} \tag{30′}$$

If we write $f(z)=|z|\,g(z)$, where $g(z)$ is an entire function with lowest power of $z\geq\alpha$, then it can be easily seen that

$$\|f\|=\|g\|'$$

where $\|\ \|'$ is obtained by multiplying ρ_{ij} by $|z|^2$, and this is just (17). If, furthermore, the factor $|z|$ is pulled through (21′), the expression (19) is obtained, and the correspondence is complete. So the factor $(\bar{z}z)^{1/2}$ must be interpreted as $|z|$, i.e., the determinations of the two square roots are always chosen so as to give a positive real number for their product. The operators d/dz and P in (21′) are then defined by their effect on $g(z)$. With this understanding, in the rest of this paper we will use (21′).

From the definition $\|f\|=(f,f)^{1/2}$ and the decomposition of the space into even and odd functions, it follows that

$$\|f\|^2=\|f_e\|^2+\|f_0\|^2,$$

where for p odd from the expression

$$f_+(z)=z^{\alpha}f_e(z)=\sum_m a_{2m}\ z^{\alpha+2m}$$

it can be shown that

$$\|f_+\|^2=\sum_m |a_{2m}|^2 I_{2m}, \tag{32}$$

with

$$I_{2m}=\frac{2^{\alpha+2m}}{\sqrt{\pi}}\ \Gamma(m+\alpha+\tfrac{1}{2})\,\Gamma(m+1),$$

and from

$$f_-(z)=z^{\alpha}f_0(z)=\sum_m a_{2m+1}z^{2m+\alpha+1},$$

that

$$\|f_-\|^2=\sum_m |a_{2m+1}|^2\,I_{2m+1} \tag{33}$$

with

$$I_{2m+1}=\frac{2^{\alpha+2m+1}}{\sqrt{\pi}}\ \Gamma(m+\alpha+\tfrac{3}{2})\,\Gamma(m+1).$$

In evaluating the expressions the result

$$\int_0^{\infty} K_{\nu}(t)\ t^{\mu-1}\,dt=2^{\mu-2}\,\Gamma\left(\frac{\mu+\nu}{2}\right)\,\Gamma\left(\frac{\mu-\nu}{2}\right)$$

has been used.[8]

By Schwarz's inequality

$$|f_{\pm}(z)|^2 \leq \{I_{\pm}(|z|^2)\}\, \|f_e\|_0^2, \tag{34}$$

where

$$I_{+}(|z|^2) = \sum_m |z^{\alpha+2m}|^2 / I_{2m} \tag{35}$$

and

$$I_{-}(|z|^2) = \sum_m |z^{\alpha+2m+1}|^2 / I_{2m+1}.$$

For p even the same inequalities are obtained except for the inclusion of an additional factor $|z|^2$ on the right-hand side of (34). So in all cases we have an inequality of the form

$$|f(z)| \leq w(z)\, \|f\|.$$

The usefulness of a relation of this form, apart from showing the equivalence of strong and pointwise convergence as discussed by Bargmann[2] is that it enables a set of principal vectors to be defined. $w(z)$ can be interpreted as "the reproducing kernel" as will be shown now, for the case of p odd and I_+. The other cases can be treated similarly. The "reproducing kernel" I_+ is defined by

$$f_{+}(w) = \int I_{+}(w, z)\, f_{+}(z)\, d\rho(z\bar{z}),$$

and is denoted by $e_w^+(z)$, a "principal vector." It is the representation of the unit operator and in terms of any complete orthonormal system v_k^+:

$$e_w^+(z) = \sum_k \bar{v}_k^+(w)\, v_k^+(z). \tag{36}$$

Using the orthonormal system (9), we have

$$\begin{aligned} I_{+}(wv) &= \sum_{n=0}^{n} \frac{(\bar{w}v)^{\alpha+2n}}{I_{2n}}, \\ &= \sqrt{\pi}\left(\frac{\bar{w}v}{2}\right)^{\alpha} \sum_{n=0}^{\infty} \frac{1}{n!}\, \frac{1}{\Gamma(n+\alpha+\frac{1}{2})} \left(\frac{\bar{w}v}{2}\right)^{2n}, \\ &= (\pi/2)^{1/2} (\bar{w}v)^{1/2}\, I_{\alpha-1/2}(\bar{w}v), \end{aligned} \tag{37}$$

where I_ν denotes a modified Bessel function of the first kind. Similar results can be found for I_- and for p even.

3. HARMONIC OSCILLATOR REPRESENTATIONS OF A PARABOSE OPERATOR

The third form of explicit representation of parabose operators is the harmonic oscillator or Schrödinger representation, and this is constructed by making suitable generalizations of the usual methods. These results were first obtained by Yang,[9] although he did not give their relationship with a Bargmann representation. In order to distinguish between the two representations, we denote the annihilation and creation operators by ξ and η respectively where

$$\xi = -(1/\sqrt{2})(q + ip), \quad \eta = -(1/\sqrt{2})(q - ip), \tag{38}$$

and

$$q = x, \quad p = -i\left(\frac{d}{dx} + \frac{\tau}{x} R\right). \tag{39}$$

We use R to denote the operation $x \to -x$ in order to distinguish it from P. If we decompose the representation space into even and odd functions, the metric is given by

$$(f, g) = \int_{-\infty}^{\infty} (\bar{f}_e g_e + \bar{f}_0 g_0)\, dx, \tag{40}$$

and ξ and η are mutually adjoint. The vacuum states are the solutions of $\xi\psi_0 = 0$, and are given by

$$\begin{aligned} \psi_0 &= \frac{1}{[\Gamma(\alpha+\frac{1}{2})]^{1/2}}\, x^{\alpha} \exp(-\tfrac{1}{2}x^2), && p = 2\alpha + 1, \text{ odd}, \\ &= \frac{1}{[\Gamma(\alpha+1)]^{1/2}}\, |x|^{1/2} x^{\alpha} \exp(-\tfrac{1}{2}x^2), && p = 2\alpha + 2, \text{ even}. \end{aligned} \tag{41}$$

The coefficients α and τ are again related by $\tau = (-1)^{\alpha+1}\alpha$ for p odd and $\tau = (-1)^{\alpha+1}(\alpha + \frac{1}{2})$ for p even. In order to construct orthonormal bases for the representation spaces, one could apply the raising operator η to ψ_0, but because of the presence of the operator R this is not so simple, and instead we consider the "Hamiltonian"

$$H = \tfrac{1}{2}[\eta, \xi]_+ - p/2, \tag{42}$$

which satisfies

$$[H, \eta]_- = \eta, \tag{43}$$

so that it can be interpreted as counting the power of η in an arbitrary state since $H\psi_0 = 0$. The eigenvalue equation $H\psi_\lambda = \lambda\psi_\lambda$ can be written

$$\frac{d^2\psi_\lambda}{dx^2} + \left(\lambda' - x^2 - \frac{\tau(\tau+R)}{x^2}\right)\psi_\lambda = 0, \tag{44}$$

where $\lambda' = 2\lambda + p$.

Since $[H, R] = 0$, it follows that H and R form a complete set of commuting operators. The eigenstates of H can thus be classified according to their parity and are denoted by $\psi_\pm$ according to the equation

$$R\,\psi_{\pm\lambda} = \pm\,\psi_{\pm\lambda}.$$

The differential–difference equation then becomes a second order differential equation which can be transformed to the confluent hypergeometric equation. The normalized solutions for the various cases are

$p = 2\alpha + 1$ odd: $\tau = (-1)^{\alpha+1}\alpha$,

$$\begin{aligned} \phi_{2n}^{\alpha}(x) &= (-1)^n \left(\frac{n!}{\Gamma(n+\alpha+\frac{1}{2})}\right)^{1/2} x^{\alpha} \exp(-x^2/2)\, L_n^{\alpha-1/2}(x^2), \\ \phi_{2n+1}^{\alpha}(x) &= (-1)^n \left(\frac{n!}{\Gamma(n+\alpha+\frac{3}{2})}\right)^{1/2} \\ &\qquad x^{\alpha+1} \exp(-x^2/2)\, L_n^{\alpha+1/2}(x^2), \end{aligned} \tag{45}$$

$p = 2\alpha + 2$ even: $\tau = (-1)^{\alpha+1}(\alpha + \frac{1}{2})$,

$$\begin{aligned} \phi_{2n}^{\alpha}(x) &= (-1)^n \left(\frac{n!}{\Gamma(n+\alpha+1)}\right)^{1/2} \\ &\qquad x^{\alpha}\, |x|^{1/2} \exp(-x^2/2)\, L_n^{\alpha}(x^2). \end{aligned}$$

$$\phi^{\alpha}_{2n+1}(x)=(-1)^n\left|\frac{n!}{\Gamma(n+\alpha+2)}\right|^{1/2}x^{\alpha+1}\sqrt{|x|}$$

$$\times\exp(-x^2/2)\,L_n^{\alpha+1}(x^2), \tag{46}$$

and these form a complete orthonormal basis, which could also be constructed by applying raising operators to ψ_0. The complete analysis of the various cases is given in Ref. 10.

4. EQUIVALENCE OF THE BARGMANN AND HARMONIC OSCILLATOR REPRESENTATIONS

Since both the Bargmann space F and harmonic oscillator space H are representations of the parabose algebra, it immediately follows from general theory that they are unitarily equivalent, and following Bargmann's analysis, the kernel of the integral transform $A(z,x)$ can be found by requiring that $\eta\psi$ be mapped into a^*f and $\xi\psi$ into af. The most general form of $A(z,x)$ is a 2×2 matrix $A_{ij}(z,x)$, where the indices i,j take the values e and o and the mapping requires that

$$a^*f=\int(zA)\psi\,dx=\int A(\overrightarrow{\eta}\psi)\,dx=\int(A\overleftarrow{\eta})\psi\,dx. \tag{47}$$

The arrow over η indicates the direction in which $d_x\equiv d/dx$ acts. Using the matrix representations, we can write the above equations determining A as

$$\begin{pmatrix}(x+d_x-\tau/x)A_{12} & (x+d_x+\tau/x)A_{11}\\ (x+d_x-\tau/x)A_{22} & (x+d_x+\tau/x)A_{21}\end{pmatrix}$$

$$=-\sqrt{2}\begin{pmatrix}zA_{21} & zA_{22}\\ zA_{11} & zA_{12}\end{pmatrix} \tag{48}$$

and similarly from the connection between af and $\xi\psi$ we obtain the equations

$$\begin{pmatrix}(x-d_x+\tau/x)A_{12} & (x-d_x-\tau/x)A_{11}\\ (x-d_x+\tau/x)A_{22} & (x-d_x-\tau/x)A_{21}\end{pmatrix}$$

$$=-\sqrt{2}\begin{pmatrix}(d_z-\tau/z)A_{21} & (d_z-\tau/z)A_{22}\\ (d_z+\tau/z)A_{11} & (d_z+\tau/z)A_{12}\end{pmatrix}. \tag{49}$$

In the Appendix it is shown that the off-diagonal elements are zero and that

$$A_{\alpha-1/2}(z,x)=C_{\alpha-1/2}(\sqrt{2}\,zx)^{1/2}\times\exp[-(x^2+z^2)/2]J_{\alpha-1/2}(i\sqrt{2}\,zx),$$

$$A_{\alpha+1/2}(z,x)=C_{\alpha+1/2}(\sqrt{2}\,zx)^{1/2}\times\exp[-(x^2+z^2)/2]J_{\alpha+1/2}(i\sqrt{2}\,zx), \tag{50}$$

for p odd, and

$$A_{\alpha}(z,x)=C'_{\alpha}(\sqrt{2}\,z\bar{z}|x|)^{1/2}\exp[-(x^2+z^2)/2]J_{\alpha}(i\sqrt{2}\,zx),$$

$$A_{\alpha+1}(z,x)=C'_{\alpha+1}(\sqrt{2}\,z\bar{z}|x|)^{1/2}\exp[-(x^2+z^2)/2]\times J_{\alpha+1}(i\sqrt{2}\,zx), \tag{51}$$

for p even.

The above notation means that we replace A_{ij} by A_ν with states of parity $(-1)^{\nu+1/2}$ in H going into states of the same parity in F. For example, $i=e$ corresponds to states of even parity so that $\nu+\frac{1}{2}$ must be even. Hence for p odd and α odd we have $\nu=\alpha+\frac{1}{2}$ while for α even we have $\nu=\alpha-\frac{1}{2}$, and similarly for $i=o$ or p even.

The equivalence of the two representations requires $A(z,x)$ to be unitary. This is equivalent to the following conditions:

$$\int A_\nu(z,x)\bar{A}_\nu(w,x)\,dx=I_+(zw), \tag{52}$$

where $I_+(z,w)$ is the representation of the unit element in even or odd subspaces of F, and

$$\int A_\nu(z,x)\bar{A}_\nu(z,y)\,d\rho_\nu(z\bar{z})=\delta_+(x-y), \tag{53}$$

where $\delta_+(x-y)$ is the decomposition of the δ function into its even or odd components in H. Since it is more convenient to work in terms of well defined integrals, (53) is replaced by

$$\lim_{\lambda\to1}\int A_\nu(\lambda z,x)\bar{A}_\nu(\lambda z,y)\,d\rho_\nu(z\bar{z})=\delta_+(x-y). \tag{53'}$$

For the case of p odd, we have on substituting from (49) in (51) that the integral becomes

$$\sqrt{2}\;C_\nu\bar{C}_\nu\int_{-\infty}^{\infty}(z\bar{w})^{1/2}\exp[-(z^2+w^2)/2]\times\exp(-x^2)\times J_\nu(i\sqrt{2}\,zx)J_\nu(-i\sqrt{2}\,\bar{w}x)\,dx$$

$$=(\pi/2)^{1/2}(z\bar{w})^{1/2}I_\nu(z\bar{w}),$$

which is $I_+(zw)$ for $\nu=\alpha+\frac{1}{2}$ and $I_-(zw)$ for $\nu=\alpha-\frac{1}{2}$. Some identities in Ref. 8 have been used. Since $I_\pm(zw)$ are the unit elements already calculated, (51) is verified. Similar calculation can be made for p even.

The verification of (53') is less straightforward. As is shown in the Appendix, (53') has the form

$$\frac{\lambda}{1-\lambda^4}\sqrt{xy}\;\exp[-(x^2+y^2)/2]\exp[-\lambda^4(x^2+y^2)/(1-\lambda^4)]\times I_\nu\left(\frac{2\lambda^2xy}{1-\lambda^4}\right) \tag{54}$$

for $\nu=\alpha\pm\frac{1}{2}$. Since the expression (54) for $\nu=\alpha$ or $\alpha+1$ is the same expression except for a factor λ, and since $\lambda\to1$, only (54) need be discussed. By inspection the argument of I_ν approaches infinity as $\lambda\to1$. An asymptotic expression for $I_\nu(z)$ is

$$I_\nu(z)\sim\frac{1}{(2\pi z)^{1/2}}\left(e^z\sum_{m=0}^{\infty}(-1)^m(\nu,m)(2z)^{-m}\right)$$

$$+i\exp(-z+i\nu\pi)\sum_{m=0}^{M-1}(\nu,m)(2z)^{-m}\ .$$

Considering only the first order terms in the expression, we see that the asymptotic form of $I_\nu(z)$ will be

$$(2\pi z)^{-1/2}e^z\quad\text{as}\quad z\to\infty$$

and

$$i(2\pi z)^{-1/2}\ \exp(-z+i\nu\pi)\quad \text{as}\quad z\to-\infty.$$

Since in the domain where one of these is very large the other is very small, it is convenient to incorporate both forms in the one expression:

$$I_\nu(z)\sim(2\pi z)^{-1/2}\,(e^z+i\exp(-z+i\nu\pi)), \tag{55}$$

with the understanding that only the dominant term is considered in each case. As $z=2\lambda^2xy/(1-\lambda^4)$, it can be shown that in the limit as $\lambda\to 1$, that (53′) can be written

$$\tfrac{1}{2}\lim_{\epsilon\to 0}\frac{1}{2\epsilon\sqrt{\pi}}\ [\exp(-t^2/\epsilon^2)+i(-1)^\nu\exp(-s^2/\epsilon^2)], \tag{56}$$

where

$$\epsilon=\left(\frac{1-\lambda^2}{1+\lambda^2}\right)^{1/2},\quad s=\frac{x+y}{2},\quad t=\frac{x-y}{2}.$$

This means that (53′) has the form

$$\tfrac{1}{2}[\delta(x-y)\pm(-1)^\alpha\,\delta(x+y)]$$

for $\nu=\alpha\pm\frac{1}{2}$, and this is just the decomposition of the unit operator $\delta(x-y)$ into the space of even and odd functions. A similar conclusion is obtained for p even.

5. CONCLUSIONS

It has been shown that the Bargmann construction for a single boson operator can be extended to the case of a single parabose operator with only increased computational complexity, but no real change in principle. However, the extension to several parabose operators has not yet been carried out, and it does not appear to be particularly straightforward. An approach using the representations of $Sp(2n)$ is being investigated, but it appears that a discussion of the representations of parabose fields is likely to be very involved. It can be remarked that this approach is in a sense the converse of that adopted by Alabiso and Duimio,[11] who have used parabose operators to construct representations of $Sp(4,R)$.

APPENDIX

Determination of the integral transform $A(z,x)$

The matrix elements A_{11} and A_{22} satisfy the two pairs of coupled partial differential equations

$$(x+d_x+\tau/x)A_{11}=\sqrt{2}\ zA_{22}, \tag{A1a}$$

$$(x+d_x-\tau/x)A_{22}=-\sqrt{2}\ zA_{11}, \tag{A1b}$$

$$(x-d_x-\tau/x)A_{11}=-\sqrt{2}\,(d_z-\tau/z)A_{22}, \tag{A1c}$$

$$(x-d_x+\tau/x)A_{22}=-\sqrt{2}\,(d_z+\tau/z)A_{11}. \tag{A1d}$$

If we differentiate (A1a) and substitute for A_{22} and d_xA_{22} from (A1b), the following equation is obtained:

$$[d_x^2+2xd_x+1-2z^2+x^2-\tau(\tau+1)/x^2]A_{11}=0, \tag{A2}$$

while (A1c) and (A1d) reduce to

$$[d_z^2+2zdz+1-2x^2+z^2-\tau(\tau+1)/z^2]A_{11}=0. \tag{A3}$$

The substitution

$$A_{11}=\exp[-(x^2+z^2)/2]\sqrt{\xi}\ u(\xi),$$

with $\xi=(\sqrt{2}\ zx)$ reduces both these equations to the form

$$\xi^2u''+\xi u'(\xi)-[\xi^2+(\tau+\tfrac{1}{2})^2]u(\xi)=0, \tag{A4}$$

which is Bessel's equation with imaginary argument. The solution is

$$A_{11}=a_{\tau+1/2}\exp[-(x^2+z^2)/2](\sqrt{2}\ zx)^{1/2}\ Q_{\tau+1/2}(\sqrt{2}\ zx),$$

where $Q_{\tau+1/2}(\sqrt{2}\ zx)$ is any combination of Bessel functions with imaginary argument and index $\tau+\frac{1}{2}$ or $-(\tau+\frac{1}{2})$ and a_ν is an arbitrary coefficient. The solution for A_{22} can be obtained in a similar manner.

The solution for the off-diagonal elements can also be found and they are zero. The reason for this can be shown by a more general argument as follows. The mapping (47) is equivalent to requiring that P is mapped into R, i.e.,

$$f(-z)=\int A(z,x)\,\psi(-x)dx,$$

which implies that

$$PA(z,x)=A(z,x)R,$$

and hence $A(z,x)$ is diagonal.

Evaluation of the integral (53′)

If we substitute (30) and (49) in (53′), the latter becomes

$$\begin{aligned}&(\lambda/\pi)\sqrt{xy}\ \exp[-(x^2+y^2)/2]\int dzd\bar z\ z\bar z\ \exp[-\lambda^2(z^2+\bar z^2)/2]\\&\quad\times J_\nu(i\sqrt{2}\ \lambda zx)J_\nu(-i\sqrt{2}\ \lambda\bar zy)K_\nu(z\bar z)\\&=\frac{\lambda^{2\nu+1}}{\pi}\ (xy)^{\nu+1/2}\exp[-(x^2+y^2)/2]\sum_n\sum_m(-1)^{n+m}\\&\quad\times 2^{-\nu-m-n}\ \frac{L_n^\nu(x^2)L_m^\nu(y^2)}{\Gamma(n+\nu+1)\Gamma(m+\nu+1)}\\&\quad\times\int dz\,d\bar z\,z\bar z\,\bar z^{2m+\nu}\ z^{2n+\nu}\,K_\nu(z\bar z),\end{aligned} \tag{A5}$$

since[8]

$$J_\nu(2\sqrt{xz})=(xz)^{\nu/2}\exp(-z)\sum_{n=0}^{\infty}\frac{L_n^\nu(x)z^n}{\Gamma(n+\nu+1)}.$$

If the integral is evaluated, the expression (A5) becomes

$$\begin{aligned}&\lambda^{2\nu+1}(xy)^{\nu+1/2}\exp[-(x^2+y^2)/2]\sum_n\frac{L_n^\nu(x^2)L_n^\nu(y^2)\lambda^{4n}n!}{\Gamma(n+\nu+1)}\\&=\frac{\lambda}{1-\lambda^4}\sqrt{xy}\ \exp[-(x^2+y^2)/2]\\&\quad\times\exp\left(-\frac{\lambda^4(x^2+y^2)}{(1-\lambda^4)}\right)I_\nu\left(\frac{2\lambda^2xy}{1-\lambda^4}\right).\end{aligned}$$

The last line follows from the fact that I_ν is a generating function for products of Laguerre polynomials. Substitution of $\nu=\alpha\pm\frac{1}{2}$ gives the required result (54).

[1]E. M. Loebl, *Group Theory and Its Application* (Academic, New York, 1968), Vol. I.
[2]V. Bargmann, Comm. Pure Appl. Math. 14, 187 (1961). I. E. Segal, Ill. J. Math. 6, 500 (1962), and also *Mathematical Problems of Relativistic Physics* (American Mathematical Society, Providence, R. I., 1963).
[3]P. Stehle, *Quantum Mechanics* (Holden-Day, San Francisco, 1966).
[4]S. Kamefuchi and Y. Takahashi, Nucl. Phys. 36, 177 (1962).
[5]T. F. Jordan, N. Mukunda, and S. V. Pepper, J. Math. Phys. 4, 1089 (1963).
[6]A. O. Barut, in *Lectures in Theoretical Physics, Vol. IXA*, Boulder Lectures 1966 (Gordon and Breach, New York, 1967). Also, A. O. Barut and C. Fronsdal, Proc. Roy. Soc. A 287, 582 (1965).
[7]H. J. Lipkin, *Lie Groups for Pedestrians* (North-Holland, Amsterdam, 1965).
[8]H. Bateman, *Higher Transcendental Functions, Vol.* 2, edited by Erdelyi, Magnus, Oberhettinger, and Tricomi (McGraw-Hill, New York, 1953). The relevant equations are on pages 50, 51, 86, and 189.
[9]L. M. Yang, Phys. Rev. 84, 788 (1951).
[10]D. A. Gray, Ph. D. Thesis, University of Adelaide (1973).
[11]C. Alabiso and F. Duimio, "Irreducible Representations of Parabose Operators," University of Parma Preprint IFPR-T-020 (1971).

Formula for the computation of the representation matrix elements of the group $SO(n)$

Takayoshi Maekawa

Department of Physics, Kumamoto University, Kumamoto, Japan
(Received 5 July 1974)

The formulas of expressing the infinitesimal operators of the parameter groups of $SO(n)$ in terms of the Euler angles are given. By using these, recursion relations which are useful for calculations of the representation matrix elements of $SO(n)$ are obtained. Expressions for the d functions with the highest and some weights are obtained explicitly, and it is shown that the d functions of $SO(2j+1)$ or $SO(2j)$ with special weights agree with those of $SO(3)$ or $SO(4)$ respectively. By using the results, a formula of computing the D-matrix elements of $SO(n)$ are given in terms of lowering operators corresponding to those of Pang and Hecht.

1. INTRODUCTION

It is important to construct the representation matrix elements of $SO(n)$ due to their applicability to physical problems. For $SO(3)$, they are well-known $D^{(j)}_{m',m}(\theta, \Phi, \Psi)$ functions which play an important role in angular momentum theory. The d functions of $SO(4)$ which are intimately connected with the D functions of $SO(4)$ are used in particle physics by Freedman and Wang.[1] The d functions of $SO(5)$ [2] and the relation between the d functions of $SO(4)$ and the boost matrix elements[3] have also been studied. The groups $SO(5)$ and $SO(8)$ are applied in nuclear physics.[4] Further, the group $SO(n)$ is also important in physical problems.[5]

The irreducible representation matrix elements of $SO(n)$ have been studied by many authors.[6] These, however, are restricted to special cases and functions. Among these, in the case of a single-valued representation Wolf[7] gives a recursion formula by which the d functions of $SO(n)$ are obtained as an integral of the d functions of $SO(n-1)$. On the other hand, the d functions of $SO(3)$ are well known and can be given by various methods, one of which uses recursion relations and gives an explicit expression to the d functions by operating lowering operators on the d function with the highest weight.[8] In this case, it is essential to use the matrix form for the generators D_j and the expressions for the corresponding operators of the parameter groups of $SO(3)$, i.e., differential operators with respect to the Euler angles.

The matrix elements of the infinitesimal generators of $SO(n)$ have been known since Gel'fand and Zetlin,[9] whose bases are characterized by the group chain $SO(n) \supset SO(n-1) \supset \cdots \supset SO(2)$. Therefore, if we can give to the operators of the parameter groups[10] of $SO(n)$ explicit expressions in terms of the differential operators with respect to the Euler angles, we will be able to construct recursion relations among the d functions of $SO(n)$ as in the three-dimensional case. By using these relations, the d functions will be determined. The purpose of this article is to give explicit formulas, by which the differential operators of the parameter groups of $SO(n)$ are given, and further to give a useful formula, by which the representation matrix elements of $SO(n)$ are calculated through successive applications of lowering operators to the highest weight d function.

In Sec. 2, a few results of a previous paper[11] are summarized and notations used in the following sections are introduced. In Sec. 3, two formulas are given by which the differential operators of the parameter groups of $SO(n)$ are expressed in terms of the Euler angles. And the operators are given explicitly for $n=3, 4$ and for general n two useful recursion relations for the differential operators are given explicitly. In Sec. 4, the d functions of $SO(n)$ with the highest and some weights are obtained using the results in Sec. 3. It is shown that d functions of $SO(n)$ with some weights correspond to the three- or four-dimensional d functions according to n odd or even respectively. And a formula of computing the representation matrix elements of $SO(n)$ is given in terms of lowering operators corresponding to those of Pang and Hecht.[12]

2. PRELIMINARIES

In this section, some notations are introduced and a brief review of Ref. 11 is given.

Orthogonal coordinate systems E_n, $\bar{E}_n$, $\Lambda^2 E_n$ and $\Lambda^2 \bar{E}_n$ are considered. The systems E_n and $\bar{E}_n$, whose orthonormal bases are $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ and $(\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \ldots, \bar{\mathbf{e}}_n)$ respectively, are n-dimensional. The systems $\Lambda^2 E_n$ and $\Lambda^2 \bar{E}_n$, whose orthonormal bases are $(\mathbf{E}_{12}, \ldots, \mathbf{E}_{n-1n})$ and $(\bar{\mathbf{E}}_{12}, \ldots, \mathbf{E}_{n-1n})$, are $[n(n-1)/2]$-dimensional. The bases $\mathbf{E}_{jk} = \mathbf{e}_j \wedge \mathbf{e}_k$, $\bar{\mathbf{E}}_{jk} = \bar{\mathbf{e}}_j \wedge \bar{\mathbf{e}}_k$ have the inner products:

$$\mathbf{E}_{jk} \cdot \mathbf{E}_{lm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}, \quad \bar{\mathbf{E}}_{jk} \cdot \bar{\mathbf{E}}_{lm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}. \tag{2.1}$$

The bases $\mathbf{e}_j$ and $\bar{\mathbf{e}}_j$ are connected as follows:

$$\bar{\mathbf{e}}_j = \sum_k \alpha_{jk} \mathbf{e}_k, \tag{2.2}$$

where the α_{jk}'s are given in terms of the Euler angles θ_{jk} $(j=2, 3, \ldots, n;\ k=1, 2, \ldots, j-1)$:

$$\alpha_{jk} = \left[\prod_{p=2}^{n} \left\{ \left(\prod_{q=p}^{4} t^{(p)}_{q\,q-1}(\theta_{p\,p-q+1}) \right) \times t^{(p)}_{31}(\theta_{p\,p-2})\, t^{(p)}_{12}(\theta_{p\,p-1}) \right\} \right]_{jk}. \tag{2.3}$$

It is noted that the rotations $t^{(p)}_{12}(\theta)$, $t^{(p)}_{31}(\theta)$ are considered instead of those $t^{(p)}_{21}(\theta)$, $t^{(p)}_{32}(\theta)$ in Ref. 11 in order to make the definition of angles consistent with the usual one in a special case of $n=3$. These $t^{(p)}_{jk}(\theta)$ are $n \times n$ matrices and have the following nonzero elements:

$$(t^{(p)}_{q\,q-1}(\theta))_{q\,q-1} = -(t^{(p)}_{q\,q-1}(\theta))_{q-1\,q} = \sin\theta$$

$$
\begin{aligned}
&(t^{(\rho)}_{q\,q-1}(\theta))_{qq} = (t^{(\rho)}_{q\,q-1}(\theta))_{q-1\,q-1} = \cos\theta \qquad (q \geq 4),\\
&(t^{(\rho)}_{31}(\theta))_{31} = -(t^{(\rho)}_{31}(\theta))_{13} = \sin\theta,\\
&(t^{(\rho)}_{31}(\theta))_{11} = (t^{(\rho)}_{31}(\theta))_{33} = \cos\theta,\\
&(t^{(\rho)}_{12}(\theta))_{12} = -(t^{(\rho)}_{12}(\theta))_{21} = \sin\theta,\\
&(t^{(\rho)}_{12}(\theta))_{11} = (t^{(\rho)}_{12}(\theta))_{22} = \cos\theta,\\
&(t^{(\rho)}_{qq'}(\theta))_{rs} = \delta_{rs} \quad (r, s \neq q, q').
\end{aligned} \tag{2.4}
$$

In order to clarify the relation of the systems E_n and $\bar{E}_n$ ($\Lambda^2 E_n$ and $\Lambda^2 \bar{E}_n$), we, hereafter, assume that the bases of E_n ($\Lambda^2 E_n$) are fixed in space and those of $\bar{E}_n$ ($\Lambda^2 \bar{E}_n$) are rotating with respect to E_n ($\Lambda^2 E_n$).

An n-dimensional rotation vector $\mathbf{D}$ corresponding to the three-dimensional one can be written as a vector in $\Lambda^2 E_n$ as follows:

$$
\mathbf{D} = \tfrac{1}{2} \sum_{j,k} \mathbf{E}_{jk} D_{jk}, \tag{2.5}
$$

where the $\mathbf{E}_{jk}$ component D_{jk} is a generator of the representation corresponding to the rotation in the $(\mathbf{e}_j - \mathbf{e}_k)$ plane. The generators D_{jk} satisfy the following commutation relations:

$$
\begin{aligned}
&[D_{jk}, D_{lm}]\\
&\quad = i(\delta_{jl} D_{km} + \delta_{km} D_{jl} - \delta_{kl} D_{jm} - \delta_{jm} D_{kl}).
\end{aligned} \tag{2.6}
$$

The representation matrices corresponding to the rotation of (2.3) is given as follows

$$
\begin{aligned}
D^{(n)} &= D^{(n)}(\theta_{21}, \ldots, \theta_{n\,n-1})\\
&= R_{12}(\theta_{n\,n-1})\, R_{31}(\theta_{n\,n-2}) \left(\prod_{k=4}^{n} R_{k\,k-1}(\theta_{n\,n-k+1}) \right) D^{(n-1)}\\
&= \prod_{j=0}^{n-2} \left(R_{12}(\theta_{n-j\,n-j-1})\, R_{31}(\theta_{n-j\,n-j-2}) \right.\\
&\quad \left. \times \prod_{k=4}^{n-j} R_{k\,k-1}(\theta_{n-j\,n-j-k+1}) \right),
\end{aligned} \tag{2.7}
$$

where $R_{jk}(\theta)$ is given by $\exp(i\theta D_{jk})$.[11]

3. OPERATORS IN TERMS OF THE EULER ANGLES

In this section, we give formulas by which the infinitesimal operators of the rotation (angular momentum) are expressed in terms of the Euler angles.

It is known that the action of D_{jk} on the $D^{(n)}$ which is given by (2.7) leads to the following relations[10]:

$$
D^{(n)} D_{jk} = J_{jk} D^{(n)}, \quad D_{jk} D^{(n)} = \bar{J}_{jk} D^{(n)}, \tag{3.1}
$$

where J_{jk} and $\bar{J}_{jk}$ are differential operators corresponding to the second and the first parameter groups. Having obtained the relation (2.2) between the bases of E_n and $\bar{E}_n$, we can express these J_{jk} and $\bar{J}_{jk}$ in terms of the Euler angles and $p_{jk} = -i\partial/\partial\theta_{jk}$ as in the three-dimensional case. It follows that it is sufficient for us to give the expressions for J_{nj} and $\bar{J}_{nj}$ $(j = 1, 2, \ldots, n-1)$. As the calculation is elementary and straightforward, we give only the results.

For J_{nj} $(j = 1, 2, \ldots, n-1)$,

$$
\begin{aligned}
J_{nj} &= (-1)^{n+j-1} \Delta_{n-1\,j}\, b_{n-1} + \sum_{l=3}^{n-2} (-1)^{n+j-1} \frac{\Delta_{l\,j}}{\sin\theta_{n\,1} \sin\theta_{n-1\,1} \cdots \sin\theta_{l+2\,1}}\, b_l + (-1)^{n+j} \frac{\Delta_{1\,j}}{\sin\theta_{n\,1} \cdots \sin\theta_{41}}\, b_1\\
&\quad + (-1)^{n+j} \frac{\Delta_{2\,j}}{\sin\theta_{n\,1} \cdots \sin\theta_{31}}\, b_2,
\end{aligned} \tag{3.2}
$$

where

$$
b_1 = p_{31} + \sum_{p,q}^{n-1} (M_3)_{1\,p} (M_3)_{3\,q} J_{pq}, \quad b_2 = p_{21} + \sum_{p,q}^{n-1} (M_2)_{2\,p} (M_2)_{1\,q} J_{pq}, \quad b_j = p_{j+1\,1} + \sum_{p,q}^{n-1} (M_{j+1})_{jp} (M_{j+1})_{j+1\,q} J_{pq}
$$

$$
(3 \leq j \leq n-2), \quad b_{n-1} = p_{n1},
$$

$$
\Delta_{l\,j} = \begin{vmatrix}
(M_3)_{11} & \cdots & (M_3)_{1\,j-1} & (M_3)_{1\,j+1} & \cdots & (M_3)_{1\,n-1}\\
(M_2)_{21} & \cdots & (M_2)_{2\,j-1} & (M_2)_{2\,j+1} & \cdots & (M_2)_{2\,n-1}\\
(M_4)_{31} & \cdots & (M_4)_{3\,j-1} & (M_4)_{3\,j+1} & \cdots & (M_4)_{3\,n-1}\\
\vdots & & \vdots & \vdots & & \vdots\\
(M_l)_{l-1\,1} & \cdots & (M_l)_{l-1\,j-1} & (M_l)_{l-1\,j+1} & \cdots & (M_l)_{l-1\,n-1}\\
(M_{l+2})_{l+1\,1} & \cdots & (M_{l+2})_{l+1\,j-1} & (M_{l+2})_{l+1\,j+1} & \cdots & (M_{l+2})_{l+1\,n-1}\\
\vdots & & \vdots & \vdots & & \vdots\\
(M_n)_{n-1\,1} & \cdots & (M_n)_{n-1\,j-1} & (M_n)_{n-1\,j+1} & \cdots & (M_n)_{n-1\,n-1}
\end{vmatrix},
$$

$M_j = F_j G^{(j+1,n)}$, $F_j = \left(\prod_{l=1}^{j-4} t^{(j)}_{j-l\,j-l-1}(\theta_{j\,l+1}) \right) t^{(j)}_{31}(\theta_{j\,j-2})\, t^{(j)}_{12}(\theta_{j\,j-1})$ $(j \geq 4)$, $F_3 = t^{(3)}_{12}(\theta_{32})$, $F_2 = 1$, $G^{(j,k)} = \Gamma^{(j)} \Gamma^{(j+1)} \cdots \Gamma^{(k)}$, $\Gamma^{(j)} = t^{(j)}_{j\,j-1}(\theta_{j1})\, F_j$, $\Gamma^{(3)} = t^{(3)}_{31}(\theta_{31})\, F_3$, $\Gamma^{(2)} = t^{(2)}_{12}(\theta_{21})$.

The expressions for J_{jk} $(j,k \leq n-1)$ which also appear on the right-hand side of (3.2) are obtained from those of the $(n-1)$-dimensional case by the substitutions $\theta_{lm} \to \theta_{l+1\,m+1}$, $p_{lm} \to p_{l+1\,m+1}$ $(l, m \leq n-1)$.

For $\bar{J}_{nj}$ $(j=1, 2, \ldots, n-1)$,

$$\bar{J}_{nj} = (-1)^{n+j-1}\bar{\Delta}_{n-1\,j}\,c_{n-1} + \sum_{l=3}^{n-2}(-1)^{l+j}\frac{\bar{\Delta}_{lj}}{\sin\theta_{n1}\cdots\sin\theta_{n\,n-l-1}}c_l + (-1)^{j+1}\frac{\bar{\Delta}_{1j}}{\sin\theta_{n1}\cdots\sin\theta_{n\,n-3}}c_1 + (-1)^j\frac{\bar{\Delta}_{1j}}{\sin\theta_{n1}\cdots\sin\theta_{n\,n-2}}c_2,$$

where (3.3)

$$c_1 = p_{n\,n-2} - \sum_{p,q}^{n-1}(H_{n-3})_{p3}(H_{n-3})_{q1}\bar{J}_{pq},\quad c_2 = p_{n\,n-1} - \sum_{p,q}^{n-1}(H_{n-2})_{p1}(H_{n-2})_{q2}\bar{J}_{pq},\quad c_j = p_{n\,n-j} - \sum_{p,q}^{n-1}(H_{n-j-1})_{pj+1}(H_{n-j-1})_{qj}\bar{J}_{pq}$$

$(3 \leq j \leq n-2)$, $c_{n-1} = p_{n1}$,

$$\bar{\Delta}_{lj} = \begin{vmatrix} (H_{n-3})_{11} & \cdots & (H_{n-3})_{j-1\,1} & (H_{n-3})_{j+1\,1} & \cdots & (H_{n-3})_{n-1\,1} \\ (H_{n-2})_{12} & \cdots & (H_{n-2})_{j-1\,2} & (H_{n-2})_{j+1\,2} & \cdots & (H_{n-2})_{n-1\,2} \\ (H_{n-4})_{13} & \cdots & (H_{n-4})_{j-1\,3} & (H_{n-4})_{j+1\,3} & \cdots & (H_{n-4})_{n-1\,3} \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ (H_{n-l})_{1\,l-1} & \cdots & (H_{n-l})_{j-1\,l-1} & (H_{n-l})_{j+1\,l-1} & \cdots & (H_{n-l})_{n-1\,l-1} \\ (H_{n-l-2})_{1\,l+1} & \cdots & (H_{n-l-2})_{j-1\,l+1} & (H_{n-l-2})_{j+1\,l+1} & \cdots & (H_{n-l-2})_{n-1\,l+1} \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ (H_0)_{1\,n-1} & \cdots & (H_0)_{j-1\,n-1} & (H_0)_{j+1\,n-1} & \cdots & (H_0)_{n-1\,n-1} \end{vmatrix},$$

$$H_j = G^{(2,n-1)}Y_j,\quad Y_j = t^{(n)}_{n\,n-1}(\theta_{n1})\cdots t^{(n)}_{n-j+1\,n-j}(\theta_{nj})\quad (1 \leq j \leq n-3),\quad Y_0 = 1,\quad Y_{n-2} = Y_{n-3}t^{(n)}_{31}(\theta_{n\,n-2}).$$

The expressions for $\bar{J}_{jk}$ $(j,k \leq n-1)$ which also appear on the right-hand side of (3.3) are the same as those of the $(n-1)$-dimensional case.

It is easy to give expressions for J_{jk} and $\bar{J}_{jk}$ from (3.2) and (3.3). The expressions for $n=3, 4$ are given below.

For $n=3$,

$$J_{12} = p_{32},$$

$$J_{23} = -s_{32}p_{31} + \frac{c_{32}}{s_{31}}p_{21} - c_{32}\frac{c_{31}}{s_{31}}p_{32}, \tag{3.4}$$

$$J_{31} = c_{32}p_{31} + \frac{s_{32}}{s_{31}}p_{21} - s_{32}\frac{c_{31}}{s_{31}}p_{32},$$

$$\bar{J}_{12} = p_{21},$$

$$\bar{J}_{23} = s_{21}p_{31} - \frac{c_{21}}{s_{31}}p_{32} + c_{21}\frac{c_{31}}{s_{31}}p_{21}, \tag{3.5}$$

$$\bar{J}_{31} = c_{21}p_{31} + \frac{s_{21}}{s_{31}}p_{32} - s_{21}\frac{c_{31}}{s_{31}}p_{21}.$$

For $n=4$,

$$J_{12} = p_{43},$$

$$J_{23} = -s_{43}p_{42} + \frac{c_{43}}{s_{42}}p_{32} - c_{43}\frac{c_{42}}{s_{42}}p_{43},$$

$$J_{31} = c_{43}p_{42} + \frac{s_{43}}{s_{42}}p_{32} - s_{43}\frac{c_{42}}{s_{42}}p_{43},$$

$$J_{41} = c_{43}s_{42}p_{41} + \frac{c_{43}c_{42}c_{41}}{s_{41}}p_{42} - \frac{s_{43}c_{41}}{s_{42}s_{41}}p_{43} + \frac{1}{s_{41}}(s_{43}s_{32} - c_{43}c_{42}c_{32})p_{31} + \frac{1}{s_{42}s_{41}s_{31}}(s_{43}c_{42}c_{41}s_{31} + s_{43}s_{42}c_{32}c_{31} + c_{43}c_{42}s_{42}s_{32}c_{31})p_{32} - \frac{1}{s_{41}s_{31}}(s_{43}c_{32} + c_{43}c_{42}s_{32})p_{21}, \tag{3.6}$$

$$J_{42} = s_{43}s_{42}p_{41} + \frac{s_{43}c_{42}c_{41}}{s_{41}}p_{42} + \frac{c_{43}c_{41}}{s_{42}s_{41}}p_{43} - \frac{1}{s_{41}}(c_{43}s_{32} + s_{43}c_{42}c_{32})p_{31} - \frac{1}{s_{42}s_{41}s_{31}}(c_{43}c_{42}c_{41}s_{31} + c_{43}s_{42}c_{32}c_{31} - s_{43}c_{42}s_{42}s_{32}c_{31})p_{32} + \frac{1}{s_{41}s_{31}}(c_{43}c_{32} - s_{43}c_{42}s_{32})p_{21},$$

$$J_{43} = c_{42}p_{41} - \frac{s_{42}c_{41}}{s_{41}}p_{42} + \frac{s_{42}c_{32}}{s_{41}}p_{31} - \frac{s_{42}s_{32}c_{31}}{s_{41}s_{31}}p_{32} + \frac{s_{42}s_{32}}{s_{41}s_{31}}p_{21},$$

$\bar{J}_{12}, \bar{J}_{23}, \bar{J}_{31}$; the same expressions for $n=3$,

$$\bar{J}_{41} = -s_{31}c_{21}p_{41} + \frac{1}{s_{41}}(c_{32}c_{31}c_{21} - s_{32}s_{21})p_{42} + \frac{1}{s_{42}s_{41}}(s_{32}c_{31}c_{21} + c_{32}s_{21})p_{43} - \frac{c_{41}c_{31}c_{21}}{s_{41}}p_{31} - \frac{1}{s_{42}s_{41}s_{31}}(s_{42}c_{41}c_{31}s_{21} + c_{42}c_{32}s_{31}s_{21} + c_{42}s_{32}s_{31}c_{31}c_{21})p_{32} + \frac{c_{41}s_{21}}{s_{41}s_{31}}p_{21}, \tag{3.7}$$

$$\bar{J}_{42} = s_{31}s_{21}p_{41} - \frac{1}{s_{41}}(c_{32}c_{31}s_{21} + s_{32}c_{21})p_{42} + \frac{1}{s_{42}s_{41}}(-s_{32}c_{31}s_{21} + c_{32}c_{21})p_{43} + \frac{c_{41}c_{31}s_{21}}{s_{41}}p_{31} - \frac{1}{s_{42}s_{41}s_{31}}(s_{42}c_{41}c_{31}c_{21}$$

$$+c_{42}c_{32}s_{31}c_{21}-c_{42}s_{32}s_{31}c_{31}s_{21})p_{32}+\frac{c_{41}c_{21}}{s_{41}s_{31}}p_{21},$$

$$\bar{J}_{43}=c_{31}p_{41}+\frac{c_{32}s_{31}}{s_{41}}p_{42}+\frac{s_{32}s_{31}}{s_{42}s_{41}}p_{43}-\frac{c_{41}s_{31}}{s_{41}}p_{31}$$

$$-\frac{c_{42}s_{32}s_{31}}{s_{42}s_{41}}p_{32},$$

where c_{jk} and s_{jk} mean $\cos\theta_{jk}$ and $\sin\theta_{jk}$ respectively. It is easy to see that they satisfy the following commutation relations:

$$[J_{jk},J_{lm}]=i(\delta_{jl}J_{km}+\delta_{km}J_{jl}-\delta_{kl}J_{jm}-\delta_{jm}J_{kl}),$$
$$[\bar{J}_{jk},\bar{J}_{lm}]=-i(\delta_{jl}\bar{J}_{km}+\delta_{km}\bar{J}_{jl}-\delta_{kl}\bar{J}_{jm}-\delta_{jm}\bar{J}_{kl}),$$
$$[J_{jk},\bar{J}_{lm}]=0.$$

It is noted that only the signs on the right-hand sides of the first two relations differ from each other. This fact is known in the case of $n=3$. It is, however, seen that these relations must hold in general because of (3.1) or the relations $\bar{J}_{jk}=\sum_{l,m}\alpha_{jl}\alpha_{km}J_{lm}$ and $[J_{jk},a_l]=i(\delta_{jl}a_k-\delta_{kl}a_j)$, where a_j is a component of a vector $\mathbf{a}$. Therefore, it follows that J_{jk} and $\bar{J}_{jk}$ correspond to the angular momentum and are the $\mathbf{E}_{jk}$ and $\bar{\mathbf{E}}_{jk}$ components of an angular momentum vector $\mathbf{J}=\sum_{j,k}\mathbf{E}_{jk}J_{jk}/2$ in Λ^2E_n.

Finally, we give useful expressions for $J_{n\,n-1}$ and $\bar{J}_{n\,n-1}$ which will be sufficient to determine the d functions of $SO(n)$:

$$J^{(n)}_{n\,n-1}=\cos\theta_{n2}p_{n1}-\frac{\cos\theta_{n1}}{\sin\theta_{n1}}\sin\theta_{n2}p_{n2}+\frac{\sin\theta_{n2}}{\sin\theta_{n1}}J^{(n-1)}_{n-1\,n-2},$$

$$\bar{J}^{(n)}_{n\,n-1}=\cos\theta_{n-1\,1}p_{n1}-\frac{\cos\theta_{n1}}{\sin\theta_{n1}}\sin\theta_{n-1\,1}p_{n-1\,1}$$

$$+\frac{\sin\theta_{n-1\,1}}{\sin\theta_{n1}}\bar{J}^{(n-1)\prime}_{n-1\,n-2}\quad(n>4),$$

$$J^{(3)}_{31}=\cos\theta_{32}p_{31}-\frac{\cos\theta_{31}}{\sin\theta_{31}}\sin\theta_{32}p_{32}$$

$$+\frac{\sin\theta_{32}}{\sin\theta_{31}}J^{(2)}_{12}\quad(J^{(2)}_{12}=p_{21}),\tag{3.8}$$

$$\bar{J}^{(3)}_{31}=\cos\theta_{21}p_{31}-\frac{\cos\theta_{31}}{\sin\theta_{31}}\sin\theta_{21}p_{21}$$

$$+\frac{\sin\theta_{21}}{\sin\theta_{31}}\bar{J}^{(2)\prime}_{12}\quad(J^{(2)\prime}_{12}=p_{32}),$$

$$J^{(4)}_{43}=\cos\theta_{42}p_{41}-\frac{\cos\theta_{41}}{\sin\theta_{41}}\sin\theta_{42}p_{42}$$

$$+\frac{\sin\theta_{42}}{\sin\theta_{41}}J^{(3)}_{31},$$

$$\bar{J}^{(4)}_{43}=\cos\theta_{31}p_{41}-\frac{\cos\theta_{41}}{\sin\theta_{41}}\sin\theta_{31}p_{31}+\frac{\sin\theta_{31}}{\sin\theta_{41}}\bar{J}^{(3)\prime}_{31}.$$

Here we introduced the superscript n such as $J^{(n)}_{jk}$ in order to distinguish their spaces to which they belong. Therefore, for example, $J^{(n-1)}_{jk}$ is a generator in the $(n-1)$-dimensional space, and primes on $\bar{J}'$ mean the substitutions $\theta_{jk},p_{jk}\to\theta_{j+1\,k+1},p_{j+1\,k+1}$ in the expression of the corresponding $\bar{J}$.

4. SOME d FUNCTIONS AND FORMULA FOR THE D MATRIX ELEMENTS

In this section, explicit expressions of the d functions corresponding to the highest and some weights are obtained, and then a formula to calculate the representation matrix elements is given.

Let us first introduce the Gel'fand and Zetlin bases[9] of the unitary irreducible representations of $SO(n)$ which are classified by the canonical chain of subgroups $SO(n)\supset SO(n-1)\supset\cdots\supset SO(2)$. They are given by

$$|m_{jk}\rangle=\left|\begin{array}{cccc} m_{n1}, & m_{n2}, & \ldots, & m_{n[n/2]} \\ m_{n-1\,1}, & m_{n-1\,2}, & \ldots, & m_{n-1\,[(n-1)/2]} \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ m_{41} & m_{42} & & \\ m_{31} & & & \\ m_{21} & & & \end{array}\right\rangle,\tag{4.1}$$

where $[j/2]$ is the largest integer smaller than or equal to $j/2$. The numbers m_{jk} are simultaneously either integers or half-integers with restrictions

$$m_{2j+1,i}\geq m_{2j,i}\geq m_{2j+1,i+1}\quad(i=1,2,\ldots,j),$$
$$m_{2j,i}\geq m_{2j-1,i}\geq m_{2j,i+1}\quad(i=1,2,\ldots,j-1),$$
$$m_{2j+1,j}=|m_{2j,j}|.\tag{4.2}$$

The action of the generators $D_{j,j+1}$ $(1\leq j\leq n-1)$ on the bases is as follows:

$$D_{2k,2k+1}|m_{ij}\rangle=\sum_{j=1}^{k}A(m_{2k,j})|m_{2k,j}+1\rangle$$
$$-\sum_{j=1}^{k}A(m_{2k,j}-1)|m_{2k,j}-1\rangle,$$
$$D_{2k-1,2k}|m_{ij}\rangle=\sum_{j=1}^{k}B(m_{2k-1,j})|m_{2k-1,j}+1\rangle\tag{4.3}$$
$$-\sum_{j=1}^{k}B(m_{2k-1,j}-1)|m_{2k-1,j}-1\rangle+E_{2k}|m_{ij}\rangle.$$

In these equations, the matrix elements A, B, and E are given by

$$A(m_{2k,j})=-\frac{i}{2}\left(\frac{\prod_{r=1}^{k-1}(l_{2k-1\,r}-l_{2k\,j}-1)(l_{2k-1\,r}+l_{2k\,j})}{\prod_{i\neq j}^{k}(l^2_{2k\,i}-l^2_{2k\,j})[l^2_{2k\,i}-(l_{2k\,j}+1)^2]}\prod_{s=1}^{k}(l_{2k+1\,s}-l_{2k\,j}-1)(l_{2k+1\,s}+l_{2k\,j})\right)^{1/2},$$

$$B(m_{2k-1,j})=-i\left(\frac{\prod_{r=1}^{k-1}(l^2_{2k-2\,r}-l^2_{2k-1\,j})\prod_{s=1}^{k}(l^2_{2k\,s}-l^2_{2k-1\,j})}{l^2_{2k-1\,j}(4l^2_{2k-1\,j}-1)}\frac{1}{\prod_{i\neq j}^{k-1}(l^2_{2k-1\,i}-l^2_{2k-1\,j})[(l_{2k-1\,i}-1)^2-l^2_{2k-1\,j}]}\right)^{1/2},$$

$$E_{2k}=\frac{\prod_{r=1}^{k-1}l_{2k-2\,r}\prod_{s=1}^{k}l_{2k\,s}}{\prod_{i=1}^{k-1}l_{2k-1\,i}(l_{2k-1\,i}-1)}.\tag{4.4}$$

where l_{jk} is defined by

$$l_{jk}=m_{jk}+[(j+1)/2]-k.\tag{4.5}$$

By taking the matrix elements of (3.1) after putting $j\to n$, $k\to n-1$ and taking into account (3.8), the following relations are obtained for $n>4$:

$$\sum_{\lambda''_{n-1}}d^{(\lambda_n)}_{\lambda'_{n-1}(\lambda'_{n-2})\lambda''_{n-1}}(\theta_{n1})d^{(\lambda''_{n-1})}_{\lambda''_{n-2}(\lambda_{n-3})\lambda_{n-2}}(\theta_{n2})$$

$$\times\langle\lambda_n\lambda''_{n-1}\ \lambda_{n-2}\cdots\lambda_2\,|\,D_{n\,n-1}\,|\,\lambda_n\lambda_{n-1}\cdots\lambda_2\rangle$$
$$=\cos\theta_{n2}d^{(\lambda_{n-1})}_{\lambda'_{n-2}(\lambda_{n-3})\lambda_{n-2}}(\theta_{n2})p_{n1}d^{(\lambda_n)}_{\lambda'_{n-1}(\lambda'_{n-2})\lambda_{n-1}}(\theta_{n1})$$
$$+\frac{\cos\theta_{n1}}{\sin\theta_{n1}}\sin\theta_{n2}\sum_{\lambda''_{n-2}}\langle\lambda_{n-1}\lambda'_{n-2}\lambda_{n-3}\cdots\lambda_2$$
$$\times\,|\,D_{n-1\,n-2}\,|\,\lambda_{n-1}\lambda''_{n-2}\lambda_{n-3}\cdots\lambda_2\rangle$$
$$\times d^{(\lambda_{n-1})}_{\lambda''_{n-2}(\lambda_{n-3})\lambda_{n-2}}(\theta_{n2})d^{(\lambda_n)}_{\lambda'_{n-1}(\lambda'_{n-2})\lambda_{n-1}}(\theta_{n1})+\frac{\sin\theta_{n2}}{\sin\theta_{n1}}$$
$$\times\sum_{\lambda''_{n-2}}\langle\lambda'_{n-1}\lambda'_{n-2}\lambda_{n-3}\cdots\lambda_2\,|\,D_{n-1\,n-2}\,|\,\lambda'_{n-1}\lambda''_{n-2}\lambda_{n-3}\cdots\lambda_2\rangle$$
$$\times d^{(\lambda_{n-1})}_{\lambda''_{n-2}(\lambda_{n-3})\lambda_{n-2}}(\theta_{n2})d^{(\lambda_n)}_{\lambda'_{n-1}(\lambda''_{n-2})\lambda_{n-1}}(\theta_{n1}),\qquad(4.6)$$

$$\sum_{\lambda''_{n-1}}\langle\lambda_n\lambda'_{n-1}\cdots\lambda_2\,|\,D_{n\,n-1}\,|\,\lambda_n\lambda''_{n-1}\lambda'_{n-2}\cdots\lambda_2\rangle$$
$$\times d^{(\lambda''_{n-1})}_{\lambda'_{n-2}(\lambda_{n-3})\lambda_{n-2}}(\theta_{n-11})d^{(\lambda_n)}_{\lambda''_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\theta_{n1})$$
$$=\cos\theta_{n-1\,1}d^{(\lambda'_{n-1})}_{\lambda'_{n-2}(\lambda_{n-3})\lambda_{n-2}}(\theta_{n-1\,1})p_{n1}d^{(\lambda_n)}_{\lambda''_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\theta_{n1})$$
$$+\frac{\cos\theta_{n1}}{\sin\theta_{n1}}\sin\theta_{n-1\,1}\sum_{\lambda''_{n-2}}d^{(\lambda'_{n-1})}_{\lambda'_{n-2}(\lambda_{n-3})\lambda''_{n-2}}(\theta_{n-1\,1})$$
$$\times\langle\lambda'_{n-1}\lambda''_{n-2}\lambda_{n-3}\cdots\lambda_2\,|\,D_{n-1\,n-2}\,|\,\lambda'_{n-1}\lambda_{n-2}\cdots\lambda_2\rangle$$
$$\times d^{(\lambda_n)}_{\lambda'_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\theta_{n1})+\frac{\sin\theta_{n-1\,1}}{\sin\theta_{n1}}\sum_{\lambda''_{n-2}}d^{(\lambda'_{n-1})}_{\lambda'_{n-2}(\lambda_{n-3})\lambda''_{n-2}}(\theta_{n-1\,1})$$
$$\times\langle\lambda_{n-1}\lambda''_{n-2}\lambda_{n-3}\cdots\lambda_2\,|\,D_{n-1\,n-2}\,|\,\lambda_{n-1}\cdots\lambda_2\rangle$$
$$\times d^{(\lambda_n)}_{\lambda'_{n-1}(\lambda''_{n-2})\lambda_{n-1}}(\theta_{n1}),\qquad(4.7)$$

where the d functions are defined as follows:

$$d^{(\lambda_n)}_{\lambda'_{n-1}(\lambda_{n-2})\lambda_{n-1}}(\theta)$$
$$=\langle\lambda_n\lambda'_{n-1}\lambda_{n-2}\cdots\lambda_2\,|\,R_{n\,n-1}(\theta)\,|\,\lambda_n\lambda_{n-1}\cdots\lambda_2\rangle,$$
$$\lambda'_{n-1}\equiv(m'_{n-1\,1},m'_{n-1\,2},\ldots,m'_{n-1\,[(n-1)/2]}),$$
$$\lambda'_{n-2}\equiv(m'_{n-2\,1},m'_{n-2\,2},\ldots,m'_{n-2\,[(n-2)/2]}),\qquad(4.8)$$

and λ_j stands for the row $(m_{j1},m_{j2},\ldots,m_{j[j/2]})$ in the base (4.1) which is now written in a row on λ_j's. The summations over λ''_j in (4.6) and (4.7) are to be taken over all possible numbers determined by the relations (4.2) and (4.3). In order to derive (4.7), the following relation has been used:

$$D^{(n)}=D^{(n-1)'}\left(\prod_{k=n}^{4}R_{k\,k-1}(\theta_{k1})\right)R_{31}(\theta_{31})R_{12}(\theta_{21}),\qquad(4.9)$$

where $D^{(n-1)'}$ is given by the substitutions $\theta_{jk}\to\theta_{j+1\,k+1}$ from $D^{(n-1)}$.

If the d functions of $SO(n-1)$ are known, (4.6) and (4.7) may be rewritten as the recursion relations for the d functions of $SO(n)$ by using the orthogonal relations of the d functions.[7] They, however, can be used to determine the d functions with the highest and some special weights. If we can determine the d function with the highest weight, we will be able to express any d functions and therefore the representation matrix elements by operating the lowering operators on the d function. In this sense, the d function with the highest weight is important.

Let us, therefore, determine an expression for the d function with the highest weight. For $n=3$, it is well known that the relations hold:

$$[m_{31}\mp m_{21})(m_{31}\pm m_{21}+1)]^{1/2}d^{(m_{31})}_{m'_{21}m_{21}\pm1}(\theta)$$
$$=\left(\frac{m'_{21}-m_{21}\cos\theta}{\sin\theta}\pm\frac{d}{d\theta}\right)d^{(m_{31})}_{m'_{21}m_{21}}(\theta),$$
$$[(m_{31}\pm m'_{21})(m_{31}\mp m'_{21}+1)]^{1/2}d^{(m_{31})}_{m'_{21}\mp1\,m_{21}}(\theta)\qquad(4.10)$$
$$=\left(\frac{m'_{21}\cos\theta-m_{21}}{\sin\theta}\pm\frac{d}{d\theta}\right)d^{(m_{31})}_{m'_{21}m_{21}}(\theta).$$

From these, we obtain

$$d^{(m_{31})}_{m_{31}m_{31}}(\theta)=(\cos\theta/2)^{2m_{31}},$$
$$d^{(m_{31})}_{m_{31}-1\,m_{31}}(\theta)=-(2m_{31})^{1/2}(\sin\theta/2)(\cos\theta/2)^{2m_{31}-1},\qquad(4.11)$$
$$d^{(m_{31})}_{m_{31}\,m_{31}-1}(\theta)=(2m_{31})^{1/2}(\sin\theta/2)(\cos\theta/2)^{2m_{31}-1}.$$

For $n=4$, the following relations are easily obtained by taking the matrix elements of (3.1) and considering (3.8):

$$\sum_{m''_{31}}d^{(m_{41}m_{42})}_{m'_{31}(m'_{21})m''_{31}}(\theta_{41})d^{(m''_{31})}_{m'_{21}m_{21}}(\theta_{42})$$
$$\times\langle m_{41}m_{42};m''_{31}m_{21}\,|\,D_{43}\,|\,m_{41}m_{42};m_{31}m_{21}\rangle$$
$$=\cos\theta_{42}d^{(m_{31})}_{m'_{21}m_{21}}(\theta_{42})p_{41}d^{(m_{41}m_{42})}_{m''_{31}(m'_{21})m_{31}}(\theta_{41})$$
$$-\frac{\cos\theta_{41}}{\sin\theta_{41}}\sin\theta_{42}d^{(m_{41}m_{42})}_{m'_{31}(m''_{21})m_{31}}(\theta_{41})$$
$$\times\sum_{m''_{21}}\langle m_{31}m'_{21}\,|\,D_{31}\,|\,m_{31}m''_{21}\rangle d^{(m_{31})}_{m''_{21}m_{21}}(\theta_{42})$$
$$+\frac{\sin\theta_{42}}{\sin\theta_{41}}\sum_{m''_{21}}\langle m'_{31}m'_{21}\,|\,D_{31}\,|\,m'_{31}m''_{21}\rangle$$
$$\times d^{(m_{41}m_{42})}_{m'_{31}(m''_{21})m_{31}}(\theta_{41})d^{(m_{31})}_{m''_{21}m_{21}}(\theta_{42}),\qquad(4.12)$$

$$\sum_{m''_{31}}\langle m_{41}m_{42};m'_{31}m'_{21}\,|\,D_{43}\,|\,m_{41}m_{42};m''_{31}m'_{21}\rangle$$
$$\times d^{(m''_{31})}_{m''_{21}m_{21}}(\theta_{42})d^{(m_{41}m_{42})}_{m''_{31}(m_{21})m_{31}}(\theta_{41})$$
$$=\cos\theta_{31}d^{(m'_{31})}_{m'_{21}m_{21}}(\theta_{31})d^{(m_{41}m_{42})}_{m''_{31}(m_{21})m_{31}}(\theta_{41})$$
$$-\frac{\cos\theta_{41}}{\sin\theta_{41}}\sin\theta_{31}\sum_{m''_{21}}d^{(m'_{31})}_{m'_{21}m''_{21}}(\theta_{31})$$
$$\times\langle m'_{31}\,m''_{21}\,|\,D_{31}\,|\,m'_{31}m_{21}\rangle d^{(m_{41}m_{42})}_{m'_{31}(m_{21})m_{31}}(\theta_{41})$$
$$+\frac{\sin\theta_{31}}{\sin\theta_{41}}\sum_{m''_{21}}d^{(m'_{31})}_{m'_{21}m''_{21}}(\theta_{31})d^{(m_{41}m_{42})}_{m'_{31}(m''_{21})m_{31}}(\theta_{41})$$
$$\times\langle m_{31}m''_{21}\,|\,D_{31}\,|\,m_{31}m_{21}\rangle.\qquad(4.13)$$

The relations (4.12) and (4.13) are sufficient to determine the d functions of $SO(4)$. If we put $m_{31}=m_{21}=m'_{31}=m'_{21}=m_{41}$ and take into account (4.11), we obtain the following results:

$$d^{(m_{41}m_{42})}_{m_{41}(m_{41})m_{41}}(\theta)=\exp(-im_{42}\theta),$$
$$d^{(m_{41}m_{42})}_{m_{41}(m_{41}-1)m_{41}}(\theta)\qquad(4.14)$$
$$=(1/m_{41})(m_{41}\cos\theta+im_{42}\sin\theta)\exp(-im_{42}\theta).$$

Before determining the d functions of $SO(n)$, we introduce some notations to express the matrix elements in compact form:

$$\Lambda_j\equiv(m_{n1},m_{n2},\ldots,m_{n\,[(j-2)/2]}),$$
$$\Lambda_{j\,[j/2]}\equiv(m_{n1},m_{n2},\ldots,m_{n\,[(j-2)/2]},m_{j\,[j/2]}),$$
$$\Lambda_{j\,[j/2]'}\equiv(m_{n1},m_{n2},\ldots,m_{n\,[(j-2)/2]},m'_{j\,[j/2]}),$$
$$\Lambda_{(j)}\equiv(m_{n1},m_{n2},\ldots,m_{n\,[(j-2)/2]},m_{n\,[j/2]}-1).$$

That is, Λ_j means that the numbers m_{jk} in λ_j take maximum values which characterize the representa-

tions, while $\Lambda_{j\,[j/2]}$, $\Lambda_{j\,[j/2]'}$, and $\Lambda_{(j)}$ mean that these numbers except for the last one take maximum values. It is noted that we must take $n \to n+1$ in these Λ's when $SO(n+1)$ is considered.

Now let us determine some d functions by induction. First we assume the d functions with the highest and other two weights for n odd as follows:

$$d^{(\lambda_n)}_{\Lambda_{n-1}(\Lambda_{n-2})\Lambda_{n-1}}(\theta) = (\cos\theta/2)^{2m_{n(n-1)/2}},$$

$$d^{(\lambda_n)}_{\Lambda_{(n-1)}(\Lambda_{n-2})\Lambda_{n-1}}(\theta)$$

$$= -(2m_{n(n-1)/2})^{1/2}\sin\theta/2\,(\cos\theta/2)^{2m_{n(n-1)/2}-1}, \qquad (4.15)$$

$$d^{(\lambda_n)}_{\Lambda_{n-1}(\Lambda_{n-2})\Lambda_{(n-1)}}(\theta)$$

$$= (2m_{n(n-1)/2})^{1/2}\sin\theta/2\,(\cos\theta/2)^{2m_{n(n-1)/2}-1}.$$

The following relation is obtained from (4.7) for $n+1$ even:

$$\langle\lambda_{n+1}\Lambda_n\cdots\Lambda_2 \mid D_{n+1\,n} \mid \lambda_{n+1}\Lambda_n\cdots\Lambda_2\rangle$$

$$\times d^{(\Lambda_n)}_{\Lambda_{n-1}(\Lambda_{n-1})}(\theta_{n+1\,2})\, d^{(\lambda_{n+1})}_{\Lambda_n(\Lambda_{n-1})\Lambda_n}(\theta_{n+1\,1})$$

$$= \cos\theta_{n+1\,2}\, d^{(\Lambda_n)}_{\Lambda_{n-1}(\Lambda_{n-2})\Lambda_{n-1}}(\theta_{n+1\,2}) \times p_{n+1\,1}\, d^{(\lambda_{n+1})}_{\Lambda_n(\Lambda_{n-1})\Lambda_n}(\theta_{n+1\,1})$$

$$+ \frac{\cos\theta_{n+1\,1}}{\sin\theta_{n+1\,1}}$$

$$\times \sin\theta_{n+12}\, d^{(\Lambda_n)}_{\Lambda_{n-1}(\Lambda_{n-2})\Lambda_{(n-1)}}(\theta_{n+1\,2})$$

$$\times \langle\Lambda_n\Lambda_{(n-1)}\Lambda_{n-2}\cdots\Lambda_2 \mid D_{n\,n-1} \mid \Lambda_n\Lambda_{n-1}\cdots\Lambda_2\rangle$$

$$\times d^{(\lambda_{n+1})}_{\Lambda_n(\Lambda_{n-1})\Lambda_n}(\theta_{n1}) + \frac{\sin\theta_{n+1\,2}}{\sin\theta_{n+1\,1}}\, d^{(\Lambda_n)}_{\Lambda_{n-1}(\Lambda_{n-2})\Lambda_{(n-1)}}(\theta_{n+1\,2})$$

$$\times d^{(\lambda_{n+1})}_{\Lambda_n(\Lambda_{(n-1)})\Lambda_n}(\theta_{n+1\,1})$$

$$\times \langle\Lambda_n\Lambda_{(n-1)}\Lambda_{n-2}\cdots\Lambda_2 \mid D_{n\,n-1} \mid \Lambda_n\Lambda_{n-1}\cdots\Lambda_2\rangle, \qquad (4.16)$$

where the restrictions (4.2) for the numbers m_{j_k} and the relations (4.3) are taken into account. From (4.16), we obtain the following results by considering (4.15) and the explicit expressions (4.4):

$$d^{(\lambda_{n+1})}_{\Lambda_n(\Lambda_{n-1})\Lambda_n}(\theta) = \exp(-i\theta m_{n+1\,(n+1)/2}),$$

$$d^{(\lambda_{n+1})}_{\Lambda_n(\Lambda_{(n-1)})\Lambda_n}(\theta) = \frac{1}{m_{n+1\,(n-1)/2}}\,(m_{n+1\,(n-1)/2}\cos\theta$$

$$+ im_{n+1\,(n+1)/2}\sin\theta)\exp(-i\theta m_{n+1\,(n+2)/2}). \qquad (4.17)$$

(4.6) gives the same expressions as (4.17). Similarly, it is easy to obtain the corresponding expressions (4.15) by assuming the forms of (4.17).

Therefore, by induction we could obtain the expressions for the d functions with the highest and some weights. Similarly, we can obtain the following results from (4.6) and (4.7) by using the expressions (4.15) and (4.17).

For n even,

$$d^{(\lambda_n)}_{\Lambda_{n-1}\,(n-2)'/2(\Lambda_{n-2(n-2)/2})\Lambda_{n-1}(n-2)/2}(\theta)$$

$$= \sum_m (j_1m_1, j_2m_2; m'_{n-1\,(n-2)/2}m_{n-2(n-2)/2})$$

$$\times (j_1m_1, j_2m_2; m_{n-1\,(n-2)/2}m_{n-2(n-2)/2})\exp(-i\theta m), \qquad (4.18)$$

where

$$j_1 = \tfrac{1}{2}(m_{n\,(n-2)/2} + m_{n\,n/2}),\quad j_2 = \tfrac{1}{2}(m_{n\,(n-2)/2} - m_{n\,n/2}),$$

$$m_1 = \tfrac{1}{2}(m_{n-2\,(n-2)/2} + m),\quad m_2 = \tfrac{1}{2}(m_{n-2(n-2)/2} - m),$$

and $(j_1m_1,\ j_2m_2;jm)$ is the Clebsch–Gordan coefficient.

For n odd,

$$d^{(\lambda_n)}_{\Lambda_{n-1}\,(n-1)'/2(\Lambda_{n-2})\Lambda_{n-1}\,(n-1)/2}(\theta)$$

$$= d^{(m_{n(n-1)/2})}_{m'_{n-1}\,(n-1)/2\,m_{n-1}\,(n-1)/2}(\theta), \qquad (4.19)$$

where $d^{(j)}_{m',m}(\theta)$ is the usual d function for the three-dimensional case. (4.18) and (4.19) agree with the d functions of the four- and three-dimensional cases respectively.

Finally, we give a useful formula of computing the representation matrix elements $(D^{(n)})$ of $SO(n)$. It follows from the above discussions that the $D^{(n)}$-matrix element with the highest weight is given by

$$\langle\lambda_n\Lambda_{n-1}\circ\circ\circ\Lambda_2 \mid D^{(n)} \mid \lambda_n\Lambda_{n-1}\cdots\Lambda_2\rangle$$

$$= d^{(m_{n1})}(\theta_{n\,n-1})\, d^{(m_{n1})}_{m_{n1}m_{n1}}(\theta_{n\,n-2})\times\cdots\times d^{(\lambda_n)}_{\Lambda_{n-1}(\Lambda_{n-2})\Lambda_{n-1}}(\theta_{n1})\langle\lambda_n\Lambda_{n-1}$$

$$\circ\circ\circ\Lambda_2 \mid D^{(n-1)} \mid \lambda_n\Lambda_{n-1}\circ\circ\circ\Lambda_2\rangle$$

$$= d^{(\lambda_n)}_{\Lambda_{n-1}(\Lambda_{n-2})\Lambda_{n-1}}(\theta_{n1}) \prod_{j=2}^{n-1} d^{(\Lambda_j)}_{\Lambda_{j-1}(\Lambda_{j-2})\Lambda_{j-1}}(\theta_{n\,n-j+1})$$

$$\times \prod_{j=n-1}^{2}\left(\prod_{i=2}^{j} d^{(\Lambda_i)}_{\Lambda_{i-1}(\Lambda_{i-2})\Lambda_{i-1}}(\theta_{j\,j-i+1})\right), \qquad (4.20)$$

where

$$d^{(\Lambda_2)}_{\Lambda_1(\Lambda_0)\Lambda_1}(\theta) = d^{(\Lambda_2)}(\theta) = d^{(m_{n1})}(\theta) = \exp(i\theta m_{n1}),$$

$$d^{(\Lambda_3)}_{\Lambda_2(\Lambda_1)\Lambda_2}(\theta) = d^{(\Lambda_3)}_{\Lambda_2\Lambda_2}(\theta),$$

and the explicit expression for (4.20) can be easily given by using (4.15) and (4.17). Therefore, the general $D^{(n)}$-matrix elements are obtained by operating the lowering operators on the $D^{(n)}$-matrix element (4.20):

$$D^{(\lambda_n)}_{(\lambda'_{n-1}\cdots\lambda'_2)(\lambda_{n-1}\cdots\lambda_2)}(\theta_{21},\theta_{31},\theta_{32},\ldots,\theta_{n1},\ldots,\theta_{n\,n-1})$$

$$= \langle\lambda_n\lambda'_{n-1}\cdots\lambda'_2 \mid D^{(n)} \mid \lambda_n\lambda_{n-1}\cdots\lambda_2\rangle$$

$$= \left(\prod_{s=3}^{n}\prod_{t=1}^{[(s-1)/2]} \bar{U}^{m_{s-t}m'_{s-1},t}_{s,-t}\right)\left(\prod_{i=3}^{n}\prod_{j=1}^{[(i-1)/2]} U^{m_{ij}-m_{i-1},j}_{i,-j}\right)$$

$$\times\langle\lambda_n\Lambda_{n-1}\cdots\Lambda_2 \mid D^{(n)} \mid \lambda_n\Lambda_{n-1}\circ\circ\circ\Lambda_2\rangle, \qquad (4.21)$$

where $\lambda'_j = (m'_{j1}, m'_{j2}, \ldots, m'_{j\,[j/2]})$, $\lambda_n = \lambda'_n$ and the normalized lowering operators $U_{j,-k}$ and $\bar{U}_{j,-k}$ are obtained respectively by replacing the generators D_{j_k} appearing in the definition of the lowering and raising operators of Pang and Hecht[12] by the differential operators J_{j_k} and $\bar{J}_{j_k}$ given in Sec. 3.

Thus we conclude that the representation matrix elements of $SO(n)$ are calculated by (4.21) through the action of the lowering operators on the D-matrix element (4.20) with the highest weight. The only task for us is to express J_{jk} and $\bar{J}_{jk}$ of (3.2) and (3.3) in terms of the differential operators and then to construct the lowering operators by these. The calculation of (3.2) and (3.3) is elementary but lengthy. It is, however, expected that the representation matrix elements and d functions in general have compact expressions, for the recursion relations (4.6) and (4.7) have the similar and simple forms.

[1]D. Z. Freedman and J. M. Wang, Phys. Rev. 160, 1560 (1967).
[2]J. W. Holman III, J. Math. Phys. 10, 1710 (1969).
[3]Ya. A. Smorodinskii and G. I. Shepelev, Yad. Fiz. 13, 441 (1971), [Sov. J. Nucl. Phys. 13, 248 (1971)].
[4]B. H. Flowers and S. Szpikowski, Pro. Phys. Soc. (London) 84, 193 (1964); J. C. Parikh, Nucl. Phys. 63, 214 (1965); J. N. Ginsecchio, *ibid.* 74, 321 (1965); K. T. Hecht, Phys. Rev. 139, B794 (1965).
[5]P. Kramer and M. Moshinsky, Nucl. Phys. 82, 241 (1966); J. D. Louck and J. W. Galbraith, Rev. Mod. Phys. 44, 540 (1972).
[6]N. Ja. Vilenkin, Dokl. Akad. Nauk SSSR 113, 16 (1957); E. N. Lambina, Dokl. Akad. Nauk BSSR 9, 777 (1965); N. Ja. Vilenkin, *Special Functions and the Theory of Group Representations*, Math. Monographs Vol. 22, transl. by V. N. Singh (Am. Math. Soc., Providence, R. I., 1968); J. Fisher, J. Niederle, and R. Raczka, J. Math. Phys. 7, 816 (1966); R. Raczka, N. Limic, and J. Niederle, *ibid.*, 1861 (1966); N. Limic, J. Niederle, and R. Raczka, *ibid.*, 2026 (1966); 8, 1079 (1967); R. Delbourgo, J. Koller, and R. M. Williams, *ibid.* 10, 957 (1969).
[7]K. B. Wolf, J. Math. Phys. 12, 197 (1971).
[8]For example, U. Fano, and G. Racah, *Irreducible Tensorial Sets* (Academic, New York, 1958).
[9]I. M. Gel'fand and M. L. Zetlin, Dokl. Akad. Nauk SSSR 71, 1017 (1950).
[10]G. Racah, Group Theory and Spectroscopy, lecture notes (Princeton, 1951).
[11]T. Maekawa (to be published).
[12]S. C. Pang and K. T. Hecht, J. Math. Phys. 8, 1233 (1967); M. K. F. Wong, J. Math. Phys. 8, 1899 (1967).

The Sagnac effect in general relativity

Abhay Ashtekar*† and Anne Magnon‡

Enrico Fermi Institute, University of Chicago, Chigaco, Illinois
(Received 11 March 1974; final revised manuscript received 13 May 1974)

The Sagnac effect is a phase shift observed between two beams of light traversing in opposite directions the same closed path around a rotating object. A description of this experiment is obtained within the context of general relativity. In this context the effect provides an operational definition of rotation. An expression for the magnitude of the phase shift is derived under fairly general conditions. The general definition of rotation provided by this experiment is shown to reduce, in certain particular cases, to the usual definitions available. It is observed that the Sagnac effect represents a gravitational analog of the Aharanov–Bohm effect in electrodynamics.

1. INTRODUCTION

Consider a hollow toroidal tube of glass with perfectly reflecting internal walls. Let a half-silvered mirror be fixed to this tube at some point. Through this mirror, shine, around the tube, two rays of monochromatic light in opposite directions. Observe the interference pattern of the rays emerging out of the half-silvered mirror. It is clear, e.g., from the symmetry considerations, that when the two rays meet again at the mirror, they will be in phase. Let the tube now rotate with a certain angular speed. Then the two rays will be no longer in phase when they meet, and the phase difference will be a measure of the angular speed of the tube. Such an experiment was first performed by Sagnac.[1]

Within the framework of Newtonian mechanics[2] it is straightforward to obtain an expression for the Sagnac phase shift: if R denotes the radius of the tube, ω its angular speed, and ν the frequency of the light used, then this phase shift is $8\pi^2\nu\omega R^2$.[3] Thus in the Newtonian interpretation the Sagnac shift is directly proportional to the angular speed of the tube.

What is the situation in general relativity? Note, first, that it is meaningful to analyze this experiment in the framework of general relativity, since rotation is an unambiguous and absolute concept in this theory. However, most of the usual descriptions of this concept within the theory refer to particular contexts and their respective descriptions often seem to be unrelated. For example, one describes the rotation of a stationary massive object via the twist of the timelike Killing vector field associated with the object, whereas the rotation of an infinitesimal object is described by the fact that the spatial triad fixed to the object is not Fermi transported along its worldline. An analysis of the Sagnac experiment in the framework of general relativity would provide a general description of rotation within the theory; e.g., to know whether or not a given object rotates one can fix a Sagnac tube to the object and look for the phase shift. In particular, by performing the Sagnac experiment in the appropriate contexts one can compare the usual description of rotation which arises in each of these contexts with the one obtained from the Sagnac experiment.

In the next section, we begin with a geometrical description of the experiment, a description which is suitable for its analysis in general relativity. The analysis itself is carried out only under certain restrictive conditions. However, these conditions are sufficiently general so that the results of the analysis are valid in the usual contexts in which one describes rotating objects in general relativity. By choosing suitable special cases we compare the Sagnac criterion for presence or absence of rotation with the usual criteria available in the respective contexts. Finally, we point out a formal analogy between the Sagnac experiment and the Aharanov–Bohm experiment.

2. ANALYSIS IN GENERAL RELATIVITY

Let $(\mathcal{M}, g_{ab})$ be a space–time, i.e., let $\mathcal{M}$ be a 4-manifold with a metric g_{ab} of signature $(-+++)$. The Sagnac tube is represented by a two-dimensional timelike submanifold μ of $\mathcal{M}$ (see figure). On this μ, the mirror is represented by a timelike curve M, the event at which the light rays are first shone by a point p of this curve M, and the two light rays by the null curves C^+ and C^- through this p. The fact that the two rays are of frequency ν with respect to the mirror means $\nu = \kappa_a\xi^a = \kappa'_a\xi^a$, where ξ^a, κ^a, and κ'^a are, respectively, the tangent vectors to the curves M, C^+, and C^-. The events m and m' at which the two rays meet the mirror after once traveling around the tube are represented by the points where C^+ and C^- first intersect the curve M. The Sagnac shift $\Delta\tau$ in this space–time picture is given by the distance between the events m and m', measured along the world line M of the mirror.

Given an arbitrary Sagnac tube in any particular space–time, the Sagnac shift $\Delta\tau$ can always be computed. However, this shift will generally exhibit a complicated dependence upon the details of the geometry of the 2-manifold μ representing the tube. This is analogous to the following situation in the Newtonian analysis: Given any arbitrary Sagnac tube, not necessarily in uniform rotation, one can always compute the Sagnac shift, although the analysis becomes somewhat complicated. A convenient simplification arises in the Newtonian analysis if one restricts oneself to uniformly rotating tubes. Similarly, in the general relativistic framework one may expect a simplification to occur if the tube is in a "stationary" motion. This expectation is indeed correct. More precisely, it is easy to obtain a simple explicit formula for the Sagnac shift if the 2-manifold μ admits a timelike Killing vector field and if the tube moves along the trajectories of this Killing field.[4] Throughout this paper, we shall restrict ourselves to such Sagnac tubes.

Let h_{ab} denote the metric induced on μ by the metric g_{ab} on $\mathcal{M}$. Denote by t^a the timelike Killing vector field on (μ, h_{ab}) (Ref. 5) whose integral curves are being followed

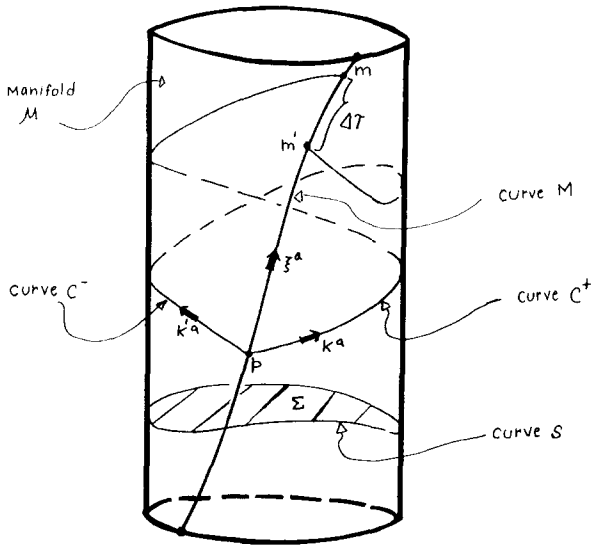


FIG. 1. The two-dimensional manifold μ represents the Sagnac tube, the curve M is the worldline of the mirror and the curves C^+ and C^- are the two light rays. The closed curve S on μ is the boundary of the 2-surface Σ in the space–time manifold $\mathcal{M}$.

by the tube. Thus, if $l^a l_a = -\lambda$, then the 4-velocity ξ^a of the tube is given by $\lambda^{-1/2} l^a$. Since μ is a 2-manifold and h_{ab} a metric of signature $(-+)$, at each point of μ there are exactly two null directions. Hence, once the event p is chosen, the null curves C^+ and C^-, events m and m', and the Sagnac shift $\Delta\tau$ are completely determined. However, it is convenient to introduce, in addition, the null vector fields κ^a and κ'^a tangent to C^+ and C^-. Choose κ^a and κ'^a such that $\kappa^a l_a = \kappa'^a l_a = -1$.[6] It is easy to check that these vector fields are curl-free; hence the integrals[7] $I = \oint_c \kappa_a dS^a$ and $I' = \oint_c \kappa'_a dS^a$ are independent of the particular choice of the closed curve[8] C on μ. Choose for C, in the definition of I, the closed curve pmp obtained by moving along C^+ from p to m and along M from m to p, and, in the definition of I', the closed curve $pm'p$ obtained by moving along M from p to m' and along C^- from m' to p (see figure). Then, since κ^a and κ'^a are null, $I = \int_m^p \kappa_a dS^a$ and $I' = \int_p^{m'} \kappa'_a dS^a$, the integrals being evaluated along M. Hence $I + I' = (\lambda_M)^{-1/2} \Delta\tau$, where λ_M is the value of the scalar field λ along the worldline M of the mirror, and $\Delta\tau$ is the distance between the events m and m' evaluated also along M. Now choose for C, in the definition of I and I', any closed curve S on μ (see figure).[8] Then $I + I' = 2\oint_S \lambda^{-1} l_a dS^a$. Hence the Sagnac shift is given by

$$\Delta\tau = 2(\lambda_M)^{1/2} \oint_S \lambda^{-1} l_a dS^a \tag{1}$$

$$= 2(\nu_M)^{-1} \oint_S \nu \xi_a dS^a \tag{2}$$

where $\nu = (\kappa^a \xi_a = \kappa'^a \xi_a)$ is the frequency of the light as seen by ξ^a, and ν_M its value along the worldline M of the mirror.[9] The integrals in the above expressions are independent of the particular choice of the closed curve S. Thus, the Sagnac shift $\Delta\tau$ is completely determined by the value ν_M of the frequency of the light as seen by the mirror. Thus, if we change the location of the mirror on the tube, the value of the Sagnac shift will also change. In the Newtonian analysis, on the other hand, this shift is independent of the location of the mirror on the tube. The difference in the two predictions arises as follows: Whereas in general relativity the frequency of the light rays remains constant only along the worldline of a point on the tube, but changes from one point of the tube to another, in the Newtonian framework the frequency is constant everywhere on the tube.

We now consider certain situations which arise when describing rotating objects in general relativity, and using Eq. (1) obtain an expression for the Sagnac shift in each of these cases. These expressions will show explicitly how the Sagnac criterion for rotation is related to the usual ones.

Case 1: Consider, as the first example, a stationary object. In general relativity such an object is described by a space–time $(\mathcal{M}, g_{ab})$ with a timelike Killing vector field l^a and the state of rotation of this object by the twist of this l^a. To relate the Sagnac criterion of rotation with the presence or absence of twist, it is convenient to let each point of the tube follow a trajectory of this Killing field.[10] For such a tube our previous analysis holds and the Sagnac shift is given by Eq. (1).

Let Σ denote any two-dimensional surface in space–time with boundary S (see figure). Using Stoke's theorem, it then follows from Eq. (1) that the Sagnac shift is given by

$$\Delta\tau = 2(\lambda_M)^{1/2} \int_\Sigma \nabla_{[a} \lambda^{-1} l_{b]} dS^{ab}$$

$$= (\lambda_M)^{1/2} \int_\Sigma \lambda^{-3/2} \omega^a \epsilon_{abc} dS^{bc} \tag{3}$$

where ∇_a denotes any derivative operator on $\mathcal{M}$, ϵ_{abc} is defined in terms of the alternating tensor ϵ_{abcd} on $(\mathcal{M}, g_{ab})$ by $\epsilon_{abc} = \epsilon_{abcd} \lambda^{-1/2} l^d$, and $\omega^a = \epsilon^{abcd} l_b \nabla_c l_d$ is the twist of the Killing field l^a. Thus the Sagnac shift may be regarded as a measure of the flux of ($\lambda^{-3/2}$ times) the twist of the Killing field through the tube. Unfortunately, within the exact framework of general relativity, there is no simple, direct, quantitative relationship between the geometrical properties of the Killing field, e.g., its twist, and the properties of the source, e.g., its angular momentum.[11] Hence, in general, we can draw only qualitative conclusions about the state of rotation of the source: If the source is static, the Killing field is twist free hence the Sagnac shift vanishes and, conversely, if the shift vanishes for arbitrary Sagnac tubes following Killing trajectories, the twist must vanish, and therefore the source must be static.

Case 2: Consider, as a second example, a stationary axisymmetric object. In general relativity one describes such an object by a space–time with two commuting Killing vector fields; one timelike (in some neighborhood of spatial infinity) and one rotational. Let these fields be denoted by T^a and R^a, respectively.[12] The state of rotation of the object is described by the scalar product $T^a R_a$, e.g., this scalar product vanishes if and only if the object is static. To compare the Sagnac criterion for presence of rotation for such an object with this usual one, it is convenient to arrange the experiment as follows: Let the 2-manifold μ representing the Sagnac tube be an integral manifold of the two Killing fields T^a and R^a, and let each point of the tube follow a trajectory of T^a.[13] Then our previous analysis holds and the Sagnac shift $\Delta\tau$ is given by Eq. (1). Choosing for the closed curve S in Eq. (1) an integral curve of the rota-

tional Killing field R^a, one obtains

$$\Delta\tau = [2\lambda^{-1/2}(R^aR_a)^{-1/2}\ (\text{length of } S)](T^aR_a). \tag{4}$$

It is clear from Eq. (4) that the Sagnac shift vanishes if and only if the object is static (i.e., $T^aR_a = 0$), and thus the Sagnac criterion for rotation of this object agrees with the usual one.

Consider a rotating object, i.e., let $T^aR_a \neq 0$. In the gravitational field of such an object, how would a Sagnac tube have to move in order to register zero shift? We claim that it must move along the integral curves of the locally nonrotating vector field[14] $l^a = T^a - [(T\cdot R)/(R\cdot R)]R^a$. Note first that this l^a is a timelike Killing field on the 2-manifold (μ, h_{ab}). Hence the tube moving along l^a satisfies the assumptions we made in the beginning of this section, so that $\Delta\tau = [2\lambda^{-1/2}(R\cdot R)^{-1/2}$ (length of S)$](l^aR_a)$. But the only direction orthogonal to R^a in the 2-manifold μ is that of l^a. Thus $\Delta\tau = 0$ if and only if the tube follows the locally nonrotating vector field. This result reflects the fact that rotation is a "local" concept in general relativity: the shift vanishes only when the tube is at rest with respect to the local zero angular momentum observers and not when it is at rest with respect to the static observer at infinity.

As particular examples of this case, we consider a Sagnac tube in a Kerr space—time and in the Minkowski space—time. In a Kerr space—time, choose for μ the integral manifold of T^a and R^a defined by $r = r_0$, $\theta = \theta_0$ in Boyer—Lindquist coordinates. If we further require that the 4-velocity of the tube be a multiple of T^a, then we obtain

$$\Delta\tau = 4\pi(4A_0M_0r_0\sin^2\theta_0)(r_0^2 + A_0^2\cos^2\theta_0 - 2M_0r_0)^{-1/2} \times (r_0^2 + A_0^2\cos^2\theta_0)^{-1/2}.$$

Note that this shift is nonzero although the tube is at rest with respect to the static observer at infinity. In Minkowski space—time, if $r = r_0$ and $\theta = \theta_0$ define the 2-manifold μ, and if the tube rotates with a uniform angular speed ω, then Eq. (4) reduces to

$$\Delta\tau = 4\pi\omega[r_0\sin\theta_0]^2[1 - \omega^2r_0^2\sin^2\theta_0]^{-1/2}.$$

Note that, since the nonrotating frames in Minkowski space are global, the tube is in a state of rotation with respect to both the local zero angular momentum observer and the static observer at infinity. This reflects the fact that rotation is a global concept in special relativity. (In the limit $\omega^2r_0^2 \ll 1$ we recover from the last formula the Newtonian expression $\Delta\tau = 4\pi\omega r_0^2\sin^2\theta_0$).

Case 3: In a given space—time $(\mathcal{M}, g_{ab})$ consider as the last example a test body whose dimensions are small compared to the radius of curvature of this space—time.[15] The notion of Fermi transport enables one to decide whether or not such a body rotates. We now assume that this space—time admits a timelike Killing vector field l^a, choose for the test body a small Sagnac tube, and obtain an expression for the shift.

Let the tube have radius δ and let it follow the integral curves of the timelike Killing field l^a, i.e., if $\lambda = -l^at_a$, the 4-velocity field ξ^a of the tube is given by $\xi^a = \lambda^{-1/2}l^a$. Fix any two points on the tube and denote by U^a and V^a the "vector fields"[16] joining the center of the tube to these two points. Then $\mathcal{L}_\xi U^a = \mathcal{L}_\xi V^a = 0$. It is convenient to orient the tube such that U^a and V^a are orthogonal to $\nabla^a\lambda$ (in addition to being orthogonal to ξ^a). In this case the Sagnac shift of Eq. (3) reduces to

$$\Delta\tau = 2\pi\delta^2\epsilon_{abcd}W^bU^c\xi^d(\xi^m\nabla_mU^a) \tag{5}$$

where $W^a = [(\nabla^m\lambda)(\nabla_m\lambda)]^{-1/2}\nabla^a\lambda$, and ∇_a is the derivative operator on $(\mathcal{M}, g_{ab})$. It is clear from Eq. (5) that the shift vanishes if U^a is Fermi transported, i.e., if $\xi^m\nabla_mU^a = \xi^a(U^m\xi^b\nabla_b\xi_m)$. We now consider the converse. Let $\Delta\tau = 0$. Then it follows from Eq. (3) that the twist ω_a of the Killing field l^a is orthogonal to $\nabla^a\lambda$ and hence ω_a is both Lie and Fermi transported by ξ^a. Using this fact and the linear independence of U^a and V^a, it is straightforward to show that both these vectors are Fermi transported along ξ^a. Thus, the Sagnac test of rotation, a global test in general,[17] reduces, on choosing an infinitesimal Sagnac tube, to the usual local test of rotation provided by Fermi transport.

Remark: There is a sense in which the Sagnac experiment represents a gravitational analog of Aharanov—Bohm experiment in electrodynamics. Let $(\mathcal{M}, g_{ab})$ be a stationary space—time with a timelike Killing vector field l^a. One often thinks of the gradient of the norm $\nabla_a\lambda$ of this l^a and of its twist ω_a as representing, respectively, the electric and the magnetic parts of a natural, test, source-free Maxwell field $F_{ab} = \nabla_al_b$ on $\mathcal{M}$.[18] From this viewpoint, the Sagnac shift may be regarded as a measure of the flux, through the tube, of the "natural magnetic field" associated with the Killing vector [see Eq. (3)].

The following experiment may be regarded as an electromagnetic analog of the present gravitational experiment. In Minkowski space—time shine a coherent beam of electrons. Put an obstacle in their path, splitting the beam into two parts, and observe the usual (double slit) interference pattern. Then switch on a magnetic field in a region of space—time forbidden to the electrons by the obstacle. The interference pattern will be shifted. This, which was first predicted by Aharanov and Bohm,[19] may also be regarded as a measure of the flux of the magnetic field through the loop determined by the two beams.

Although the two experiments are intuitively similar, their respective analyses in general relativity and in quantum mechanics are very different from one another. It is the final expressions for the shifts which are formally analogous.

ACKNOWLEDGMENTS

We are grateful to R. Geroch for his constant encouragement and for suggesting several coordinate-free arguments. We wish to thank S. Chandrasekhar and J. Ipser for reading the manuscript and for their valuable comments. One of us (A.M.) thanks V.L. Telegdi for fruitful discussions.

*Supported in part by the National Science Foundation contract GP34721X1.

†Present address: Mathematical Institute, Oxford University, Oxford, England.

†Present address: Département de Mathématiques, Complexe Universitaire des Cézeaux, Clermont-Ferrand, France.

M.G. Sagnac, C.R. Acad. Sci. (Paris), **157**, 708, 1410. (1913). This experiment is also described in most text books on relativity theory. See, e.g., Landau & Lifshitz, *Classical Theory of Fields* (Addison-Wesley, Reading, Mass., 1969).

[2]We mean the framework of Newtonian mechanics together with the hypothesis of constancy of speed of light.

[3]Throughout this paper we use units where $c = G = 1$.

[4]Since the glass tube is rigid, it is natural to demand that the 4-velocity vector field, ξ^a of the tube be Born-rigid on μ[i.e., $\mathcal{L}_\xi(g_{ab} + \xi_a\xi_b) = 0$]. This requirement is automatically satisfied if the tube moves along the trajectories of a Killing field.

[5]Note that we only require that t^a be a timelike Killing vector field on (μ, h_{ab}). The space-time $(\mathcal{M}, g_{ab})$ may not have any Killing field [e.g., the locally nonrotating vector field which we use in Case 2 is a Killing field on (μ, h_{ab}) but not on $(\mathcal{M}, g_{ab})$].

[6]Note that since $\nu = -\lambda^{-1/2}$ and since the scalar field λ remains constant along the integral curves of t^a, the frequency of the light rays represented by C^+ and C^- remains constant along the worldline of the mirror.

[7]It is convenient, on an n-manifold, only to integrate n-forms, so that the volume element is independent of the metric.

[8]The curve C appearing in the definitions of I and I' is an element of the first homology group on μ. The curve S in Eq. (1) is homologous to C.

[9]An interesting geometrical property of (μ, h_{ab}) is the following. Suppose we start at the event m' (see Fig. 1.) and move orthogonal to the Killing field t^a everywhere. Let m'' be the event (in future of m') where we would first meet the worldline M of the mirror. Then m'' is the midpoint of m and m' on M. To see this, choose for S in Eq. (1) the closed curve $m'm''m'$ obtained by moving orthogonal to t^a from m' to m'', and along M from m'' to m'. Then, since the tangent vector to the curve from m' to m'' is orthogonal to t^a, $\Delta\tau = 2\lambda_M^{1/2}\int_{m'}^{m''}\lambda^{-1}t_a dS^a = 2$ (distance between m'' and m' as measured along M). The result is immediate if one recalls that $\Delta\tau$ is the distance between m and m' measured along M.

[10]In particular, this could be done by fixing the tube to the stationary object. Note also that by requiring the tube to follow trajectories of t^a we have restricted ourselves to the experiments which can be performed only outside the ergosphere.

[11]However, in the post-Newtonian approximations there does exist such a relationship. For example, to the first post-Newtonian approximation $\Delta\tau$ is essentially the flux of the angular momentum of the source through the tube. For details see, e.g., S. Chandrasekhar in *Relativity*, edited by Carmeli, Fickler, Witten (Plenum, New York, 1970).

[12]T^a is that Killing vector field which is timelike at spatial infinity. We have assumed throughout that the two Killing vectors commute, i.e., $\mathcal{L}_R T^a = 0$. If there are no other independent Killing vectors and if there is an open region in which T^a is timelike and R^a spacelike with closed integral curves, this condition is always satisfied. Proof: let $\mathcal{L}_R T^a = aT^a + bR^a$, with a, b, constants. Consider the scalar field $\alpha = (aT^a - 2bR^a)T_a$ then $\mathcal{L}_R\alpha = 2a^2T^aT_a - 2b^2R^aR_a \leq 0$. But the integral curves of R^a are closed. Since α is a well-defined (single-valued) function, $\mathcal{L}_R\alpha = 0$, i.e., $a = b = 0$. Thus the two Killing fields commute in an open region and hence everywhere. This result is due to P.S. Jang (private communication).

[13]Note that we have let each point on the tube evolve along a trajectory of T^a only because we wish to compare the Sagnac criterion for presence of rotation of the object itself with the usual one. Note also that we have restricted ourselves to experiments which can be performed only outside the ergosphere.

[14]This vector field was first introduced by Bardeen. See, e.g., J. Bardeen, Ap. J. **162**, 71, (1970). Since the locally nonrotating vector field remains timelike down to the horizon, the "zero rotation" Sagnac experiment can be performed in the ergosphere unlike the previous ones.

[15]More precisely, consider a body which can be approximated by a worldline together with a spatial triad attached to each point of this worldline. Given a body, its description by a worldline with spatial triads becomes more and more accurate as the scalar curvature becomes smaller and smaller.

[16]Thus the "vector fields" U^a and V^a are defined only along the worldline of the center of the tube.

[17]For example, a test applicable to bodies of finite size.

[18]See, e.g., R. Geroch, J. Math. Phys. **12**, 918 (1971), especially the Appendix.

[19]Y. Aharanov and D. Bohm, Phys. Rev. **115**, 485 (1959). This prediction has been experimentally verified. See, e.g., R.G. Chambers, Phys. Rev. Lett. **5**, 3 (1960).

Spacetime symmetries and the complexion of the electromagnetic field

John R. Ray and Everette L. Thompson

Department of Physics and Astronomy, Clemson University, Clemson, South Carolina 29631
(Received 11 September 1974; revised manuscript received 24 October 1974)

Considering spacetimes admitting groups of motions, we derive the symmetry imposed on the electromagnetic field for solutions to the Einstein-Maxwell equations. It is found that the scalar products between the gradient of the complexion and the Killing vectors must be constants and that these constants enter into the symmetry condition on the electromagnetic field. We therefore obtain a geometrical interpretation of the complexion in spacetimes having symmetries.

1. INTRODUCTION

The Einstein equations are such complicated equations that in order to obtain exact solutions, one usually imposes certain restrictions on the spacetime. One useful method is to assume the spacetime allows a symmetry generated by a set of Killing vectors (ξ^i). The invariance of the spacetime with respect to the symmetry implies the Lie derivative of the metric with respect to the Killing vectors vanishes:

$$L_\xi g_{jk} = \xi_{j;k} + \xi_{k;j} = 0. \tag{1.1}$$

Since Lie differentiation commutes with ordinary differentiation,[1] we can at once prove that the Lie derivative of the Christoffel symbols Γ^i_{jk} with respect to the Killing vectors vanishes

$$L_\xi \Gamma^i_{jk} = 0. \tag{1.2}$$

Note that (1.2) is valid in any coordinate system since the difference of two Christoffel symbols is a tensor. From (1.2) and the commutation of ordinary and Lie differentiation it follows that the Lie derivative of the Einstein tensor G^{jk} with respect to the Killing vectors vanishes

$$L_\xi G^{jk} = 0. \tag{1.3}$$

If we apply Eqs. (1.3) to the Einstein field equations the Lie derivative of the energy momentum tensor with respect to the Killing vectors must also vanish

$$L_\xi T^{jk} = 0. \tag{1.4}$$

This last result gives a way of imposing the symmetry on the sources of the gravitational field.

In this paper we are interested in imposing symmetry on the Einstein—Maxwell equations

$$G^{jk} = (8\pi\kappa/c^4)\, T^{jk}, \tag{1.5a}$$

$$f^{jk}{}_{;k} = 0, \tag{1.5b}$$

$$*f^{jk}{}_{;k} = 0, \tag{1.5c}$$

where

$$T^{jk} = (1/4\pi)(f^{jl} f^k{}_l - \tfrac{1}{4} g^{jk} f_{rs} f^{rs}). \tag{1.6}$$

We shall restrict ourselves to the case of a nonnull electromagnetic field tensor. In this case the Einstein—Maxwell equations (1.5) are equivalent to the "already unified" field equations of Rainich—Misner—Wheeler.[2] In the "already unified" field theory the electromagnetic field is determined up to a constant duality rotation by the spacetime metric. This gives a simple way of imposing the Killing vector symmetry on the electromagnetic field f_{jk}.

In a recent paper[3] Woolley has discussed the same problem as we discuss in this paper. However, the results he obtained are not general.

We shall prove that the general result of imposing the symmetry on the electromagnetic field is

$$L_\xi f_{jk} = k\, *f_{jk}, \tag{1.7}$$

where k is a constant equal to the scalar product of the gradient of the complexion with the Killing vector. This gives a geometrical interpretation to the complexion in spactimes having symmetries.

2. SYMMETRIES

For a solution to the Einstein—Maxwell equations the gradient of the complexion α is given by

$$\alpha_{,i} = \eta_{ijkl} R^{jp}{}_{;q} g^{kq} R^l{}_p / R_{rs} R^{rs}. \tag{2.1}$$

From Eqs. (1.1) and (1.3) it follows that

$$L_\xi \alpha_{,i} = 0. \tag{2.2}$$

In deriving Eqs. (2.2) one must use the commutation of Lie differentiation with respect to a Killing vector and covariant differentiation.[1] From Eqs. (2.2) we have

$$(L_\xi \alpha)_{,i} = 0 \tag{2.3}$$

or

$$L_\xi \alpha = \alpha_{,i} \xi^i = k, \tag{2.4}$$

where k is a constant determined by the complexion and the Killing vector.

The electromagnetic field is given by

$$f_{ij} = f'_{ij} \cos\alpha + *f'_{ij} \sin\alpha, \tag{2.5}$$

where f'_{ij} is the extremal field.[2] The extremal field can be found from

$$f'_{ij} f'_{pq} = -\tfrac{1}{2} E_{ijpq} - \tfrac{1}{2} E_{ijlm} E_{pq}{}^{lm} / (R_{rs} R^{rs})^{1/2}, \tag{2.6}$$

where

$$E_{ij}{}^{kl} = \tfrac{1}{2}(-\delta^k_i R^l_j + \delta^k_j R^l_i - \delta^l_j R^k_i + \delta^l_i R^k_j). \tag{2.7}$$

From Eqs. (2.6) and (2.7) we find

$$L_\xi f'_{ij} = 0. \tag{2.8}$$

Now, calculating the Lie derivative of f_{ij} from Eqs. (2.5), we find

$$L_\xi f_{ij} = {}^*f_{ij} L_\xi \alpha, \tag{2.9}$$

or, using Eqs. (2.4),

$$L_\xi f_{ij} = k\, {}^*f_{ij}, \tag{2.10}$$

which is the general result of imposing symmetries on the electromagnetic field. If there are several Killing vectors labeled by μ, ξ^i_μ, then the result is

$$L_{\xi_\mu} f_{ij} = k_\mu {}^*f_{ij}, \tag{2.11}$$

where the k_μ are constants defined by (2.4) for each Killing vector.

3. DISCUSSION

Often when one solves the Einstein–Maxwell equations one assumes that the Lie derivative of the electromagnetic field vanishes, e.g., Letelier and Tabensky.[4] From Eqs. (2.11) we see that this will be correct only if the k_μ all vanish. In particular this will be true if the complexion is constant. The complexion is in fact constant for the plane-symmetric metric studied by Letelier and Tabensky.

In general, however, the complexion will not be constant and the symmetry on the electromagnetic field must be imposed by using Eqs. (2.11). In practice one would start with a spacetime that allows a certain symmetry (ξ^i). This would specify the form of the metric. One could then calculate $\alpha_{,i}$, using Eqs. (2.1). If $\alpha_{,i}$ as calculated from the right-hand side of Eqs. (2.1) is the gradient of a scalar then the complexion α is determined up to a constant. The Lie derivative of α with respect to the Killing vectors then gives the k_μ. Next we impose the symmetry on the electromagnetic field by using Eqs. (2.11). Finally after imposing the symmetry one hopes to find a solution to the Einstein–Maxwell equations.

It is interesting to note that if one solves the problem using the "already unified" field equations, then the symmetry is imposed on the electromagnetic field automatically. What we have done in this paper is to determine the effect of this symmetry on the electromagnetic field.

The fact that the complexion enters into the symmetry condition is a bonus since the physical meaning of the complexion is obscure. Geometrically we can say that the scalar products between the Killing vectors and the gradient of the complexion are constants in solutions to the Einstein–Maxwell equations and that these constant angles enter the symmetry conditions on the electromagnetic field. Although this is not much of an interpretation of the complexion, it is at least some further information concerning its meaning.

It would be valuable to find a solution with $k_\mu \neq 0$; however, we have not yet found such a solution.

ACKNOWLEDGMENT

We thank Dr. John Safko of the University of South Carolina for helping with the computation of the complexion for the plane symmetric metric.

Noted added in proof: These same results have been discussed by H. Michalski and J. Wainwright in *General Relativity and Gravitation Journal*, to be published. These authors present an example. We thank J. Wainwright for a copy of their paper prior to publication.

[1] K. Yano, *The Theory of the Lie Derivative* (North-Holland, Amsterdam, 1955).
[2] C. Misner and J. Wheeler, Ann. Phys. (N.Y.) **2**, 525 (1957).
[3] M. Woolley, Comm. Math. Phys. **31**, 75 (1973).
[4] P. Letelier and R. Tabensky, J. Math. Phys. **15**, 594 (1974).

A new approach to the stability of matter problem. I*

Paul Federbush†‡

Institute for Advanced Study, Princeton, New Jersey 08540
(Received 21 June 1974; revised manuscript 20 October 1974)

The stability of matter problem solved by Dyson and Lenard is studied by more field theoretic techniques. Stability is proven for matter in a periodic cube.

INTRODUCTION

We here study the problem of stability of matter in the following form. There are N positive and N negative charges, one species fermion the other boson, in a periodic cube. We establish for this system that the energy goes to $+\infty$ as N increases. The present result is apparently weaker than the classic result of Dyson and Lenard.[1,2] The present research we feel is interesting for a number of reasons.

Firstly, the present paper is much easier than the papers of Dyson and Lenard. It introduces to the subject techniques developed in constructive quantum field theory, and in a subsequent paper we will show the present approach can be extended to the Dyson–Lenard result in a still simple form. (The latter result is a lower bound on the energy in the infinite volume proportional to N.) We entered the present calculation largely to test the techniques of present field theory, trying to obtain lower bound estimates in a hard problem from many-body physics. This has been rewarding, an easy—though recent—estimate from field theory is very effective on a hard problem in many-body physics. This indicates the value of pushing techniques from constructive quantum field theory into traditional fields of physics and that the field is coming of age with useful, nontrivial, techniques. This paper is self-contained; no knowledge of constructive quantum field theory is necessary.

The paper is organized as follows. In Sec. 1 the notation is presented. In Sec. 2 the results are presented. This is the statement that the lower bound on the energy increases with N, and two technical results used in the proof of this fact. The first technical result is a generalization of the so-called N_τ estimates. It requires the anticommutivity of the fermion variables and substitutes for complicated arguments on antisymmetric wave functions in Dyson and Lenard.[3] It is proved in Sec. 3.

The other technical result, proved in Sec. 4, is an inequality two-point distribution functions must satisfy. It is a geometric requirement following from the three-dimensional nature of space and has nothing to do with interactions. It must be known by someone, but we have found no reference for it. It is not from constructive quantum field theory. It is not surprising when thought about, but we think it is amusing. Roughly stated it says: If each particle on the average has M particles within distance R of it, and $R' < R$, then each particle has on the average $\gtrsim M\,(R'/R)^3$ particles within distance R' of it.

The energy bounds are tied together in the final sections. As in Lenard and Dyson, basically one calculates the energy of fermions moving in a fixed distribution of bosons, and compensates this by the repulsion energy of the bosons. Field theory ideas are here evident. We note that neither the mass nor the statistics of the positive charges enter our estimates.

1. NOTATION

We shall work in a unit periodic cube containing N fermions of charge $-q$ and N bosons of charge $+q$. The mass of the fermions will be m, the mass of the bosons will not be of interest and might be infinite. The fermions are described by fields ψ and $\bar{\psi}$,

$$\psi = \sum \exp(ikx)\, b_k, \qquad \bar{\psi} = \sum \exp(-ikx)\, b_k^* \tag{1.1}$$

with k summed over modes of the unit cube with periodic boundary conditions. The bosons are described by fields ϕ and $\bar{\phi}$. We use V for the basic interaction,

$$V = q^2/r \tag{1.2}$$

and V_n for a "cutoff" interaction

$$V_n = q^2 \exp(-nr)/r. \tag{1.3}$$

By defining $\omega_k = k^2/2m$, the fermion kinetic energy is given by

$$H_{0F} = \sum \omega_k b_k^* b_k. \tag{1.4}$$

It is also convenient to define F, a rescaled kinetic energy:

$$F = H_{0F} + N = \sum (\omega_k + 1)\, b_k^* b_k = \sum \omega_k' b_k^* b_k. \tag{1.5}$$

The total Hamiltonian may be written

$$H = H_{0F} + H_{0B} + \tfrac{1}{2} \int : (\bar{\psi}(x)\psi(x) - \bar{\phi}(x)\phi(x)) \times V(x-y)(\bar{\psi}(y)\psi(y) - \bar{\phi}(y)\phi(y)) : . \tag{1.6}$$

We shall use the two-point distribution function for the boson field,

$$\rho(x,y) = \langle : \bar{\phi}(x)\phi(x)\bar{\phi}(y)\phi(y) : \rangle. \tag{1.7}$$

A number of constants will arise in estimates, which do not depend on any variables of interest; they will be denoted by c_i, and satisfy $0 < c_i < \infty$.

The matter problem in a periodic cube is somewhat ambiguous. We here will work with periodic boundary conditions on the wavefunctions, but the interaction q^2/r is not taken periodic. One could instead choose the interaction to be $\sum_k (q^2 4\pi/k^2) \exp(ik \cdot r)$, a periodic interaction corresponding to periodic matter. Whereas this latter choice is more interesting, and all the results of the present paper remain true for this choice also (requiring only simple modifications of the proof), the procedure we follow is more in tune with the infinite volume problem pursued in a sequel paper.

2. RESULTS

We summarize the results of the present paper in three facts. The first two are technical results, interesting in their own right, used in the proof of the third. Fact 1 is proven in III, Fact 2 in IV, and Fact 3 in V and VI.

Fact 1

Let $\alpha \geq \delta$, $\beta \geq \delta$, $\alpha + \beta - 1/2 = \delta > 0$, and $0 < \delta' < \delta$, then there is a constant $c(\delta, \delta')$ such that for all constants d_{ik}

$$\left\| \frac{1}{F^\alpha} \sum d_{ik} b_i^* b_k \frac{1}{F^\beta} \right\| \leq c(\delta, \delta').$$

$$\left(\sum |d_{ik}|^2 \frac{1}{\tilde{\omega}_{i,k}} \cdot \frac{1}{(\tilde{\tilde{\omega}}_{i,k})^{2\delta'}} \right)^{1/2} \tag{2.1}$$

where $\tilde{\omega}_{i,k} = \min\{\omega_i', \omega_k'\}$, $\tilde{\tilde{\omega}}_{i,k} = \max\{\omega_i', \omega_k'\}$, and the double bars indicate operator norm.

Fact 2

There is a constant $c_2 > 0$ such that if $\rho'(x,y) = \langle \bar{\phi}(x)\phi(x)\bar{\phi}(y)\phi(y) \rangle$, $F(x) \geq 0$, and $0 \leq R' \leq R$ then

$$\int_{|x-y|<R'} F(x)\rho'(x,y)F(y) \geq c_2 \left(\frac{R'}{R}\right)^3 \int_{|x-y|<R} F(x)\rho'(x,y)F(y) \tag{2.2}$$

Fact 3

If E_N is the minimum energy for N positive and N negative charges (one boson, one fermion) in a periodic cube, then

$$E_N \underset{N\to\infty}{\longrightarrow} \infty. \tag{2.3}$$

If Fact 2 is specialized to a situation where $F(x) = 1$ and $\rho'(x,y)$ is given by

$$\rho'(x,y) = \sum_{\substack{x_i \in S \\ y_i \in S}} \delta(x - x_i)\delta(y - y_i) \tag{2.4}$$

for S set of points in R^3 (a classical limit for a distribution of particles located at the points in S), one obtains Fact 2′.

Fact 2′

There is a constant $c_2 > 0$ such that if S is a set of points in R^3 and $0 \leq R' \leq R$, then

$$\mathrm{Card}\{(x,y) \in S \times S \mid |x-y| < R'\} \geq c_2 \left(\frac{R'}{R}\right)^3 \mathrm{Card}\{(x,y) \in S \times S \mid |x-y| < R\}. \tag{2.5}$$

Fact 2′ is the form of Fact 2 used in this paper. Finally we give another specialization of Fact 2 for amusement.

Fact 2″

There is a constant $c_2 > 0$ such that if $F \geq 0$ is a function on R^3 and $0 \leq R' \leq R$, then

$$\int_{|x-y|<R'} F(x)F(y) \geq c_2 \left(\frac{R'}{R}\right)^3 \int_{|x-y|<R} F(x)F(y). \tag{2.6}$$

3. GENERALIZED N_τ ESTIMATES

N_τ estimates were derived in Ref. 3 and generalized in the direction of our interest in Refs. 4 and 5. Our derivation of (2.1) is rather complete, since we are more interested in presenting methods than results. We begin with a classical result following from the anticommutivity of the b's.

Lemma 1: Let $A = \sum d_i b_i$ (the d_i constants); then $\|A\| \leq (\sum |d_i|^2)^{1/2}$.

Proof: By anticommutivity of the b's,

$$AA^* + A^*A = \sum |d_i|^2$$

$$\Rightarrow A^*A \leq \sum |d_i|^2 \tag{3.1}$$

$$\Rightarrow \|A\| \leq (\sum |d_i|^2)^{1/2}$$

We next prove a slightly strengthened form of one of the original N_τ estimates.[3]

Lemma 2: If the d_{ik} are constants and $\|B\| \leq 1$,

$$\|\sum d_{ik} b_i^* B b_k / F^{1/2}\| \leq (\sum |d_{ik}|^2/\omega_k')^{1/2}. \tag{3.2}$$

Proof: Let $|s>$ and $|t>$ be normalized states:

$$\left| \left\langle s \middle| \sum d_{ik} b_i^* B b_k \frac{1}{F^{1/2}} \middle| t \right\rangle \right| = \left| \sum_k \left\langle \sum_i \frac{d_{ik}^*}{\omega_k'^{1/2}} b_i s \middle| B b_k \omega_k'^{1/2} \frac{1}{F^{1/2}} t \right\rangle \right|. \tag{3.3}$$

By the Schwartz inequality

$$\leq \left(\sum_k \left| \left\langle \sum_i d_{ik}^* \frac{1}{\omega_k'^{1/2}} b_i s \right| \right|^2 \right)^{1/2} \cdot \left(\sum_k \left\langle t \middle| \frac{1}{F^{1/2}} \omega_k' b_k^* b_k \frac{1}{F^{1/2}} \middle| t \right\rangle \right)^{1/2}, \tag{3.4}$$

using Lemma 1, we have

$$\leq \left[\sum_k \left| \left(\sum_i |d_{ik}|^2 \frac{1}{\omega_k'} \right)^{1/2} \right|^2 \right]^{1/2} \leq \left(\sum |d_{ik}|^2 \frac{1}{\omega_k'} \right)^{1/2}. \tag{3.5}$$

Now to the proof of Fact 1. We consider

$$\frac{1}{F^\alpha} \sum d_{ik} b_i^* b_k \frac{1}{F^\beta}. \tag{3.6}$$

We break up $\{d_{ik}\}$ into two sets of terms, determined by the following inequalities:

$$\omega_i - \omega_k \geq 0, \quad \omega_i - \omega_k < 0$$

It is clearly sufficient to prove the inequality with the d's in each of these sets separately. We assume $\omega_i - \omega_k \geq 0$. We note the equality

$$F^{-\delta} = \int_0^\infty \exp(-Ft) \frac{dt}{t^{1-\delta}} \Big/ \int_0^\infty \exp(-x) \frac{dx}{x^{1-\delta}} \tag{3.7}$$

$$= c_\delta \int_0^\infty \exp(-Ft)\frac{dt}{t^{1-\delta}}.$$

(3.6) becomes

$$c_\delta \int_0^\infty \frac{dt}{t^{1-\delta+\delta'}} \frac{1}{F^{\alpha-\delta}} \sum d_{ik} t^{\delta'} \exp(-\omega_i t) b_i^* \exp(-Ft) b_k \frac{1}{F^\delta}. \tag{3.8}$$

Notice the fancy footwork on the $t^{\delta'}$. We now use

$$|\exp(-\omega t)t^\alpha| \leq c/\omega^\alpha \tag{3.9}$$

and the following lemma.

Lemma 3: Let A be a positive operator, B an arbitrary operator, and $r \geq 0$, $s \geq 0$; then

$$\left\| \frac{1}{A^r} B \frac{1}{A^s} \right\| \leq \left\| B \frac{1}{A^{r+s}} \right\|^{s'} \cdot \left\| \frac{1}{A^{r+s}} B \right\|^{r'} \tag{3.10}$$

with $r' = r/(r+s)$ and $s' = s/(r+s)$.

We omit a proof of this lemma, an elementary type of interpolation theorem based on the three lines theorem.[6]

Fact 1 follows from (3.8), (3.9), (3.10), Lemma 2, and the inequality

$$\| \int g(t) O(t) dt \| \leq \int |g(t)| \cdot \| O(t) \| \, dt \tag{3.11}$$

with $g(t)$ a numerical function and $O(t)$ an operator depending on the parameter t. Fact 1 first appeared in an unpublished paper by the author.

Many variations of Fact 1 are possible. We now restrict ourselves to noting the following variation: Using a result of McBryan,[7] we may in (2.1) pick $\delta' = \delta$ and $c(\delta, \delta') = 1$ if $\tilde{\tilde{\omega}}_{i,k}$ is replaced by $\tilde{\omega}_{i,k}$.

4. A PACKING INEQUALITY FOR $\rho_2(x,y)$

We reduce the proof of Fact 2 to the proof of Fact 2′. Let the N-particle wavefunction of the boson field be given by the symmetric function $\psi(x_1, \cdots, x_N)$. Then

$$\rho'(x,y) = \langle \bar{\phi}(x)\phi(x)\bar{\phi}(y)\phi(y) \rangle$$
$$= \int dx_1 \cdots dx_N \bar{\psi}(x_i)\psi(x_i) \times \sum_{j,k=1}^{N} \delta(x - x_j)\delta(y - x_k). \tag{4.1}$$

Thus $F(x)\rho'(x,y)F(y)$ is a (continuous) positive linear combination of terms of the form

$$F(x)\left(\sum_{j,k=1}^{N} \delta(x - x_j)\delta(y - x_k) \right) F(y). \tag{4.2}$$

Fact 2 would follow from

$$\int_{|x-y|<R'} F(x)\left(\sum \delta(x - x_j)\delta(y - x_k) \right) F(y)$$
$$\geq c_2 (R'/R)^3 \int_{|x-y|<R} F(x)\left(\sum \delta(x - x_j)\delta(y - x_k) \right) F(y) \tag{4.3}$$

by taking positive linear combinations, an operation preserving the inequality. It is an easy deduction that (4.3) is implied by Fact 2′.

Proof of fact 2′

Fill space with a lattice of small cubes $\{\Delta_i\}$, n_i points of S in Δ_i. There are eight lattices of large cubes $\{H_i\}$, N_i points of S in H_i. Each small cube is in eight large cubes. The eight lattices of large cubes are obtained by starting with one of these lattices and displacing in some number of coordinate directions by one-half a cube side.

We require the following additional properties:

(A) The diameter of the small cubes is less than R'.

(B) Every sphere of radius R is contained in some large cube.

We can set up these lattices of cubes such that if there are Q small cubes in each large cube

$$Q \leq c_1 (R/R')^3 \tag{4.4}$$

for some universal constant c_1.

There follows

$$\text{Card}\{(x,y) \in S \times S \mid |x - y| < R\}$$
$$\leq \sum N_i^2$$
$$\leq 8Q \sum n_i^2$$
$$\leq 8Q \, \text{Card}\{(x,y) \in S \times S \mid |x - y| < R'\}. \tag{4.5}$$

This yields Fact 2′ with $c_2 = 1/8c_1$. It would be of interest to find a best possible value of c_2.

Although as we said before we could find no reference for this specific result, Lenard, in Ref. 8, gives the most general conditions on distribution functions, so that this inequality must be contained *implicitly* there.

5. THE BASIC PROBLEM

We recall H:

$$H = H_{0F} + H_{0B} + \tfrac{1}{2}\int :(\bar{\psi}\psi - \bar{\phi}\phi)V(\bar{\psi}\psi - \bar{\phi}\phi):. \tag{5.1}$$

We will make no use of the (positive) boson kinetic energy in obtaining our bound. By a standard argument it is sufficient to obtain a uniform bound for the bosons in classical configurations. We will pursue our proof with the bosons in a fixed configuration. We divide H into seven pieces

$$H = H_1 + H_2 + \cdots + H_6 - N, \tag{5.2}$$
$$H_1 = H_{0B}, \tag{5.3}$$
$$H_2 = \tfrac{1}{2}F, \tag{5.4}$$
$$H_3 = \tfrac{1}{2}\int :\bar{\psi}\psi V_n \bar{\psi}\psi:, \tag{5.5}$$
$$H_4 = \tfrac{1}{2}\int :(\bar{\psi}\psi - \bar{\phi}\phi)(V - V_n)(\bar{\psi}\psi - \bar{\phi}\phi):, \tag{5.6}$$
$$H_5 = \tfrac{1}{2}F - \int :\bar{\psi}\psi V_n \bar{\phi}\phi: \tag{5.7}$$
$$H_6 = \tfrac{1}{2}\int :\bar{\phi}\phi V_n \bar{\phi}\phi: \tag{5.8}$$

The value of n will be picked later, depending on the boson configuration. We procede to bound the terms individually.

We will not detail dependences on q and m. The Fermi sea energy gives

$$H_2 = \tfrac{1}{2}F \geq cN^{5/3}. \tag{5.9}$$

Unsubscripted c's will not be required to have the same value in different equations—c may be read as $O(1)$ to those who prefer that notation.

Consider H_3:

$$H_3 = \tfrac{1}{2}\int :\bar{\psi}(x)\psi(x)V_n(x-y)\bar{\psi}(y)\psi(y):$$

We now observe

$$V_n(r) = q^2\exp(-nr)/r$$

is a positive potential. But H_3 then merely multiplies the many-body wavefunction by the potentials. Thus, being a positive multiplication operator,

$$H_3 \geq 0 \tag{5.10}$$

H_4 looks alot like H_3 in form, but H_4 is not positive. We write

$$H_4 = \tfrac{1}{2}\int (\bar{\psi}\psi - \bar{\phi}\phi)(V - V_n)(\bar{\psi}\psi - \bar{\phi}\phi) + \tfrac{1}{2}\int [:(\bar{\psi}\psi - \bar{\phi}\phi) \times (V - V_n)(\bar{\psi}\psi - \bar{\phi}\phi): - (\bar{\psi}\psi - \bar{\phi}\phi)(V - V_n)(\bar{\psi}\psi - \bar{\phi}\phi)] \tag{5.11}$$

and observe

$$(V - V_n) = \int dk \frac{1}{2\pi^2} q^2\left(\frac{1}{k^2} - \frac{1}{k^2+n^2}\right)\exp(ik\cdot x)\exp(-ik\cdot y). \tag{5.12}$$

The first term in the right side of (5.11) is positive, as the (continuous) positively weighted sum of operators times their conjugates. The second term is evaluable and yields

$$H_4 \geq -Nn. \tag{5.13}$$

H_5 requires some study. Recalling

$$H_5 = \tfrac{1}{2}F - \int :\bar{\psi}\psi V_n\bar{\phi}\phi :.$$

Let

$$W = \int :\bar{\psi}\psi V_n\bar{\phi}\phi :. \tag{5.14}$$

Let $\epsilon > 0$ be a small number to be further specified later and consider

$$\left\| \frac{1}{F^{3/8+\epsilon}} W \frac{1}{F^{3/8+\epsilon}} \right\|_F = A(\bar{\phi}, \phi). \tag{5.15}$$

Here the subscripted double bars indicate the operator norm of the expression inside as an operator in the fermion variables, treating the boson operators as numerical quantities. In fact in W the boson operators occur only in the combination $\bar{\phi}(x)\phi(x)$, a numerical quantity for a classical boson distribution. Using Fact 1, we get

$$A \leq c_1(\epsilon)\left(\sum_k \sigma_k\sigma_{-k}\right)^{1/2} \tag{5.16}$$

with

$$\sigma_k = \int \bar{\phi}(x)\phi(x)V(x-y)\exp(ik\cdot y). \tag{5.17}$$

To see where (5.16) comes from, we write

$$W = \sum_{k_1,k_2} \sigma_{k_1-k_2} b_{k_2}^* b_{k_1} \tag{5.18}$$

and apply (2.1) with

$$d_{k_2,k_1} = \sigma_{k_1-k_2}, \tag{5.19}$$

remembering that we do not require that the right side of (2.1) equal the right side of (5.16), but only that the right side of (5.16) is greater than the right side of (2.1). This is easy.

(5.15) and (5.16) yield

$$W \leq c_1(\epsilon)\left(\sum_k \sigma_k\sigma_{-k}\right)^{1/2} F^{3/4+2\epsilon}. \tag{5.20}$$

We now use the numerical estimate

$$xy \leq x^p/p + y^q/q$$

for $1/p + 1/q = 1$, $p > 0$, $q > 0$, $x > 0$, $y > 0$ with

$$q = (\tfrac{3}{4} + 2\epsilon)^{-1}, \quad p = (\tfrac{1}{4} - 2\epsilon)^{-1}, \quad y \sim F^{3/4} + 2\epsilon, \tag{5.21}$$

and, defining $\epsilon' = 2/(1-8\epsilon) - 2$,

$$W \leq c_2(\epsilon)\left(\sum_k \sigma_k\sigma_{-k}\right)^{2+\epsilon'} + 1/2F. \tag{5.22}$$

The fact that F and the expression in parentheses are positive commuting operators allows the use of the numerical estimate. This gives directly

$$H_5 \geq -c_2(\epsilon)\left(\sum_k \sigma_k\sigma_{-k}\right)^{2+\epsilon'}. \tag{5.23}$$

The sequence of steps (5.15)–(5.23) is exactly patterned on calculations in Refs. 4 and 5.

We now observe

$$\sum_k \sigma_k\sigma_{-k} = \int \bar{\phi}\phi V\chi V\bar{\phi}\phi, \tag{5.24}$$

where χ is the characteristic function of the unit cube,

$$\begin{aligned} &\leq \int \bar{\phi}\phi V\, V\bar{\phi}\phi \\ &= c\int \bar{\phi}\phi[\exp(-nr)/n]\bar{\phi}\phi \\ &= c[\int :\bar{\phi}\phi(\exp(-nr)/n)\bar{\phi}\phi : + N/n] \end{aligned} \tag{5.25}$$

so that

$$H_5 \geq -c(\epsilon)[\int (\rho/n)\exp(-nr)]^{2+\epsilon'} - c(\epsilon)(N/n)^{2+\epsilon'}. \tag{5.26}$$

H_6 may be written as

$$H_6 = \tfrac{1}{2}\int \rho\exp(-nr)/r. \tag{5.27}$$

Observe that

$$|\int \bar{\phi}\phi\exp(-nr)(1/n)\bar{\phi}\phi| \leq N^2/n. \tag{5.28}$$

We collect terms to achieve an estimate for H:

$$\begin{aligned} H \geq\ & c_A N^{5/3} - c_B Nn - c_C(\epsilon)(N/n)^{2+\epsilon'} - N \\ & - c_D(\epsilon)(N^2/n)^{\epsilon'}[\int \rho/n\exp(-nr)]^2 + c_E\int \rho\exp(-nr)/r. \end{aligned} \tag{5.29}$$

The remaining program is to pick an ϵ, and a boson distribution dependent value of n, to show the bound (5.29) forces $E_{N\overrightarrow{N}\infty}\infty$.

6. COMPLETION OF ESTIMATES, INCORPORATION OF BOSON REPULSION

In this final section we study (5.29). Basically the last term arising from boson repulsion is used to can-

cel the next to last term. Two cases are considered, depending on the boson configuration. In the first case n is picked $\sim N^{2/3}$ and includes all configurations except a very unusual set where the bosons are extremely clustered.

We pick ϵ and ϵ' of the last section and two new parameters γ and γ satisfying

$$\gamma' > 0, \quad \gamma > 0, \quad \epsilon > 0, \quad \epsilon' > 0,$$
$$\epsilon' = 2/(1-8\epsilon) - 2,$$
$$\gamma' > \gamma + \tfrac{4}{3}\epsilon',$$
$$\gamma > \epsilon',$$
$$\gamma' < \tfrac{1}{10}, \quad \gamma < \tfrac{1}{10}, \quad \epsilon' < \tfrac{1}{10}.$$

We have not sought optimal estimates in making these choices.

The two cases are given by the following:

Case I:

$$\int_{|x-y|<N^{-2/3+\gamma}} \rho < N \cdot N^{1-\gamma'}. \tag{6.1}$$

Case II:

$$\int_{|x-y|<N^{-2/3+\gamma}} \rho \geq N \cdot N^{1-\gamma'}. \tag{6.2}$$

We apply the Schwartz inequality to the next to last term in (5.29) to obtain

$$H \geq Q + R,$$

$$Q = c_A N^{5/3} - c_B N n - c_C \left(\frac{N}{n}\right)^{2+\epsilon'} - N, \tag{6.3}$$

$$R = c_D \int \rho \frac{\exp(-nr)}{r} \left[\frac{c_E}{c_D} - \left(\frac{N^2}{n}\right)^{\epsilon'} \int \rho \frac{\exp(-nr)}{n^2} r\right]. \tag{6.4}$$

To handle Case I, we pick $n = \lambda N^{2/3}$ with $\lambda < c_A/c_B$. Then clearly Q approaches ∞ as $N \to \infty$. Consider the second term in brackets in (6.4); using (6.1), we get

$$\left(\frac{N^2}{n}\right)^{\epsilon'} \int \rho \frac{\exp(-nr)}{n^2} r$$
$$\leq c \left(\frac{N^2}{N^{2/3}}\right)^{\epsilon'} N \cdot N^{1-\gamma'} \frac{1}{N^{4/3}} \cdot \frac{1}{N^{2/3-\gamma}} \tag{6.5}$$

when N is large. The conditions on the parameters imply this last term in (6.5) goes to zero as $N \to \infty$ and thus $E_N \to \infty$. This estimate on the integral in (6.5) may be obtained by separately considering the regions where r is greater or less than $N^{-2/3+\gamma}$. In the first region the exponential $\exp(-nr)$ is so small this region contributes negligibly.

Turning to Case II, we now choose $n = N^{2/3+\gamma}$. The second integral in (6.4) is small similar to the above estimate. For N large we have

$$R > c \int \rho \exp(-nr)/r. \tag{6.6}$$

Looking at the contribution to the integral in (6.6) from the region $|x-y| < N^{-2/3-\gamma}$, we get

$$\int \rho \frac{\exp(-nr)}{r} \geq \int_{|x-y|<N^{-2/3-\gamma}} \cdot N^{2/3+\gamma} \cdot e^{-1}$$
$$\geq c N^{2/3+\gamma} \int_{|x-y|<N^{-2/3-\gamma}} \rho \tag{6.7}$$
$$\geq c N^{2/3+\gamma} \cdot N \cdot N^{1-\gamma'} \cdot \left(\frac{N^{-2/3-\gamma}}{N^{-2/3+\gamma}}\right)^3 \tag{6.8}$$
$$\geq c N^{8/3} \cdot N^{5\gamma-\gamma'} \geq c N^2 \tag{6.9}$$

(6.8) follows by using Fact 2. (6.9) implies $R \gtrsim N^2$ for N large, which dominates Q and implies $E_N \to \infty$. To this end it is important but trivial to note that in each of the two classes estimates can be obtained uniformly over all boson configurations (in each class estimates depend on N but not the distribution).

ACKNOWLEDGMENT

I am grateful to Dr. Carl Kaysen for the hospitality at the Institute for Advanced Study.

*This work was supported in part by NSF Grant GP-17523.
†John Simon Guggenheim Memorial Fellow.
‡Permanent address: Department of Mathematics, The University of Michigan, Ann Arbor, Michigan 48104,
[1]F. J. Dyson and A. Lenard, J. Math. Phys. **8**, 423 (1967).
[2]F. J. Dyson and A. Lenard, J. Math. Phys. **9**, 698 (1968).
[3]J. Glimm, Comm. Math. Phys. **5**, 343 (1967).
[4]D. Brydges, "Boundedness Below for Fermion Model Theories," (to be published).
[5]P. Federbush, J. Math. Phys. **14**, 1532 (1973).
[6]N. Dunford and J. T. Schwartz, *Linear Operators, Part* 1 (Interscience, New York, 1958), p. 520.
[7]O. A. McBryan, "The SU_3-Invariant Yukawa$_2$ Quantum Field Theory," preprint.
[8]A. Lenard, "Characterization of Correlation Measures," (to be published).

Statistical theory of effective electrical, thermal, and magnetic properties of random heterogeneous materials. IV. Effective-medium theory and cumulant expansion method

Motoo Hori

Department of Applied Physics, Tokyo Institute of Technology, Meguroku, Tokyo 152, Japan

Fumiko Yonezawa

Belfer Graduate School of Science, Yeshiva University, New York, New York 10033

Department of Applied Physics, Tokyo Institute of Technology, Meguroku, Tokyo 152, Japan
(Received 28 December 1973; revised manuscript received 16 September 1974)

Several perturbation solutions for the effective permittivity in a completely random medium are evaluated and the validity of the approximations is discussed. It is shown through a diagram technique that the effective-medium theory in a macroscopically inhomogeneous material is equivalent to the coherent-potential approximation in a disordered binary alloy not only in its physical concept but also in its mathematical structure. A cumulant expansion method which substantially takes into account the clustering effects is proposed while the effective-medium theory or the coherent-potential approximation is, by its nature, a single-site approximation and neglects the clustering effects. The numerical results obtained by the effective-medium theory and by the cumulant method for a binary mixture of a conducting material and an insulating material are compared with the computer simulation data on the effective conductivity of a three-dimensional random network. The solution of the cumulant method gives a remarkably good agreement with the computer simulation for the whole range of parameters. An important point is that the cumulant expansion theory holds excellently even near the critical percolation concentration where the clustering plays an essential role and where the effective-medium theory fails to work.

1. INTRODUCTION

This is part of a series of work on the study of the effective permittivity of a disordered inhomogeneous material whose local permittivity is given as a random function of position. The whole discussion in this series of work holds for other physical constants such as magnetic permeability, electrical and thermal conductivity, and diffusion constant. In the first paper I,[1] a formal perturbation solution was derived for a cell material where constituent cells are distributed in a statistically homogeneous manner and the property of a particular cell is statistically independent of that of any other cell. In the second paper II,[2] upper and lower bounds on the effective permittivity were expressed in terms of the three-point correlation functions. In the preceding paper III[3] and the present paper, we are concerned with the perturbation treatment for completely random systems. A usual procedure of the perturbation technique for obtaining an effective property in a random medium is stated as follows:

(i) Expand a local field in a perturbation series;

(ii) average in the ensemble sense each term of the expansion series;

(iii) resum the averaged perturbation terms and determine the effective constant.

The first and second steps have already been investigated in I and III. Particularly, the concept and importance of the exclusion effect[4] were discussed, on the basis of which the prescription for constructing the expansion coefficients of all orders was given and the explicit forms of the leading terms were calculated. In order to get a general idea about the whole scheme of the work, it would help if readers could briefly study at least the introduction and summary of III.

The present paper will concentrate on the third step of resummation and the entire procedure will be completed. Especially, we have two significant purposes in this article. The first aim is to prove by means of a diagram technique that the effective-medium (EM) theory is equivalent to the coherent-potential approximation (CPA) not only in its physical concept but also in its mathematical structure. The EM theory, which is sometimes called just the self-consistent theory, has attained some success in the understanding of the average properties in classical inhomogeneous materials,[5-25] while the CPA has been proposed to give a good overall explanation of quantum-mechanical quasiparticle properties in substitutionally disordered alloys.[26] Although it has been pointed out that the philosophy underlying the EM theory for classical mixtures is analogous to that leading to the CPA in solid state physics,[27-29] there is no detailed proof of the mathematical equivalence between the EM theory and the CPA. Showing the indicated equivalence is important in the sense that this guarantees the validity of the EM theory since the CPA has been studied intensively and its usefulness and adequacy have been well-accepted.

The EM theory for classical inhomogeneous mixtures has recently been revalued mainly because the theory seems to work well for describing classical aspects of the problem of electron localization in some disordered systems.[13,30] Actually, it has turned out that the EM theory is a far better approach to this problem of electron localization than it might have been expected to be. The breakdown of the EM theory is observed in the concentration region near the critical percolation concentration for a mixture of a conducting material and an insulating material.[27,29,31] This breakdown is due to the fact that the EM theory or the CPA is essentially a

single-site approximation and ignores the effect of clusters, while the region near the critical percolation concentration is the very range where the clustering plays a crucial role. Therefore, another purpose of the present article is to propose an approximate solution which takes some important clustering effects into account and works as a more adequate theory for a critical concentration region. We argue that the cumulant expansion method[32] serves this purpose of a better approximation because it can be shown that the cumulant solution picks up the contributions from most important clusters. This argument is supported by the numerical agreement of the cumulant solution with the result of a computer simulation.[27,29,31]

2. DIAGRAMMATIC REPRESENTATION OF THE PERTURBATION EXPANSION

In the perturbation theory of quantum mechanics and other fields, the third procedure (as explained in Sec. 1) of resumming the averaged perturbation terms is quite often facilitated by the use of diagram techniques.[33] By means of diagrams, the complicated perturbation integrals of higher orders are visualized and intuition helps to understand the structure of higher-order terms and to pick up terms more important than others. In view of introducing appropriate diagrams for our perturbation series, let us first review how the expansion terms of general orders are constructed.

The problem is to obtain the effective scalar permittivity ϵ^* for a completely random inhomogeneous system where a local permittivity $\epsilon(\mathbf{r})$ at a point $\mathbf{r}$ is statistically independent of permittivities at other points. As shown in I, ϵ^* is expressed as a sum of infinite perturbation series in the form

$$\epsilon'^* \equiv \epsilon^* - \langle\epsilon\rangle = \sum_{n=2}^{\infty} \epsilon^{(n)}, \tag{2.1}$$

with

$$\epsilon^{(n)} = \left(\frac{1}{4\pi\langle\epsilon\rangle}\right)^{n-1} \int_V d\omega_{12} \int_V d\omega_{23} \cdots \int_V d\omega_{n-1,n} \frac{x_{12,i}}{r_{12}^3} \times \frac{x_{23,k}}{r_{23}^3} \cdots \frac{\partial^{n-1}\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\cdots\epsilon'(\mathbf{r}_n)\rangle}{\partial x_{12,k}\partial x_{23,h}\cdots\partial x_{n-1,n;(i)}}. \tag{2.2}$$

Here the angular brackets indicate the ensemble average and the summation convention is used. Under appropriate boundary conditions we can transform Eq. (2.2) into

$$\epsilon^{(n)} = \int_V d\omega_{12} \int_V d\omega_{23} \cdots \int_V d\omega_{n-1,n} G_{ik}(\mathbf{r}_{12}) \times G_{kh}(\mathbf{r}_{23}) \cdots \langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\cdots\epsilon'(\mathbf{r}_n)\rangle, \tag{2.3}$$

where

$$G_{ik}(\mathbf{r}_{12}) = \frac{1}{\langle\epsilon\rangle} \frac{\partial^2 (1/4\pi r_{12})}{\partial x_{12,i}\partial x_{12,k}}. \tag{2.4}$$

It has been shown in III that the concept of complete randomness leads us to a more compact expression

$$\epsilon'^* = \sum_{n=2}^{\infty} \frac{\sum_{\{\nu_m\}} B^{(n)}_{\{\nu_m\}} \epsilon(n, \{\nu_m\})}{(-\langle\epsilon\rangle)^{n-1}}. \tag{2.5}$$

The symbol $\sum_{\{\nu_m\}}$ denotes the sum over all possible partitions of n variables $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_n$ into m subsets of at least two variables each. The factors $B^{(n)}_{\{\nu_m\}}$ and $\epsilon(n, \{\nu_m\})$ appearing in the nth-order term are defined by

$$B^{(n)}_{\{\nu_m\}} = \left(-\frac{1}{4\pi}\right)^{n-1} \int_V d\omega_{12} \int_V d\omega_{23} \cdots \int_V d\omega_{n-1,n} \frac{x_{12,i}}{r_{12}^3} \times \frac{x_{23,k}}{r_{23}^3} \cdots \frac{\partial^{n-1} F_n(\{\nu_m\}; \mathbf{r}_{12}, \mathbf{r}_{23}, \cdots, \mathbf{r}_{n-1,n})}{\partial x_{12,k}\partial x_{23,h}\cdots\partial x_{n-1,n;(i)}}$$

$$= (-\langle\epsilon\rangle)^{n-1} \int_V d\omega_{12} \int_V d\omega_{23} \cdots \int_V d\omega_{n-1,n} G_{ik}(\mathbf{r}_{12}) \times G_{kh}(\mathbf{r}_{23}) \cdots F_n(\{\nu_m\}; \mathbf{r}_{12}, \mathbf{r}_{23}, \cdots, \mathbf{r}_{n-1,n}), \tag{2.6}$$

$$\epsilon(n, \{\nu_m\}) = \langle\epsilon^{\nu_1}\rangle_c \langle\epsilon^{\nu_2}\rangle_c \cdots \langle\epsilon^{\nu_m}\rangle_c. \tag{2.7}$$

Here $\{\nu_m\}$ denotes the way of partitioning n variables into m subsets of at least two elements and $\langle\epsilon^\nu\rangle_c$ stands for the νth-order cumulant of $\epsilon(\mathbf{r})$. Besides, $F_n(\{\nu_m\}; \mathbf{r}_{12}, \mathbf{r}_{23}, \cdots, \mathbf{r}_{n-1,n})$ means the product of null functions $\delta_{\mathbf{r}_{ij}}$ corresponding to the partition $\{\nu_m\}$ and expresses the condition that variables belonging to the same group all coincide. Notice that we have employed the relation

$$\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\cdots\epsilon'(\mathbf{r}_n)\rangle = \sum_{\{\nu_m\}} F_n(\{\nu_m\}; \mathbf{r}_{12}, \mathbf{r}_{23}, \cdots \mathbf{r}_{n-1,n}) \epsilon(n, \{\nu_m\}). \tag{2.8}$$

In addition to $\epsilon(n, \{\nu_m\})$ we introduce quantities $\gamma(n, \{\nu_m\})$ such that

$$\gamma(n, \{\nu_m\}) = \langle\epsilon'^{\nu_1}\rangle\langle\epsilon'^{\nu_2}\rangle \cdots \langle\epsilon'^{\nu_m}\rangle. \tag{2.9}$$

Then Eq. (2.8) is rearranged as

$$\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\cdots\epsilon'(\mathbf{r}_n)\rangle = \sum_{\{\nu_m\}} H_n(\{\nu_m\}; \mathbf{r}_{12}, \mathbf{r}_{23}, \cdots, \mathbf{r}_{n-1,n}) \gamma(n, \{\nu_m\}), \tag{2.10}$$

where $H_n(\{\nu_m\}; \mathbf{r}_{12}, \mathbf{r}_{23}, \cdots, \mathbf{r}_{n-1,n})$ signifies the condition that variables belonging to the same group all coincide and variables belonging to different groups never coincide. Similarly, Eq. (2.5) is rewritten in the form

$$\epsilon'^* = \sum_{n=2}^{\infty} \frac{\sum_{\{\nu_m\}} A^{(n)}_{\{\nu_m\}} \gamma(n, \{\nu_m\})}{(-\langle\epsilon\rangle)^{n-1}}. \tag{2.11}$$

Needless to say, $A^{(n)}_{\{\nu_m\}}$ is obtained by substituting $H_n(\{\nu_m\}; \mathbf{r}_{12}, \mathbf{r}_{23}, \cdots, \mathbf{r}_{n-1,n})$ for $F_n(\{\nu_m\}; \mathbf{r}_{12}, \mathbf{r}_{23}, \cdots, \mathbf{r}_{n-1,n})$ in Eq. (2.6).

The leading terms in Eq. (2.5) or (2.11) are evaluated as follows:

(second-order term)

(i) $F_2(1; \mathbf{r}_{12}) = H_2(1; \mathbf{r}_{12}) = \delta_{\mathbf{r}_{12}}$, (2.12a)

(ii) $\epsilon(2,1) = \langle\epsilon^2\rangle_c$, $\gamma(2,1) = \langle\epsilon'^2\rangle$, (2.12b)

(iii) $B^{(2)} = A^{(2)} = \frac{1}{3}$, (2.12c)

(iv) $\epsilon^{(2)} = -\frac{\langle\epsilon^2\rangle_c}{3\langle\epsilon\rangle} = -\frac{\langle\epsilon'^2\rangle}{3\langle\epsilon\rangle}$. (2.12d)

(third-order term)

(i) $F_3(1; \mathbf{r}_{12}, \mathbf{r}_{23}) = H_3(1; \mathbf{r}_{12}, \mathbf{r}_{23}) = \delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}$, (2.13a)

(ii) $\epsilon(3,1) = \langle\epsilon^3\rangle_c$, $\gamma(3,1) = \langle\epsilon'^3\rangle$, (2.13b)

(iii) $B^{(3)} = A^{(3)} = \frac{1}{9}$, (2.13c)

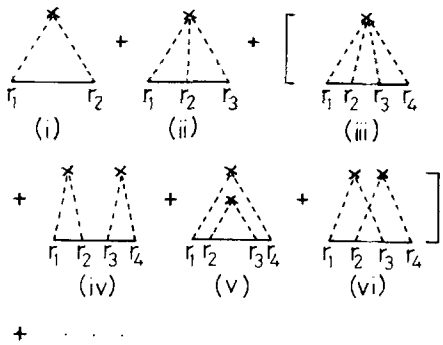


FIG. 1. Diagrammatic representation of Eq. (2.5).

(iv) $$\epsilon^{(3)}=\frac{\langle\epsilon^3\rangle_c}{9\langle\epsilon\rangle^2}=\frac{\langle\epsilon'^3\rangle}{9\langle\epsilon\rangle^2}. \tag{2.13d}$$

(fourth-order term)

(i) $$F_4(1;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34})=H_4(1;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34})=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{34}},$$
$$F_4(2;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34})=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{34}},$$
$$H_4(2;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34})=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{34}}(1-\delta_{\mathbf{r}_{23}}),$$
$$F_4(3;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34})=\delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}},$$
$$H_4(3;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34})=\delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}}(1-\delta_{\mathbf{r}_{12}}),$$
$$F_4(4;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34})=\delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{24}},$$
$$H_4(4;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34})=\delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{24}}(1-\delta_{\mathbf{r}_{12}}); \tag{2.14a}$$

(ii) $$\epsilon(4,1)=\langle\epsilon^4\rangle_c,\quad \gamma(4,1)=\langle\epsilon'^4\rangle,$$
$$\epsilon(4,2)=\epsilon(4,3)=\epsilon(4,4)=\langle\epsilon^2\rangle_c^2,$$
$$\gamma(4,2)=\gamma(4,3)=\gamma(4,4)=\langle\epsilon'^2\rangle^2; \tag{2.14b}$$

(iii) $$B_1^{(4)}=A_1^{(4)}=1/27$$
$$B_2^{(4)}=0,\quad A_2^{(4)}=-1/27,$$
$$B_3^{(4)}=1/9,\quad A_3^{(4)}=2/27,$$
$$B_4^{(4)}=0,\quad A_4^{(4)}=-1/27; \tag{2.14c}$$

(iv) $$\epsilon^{(4)}=-\frac{\langle\epsilon^4\rangle_c}{27\langle\epsilon\rangle^3}-\frac{\langle\epsilon^2\rangle_c^2}{9\langle\epsilon\rangle^3}=-\frac{\langle\epsilon'^4\rangle}{27\langle\epsilon\rangle^3}. \tag{2.14d}$$

As we shall see later, the fact that $B_4^{(4)}=0$ is significant in the sense that this provides us with a very fortunate situation in summing partially the perturbation series up to an infinite order.

(fifth-order term)

(i) $$F_5(1;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34},\mathbf{r}_{45})=H_5(1;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34},\mathbf{r}_{45})$$
$$=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{34}}\delta_{\mathbf{r}_{45}},$$
$$F_5(2;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34},\mathbf{r}_{45})=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{45}},$$
$$H_5(2;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34},\mathbf{r}_{45})=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{45}}(1-\delta_{\mathbf{r}_{34}}),$$
$$\text{etc.}; \tag{2.15a}$$

(ii) $$\epsilon(5,1)=\langle\epsilon^5\rangle_c,\quad \gamma(5,1)=\langle\epsilon'^5\rangle,$$
$$\epsilon(5,2)=\epsilon(5,3)=\cdots=\epsilon(5,11)=\langle\epsilon^3\rangle_c\langle\epsilon^2\rangle_c,$$
$$\gamma(5,2)=\gamma(5,3)=\cdots=\gamma(5,11)=\langle\epsilon'^3\rangle\langle\epsilon'^2\rangle, \tag{2.15b}$$

(iii) $$B_1^{(5)}=A_1^{(5)}=1/81,$$
$$B_2^{(5)}=B_3^{(5)}=0,\quad A_2^{(5)}=A_3^{(5)}=-1/81,$$
$$B_4^{(5)}=B_5^{(5)}=B_6^{(5)}=1/27,\quad A_4^{(5)}=A_5^{(5)}=A_6^{(5)}=2/81,$$
$$B_7^{(5)}=B_8^{(5)}=B_9^{(5)}=B_{10}^{(5)}=0,$$
$$A_7^{(5)}=A_8^{(5)}=A_9^{(5)}=A_{10}^{(5)}=-1/81; \tag{2.15c}$$

(iii′) $$B_{11}^{(5)}=0,\quad A_{11}^{(5)}=-1/81 \text{ (unproved)}; \tag{2.15c′}$$

(iv) $$\epsilon^{(5)}=\frac{\langle\epsilon^5\rangle_c}{81\langle\epsilon\rangle^4}+\frac{\langle\epsilon^3\rangle_c\langle\epsilon^2\rangle_c}{9\langle\epsilon\rangle^4}=\frac{\langle\epsilon'^5\rangle}{81\langle\epsilon\rangle^4}-\frac{\langle\epsilon'^3\rangle\langle\epsilon'^2\rangle}{81\langle\epsilon\rangle^4}. \tag{2.15d}$$

Out of the perturbation coefficients, we have failed to determine $B_{11}^{(5)}$ or $A_{11}^{(5)}$, but we assume that $B_{11}^{(5)}=0$ in analogy with $B_4^{(4)}$, $B_7^{(5)}$, $\cdots$, $B_{10}^{(5)}$.

Now, let us try and see how our perturbation series are interpreted in the language of diagrams. From the explicit expressions as indicated in the above, it is suggested that the diagrams proposed by Yonezawa and Matsubara[32] are most suitable for our purpose.[33] They were introduced to study the electronic properties in disordered binary alloys and used to derive the CPA.[4] Each diagram is composed of vertices, dashed vertical lines, and solid horizontal lines. A vertex may be a cross, an open circle, or a closed circle according to the type of diagrams. Diagrams whose vertices are crosses, open circles, or closed circles are called cumulant diagrams, restricted moment diagrams, or unrestricted moment diagrams, respectively. Using standard terminology in the quantum-mechanical diagram methods, we shall name a vertical line an interaction line and a horizontal line a propagator.

In order to represent the perturbation series in Eq. (2.5), we adopt cumulant diagrams with cross vertices. The perturbation expansion of ϵ'^* up to fourth order is expressed diagrammatically by Fig. 1. For example, the second-order diagram 1(i) is defined in Fig. 2 and equivalent to $\epsilon^{(2)}$. Cumulant diagrams of general order are constructed in accordance with the following prescription:

(i) Represent the points $\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_n$ by means of nodes on the horizontal base line.

(ii-a) Assign $\epsilon'(\mathbf{r}_i)$ to the ith interaction line and take the cumulant of the product of all $\epsilon'(\mathbf{r}_i)$ that correspond to the interaction lines starting from the same cross vertex.

(iii) Associate the ith propagator (connecting the points $\mathbf{r}_i$ and $\mathbf{r}_{i+1}$) with the tensor $G(\mathbf{r}_{i,i+1})$ defined in Eq. (2.4).

(iv) Multiply the tensor product $G(\mathbf{r}_{12})G(\mathbf{r}_{23})\cdots G(\mathbf{r}_{n-1,n})$ by the above-mentioned cumulants and

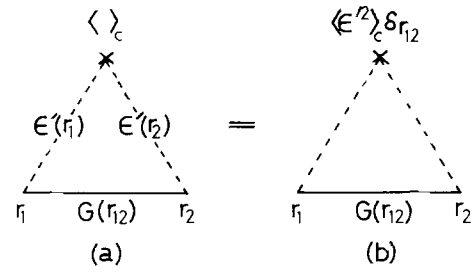


FIG. 2. Definition of the second-order cumulant diagram.

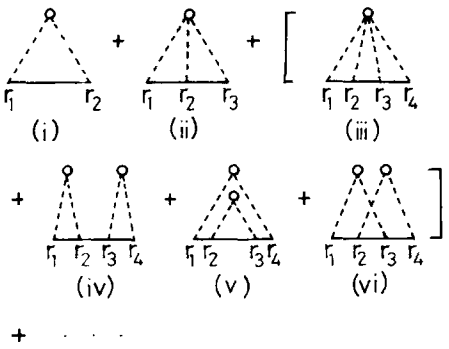


FIG. 3. Diagrammatic representation of Eq. (2.11).

integrate the result thus obtained with respect to $d\omega_{12}, d\omega_{23}, \cdots, d\omega_{n-1,n}$.

It is evident that the convention (ii-a) can be replaced by:

(ii-b) When ν interaction lines associated with $\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k, \cdots$ start from the same cross vertex, allot $\langle \epsilon'^{\nu} \rangle_c \delta_{\mathbf{r}_{ij}} \delta_{\mathbf{r}_{jk}} \cdots$ to these interaction lines.

On the other hand, restricted moment diagrams which contain open-circle vertices are introduced to represent perturbation terms in Eq. (2.11) rather than in Eq. (2.5). The diagram expression of Eq. (2.11) up to fourth order is presented in Fig. 3. For restricted moment diagrams, the following rule is framed instead of (ii-a):

(ii-a') When ν_1 interaction lines start from the first vertex, ν_2 from the second and so on (ν_m to the last), allot $H_n(\{\nu_m\}; \mathbf{r}_{12}, \mathbf{r}_{23}, \cdots, \mathbf{r}_{n-1,n}) \gamma(n, \{\nu_m\})$ to these interaction lines. Furthermore, we make another convention concerning unrestricted moment diagrams:

(ii-b′) When ν interaction lines associated with $\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k, \cdots$ start from the same closed-circle vertex, allot $\langle \epsilon'^{\nu} \rangle \delta_{\mathbf{r}_{ij}} \delta_{\mathbf{r}_{jk}} \cdots$ to these interaction lines.

We remark that the two requirements (ii-a′) and (ii-b′) are not necessarily equivalent to each other. Restricted moment diagrams obey the rules (i), (ii-a′), (iii), and (iv), while unrestricted moment diagrams obey the rules (i), (ii-b′), (iii), and (iv). Definitions of restricted and unrestricted moment diagrams of second order are illustrated in Fig. 4. Figures 5 and 6 show the relations among the three types of low-order diagrams.

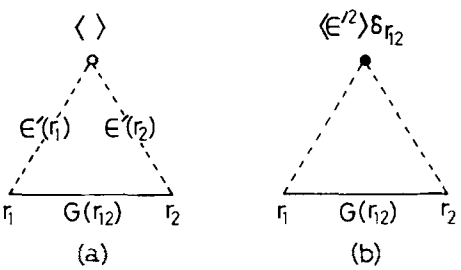


FIG. 4. Definition of the second-order moment diagrams. (a) Restricted moment diagram with an open-circle vertex. (b) Unrestricted moment diagram with a closed-circle vertex.

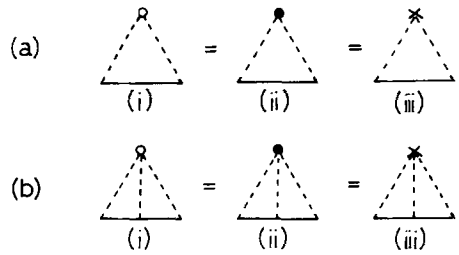


FIG. 5. Relation among the three types of second-order and third-order diagrams. The diagrams (a-i)—(a-iii) are of second order and (b-i)—(b-iii) are of third order. In these cases, restricted moment diagrams (i), unrestricted moment diagrams (ii), and cumulant diagrams (iii) coincide fortuitously with one another.

Let us again return to a discussion about cumulant diagrams. Out of these diagrams, proper or connected diagrams are defined as diagrams that cannot be divided into two separate parts by cutting a propagator *once*. It is easily seen that the diagram (iv) in Fig. 1 is unconnected. We classify proper diagrams into three categories. Figures 1 (i)—(iii), which include only one vertex, are called one-vertex or single-site diagrams. A diagram such as Fig. 1 (v) is sometimes named a nested diagram.[34] A crossed or irreducible diagram as depicted in Fig. 1 (vi) is a proper diagram where some interaction lines intersect with one another. Typical examples of unconnected, nested, and crossed diagrams representing perturbation terms of fifth and sixth order are given in Figs. 7 to 9.

With the knowledge of lower-order terms described in earlier paragraphs, we will find a simpler expedient for counting contributions from cumulant diagrams. First we investigate the behavior of improper or unconnected diagrams. As seen directly from Eq. (2.6), such diagrams correspond to terms with extra partial derivatives which lead to zero. Consequently, we have the first rule regarding unconnected diagrams:

(i) Identify unconnected diagrams with zero.

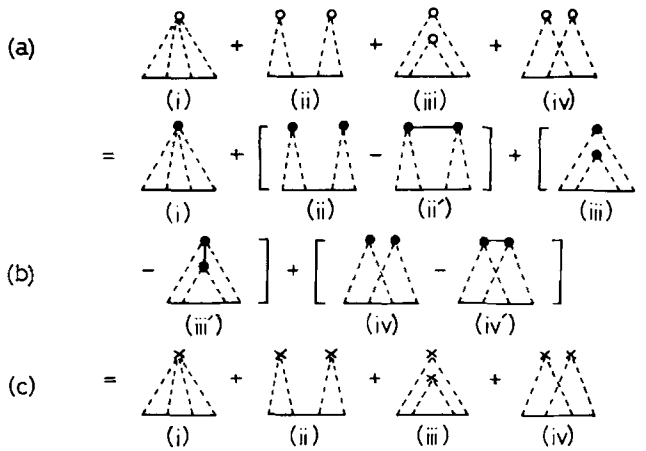


FIG. 6. Relation among the three types of fourth-order diagrams. The diagram (a-i) corresponds to $\langle \epsilon'^4 \rangle \delta_{\mathbf{r}_{12}} \delta_{\mathbf{r}_{23}} \delta_{\mathbf{r}_{34}}$, (b-i) to $\langle \epsilon'^4 \rangle \delta_{\mathbf{r}_{12}} \delta_{\mathbf{r}_{23}} \delta_{\mathbf{r}_{34}}$, (c-i) to $\langle \epsilon^4 \rangle_c \delta_{\mathbf{r}_{12}} \delta_{\mathbf{r}_{23}} \delta_{\mathbf{r}_{34}}$, (a-ii) to $\langle \epsilon'^2 \rangle^2 \delta_{\mathbf{r}_{12}} \delta_{\mathbf{r}_{34}} (1 - \delta_{\mathbf{r}_{23}})$, (b-ii) to $\langle \epsilon'^2 \rangle^2 \delta_{\mathbf{r}_{12}} \delta_{\mathbf{r}_{34}}$, (b-ii′) to $\langle \epsilon'^2 \rangle^2 \delta_{\mathbf{r}_{12}} \delta_{\mathbf{r}_{34}} \delta_{\mathbf{r}_{23}}$, (c-ii) to $\langle \epsilon^2 \rangle_c^2 \delta_{\mathbf{r}_{12}} \delta_{\mathbf{r}_{34}}$, and so on.

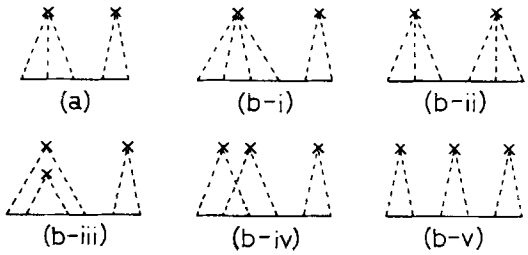


FIG. 7. Examples of unconnected diagrams. The diagram (a) is of fifth order, and (b-i)—(b-v) are of sixth order.

In other words, proper diagrams alone contribute to the effective permittivity and this situation is entirely different from the quantum-mechanical case. In our problem, ϵ'^* is nothing but the sum of all possible proper diagrams, which is called the mass operator in the quantum field theory and the effective index operator in the theory of wave propagation.[35,36]

Next we deal with one-vertex or single-site diagrams. It is obvious from Eq. (III.2.27) that the nth-order diagram with one cross vertex gives

$$\frac{B_1^{(n)}\langle\epsilon'^n\rangle_c}{(-\langle\epsilon\rangle)^{n-1}}=\frac{1}{3^{n-1}}\cdot\frac{\langle\epsilon'^n\rangle_c}{(-\langle\epsilon\rangle)^{n-1}}. \tag{2.16}$$

Therefore, the prescription for evaluating one-vertex diagrams is written out as follows:

(ii) Assign ϵ' to a dashed interaction line and take the cumulant of the product of all ϵ''s that correspond to interaction lines starting from a cross vertex. That is to say, allot the νth-order cumulant $\langle\epsilon'^\nu\rangle_c$ to ν dashed interaction lines which start from the same cross vertex. Note that $\langle\epsilon'^\nu\rangle_c=\langle\epsilon^\nu\rangle_c$ for $\nu\geq 2$.

(iii) Assign $-1/\langle\epsilon\rangle$ to each propagator.

(iv) Assign $\frac{1}{3}$ to each propagator.

(v) Calculate the product of all factors determined by the above three rules.

Nested diagrams are in general reducible to the products of single-site diagrams, although they are associated with more than one vertex. Hence nested diagrams belong to single-site diagrams in a wide sense. By way of explanation consider nested diagrams in Fig. 8. Equations (2.15) demonstrate that the terms corresponding to the diagrams (a-i) and (a-ii) are evaluated as

$$-\frac{1}{\langle\epsilon\rangle}\cdot\left(-\frac{\langle\epsilon^2\rangle_c}{3\langle\epsilon\rangle}\right)\cdot\frac{\langle\epsilon^3\rangle_c}{9\langle\epsilon\rangle^2}=\frac{\langle\epsilon^3\rangle_c\langle\epsilon^2\rangle_c}{27\langle\epsilon\rangle^4}. \tag{2.17}$$

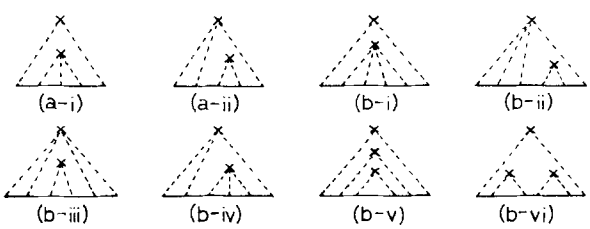


FIG. 8. Examples of nested diagrams. The diagrams (a-i) and (a-ii) are of fifth order, and (b-i)—(b-vi) are of sixth order.

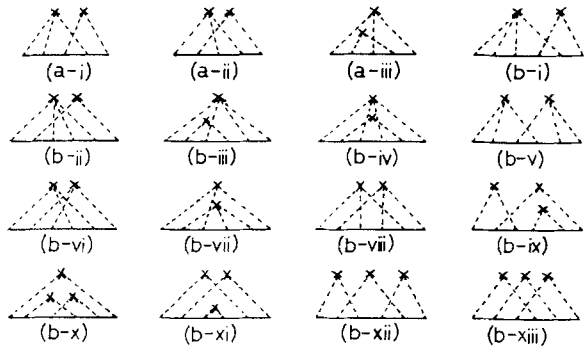


FIG. 9. Examples of crossed diagrams. The diagrams (a-i)—(a-iii) are of fifth order, and (b-i)—(b-xiii) are of sixth order.

Likewise, the diagrams (b-i)—(b-iii) are

$$-\frac{1}{\langle\epsilon\rangle}\cdot\left(-\frac{\langle\epsilon^2\rangle_c}{3\langle\epsilon\rangle}\right)\cdot\left(-\frac{\langle\epsilon^4\rangle_c}{27\langle\epsilon\rangle^3}\right)=-\frac{\langle\epsilon^4\rangle_c\langle\epsilon^2\rangle_c}{81\langle\epsilon\rangle^5}, \tag{2.18}$$

(b-iv) is

$$-\frac{1}{\langle\epsilon\rangle}\cdot\left(-\frac{\langle\epsilon^3\rangle_c}{9\langle\epsilon\rangle^2}\right)^2=-\frac{\langle\epsilon^3\rangle_c^2}{81\langle\epsilon\rangle^5}, \tag{2.19}$$

and (b-v) and (b-vi) are

$$\left(-\frac{1}{\langle\epsilon\rangle}\right)^2\cdot\left(-\frac{\langle\epsilon^2\rangle_c}{3\langle\epsilon\rangle}\right)^3=-\frac{\langle\epsilon^2\rangle_c^3}{27\langle\epsilon\rangle^5}. \tag{2.20}$$

The proof will be presented in Appendix A. The diagram equations in Fig. 10 show the process of factorization of the diagrams 8(b-i), (b-iv), and (b-vi) into single-site diagrams.

Taking these results into account, we arrive at the conclusion that for nested diagrams rule (iv) should be replaced by:

(iv′) Assign $\frac{1}{3}$ to each "independent" propagator.

For instance, the propagator lines indicated by arrows in Fig. 11 are not independent. More generally, any polygon formed in the interior of a diagram contains only one independent propagator. In the Green's function formalism and diagrams in the momentum space, this factor $\frac{1}{3}$ corresponds to the integral or summation over an inner variable **k** associated with an independent propagator (see for detail Refs. 4 and 26). The prescriptions (ii), (iii), (iv′), and (v) suffice to determine the contribution of every single-site diagram in the wide sense.

From a slightly different point of view, nested dia-

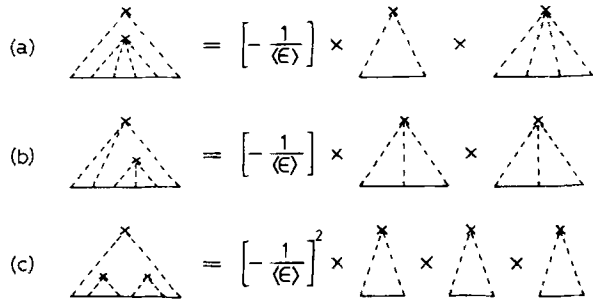


FIG. 10. Process for factorizing nested diagrams into lower-order one-vertex diagrams.

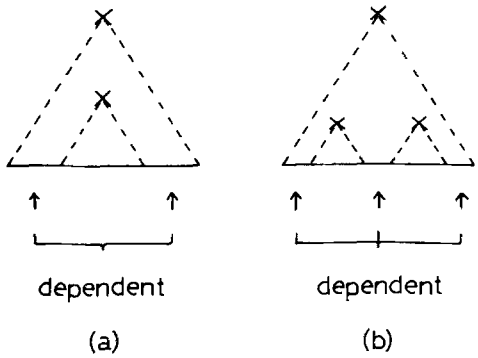


FIG. 11. Explanation for independent propagators.

grams can also be interpreted as one-vertex diagrams with renormalized propagators. Take the diagram 1(v) as an example. The process of renormalization is illustrated in Fig. 12. As stated in Fig. 12(a), the fourth-order nested diagram is equivalent to the second-order diagram with a renormalized propagator expressed by a thick line which is in turn defined by Fig. 12(b). It is interesting to note that, although all improper diagrams as shown in Fig. 7 vanish, Fig. 8 (b-vi) gives a nonzero contribution [see Fig. 10(c)] and this is the case for any nested diagram whose inner propagators involve improper diagrams.

For later convenience, we shall express the sum of all one-vertex and nested diagrams in terms of renormalized propagators. In Fig. 13(a), let a wavy interaction line denote the sum of all single-site diagrams in the wide sense. This wavy interaction line is also defined by Figs. 13(b) and (c) in a self-consistent manner. Figure 13(b) shows the sum of all one-vertex diagrams where renormalized propagators are expressed by double lines. These double-line propagators are related to the wavy interaction lines by the diagram equation in Fig. 13(c), whose second line may be regarded as a kind of Dyson equation. The equivalence of Figs. 13(b) and (c) to Fig. 13(a) is easily confirmed by repeated application of Fig. 13(c) to Fig. 13(b).

The contributions from crossed diagrams are normally difficult to compute but some of them are proved to vanish. As pointed out by Eqs. (2.14c) and (2.15c), diagrams such as Figs. 1(vi), 9(a-i), and 9(a-ii) contribute zero. For higher-order crossed diagrams we obtain an additional rule:

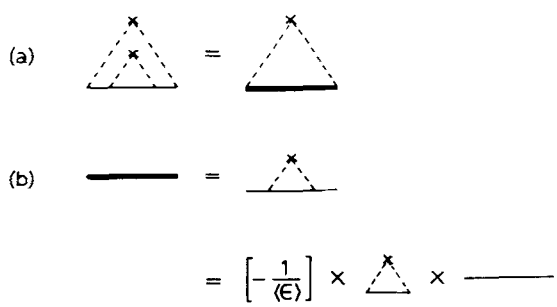


FIG. 12. Process for renormalizing an inner propagator in the fourth-order nested diagram.

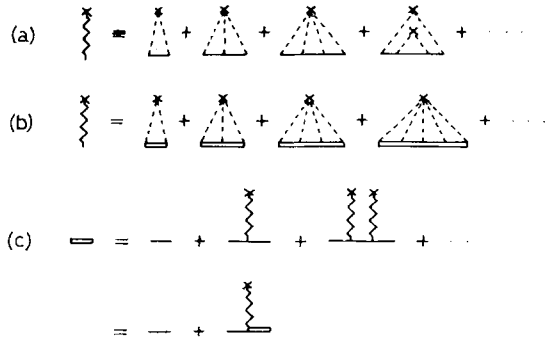


FIG. 13. Sum of all single-site diagrams in the wide sense. (a) Direct expression up to fourth order. (b) Indirect expression in terms of renormalized propagators. (c) Definition of a renormalized propagator.

(vi) Associate zero with crossed diagrams which are comprised as addends in Fig. 14(a).

As shown in Figs. 14(b) and (c), a wavy interaction line and a double-line propagator in Fig. 14(a) are the same as in Fig. 13. Actually, Fig. 14(b) and $\langle \epsilon' \rangle_c = 0$ imply Fig. 13(b). By formal iteration, the sixth-order crossed diagrams in Figs. 9(b-i), (b-ii), (b-v), (b-vi), (b-ix), and (b-xi) are found to belong to the category defined in Fig. 14(a). For the proof of the prescription (vi) we refer to Appendix B. The existence of crossed diagrams equal to zero is a characteristic feature of our problem which we do not encounter in the quantum-mechanical problems.

Concerning the rest of crossed diagrams, we have not been successful in proving zero contributions rigorously. In fact, we are not able to determine $B_{11}^{(5)}$ but conjecture that $B_{11}^{(5)} = 0$. Similarly, it is expected on inspection of mathematical structures that contributions from a number of other crossed diagrams vanish. However, there are still higher-order diagrams which give nonzero contributions. The higher the order of the term is, the more complicated the structure of the diagram becomes. Accordingly, an exact infinite summation of all nonvanishing diagrams is impossible. In the subsequent section, we shall seek a partial summation of most important terms up to an infinite order, following closely the procedures usually employed in the diagram methods.

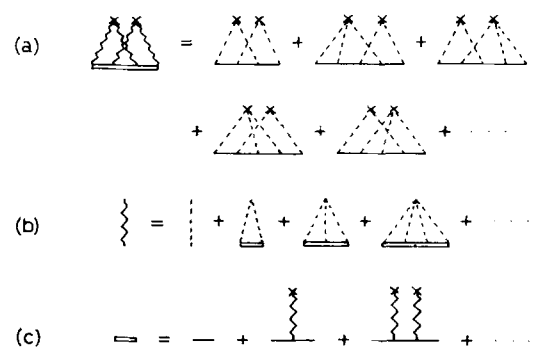


FIG. 14. Crossed diagrams whose contribution is zero. (a) Sum of vanishing crossed diagrams. (b) Definition of a wavy interaction line. (c) Definition of a double-line propagator.

Finally, we state the prescription for counting contributions from moment diagrams with open-circle or closed-circle vertices. For an unrestricted moment diagram whose vertices are indicated by closed circles, the νth-order cumulant $\langle\epsilon'^{\nu}\rangle_c$ in rule (ii) must be replaced by the νth-order moment $\langle\epsilon'^{\nu}\rangle$. The other prescriptions (i), (iii), (iv'), (v), and (vi) hold good for a diagram of this sort. As regards restricted moment diagrams with open-circle vertices, their contributions are evaluated indirectly from those of unrestricted moment diagrams.

3. VARIOUS APPROXIMATIONS TO THE PERTURBATION SERIES

A. Non-self-consistent cumulant solution

As the simplest case we take into consideration only single-site cumulant diagrams in the strict sense. The partition in this case is to group all variables in one subset, to which the factors $B_1^{(n)}=(\frac{1}{3})^{n-1}$ and $\epsilon(n,1)=\langle\epsilon^n\rangle_c$ are allotted. From Eq. (2.5) it follows that the partial summation of the perturbation series is expressed in the form

$$\begin{aligned}\epsilon_0^*&=\langle\epsilon\rangle+\sum_{n=2}^{\infty}(\tfrac{1}{3})^{n-1}\frac{\langle\epsilon^n\rangle_c}{(-\langle\epsilon\rangle)^{n-1}}\\&=\langle\epsilon\rangle-\frac{\langle\epsilon^2\rangle_c}{3\langle\epsilon\rangle}+\frac{\langle\epsilon^3\rangle_c}{9\langle\epsilon\rangle^2}-\frac{\langle\epsilon^4\rangle_c}{27\langle\epsilon\rangle^3}+\frac{\langle\epsilon^5\rangle_c}{81\langle\epsilon\rangle^4}-\cdots\\&=\langle\epsilon\rangle-\frac{\langle\epsilon'^2\rangle}{3\langle\epsilon\rangle}+\frac{\langle\epsilon'^3\rangle}{9\langle\epsilon\rangle^2}-\frac{\langle\epsilon'^4\rangle-3\langle\epsilon'^2\rangle^2}{27\langle\epsilon\rangle^3}+\frac{\langle\epsilon'^5\rangle-10\langle\epsilon'^3\rangle\langle\epsilon'^2\rangle}{81\langle\epsilon\rangle^4}-\cdots,\end{aligned}\tag{3.1}$$

which is identical with the diagram equation in Fig. 15(a). Clearly Eq. (3.1) is exact only up to third order but not for higher orders.

A trick to sum up the perturbation series in Eq. (3.1) has been given by Yonezawa and Matsubara.[32] According to this theory, we rewrite Eq. (3.1) as

$$\begin{aligned}\epsilon_0'^*&=\epsilon_0^*-\langle\epsilon\rangle=\sum_{n=1}^{\infty}\left(\frac{-1}{3\langle\epsilon\rangle}\right)^{n-1}\langle\epsilon'^n\rangle_c\\&=\left(\frac{-1}{3\langle\epsilon\rangle}\right)^{-1}\sum_{n=1}^{\infty}\left(\frac{-1}{3\langle\epsilon\rangle}\right)^{n}\langle\epsilon'^n\rangle_c\frac{1}{n!}\int_0^{\infty}z^ne^{-z}\,dz\\&=-3\langle\epsilon\rangle\int_0^{\infty}e^{-z}\sum_{n=1}^{\infty}\frac{1}{n!}\langle\epsilon'^n\rangle_c\left(\frac{-z}{3\langle\epsilon\rangle}\right)^n dz,\end{aligned}\tag{3.2}$$

where we have used $\langle\epsilon'\rangle_c=\langle\epsilon'\rangle=0$. On the other hand, the cumulant generating function is defined by

$$\log\langle\exp(-\epsilon' z/3\langle\epsilon\rangle)\rangle=\sum_{n=1}^{\infty}\frac{1}{n!}\langle\epsilon'^n\rangle_c\left(\frac{-z}{3\langle\epsilon\rangle}\right)^n,\tag{3.3}$$

so that

$$\begin{aligned}\epsilon_0'^*&=-3\langle\epsilon\rangle\int_0^{\infty}e^{-z}\log\langle\exp(-\epsilon' z/3\langle\epsilon\rangle)\rangle\,dz\\&=-3\langle\epsilon\rangle\int_0^1\log\langle z^{\epsilon'/3\langle\epsilon\rangle}\rangle\,dz.\end{aligned}\tag{3.4}$$

Integrating by parts, we have

$$\epsilon_0^*=\langle\epsilon\rangle+\epsilon_0'^*=\int_0^1\frac{\langle\epsilon z^{\epsilon/3\langle\epsilon\rangle}\rangle}{\langle z^{\epsilon/3\langle\epsilon\rangle}\rangle}dz.\tag{3.5}$$

B. Self-consistent cumulant solution

Next step usually followed in the diagram methods is to make the approximation self-consistent in the sense that an inner free propagator is renormalized. As stated in the foregoing section, the process of renormalizing inner propagators is given by Figs. 12 and 13. Let us adopt the diagram equation shown in Fig. 15(b), that is, identify a wavy interaction line with ϵ'^*. Then, all single-site diagrams in the wide sense are included in the summation. The renormalized propagator is calculated by Fig. 15(b) as

$$\begin{aligned}&\frac{-1}{\langle\epsilon\rangle}+\frac{-1}{\langle\epsilon\rangle}\epsilon'^*\frac{-1}{\langle\epsilon\rangle}+\frac{-1}{\langle\epsilon\rangle}\epsilon'^*\frac{-1}{\langle\epsilon\rangle}\epsilon'^*\frac{-1}{\langle\epsilon\rangle}+\cdots\\&\quad=\frac{-1}{\langle\epsilon\rangle}\frac{1}{1+\epsilon'^*/\langle\epsilon\rangle}=\frac{-1}{\langle\epsilon\rangle+\epsilon'^*}=-\frac{1}{\epsilon^*}.\end{aligned}\tag{3.6}$$

Consequently, a factor $-1/\epsilon^*$ is assigned to a renormalized propagator, whereas $-1/\langle\epsilon\rangle$ is associated with a free propagator. Substitution of ϵ^* for $\langle\epsilon\rangle$ on the right-hand side of Eq. (3.5) yields

$$\epsilon_C^*=\int_0^1\frac{\langle\epsilon z^{\epsilon/3\epsilon_C^*}\rangle}{\langle z^{\epsilon/3\epsilon_C^*}\rangle}dz,\tag{3.7}$$

or

$$\begin{aligned}\epsilon_C^*&=\langle\epsilon\rangle-\frac{\langle\epsilon^2\rangle_c}{3\langle\epsilon\rangle}+\frac{\langle\epsilon^3\rangle_c}{9\langle\epsilon\rangle^2}-\frac{\langle\epsilon^4\rangle_c+3\langle\epsilon^2\rangle_c^2}{27\langle\epsilon\rangle^3}\\&\quad+\frac{\langle\epsilon^5\rangle_c+9\langle\epsilon^3\rangle_c\langle\epsilon^2\rangle_c}{81\langle\epsilon\rangle^4}-\cdots\\&=\langle\epsilon\rangle-\frac{\langle\epsilon'^2\rangle}{3\langle\epsilon\rangle}+\frac{\langle\epsilon'^3\rangle}{9\langle\epsilon\rangle^2}-\frac{\langle\epsilon'^4\rangle}{27\langle\epsilon\rangle^3}+\frac{\langle\epsilon'^5\rangle-\langle\epsilon'^3\rangle\langle\epsilon'^2\rangle}{81\langle\epsilon\rangle^4}-\cdots,\end{aligned}\tag{3.8}$$

which is strictly valid up to fourth order. The physical meaning of the self-consistent cumulant solution will be discussed in the next section.

C. Kröner's approximation

In the non-self-consistent or self-consistent cumulant solution, cumulant averages play a more important role than that of ordinary moment averages. Suppose that the nth-order cumulant $\langle\epsilon'^n\rangle_c$ is approximated by the nth-order moment $\langle\epsilon'^n\rangle$. This assumption corresponds to the neglect of the exclusion effect. Now, the expansion series (3.1) becomes

$$\epsilon_K^*=\langle\epsilon\rangle+\sum_{n=2}^{\infty}(\tfrac{1}{3})^{n-1}\frac{\langle\epsilon'^n\rangle}{(-\langle\epsilon\rangle)^{n-1}}$$

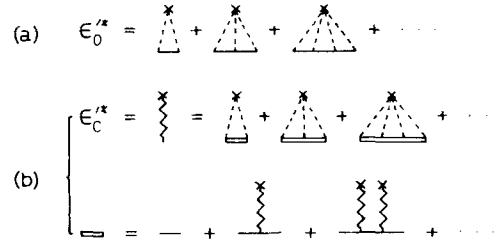


FIG. 15. Diagram equations for estimating ϵ_0^* and ϵ_C^*. (a) Non-self-consistent cumulant solution. (b) Self-consistent cumulant solution.

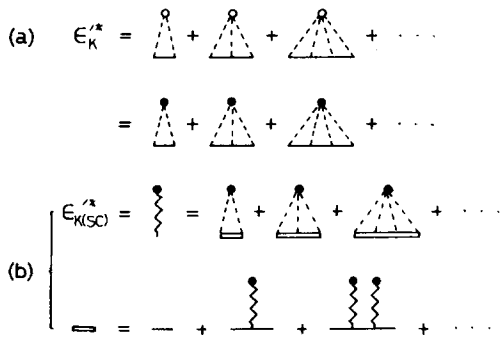


FIG. 16. Diagram equations for estimating ϵ^*_K and $\epsilon^*_{K(SC)}$. (a) Non-self-consistent Kröner solution. (b) Self-consistent Kröner solution.

$$= \langle\epsilon\rangle - \frac{\langle\epsilon'^2\rangle}{3\langle\epsilon\rangle} + \frac{\langle\epsilon'^3\rangle}{9\langle\epsilon\rangle^2} - \frac{\langle\epsilon'^4\rangle}{27\langle\epsilon\rangle^3} + \frac{\langle\epsilon'^5\rangle}{81\langle\epsilon\rangle^4} - \cdots$$

$$= \langle\epsilon\rangle - \left\langle \frac{\epsilon'^2}{3\langle\epsilon\rangle + \epsilon'} \right\rangle , \tag{3.9}$$

which is equivalent to Kröner's approximation[37-39] discussed in III. Diagrammatically, Eq. (3.9) reduces to the sum of restricted or unrestricted moment diagrams as depicted in Fig. 16(a). In the same way as employed in the previous subsection, Kröner's solution is made self-consistent by renormalizing inner propagators [see Fig. 16(b)]; the result is

$$\epsilon^*_{K(SC)} = \langle\epsilon\rangle + \sum_{n=2}^{\infty} \frac{\langle\epsilon'^n\rangle}{(-3\epsilon^*_{K(SC)})^{n-1}} = \langle\epsilon\rangle - \left\langle \frac{\epsilon'^2}{3\epsilon^*_{K(SC)} + \epsilon'} \right\rangle$$

$$= \langle\epsilon\rangle - \frac{\langle\epsilon'^2\rangle}{3\langle\epsilon\rangle} + \frac{\langle\epsilon'^3\rangle}{9\langle\epsilon\rangle^2} - \frac{\langle\epsilon'^4\rangle + 3\langle\epsilon'^2\rangle^2}{27\langle\epsilon\rangle^3}$$

$$+ \frac{\langle\epsilon'^5\rangle + 9\langle\epsilon'^3\rangle\langle\epsilon'^2\rangle}{81\langle\epsilon\rangle^4} + \cdots . \tag{3.10}$$

D. Self-contained treatment for the cumulant expansion

As can be seen from the derivation process in III, the cumulant average associated with each cross vertex is the result of corrections from more complicated diagrams. For illustration, consider the fourth-order diagrams in Fig. 6. The cumulant assigned to the diagram 6 (c-i) has the form

$$\langle\epsilon^4\rangle_c = \langle\epsilon'^4\rangle_c = \langle\epsilon'^4\rangle - 3\langle\epsilon'^2\rangle^2 . \tag{3.11}$$

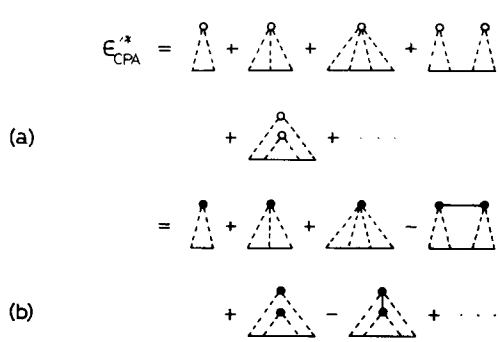


FIG. 17. Diagrammatic expansion of ϵ^*_{CPA} up to fourth order. (a) Expression by means of restricted moment diagrams. (b) Expression by means of unrestricted moment diagrams.

The second term on the right-hand side comes from the diagrams 6(b-ii'), (b-iii'), and (b-iv'), which are the correction terms attributable to the diagrams 6(c-ii), (c-iii), and (c-iv), respectively. The exclusion effect requires a factor $1-\delta_{\mathbf{r}_{23}}$ with Fig. 6(a-ii), and $1-\delta_{\mathbf{r}_{12}}$ with Figs. 6(a-iii) and (a-iv); namely,

$$\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\epsilon'(\mathbf{r}_3)\epsilon'(\mathbf{r}_4)\rangle$$

$$= \langle\epsilon'^4\rangle\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{34}} + \langle\epsilon'^2\rangle^2[\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{34}}(1-\delta_{\mathbf{r}_{23}})$$

$$+ \delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}}(1-\delta_{\mathbf{r}_{12}}) + \delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{24}}(1-\delta_{\mathbf{r}_{12}})]$$

$$= (\langle\epsilon'^4\rangle - 3\langle\epsilon'^2\rangle^2)\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{34}} + \langle\epsilon'^2\rangle^2(\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{34}}$$

$$+ \delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}} + \delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{24}}) . \tag{3.12}$$

Thus, $-3\langle\epsilon'^2\rangle^2$ is introduced in the contribution of the diagram 6(c-i).

Now, in this context, it is important to notice that, in the renormalized single-site approximation, all of the crossed diagrams such as those in Fig. 9 are discarded, and yet the cumulants included in Fig. 15(b) take care of all correction terms including those from discarded diagrams. In other words, the renormalized cumulant solution (3.7) is not self-contained since the approximation and the way of choosing corrections are not consistent.[4] The self-contained treatment is performed by discarding crossed diagrams with open-circle vertices rather than crossed diagrams with cross vertices. The corresponding diagram expansion up to fourth order is given in Fig. 17. Detailed discussion and analysis of this problem are given in Ref. 4 at full length. It is shown that the consideration on the self-containedness leads to the renormalization of interaction lines as illustrated by Fig. 18, where the renormalized interaction lines are denoted by dotted lines. In the case of disordered binary alloys treated in Ref. 4, the solution thus obtained turns out to be equivalent to the result of the CPA.

For our problem, the diagram equation in Fig. 18 is written as

$$\epsilon^*_{CPA} = \langle\epsilon\rangle + \sum_{n=2}^{\infty} \frac{\langle\epsilon'(\epsilon-\epsilon^*_{CPA})^{n-1}\rangle}{(-3\epsilon^*_{CPA})^{n-1}} , \tag{3.13}$$

which after some manipulation becomes

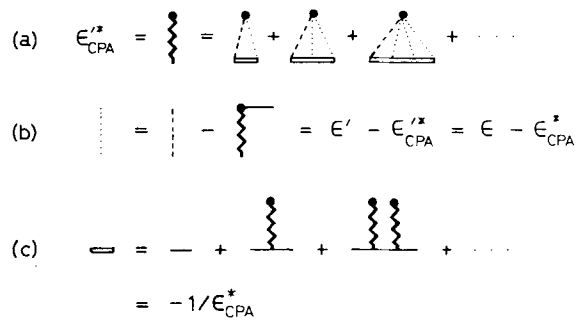


FIG. 18. Diagram equations for estimating ϵ^*_{CPA}. (a) Self-contained single-site approximation. (b) Definition of a renormalized interaction line. (c) Definition of a modified propagator (after Yonezawa[4]).

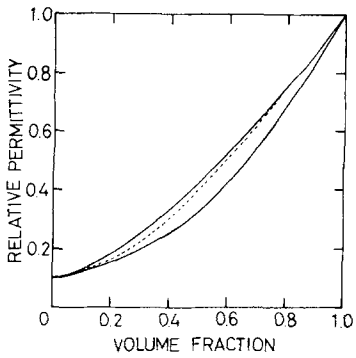


FIG. 19. Plots of the relative permittivity ϵ^*_C/ϵ_1 and $\epsilon^*_{CPA}/\epsilon_1$ against the volume fraction x for the choice of $\epsilon_2/\epsilon_1 = 0.1$. The dashed line represents ϵ^*_C and ϵ^*_{CPA}, and the solid lines indicate the bounds on ϵ^*.

$$\left\langle \frac{\epsilon - \epsilon^*_{CPA}}{\epsilon + 2\epsilon^*_{CPA}} \right\rangle = 0, \tag{3.14}$$

or

$$\epsilon^*_{CPA} = \langle \epsilon \rangle - \left\langle \frac{(\epsilon - \epsilon^*_{CPA})^2}{\epsilon + 2\epsilon^*_{CPA}} \right\rangle. \tag{3.15}$$

Equation (3.14) is nothing but the condition for the EM approximation described in the succeeding section. The explicit expansion series up to fifth order is

$$\begin{aligned}\epsilon^*_{CPA} &= \langle \epsilon \rangle - \frac{\langle \epsilon^2 \rangle_c}{3\langle \epsilon \rangle} + \frac{\langle \epsilon^3 \rangle_c}{9\langle \epsilon \rangle^2} - \frac{\langle \epsilon^4 \rangle_c + 4\langle \epsilon^2 \rangle_c^2}{27\langle \epsilon \rangle^3} \\ &\quad + \frac{\langle \epsilon^5 \rangle_c + 14\langle \epsilon^3 \rangle_c \langle \epsilon^2 \rangle_c}{81\langle \epsilon \rangle^4} - \cdots \\ &= \langle \epsilon \rangle - \frac{\langle \epsilon'^2 \rangle}{3\langle \epsilon \rangle} + \frac{\langle \epsilon'^3 \rangle}{9\langle \epsilon \rangle^2} - \frac{\langle \epsilon'^4 \rangle + \langle \epsilon'^2 \rangle^2}{27\langle \epsilon \rangle^3} \\ &\quad + \frac{\langle \epsilon'^5 \rangle + 4\langle \epsilon'^3 \rangle \langle \epsilon'^2 \rangle}{81\langle \epsilon \rangle^4} - \cdots .\end{aligned} \tag{3.16}$$

In contrast with the cumulant solution (3.8), Eq. (3.16) is not correct in fourth order. Of course, this does not contradict the assertion that Eq. (3.14) is the best single-site approximation to the effective permittivity ϵ^*.

E. Numerical calculation on two-phase materials

In order to check the validity of various approximate solutions, we shall compare some numerical results for the effective permittivity of a completely random two-phase material. Let the two constituting phases have permittivities ϵ_1 and ϵ_2 and occupy fractions of the total volume $v_1 = x$ and $v_2 = 1 - x$. In this case it is seen that

$$\langle \epsilon \rangle = \epsilon_1 v_1 + \epsilon_2 v_2 = \epsilon_2 + (\epsilon_1 - \epsilon_2)x, \tag{3.17}$$

$$\langle \epsilon'^n \rangle = (\epsilon_1 - \langle \epsilon \rangle)^n v_1 + (\epsilon_2 - \langle \epsilon \rangle)^n v_2 = (\epsilon_1 - \epsilon_2)^n R_n(x), \tag{3.18}$$

where

$$R_n(x) = x(1-x)[(1-x)^{n-1} - (-x)^{n-1}]. \tag{3.19}$$

Similarly, we can write

$$\langle \epsilon^n \rangle_c = (\epsilon_1 - \epsilon_2)^n P_n(x); \tag{3.20}$$

$P_n(x)$ has been found in Ref. 32 to possess a generating function such that

$$h(\xi; x) = \log(1 - x + xe^{\xi}) = \sum_{n=1}^{\infty} P_n(x) \frac{\xi^n}{n!}. \tag{3.21}$$

From Eqs. (3.5), (3.7), (3.9), (3.10), and (3.14) we obtain

$$\epsilon^*_0 = \epsilon_2 + (\epsilon_1 - \epsilon_2)x \times \int_0^1 dz (x + (1-x)z^{-(\epsilon_1 - \epsilon_2)/\{3[\epsilon_2 + (\epsilon_1 - \epsilon_2)x]\}})^{-1}, \tag{3.22}$$

$$\epsilon^*_C = \epsilon_2 + (\epsilon_1 - \epsilon_2)x \times \int_0^1 dz (x + (1-x)z^{-(\epsilon_1 - \epsilon_2)/3\epsilon^*_C})^{-1}, \tag{3.23}$$

$$\epsilon^*_K = [\epsilon_2 + (\epsilon_1 - \epsilon_2)x]\left(1 - 3(\epsilon_1 - \epsilon_2)^2 \times \frac{x(1-x)}{[\epsilon_1 + 2\epsilon_2 + 2(\epsilon_1 - \epsilon_2)x][3\epsilon_2 + 2(\epsilon_1 - \epsilon_2)x]}\right), \tag{3.24}$$

$$\begin{aligned}&9\epsilon^*_{K(SC)} + 3[\epsilon_1 - 4\epsilon_2 - 5(\epsilon_1 - \epsilon_2)x]\epsilon^{*2}_{K(SC)} \\ &\quad - (\epsilon_1 - \epsilon_2)[3\epsilon_2 + (\epsilon_1 - 7\epsilon_2)x - 4(\epsilon_1 - \epsilon_2)x^2]\epsilon^*_{K(SC)} \\ &\quad + (\epsilon_1 - \epsilon_2)^2 x(1-x)[\epsilon_2 + (\epsilon_1 - \epsilon_2)x] = 0,\end{aligned} \tag{3.25}$$

$$2\epsilon^{*2}_{CPA} + [\epsilon_1 - 2\epsilon_2 - 3(\epsilon_1 - \epsilon_2)x]\epsilon^*_{CPA} - \epsilon_1\epsilon_2 = 0. \tag{3.26}$$

The root of the quadratic equation (3.26) is

$$\epsilon^*_{CPA} = \tfrac{1}{4}(-\epsilon_1 + 2\epsilon_2 + 3(\epsilon_1 - \epsilon_2)x + [(\epsilon_1 + 2\epsilon_2)^2 - 6(\epsilon_1 - \epsilon_2)(\epsilon_1 - 2\epsilon_2)x + 9(\epsilon_1 - \epsilon_2)^2 x^2]^{1/2}). \tag{3.26'}$$

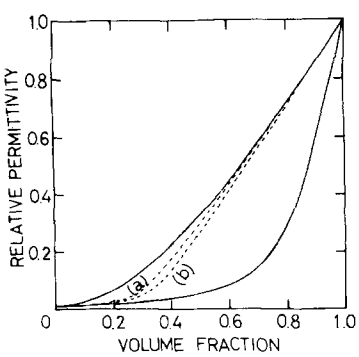


FIG. 20. Plots of the relative permittivity ϵ^*_C/ϵ_1 and $\epsilon^*_{CPA}/\epsilon_1$ against the volume fraction x for the choice of $\epsilon_2/\epsilon_1 = 0.01$. The dashed lines (a) and (b) represent ϵ^*_C and ϵ^*_{CPA}, respectively, and the solid lines indicate the bounds on ϵ^*.

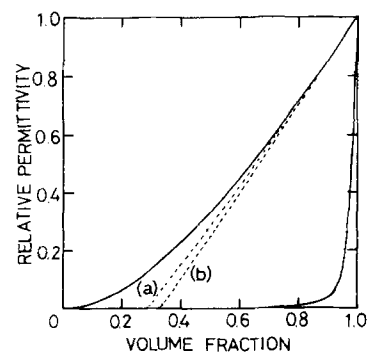


FIG. 21. Plots of the relative permittivity ϵ^*_C/ϵ_1 and $\epsilon^*_{CPA}/\epsilon_1$ against the volume fraction x for the choice of $\epsilon_2/\epsilon_1 = 2\times10^{-4}$. The dashed lines (a) and (b) represent ϵ^*_C and ϵ^*_{CPA}, respectively, and the solid lines indicate the bounds on ϵ^*.

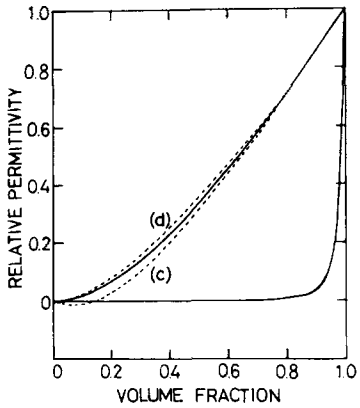


FIG. 22. Plots of the relative permittivity ϵ_K^*/ϵ_1 and ϵ_0^*/ϵ_1 against the volume fraction x for the choice of $\epsilon_2/\epsilon_1 = 2\times10^{-4}$. The dashed lines (c) and (d) represent ϵ_K^* and ϵ_0^*, respectively, and the solid lines indicate the bounds on ϵ^*.

Upper and lower bounds of ϵ^* are expressed by the inequalities (2.30) and (2.31) in III; that is,

$$\epsilon^* \leq \epsilon_2 + (\epsilon_1 - \epsilon_2)x - \frac{(\epsilon_1 - \epsilon_2)^2 x(1-x)}{\epsilon_1 + 2\epsilon_2 + (\epsilon_1 - \epsilon_2)x}, \tag{3.27}$$

$$\epsilon^* \geq \frac{\epsilon_1\epsilon_2}{\epsilon_1 - (\epsilon_1 - \epsilon_2)x - \{2(\epsilon_1 - \epsilon_2)^2 x(1-x)/[\epsilon_1 + 2\epsilon_2 + (\epsilon_1 - \epsilon_2)x]\}}. \tag{3.28}$$

It was pointed out in III that Kröner's solution cannot be used in the whole range of parameters because Eq. (3.24) fails to give a nonnegative permittivity for some concentration regions. Especially when $\epsilon_1 \to \infty$ or $\epsilon_2 = 0$, ϵ_K^* takes negative values for $x < 1/7$. This demerit is partially improved by the self-consistent treatment since Eq. (3.25) gives nonnegative solutions over all parameter regions. For some values of x and ϵ_2/ϵ_1, however, $\epsilon_{K(SC)}^*$ becomes smaller than the lower bound (3.28). The non-self-consistent cumulant solution (3.22) always exceeds the upper bound (3.27). On the other hand, ϵ_C^* and ϵ_{CPA}^* given by Eqs. (3.23) and (3.26′) are safely within the bounds. Numerical results for the three cases of $\epsilon_2/\epsilon_1 = 0.1$, 0.01, 2×10^{-4} are plotted in Figs. 19—22.

It is worth while to remark that all the approximate solutions ϵ_K^*, $\epsilon_{K(SC)}^*$, ϵ_0^*, ϵ_C^*, and ϵ_{CPA}^* obtained in this section correctly reproduce the behavior of effective permittivity in the dilute limit. According to the theory of dilute suspensions,[40] the effective permittivity of a dilute suspension where spherical particles of permittivity ϵ_1 are embedded in a matrix of permittivity ϵ_2 takes the form

$$\frac{\epsilon^* - \epsilon_2}{\epsilon^* + 2\epsilon_2} = \frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + 2\epsilon_2}x, \tag{3.29}$$

or

$$\frac{\epsilon^*}{\epsilon_2} = 1 + \frac{3(\epsilon_1 - \epsilon_2)}{\epsilon_1 + 2\epsilon_2}x + \cdots, \tag{3.29′}$$

under the condition that the inclusion concentration x is very small. Since $\langle\epsilon'^n\rangle$ or $\langle\epsilon^n\rangle_c$ has a dominant term proportional to x, perturbation expansions involving all one-vertex diagrams must be exact in the low-concentration limit. Actually, it may be readily confirmed that Eqs. (3.22)—(3.26) as well as the bounding equations (3.27) and (3.28) are identical with Eq. (3.29′) in the first order of x.

4. COMPARISON OF THE PERTURBATION SOLUTIONS WITH THE EFFECTIVE-MEDIUM THEORY

The effective-medium theory based upon the self-consistent local field concept has been widely applied in predicting the electrical[5–13] and mechanical[14–25] properties of heterogeneous materials. The essence of this method is to solve the field equation for a representative element of the material, which is taken to be embedded in an effective medium with the as yet unknown physical constant. The effective constant is determined in turn by requiring that the effects of the deviation from the true constant shall, on the average, cancel out. Therefore, the basic principle underlying the classical EM theory is closely akin to that leading to the CPA in the quantum mechanics of random alloys.[27–29]

Let ϵ_{EM}^* be the EM solution for the effective permittivity of a random mixture composed of spherical particles. Consider first a sphere of permittivity ϵ_i embedded in an infinite medium of permittivity ϵ_o under the influence of a uniform external field $\mathbf{E}_o$. As is well known, the electric field $\mathbf{E}_i$ inside the sphere is related to $\mathbf{E}_o$ by

$$\mathbf{E}_i = \left(1 - \frac{\epsilon_i - \epsilon_o}{\epsilon_i + 2\epsilon_o}\right)\mathbf{E}_o. \tag{4.1}$$

The effective-medium assumption and the ergodic hypothesis assert that for the electric field $\mathbf{E}(\mathbf{r})$ in the mixture

$$\mathbf{E}(\mathbf{r}) = \left(1 - \frac{\epsilon(\mathbf{r}) - \epsilon_{EM}^*}{\epsilon(\mathbf{r}) + 2\epsilon_{EM}^*}\right)\langle\mathbf{E}\rangle, \tag{4.2}$$

whence

$$\left\langle \frac{\epsilon - \epsilon_{EM}^*}{\epsilon + 2\epsilon_{EM}^*} \right\rangle = 0, \tag{4.3}$$

or equivalently,

$$\epsilon_{EM}^* \left\langle \frac{1}{\epsilon + 2\epsilon_{EM}^*} \right\rangle = \left\langle \frac{\epsilon}{\epsilon + 2\epsilon_{EM}^*} \right\rangle = \frac{1}{3}. \tag{4.4}$$

From Eqs. (3.14) and (4.3), we can see that the CPA result satisfies the same condition as that of the EM approximation, so that

$$\epsilon_{EM}^* = \epsilon_{CPA}^*. \tag{4.5}$$

Thus we reach the conclusion that the EM theory is equivalent to the CPA not only in its physical concept but also in its mathematical structure. In the framework of single-site approximations, the CPA is known to be the best possible from both physical and mathematical points of view.[26] The CPA or EM theory works very well for explaining the overall properties of random inhomogeneous materials. Particularly, it has been shown that the EM theory for a random mixture of conducting and insulating materials gives a remarkably good agreement with the results of the computer simulation of a simple cubic network of resistors.[27,29,31] Nevertheless, the EM solution deviates from the computer results near the critical percolation concentration.

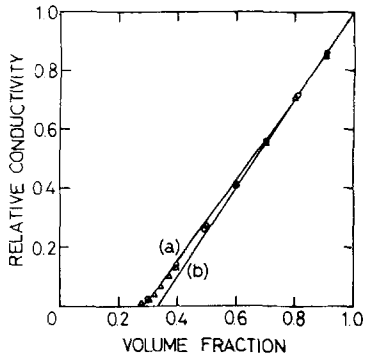


FIG. 23. Comparison of σ_C^*/σ_1 and σ_{EM}^*/σ_1 with the simulation data of Kirkpatrick.[27,29,31] The solid lines (a) and (b) represent σ_C^* and σ_{EM}^* for $\sigma_2/\sigma_1=0$, respectively. Data points indicate the results of the computer simulation for simple cubic networks which range in size from 15×15×15 to 25×25×25.

Let us treat a random binary mixture which consists of a conducting material with conductivity σ_1 and volume fraction x and an insulating material with conductivity 0 and volume fraction $1-x$. Then, the CPA or EM solution corresponding to Eq. (3.26′) is written as

$$\frac{\sigma_{EM}^*}{\sigma_1}=\begin{cases}0 & \text{for } x<\frac{1}{3},\\ (-1+3x)/2 & \text{for } x>\frac{1}{3}.\end{cases} \tag{4.6}$$

In other words, the concentration at which the effective conductivity σ_{EM}^* vanishes is $\tilde{x}_c=1/3$. Moreover, Eq. (4.6) suggests a scaling law with a critical exponent $s=1$; namely,

$$\sigma_{EM}^* \propto (x-\tilde{x}_c)^s, \quad s=1. \tag{4.7}$$

From the standpoint of percolation theory,[29,41,42] the critical concentration can be interpreted as the critical probability of bond percolation on a simple cubic lattice. The estimates obtained from the percolation theory and the computer simulation are $x_c\approx 0.25$ and $s\approx 1.6$.[27,29,31]

This breakdown of the EM theory is due to the fact that the CPA or EM approximation neglects essential clustering effects. Note that the clustering effects are most important near the critical percolation concentration where infinite islands of conducting materials begin to form. Generally speaking, however, the effects of clusters are very difficult to take into account in any sensible manner, but in our problem we are extremely fortunate because some of the most important crossed diagrams give zero contributions. For instance, all diagrams comprised in Fig. 14 drop out and the contributions from some other crossed diagrams are predicted to vanish on inspecting their mathematical structures. This implies that the self-consistent cumulant solution is superior to the EM approximation. Although the crossed diagrams themselves such as those in Fig. 14 contribute nothing, the exclusion correction terms from them are not zero and these corrections are exactly counted in the cumulant solution. Here we recall that the fourth-order term of σ_C^* is correct in contrast to σ_{EM}^*. Hence it is concluded that the self-consistent cumulant solution takes into consideration most of the clustering effects.

The above argument is supported by the numerical success of the self-consistent cumulant solution. In the present case Eq. (3.23) reduces to

$$\frac{\sigma_C^*}{\sigma_1}=x\int_0^1 \frac{z^{\sigma_1/3\sigma_C^*}}{1-x+xz^{\sigma_1/3\sigma_C^*}}dz. \tag{4.8}$$

As illustrated in Fig. 23, the numerical agreement of σ_C^* with the simulation data of Kirkpatrick[27,29,31] is truly excellent for the whole range of the concentration x including the vicinity of the critical concentration. The critical concentration calculated by Eq. (4.8) is

$$\tilde{x}_c=1-e^{-1/3}=0.283\cdots, \tag{4.9}$$

which is a little higher than the correct value of x_c (see Appendix C). The little difference is due to complicated higher-order clusters of nonzero contributions which are not included in σ_C^*. These higher-order clusters become distinguished only in the region very close to the critical percolation concentration. The behavior of σ_C^* near $\tilde{x}_c$ is expressed in the form

$$\sigma_C^*/\sigma_1=a(x-\tilde{x}_c)+o(x-\tilde{x}_c), \tag{4.10}$$

where

$$a=-\left(3\tilde{x}_c(1-\tilde{x}_c)\int_0^1 \frac{\log Z\,dZ}{1-\tilde{x}_c+\tilde{x}_cZ}\right)^{-1}=1.284\cdots. \tag{4.11}$$

ACKNOWLEDGMENT

The authors would like to thank Mr. Y. Ishida for his assistance in numerical calculations.

APPENDIX A

We shall evaluate the contributions from nested diagrams of sixth order shown in Fig. 8(b). For the diagrams (b-i)—(b-v), the procedure goes exactly as in Appendix D of III. The diagrams (b-i)—(b-iii) lead to $-\langle\epsilon^4\rangle_c\langle\epsilon^2\rangle_c/81\langle\epsilon\rangle^5$, (b-iv) to $-\langle\epsilon^3\rangle_c^2/81\langle\epsilon\rangle^5$, and (b-v) to $-\langle\epsilon^2\rangle_c^3/27\langle\epsilon\rangle^5$. To calculate the contribution of the last diagram we begin with the relation

$$\delta_{\mathbf{r}_{16}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{45}}=\delta_{\mathbf{r}_{12}+\mathbf{r}_{34}+\mathbf{r}_{56}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{45}}, \tag{A1}$$

from which we get

$$\begin{aligned}&-\frac{1}{(4\pi)^5}\int_V d\omega_{12}\int_V d\omega_{23}\int_V d\omega_{34}\int_V d\omega_{45}\int_V d\omega_{56}\\ &\times\frac{x_{12,i}}{r_{12}^3}\frac{x_{23,k}}{r_{23}^3}\frac{x_{34,k}}{r_{34}^3}\frac{x_{45,l}}{r_{45}^3}\frac{x_{56,m}}{r_{56}^3}\\ &\times\frac{\partial^5\delta_{\mathbf{r}_{16}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{45}}}{\partial x_{12,k}\partial x_{23,k}\partial x_{34,l}\partial x_{45,m}\partial x_{56,(i)}}\\ &=-\frac{1}{9(4\pi)^3}\int_V d\omega_{12}\int_V d\omega_{34}\int_V d\omega_{56}\frac{x_{12,i}}{r_{12}^3}\frac{x_{34,k}}{r_{34}^3}\frac{x_{56,l}}{r_{56}^3}\\ &\times\frac{\partial^3\delta_{\mathbf{r}_{12}+\mathbf{r}_{34}+\mathbf{r}_{56}}}{\partial x_{12,k}\partial x_{34,l}\partial x_{56,(i)}}.\end{aligned} \tag{A2}$$

By means of Eqs. (3.13), (3.14), (4.1), (4.3), and (C4) in III, Eq. (A2) is converted into

$$\lim_{\rho\to+0}\frac{1}{9(4\pi)^2}\int_V d\omega_{34}\int_V d\omega_{56}\frac{x_{34,k}}{r_{34}^3}\frac{x_{56,l}}{r_{56}^3}\frac{\partial^2 J_{\rho,ik}(\mathbf{r}_{34}+\mathbf{r}_{56})}{\partial x_{34,l}\partial x_{56,(i)}}$$

$$= \lim_{\rho \to +0} \left(-\frac{1}{9 \cdot 4\pi} \right) \int_V d\omega_{56} \frac{x_{56,i}}{r_{56}^3} \frac{\partial J_{\rho, ii}(\mathbf{r}_{56})}{\partial x_{56,(i)}}$$

$$= \lim_{\rho \to +0} \tfrac{1}{9} J_{\rho, i(i)}(0) = 1/27. \tag{A3}$$

Thus the diagram (b-vi) also gives $-\langle \epsilon^2 \rangle_c^3 / 27 \langle \epsilon \rangle^5$. In the same way, nested diagrams representing perturbation terms of arbitrary order can be reduced to products of lower-order single-site diagrams.

APPENDIX B

Out of the crossed diagrams in Fig. 9(b), the diagrams (b-i), (b-ii), (b-v), (b-vi), (b-ix), and (b-xi) belong to the category defined by Fig. 14. In a similar manner to that developed in Appendix D of III, the first four diagrams are easily proved to vanish. For the diagram (b-ix), we have

$$\delta_{\mathbf{r}_{13}} \delta_{\mathbf{r}_{26}} \delta_{\mathbf{r}_{45}} = \delta_{\mathbf{r}_{12}+\mathbf{r}_{23}} \delta_{\mathbf{r}_{23}+\mathbf{r}_{34}+\mathbf{r}_{56}} \delta_{\mathbf{r}_{45}}, \tag{B1}$$

whence

$$-\frac{1}{(4\pi)^5} \int_V d\omega_{12} \int_V d\omega_{23} \int_V d\omega_{34} \int_V d\omega_{45} \int_V d\omega_{56} \frac{x_{12,i}}{r_{12}^3} \frac{x_{23,k}}{r_{23}^3} \frac{x_{34,h}}{r_{34}^3} \times \frac{x_{45,l}}{r_{45}^3} \frac{x_{56,m}}{r_{56}^3} \frac{\partial^5 \delta_{\mathbf{r}_{13}} \delta_{\mathbf{r}_{26}} \delta_{\mathbf{r}_{45}}}{\partial x_{12,k} \partial x_{23,h} \partial x_{34,l} \partial x_{45,m} \partial x_{56,(i)}}$$

$$= \frac{1}{3(4\pi)^4} \int_V d\omega_{12} \int_V d\omega_{23} \int_V d\omega_{34} \int_V d\omega_{56} \frac{x_{12,i}}{r_{12}^3} \frac{x_{23,k}}{r_{23}^3} \frac{x_{34,h}}{r_{34}^3} \times \frac{x_{56,l}}{r_{56}^3} \frac{\partial^4 \delta_{\mathbf{r}_{12}+\mathbf{r}_{23}} \delta_{\mathbf{r}_{23}+\mathbf{r}_{34}+\mathbf{r}_{56}}}{\partial x_{12,k} \partial x_{23,h} \partial x_{34,l} \partial x_{56,(i)}}. \tag{B2}$$

The argument in Appendix A shows that

$$-\frac{1}{(4\pi)^3} \int_V d\omega_{23} \int_V d\omega_{34} \int_V d\omega_{56} \frac{x_{23,k}}{r_{23}^3} \frac{x_{34,h}}{r_{34}^3} \frac{x_{56,l}}{r_{56}^3} \times \frac{\partial^3 \delta_{\mathbf{r}_{23}+\mathbf{r}_{34}+\mathbf{r}_{56}}}{\partial x_{23,h} \partial x_{34,l} \partial x_{56,(i)}}$$

$$= \frac{1}{(4\pi)^2} \int_V d\omega_{23} \int_V d\omega_{56} \frac{x_{23,k}}{r_{23}^3} \frac{x_{56,h}}{r_{56}^3} \frac{\partial^2 \delta_{\mathbf{r}_{23}+\mathbf{r}_{56}}}{\partial x_{23,h} \partial x_{56,(i)}}. \tag{B3}$$

Consequently, Eq. (B2) becomes

$$-\frac{1}{3(4\pi)^3} \int_V d\omega_{12} \int_V d\omega_{23} \int_V d\omega_{56} \frac{x_{12,i}}{r_{12}^3} \frac{x_{23,k}}{r_{23}^3} \frac{x_{56,h}}{r_{56}^3} \times \frac{\partial^3 \delta_{\mathbf{r}_{12}+\mathbf{r}_{23}} \delta_{\mathbf{r}_{23}+\mathbf{r}_{56}}}{\partial x_{12,k} \partial x_{23,h} \partial x_{56,(i)}} = \tfrac{1}{3} B_4^{(4)} = 0. \tag{B4}$$

Now we turn to the estimation of the diagram (b-xi). The product of δ's associated with this diagram is

$$\delta_{\mathbf{r}_{15}} \delta_{\mathbf{r}_{26}} \delta_{\mathbf{r}_{34}} = \delta_{\mathbf{r}_{12}+\mathbf{r}_{23}+\mathbf{r}_{45}} \delta_{\mathbf{r}_{23}+\mathbf{r}_{45}+\mathbf{r}_{56}} \delta_{\mathbf{r}_{34}}, \tag{B5}$$

so that

$$-\frac{1}{(4\pi)^5} \int_V d\omega_{12} \int_V d\omega_{23} \int_V d\omega_{34} \int_V d\omega_{45} \int_V d\omega_{56} \frac{x_{12,i}}{r_{12}^3} \frac{x_{23,k}}{r_{23}^3} \frac{x_{34,h}}{r_{34}^3} \times \frac{x_{45,l}}{r_{45}^3} \frac{x_{56,m}}{r_{56}^3} \frac{\partial^5 \delta_{\mathbf{r}_{15}} \delta_{\mathbf{r}_{26}} \delta_{\mathbf{r}_{34}}}{\partial x_{12,k} \partial x_{23,h} \partial x_{34,l} \partial x_{45,m} \partial x_{56,(i)}}$$

$$= \frac{1}{3(4\pi)^4} \int_V d\omega_{12} \int_V d\omega_{23} \int_V d\omega_{45} \int_V d\omega_{56} \frac{x_{12,i}}{r_{12}^3} \frac{x_{23,k}}{r_{23}^3} \frac{x_{45,h}}{r_{45}^3} \frac{x_{56,m}}{r_{56}^3} \times \frac{\partial^4 \delta_{\mathbf{r}_{12}+\mathbf{r}_{23}+\mathbf{r}_{45}} \delta_{\mathbf{r}_{23}+\mathbf{r}_{45}+\mathbf{r}_{56}}}{\partial x_{12,k} \partial x_{23,h} \partial x_{45,m} \partial x_{56,(i)}}. \tag{B6}$$

From Eq. (III-C3) it follows that

$$\frac{1}{(4\pi)^2} \int_V dw_{23} \int_V d\omega_{45} \frac{x_{23,k}}{r_{23}^3} \frac{x_{45,h}}{r_{45}^3} \frac{\partial^2 \delta_{\mathbf{r}_{12}+\mathbf{r}_{23}+\mathbf{r}_{45}} \delta_{\mathbf{r}_{23}+\mathbf{r}_{45}+\mathbf{r}_{56}}}{\partial x_{23,h} \partial x_{45,m}}$$

$$= -\frac{1}{4\pi} \int_V d\omega \frac{x_k}{r^3} \frac{\partial \delta_{\mathbf{r}+\mathbf{r}_{12}} \delta_{\mathbf{r}+\mathbf{r}_{56}}}{\partial x_m}. \tag{B7}$$

Therefore, Eq. (B6) reduces to

$$-\frac{1}{3(4\pi)^3} \int_V d\omega_{12} \int_V d\omega \int_V d\omega_{56} \frac{x_{12,i}}{r_{12}^3} \frac{x_k}{r^3} \frac{x_{56,m}}{r_{56}^3} \times \frac{\partial^3 \delta_{\mathbf{r}_{12}+\mathbf{r}} \delta_{\mathbf{r}+\mathbf{r}_{56}}}{\partial x_{12,k} \partial x_m \partial x_{56,(i)}} = \tfrac{1}{3} B_4^{(4)} = 0. \tag{B8}$$

Similarly, we may demonstrate that any crossed diagram included in Fig. 14 contributes zero.

APPENDIX C

Equation (4.8) is recast as

$$\xi = x \int_0^1 \frac{z^{1/3\xi}}{1 - x + x z^{1/3\xi}} dz$$

$$= 1 - (1 - x) \int_0^1 \frac{dz}{1 - x + x z^{1/3\xi}}, \tag{C1}$$

where $\xi = \sigma_c^* / \sigma_1$. It is seen that ξ is a continuous function in the interval of $0 \leq x \leq 1$ which satisfies $\xi = 0$ for $x = 0$, $\xi = 1$ for $x = 1$, and $0 \leq \xi \leq 1$ for all x. Putting $y = z^{1/3\xi}$, we obtain

$$\frac{1}{x} = 3 \int_0^1 \frac{y^{3\xi}}{1 - x + xy} dy, \tag{C2}$$

whenever $\xi \neq 0$. Now suppose that ξ sometimes decreases with increasing x. Then, the left-hand side decreases as x increases, while the integrand on the right-hand side increases; this is a contradiction. Therefore, ξ is considered to be a nonnegative, monotone nondecreasing, and continuous function in the range of $0 \leq x \leq 1$.

To find the critical concentration let ξ approach zero as $x \to \bar{x}_c + 0$. Passage of Eq. (C2) to the limit yields

$$\frac{1}{\bar{x}_c} = 3 \int_0^1 \frac{dy}{1 - \bar{x}_c + \bar{x}_c y} = \frac{3}{\bar{x}_c} \log \frac{1}{1 - \bar{x}_c}, \tag{C3}$$

so that

$$\bar{x}_c = 1 - \exp(-1/3) = 0.283 \cdots. \tag{C4}$$

The value of $\bar{x}_c$ determined by Eq. (C4) gives a critical concentration such that the effective conductivity σ_c^* vanishes at x less than $\bar{x}_c$ and increases monotonically if x exceeds $\bar{x}_c$.

Finally, we want to expand ξ in a power series of $x - \bar{x}_c$. Formal differentiation of Eq. (C2) leads to

$$\frac{d\xi}{dx} = -\frac{1/x^2 + 3\int_0^1 [y^{3\xi}(1-y)/(1-x+xy)^2]\,dy}{9\int_0^1 [y^{3\xi}\log y/(1-x+xy)]\,dy}. \tag{C5}$$

Substituting $\bar{x}_c$ for x in Eq. (C5), we have

$$\left.\frac{d\xi}{dx}\right|_{x=\bar{x}_c} = -\left(3\bar{x}_c(1-\bar{x}_c)\int_0^1 \frac{\log y\,dy}{1-\bar{x}_c+\bar{x}_c y}\right)^{-1}. \tag{C6}$$

Since the integral on the right is expressed as

$$\int_0^1 \frac{\log y}{1-\bar{x}_c+\bar{x}_c y}\,dy = -\frac{1}{1-\bar{x}_c} + \frac{\bar{x}_c}{2^2(1-\bar{x}_c)^2} - \frac{\bar{x}_c^2}{3^2(1-\bar{x}_c)^3} + \cdots, \tag{C7}$$

the differential coefficient $d\xi/dx$ is estimated to be

$$\left.\frac{d\xi}{dx}\right|_{x=\bar{x}_c} = 1.284\cdots. \tag{C8}$$

Thus we arrive at the Taylor expansion equivalent to Eq. (4.10).

[1]M. Hori, J. Math. Phys. 14, 514 (1973).
[2]M. Hori, J. Math. Phys. 14, 1942 (1973).
[3]M. Hori and F. Yonezawa, J. Math. Phys. 15, Paper III (1974).
[4]F. Yonezawa, Prog. Theor. Phys. 40, 734 (1968).
[5]D. A. G. Bruggeman, Ann. Phys. (Leipzig) 24, 636 (1935); 25, 645 (1936).
[6]C. J. F. Böttcher, Rec. Trav. Chim. Pays-Bas 64, 47 (1945).
[7]D. Polder and J. H. van Santen, Physica 12, 257 (1946).
[8]V. I. Odelevskii, Zh. Tekh. Fiz. 21, 667, 678 (1951).
[9]E. I. Kondorskii, Dokl. Akad. Nauk SSSR 80, 197 (1951)
[10]W. Niesel, Ann. Phys. (Leipzig) 10, 336 (1952).
[11]R. Landauer, J. Appl. Phys. 23, 779 (1952).
[12]E. H. Kerner, Proc. Phys. Soc. (London) B 69, 802 (1956).
[13]M. H. Cohen and J. Jortner, Phys. Rev. Lett. 30, 696 (1973).
[14]D. A. G. Bruggeman, Ann. Phys. (Leipzig) 29, 160 (1937).
[15]A. V. Hershey, J. Appl. Mech. 21, 236 (1954).
[16]E. H. Kerner, Proc. Phys. Soc. (London) B 69, 808 (1956).
[17]E. Kröner, Z. Phys. 151, 504 (1958).
[18]R. Hill, J. Mech. Phys. Solids 13, 89 (1965).
[19]R. Hill, J. Mech. Phys. Solids 13, 189 (1965).
[20]R. Hill, J. Mech. Phys. Solids 13, 213 (1965).
[21]B. Budiansky, J. Mech. Phys. Solids 13, 223 (1965).
[22]R. Hill, J. Mech. Phys. Solids 15, 79 (1967).
[23]B. Budiansky, J. Compos. Mater. 4, 286 (1970).
[24]N. Law, J. Mech. Phys. Solids 21, 9 (1973).
[25]J. Korringa, J. Math. Phys. 14, 509 (1973).
[26]F. Yonezawa and K. Morigaki, Prog. Theor. Phys. Suppl. 53, 1 (1973).
[27]S. Kirkpatrick, Phys. Rev. Lett. 27, 1722 (1971).
[28]J. A. Krumhansl, in *Amorphous Magnetism*, edited by H. O. Hooper and A. M. de Graaf (Plenum, New York, 1973), p. 15.
[29]S. Kirkpatrick, Rev. Mod. Phys. 45, 574 (1973).
[30]M. H. Cohen and J. Jortner, Phys. Rev. Lett. 30, 699 (1973).
[31]S. Kirkpatrick, in *The Properties of Liquid Metals*, edited by S. Takeuchi (Taylor and Francis, London, 1973), p. 351.
[32]F. Yonezawa and T. Matsubara, Prog. Theor. Phys. 35, 357, 759 (1966).
[33]Some attempts have been made by O. Eberhardt (Master Thesis at Universität Stuttgart, 1973) to formulate the problem of classical mixtures by means of the diagram methods used in the tight-binding scheme of quantum mechanics. A general relationship between the diagrams we use in this paper and those tight-binding diagrams was discussed by F. Yonezawa [Prog. Theor. Phys. 31, 357 (1964)]. A more detailed study of this problem particularly concerning the CPA and the short-range correlations will be reported in Phys. Rev. by M. Watabe and F. Yonezawa.
[34]J. J. Thomson and G. Benford, J. Math. Phys. 14, 531 (1973).
[35]U. Frisch, in *Probabilistic Methods in Applied Mathematics*, edited by A. T. Bharucha-Reid (Academic, New York, 1968), Vol. 1, p. 75.
[36]D. Dence and J. E. Spence, in *Probabilistic Methods in Applied Mathematics*, edited by A. T. Bharucha-Reid (Academic, New York, 1973), Vol. 3, p. 121.
[37]E. Kröner, J. Mech. Phys. Solids 15, 319 (1967).
[38]E. Kröner, Phys. Status Solidi 41, 499 (1970).
[39]E. Kröner, Phys. Status Solidi B 49, K23 (1972).
[40]See, for example, J. A. Reynolds and J. M. Hough, Proc. Phys. Soc. B 70, 769 (1957); Z. Hashin, in *Mechanics of Composite Materials*, edited by F. W. Wendt, H. Liebowitz, and N. Perrone (Pergamon, Oxford, 1970), p. 201.
[41]V. K. S. Shante and S. Kirkpatrick, Adv. Phys. 20, 325 (1971).
[42]J. W. Essam, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, New York, 1972), Vol. 2, p. 197.

Statistical theory of effective electrical, thermal, and magnetic properties of random heterogeneous materials. V. One- and two-dimensional systems

Motoo Hori

Department of Applied Physics, Tokyo Institute of Technology, Meguroku, Tokyo 152, Japan

Fumiko Yonezawa

Belfer Graduate School of Science, Yeshiva University, New York, New York 10033

Department of Applied Physics, Tokyo Institute of Technology, Meguroku, Tokyo 152, Japan
(Received 16 July 1974; revised manuscript 16 September 1974)

A perturbation theory is developed for the effective permittivity of one- and two-dimensional random heterogeneous materials that are statistically homogeneous. Closely following the formulations for three-dimensional systems presented in Papers I–IV in this series of work, we derive the formal perturbation solutions for two-dimensional systems, evaluate the second-order and third-order terms for cell materials, and determine upper and lower bounds of the effective permittivity. Here statistical isotropy is not necessarily required. Several approximate perturbation solutions for completely random systems (which are statistically isotropic) are obtained by summing some selected partial series of the perturbation expansion up to an infinite order and numerical results are illustrated. We analyze validity of approximations by means of diagram representation of the perturbation series. The effective-medium theory serves as a good approximation for two-dimensional systems and gives the critical percolation concentration correctly. For one-dimensional materials, the effective-medium approximation turns out to be an exact solution.

1. INTRODUCTION

In the preceding papers of this series,[1–4] hereafter called I to IV, a perturbation treatment was performed for the effective permittivity of three-dimensional (3D) inhomogeneous materials whose local permittivity may be regarded as a random function of position. The whole argument holds also for other physical constants such as magnetic permeability, electrical and thermal conductivity, and diffusion constant. The purpose of the present paper V is to extend our formulations to one- and two-dimensional media and to compare the results with those in the 3D case. Especially, the meanings of the 2D problem are twofold. From a practical point of view, a 2D system serves as a simple model for fiber-reinforced composite materials.[5,6] On the other hand, the dependence of the overall behavior of random media on their dimensionality is very important from a theoretical point of view. For example, percolation thresholds in disordered systems are determined almost entirely by the dimensionality and insensitive to the details of the geometrical structure.[7,8] Therefore, we are mainly concerned with the effective permittivity of 2D heterogeneous materials.

In Sec. 2 we develop a general perturbation formulation for the effective permittivity of random heterogeneous materials of two dimensions that are statistically homogeneous but not necessarily statistically isotropic. In Sec. 3, the second-order and third-order perturbation terms are calculated explicitly for cell materials where the property of a particular cell is statistically independent of that of any other cell. Furthermore, upper and lower bounds of the effective permittivity are obtained taking account of the three-point correlation effects. Section 4 gives a prescription to construct the perturbation coefficient of an arbitrary order for completely random media. Regarding the completely random material as a limiting case of a circular-cell material, we compute the fourth-order and fifth-order perturbation terms. Section 5 shows how the perturbation expansions can be interpreted by means of diagrams. In Sec. 6, several approximate perturbation solutions are derived and some numerical results are presented. The last section is devoted to a discussion of the 1D problem where an exact solution is easily obtained, in order to check the validity of various approximations.

2. BASIC EQUATIONS, FORMAL PERTURBATION SERIES, AND BOUNDS FOR EFFECTIVE PERMITTIVITY

We deal with a 2D material with random variations in permittivity whose area S is eventually brought to infinity. Assume that the medium is locally isotropic and statistically homogeneous. Denote by $\epsilon(\mathbf{r})$ the permittivity at a point $\mathbf{r}$ and let $\epsilon'(\mathbf{r}) = \epsilon(\mathbf{r}) - \langle\epsilon\rangle$, where the brackets $\langle\cdots\rangle$ indicate the ensemble average. Then, the normalized n-point correlation function $g(\mathbf{r}_{12}, \mathbf{r}_{23}, \ldots, \mathbf{r}_{n-1,n})$ is given by

$$\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\cdots\epsilon'(\mathbf{r}_n)\rangle = \langle\epsilon'(0)\epsilon'(\mathbf{r}_{12})\cdots\epsilon'(\mathbf{r}_{1n})\rangle = \langle\epsilon'^n\rangle g(\mathbf{r}_{12}, \mathbf{r}_{23}, \ldots, \mathbf{r}_{n-1,n}), \quad (2.1)$$

in which $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$. According to the ergodic hypothesis for statistically homogeneous media, the ensemble average is supposed to be replaceable with the spatial average, so that in our 2D case

$$\left\langle f[\epsilon(\mathbf{r})]\right\rangle = \lim_{S\to\infty}(1/S)\int_S f[\epsilon(\mathbf{r})]\,d\sigma, \quad (2.2)$$

where $f[\epsilon(\mathbf{r})]$ is an arbitrary functional of $\epsilon(\mathbf{r})$ and $d\sigma$ is an area element of S at $\mathbf{r}$.

The basic equations governing the static electric field in a material with variable permittivity are

$$\frac{\partial}{\partial x_i}[\epsilon(\mathbf{r})E_i(\mathbf{r})]=0, \tag{2.3}$$

$$E_i(\mathbf{r})=-\frac{\partial\Phi(\mathbf{r})}{\partial x_i}. \tag{2.4}$$

Here $E_i(\mathbf{r})$ is the ith component of the electric field, $\Phi(\mathbf{r})$ is the electrostatic potential, and the summation convention is employed. Putting $E_i'(\mathbf{r})=E_i(\mathbf{r})-\langle E_i\rangle$, we have

$$\frac{\partial E_i'(\mathbf{r})}{\partial x_i}=-\frac{\langle E_i\rangle}{\langle\epsilon\rangle}\frac{\partial\epsilon'(\mathbf{r})}{\partial x_i}-\frac{1}{\langle\epsilon\rangle}\frac{\partial}{\partial x_i}[\epsilon'(r)E_i'(\mathbf{r})]. \tag{2.5}$$

By means of the Green's function for the 2D Laplacian operator, Eq. (2.5) leads to

$$E_i'(\mathbf{r}_1)=\frac{1}{2\pi\langle\epsilon\rangle}\left(\langle E_j\rangle\int_S d\sigma_2\frac{x_{12,i}}{r_{12}^2}\frac{\partial\epsilon'(\mathbf{r}_2)}{\partial x_{2,j}}+\int_S d\sigma_2\frac{x_{12,i}}{r_{12}^2}\frac{\partial}{\partial x_{2,j}}[\epsilon'(\mathbf{r}_2)E_j'(\mathbf{r}_2)]\right). \tag{2.6}$$

The effective permittivity tensor ϵ_{ij}^* of a heterogeneous medium is defined by the relation

$$\langle\epsilon(\mathbf{r})E_i(\mathbf{r})\rangle=\epsilon_{ij}^*\langle E_j\rangle. \tag{2.7}$$

In order to evaluate ϵ_{ij}^*, therefore, it is necessary to solve the random integral equation (2.6).

Through an analogous procedure to that presented in I, the solution of Eq. (2.6) can be expanded in a perturbation series of the form

$$E_i'(\mathbf{r})=\sum_{n=1}^{\infty}E_i^{(n)}(\mathbf{r}), \tag{2.8}$$

with

$$E_i^{(n)}(\mathbf{r}_1)=\frac{\langle E_j\rangle}{(2\pi\langle\epsilon\rangle)^n}\int_S d\sigma_{12}\int_S d\sigma_{23}\cdots\int_S d\sigma_{n,n+1}\times\frac{x_{12,i}}{r_{12}^2}\frac{x_{23,k}}{r_{23}^2}\cdots\frac{\partial^n\epsilon'(\mathbf{r}_2)\epsilon'(\mathbf{r}_3)\cdots\epsilon'(\mathbf{r}_{n+1})}{\partial x_{12,k}\partial x_{23,h}\cdots\partial x_{n,n+1;j}}. \tag{2.9}$$

As for the perturbation series of ϵ_{ij}^* corresponding to Eq. (2.8), we easily obtain

$$\epsilon_{ij}^*=\langle\epsilon\rangle\left(\delta_{ij}-\sum_{n=2}^{\infty}(-1)^nA_{ij}^{(n)}\frac{\langle\epsilon'^n\rangle}{\langle\epsilon\rangle^n}\right), \tag{2.10}$$

where δ_{ij} denotes the Kronecker delta. The nth-order perturbation coefficient $A_{ij}^{(n)}$ is related to the normalized n-point correlation function $g(\mathbf{r}_{12},\mathbf{r}_{23},\ldots,\mathbf{r}_{n-1,n})$ by

$$A_{ij}^{(n)}=\left(\frac{-1}{2\pi}\right)^{n-1}\int_S d\sigma_{12}\int_S d\sigma_{23}\cdots\int_S d\sigma_{n-1,n}\times\frac{x_{12,i}}{r_{12}^2}\frac{x_{23,k}}{r_{23}^2}\cdots\frac{\partial^{n-1}g(\mathbf{r}_{12},\mathbf{r}_{23},\ldots,\mathbf{r}_{n-1,n})}{\partial x_{12,k}\partial x_{23,h}\cdots\partial x_{n-1,n;j}}. \tag{2.11}$$

For a statistically isotropic medium the second-order tensors ϵ_{ij}^* and $A_{ij}^{(n)}$ reduce to scalars ϵ^* and $A^{(n)}$ such that

$$\epsilon^*=\langle\epsilon\rangle\left(1-\sum_{n=2}^{\infty}(-1)^nA^{(n)}\frac{\langle\epsilon'^n\rangle}{\langle\epsilon\rangle^n}\right). \tag{2.12}$$

In particular, the second-order coefficient $A_{ij}^{(2)}$ becomes

$$A_{ij}^{(2)}=-\frac{1}{2\pi\langle\epsilon'^2\rangle}\int_S d\sigma_{12}\frac{x_{12,i}}{r_{12}^2}\frac{\partial\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\rangle}{\partial x_{12,j}}=-\frac{1}{2\pi}\int_S d\sigma_{12}\frac{x_{12,i}}{r_{12}^2}\frac{\partial g(\mathbf{r}_{12})}{\partial x_{12,j}}. \tag{2.13}$$

Let us impose appropriate boundary conditions upon $g(\mathbf{r}_{12})$; for example, assume that in the polar coordinate system (r_{12},θ_{12})

$$\lim_{r_{12}\to 0}g(r_{12},\theta_{12})=1,\quad \lim_{r_{12}\to\infty}g(r_{12},\theta_{12})=0. \tag{2.14}$$

Then Eq. (2.13) is transformed into

$$A_{ij}^{(2)}=-\frac{1}{2\pi}\int_S d\sigma_{12}\frac{\partial^2\log(1/r_{12})}{\partial x_{12,i}\partial x_{12,j}}g(\mathbf{r}_{12}), \tag{2.15}$$

or

$$A_{ij}^{(2)}=\frac{1}{(2\pi)^2}\int_S d\sigma_1\int_S d\sigma_2\frac{x_{1,k}}{r_1^2}\frac{x_{2,k}}{r_2^2}\frac{\partial^2 g(\mathbf{r}_{12})}{\partial x_{1,i}\partial x_{2,j}}. \tag{2.16}$$

The proof of Eqs. (2.15) and (2.16) is given in Appendix A.

As in the 3D case, it follows from Eqs. (2.13), (2.15), and (2.16) that the second-order coefficient $A_{ij}^{(2)}$ satisfies

(i) $A_{ij}^{(2)}=A_{ji}^{(2)}$, (2.17)

(ii) $A_{ii}^{(2)}=1$, (2.18)

(iii) $A_{i(i)}^{(2)}\geq 0$; (2.19)

the parentheses around the index i indicate that summation over i is not performed. The symmetric second-order tensor $A_{ij}^{(2)}$ has eigenvalues $A_i^{(2)}$ such that

$$A_1^{(2)}+A_2^{(2)}=1, \tag{2.20}$$

and

$$A_1^{(2)},A_2^{(2)}\geq 0. \tag{2.21}$$

It is readily seen that for a statistically isotropic material

$$A^{(2)}=\tfrac{1}{2}A_{ii}^{(2)}=\tfrac{1}{2}. \tag{2.22}$$

The corresponding perturbation expansion of the effective permittivity ϵ^* up to second order is

$$\epsilon^*=\langle\epsilon\rangle\left(1-\frac{1}{2}\frac{\langle\epsilon'^2\rangle}{\langle\epsilon\rangle^2}\right). \tag{2.23}$$

Furthermore, the third-order coefficient $A_{ij}^{(3)}$ is defined as

$$A_{ij}^{(3)}=\frac{1}{(2\pi)^2}\int_S d\sigma_{12}\int_S d\sigma_{23}\frac{x_{12,i}}{r_{12}^2}\frac{x_{23,k}}{r_{23}^2}\frac{\partial^2 g(\mathbf{r}_{12},\mathbf{r}_{23})}{\partial x_{12,k}\partial x_{23,j}}. \tag{2.24}$$

In polar coordinates boundary conditions for $g(\mathbf{r}_{12},\mathbf{r}_{23})$ are

$$g(0,\theta_{12};0,\theta_{23})=1, \tag{2.25a}$$

$$g(0,\theta_{12};r_{23},\theta_{23})=\text{independent of }\theta_{12}, \tag{2.25b}$$

$$g(r_{12},\theta_{12};0,\theta_{23})=\text{independent of }\theta_{23}, \tag{2.25c}$$

$$g(\infty,\theta_{12};r_{23},\theta_{23})=g(r_{12},\theta_{12};\infty,\theta_{23})=0. \tag{2.25d}$$

Under these conditions we get

$$A_{ij}^{(3)}=\frac{1}{(2\pi)^2}\int_S d\sigma_{12}\int_S d\sigma_{23}\frac{\partial^2\log(1/r_{12})}{\partial x_{12,i}\partial x_{12,k}}\frac{\partial^2\log(1/r_{23})}{\partial x_{23,k}\partial x_{23,j}}\times g(\mathbf{r}_{12},\mathbf{r}_{23}), \tag{2.26}$$

which leads to

$$A_{ij}^{(3)} = A_{ji}^{(3)}. \tag{2.27}$$

Generally speaking, the perturbation coefficient $A_{ij}^{(n)}$ of an arbitrary order is symmetric.

In II, upper and lower bounds of the eigenvalues of the effective permittivity ϵ_i^* are obtained in terms of $A_i^{(2)}$ and $A_i^{(3)}$. The formulations given in II are mostly effective in 2D systems with the appropriate changes of quantities characteristic of dimensions such as the change of volume integrals into area integrals, $1/4\pi \to 1/2\pi$, $\frac{1}{3} \to \frac{1}{2}$, etc. Especially, Eqs. (II. 317) and (II. 417) hold for the 2D case; that is,

$$\epsilon_i^* \le \langle\epsilon\rangle\left(1 - \frac{(A_i^{(2)}\langle\epsilon'^2\rangle/\langle\epsilon\rangle^2)^2}{A_i^{(2)}\langle\epsilon'^2\rangle/\langle\epsilon\rangle^2 + A_i^{(3)}\langle\epsilon'^3\rangle/\langle\epsilon\rangle^3}\right), \tag{2.28}$$

$$\epsilon_i^* \ge \frac{1}{\langle 1/\epsilon\rangle}\left\{1 - (1 - B_i^{(2,1)})^2\left(1 - \frac{1}{\langle\kappa\rangle\langle\epsilon\rangle}\right)^2 \times\left[(1 - 2B_i^{(2,1)} + B_i^{(3,2)})\left(1 - \frac{1}{\langle\kappa\rangle\langle\epsilon\rangle}\right) + (A_i^{(2)} - B_i^{(3,2)})\frac{\langle\epsilon'^2\rangle}{\langle\epsilon\rangle^2}\right]^{-1}\right\}^{-1}. \tag{2.29}$$

The coefficients $B_i^{(2,1)}$ and $B_i^{(3,2)}$ are the eigenvalues of

$$B_{ij}^{(2,1)} = -\frac{1}{2\pi\langle\kappa'\epsilon'\rangle}\int_S d\sigma_{12}\frac{x_{12,i}}{r_{12}^2}\frac{\partial\langle\kappa'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\rangle}{\partial x_{12,j}}, \tag{2.30}$$

and

$$B_{ij}^{(3,2)} = \frac{1}{(2\pi)^2\langle\kappa'\epsilon'^2\rangle}\int_S d\sigma_{12}\int_S d\sigma_{23}\frac{x_{12,i}}{r_{12}^2}\frac{x_{23,k}}{r_{23}^2} \times\frac{\partial^2\langle\epsilon'(\mathbf{r}_1)\kappa'(\mathbf{r}_2)\epsilon'(\mathbf{r}_3)\rangle}{\partial x_{12,k}\partial x_{23,j}}, \tag{2.31}$$

where $\kappa = 1/\epsilon$ and $\kappa' = \kappa - \langle\kappa\rangle$. For statistically isotropic materials Eqs. (2. 28) and (2. 29) are simplified to

$$\epsilon^* \le \langle\epsilon\rangle\left(1 - \frac{\frac{1}{4}(\langle\epsilon'^2\rangle/\langle\epsilon\rangle^2)^2}{\frac{1}{2}\langle\epsilon'^2\rangle/\langle\epsilon\rangle^2 + A^{(3)}\langle\epsilon'^3\rangle/\langle\epsilon\rangle^3}\right), \tag{2.32}$$

$$\epsilon^* \ge \frac{1}{\langle 1/\epsilon\rangle}\{1 - \tfrac{1}{4}(1 - 1/\langle 1/\epsilon\rangle\langle\epsilon\rangle)^2[B^{(3,2)}(1 - 1/\langle 1/\epsilon\rangle\langle\epsilon\rangle) + (\tfrac{1}{2} - B^{(3,2)})\langle\epsilon'^2\rangle/\langle\epsilon\rangle^2]^{-1}\}^{-1}. \tag{2.33}$$

3. DETERMINATION OF LOW-ORDER PERTURBATION TERMS FOR CELL MATERIALS

As explained in I, the cell model introduced by Miller[9,10] is based on the assumption that the medium can be subdivided into a large number of nonoverlapping cells. The cell material is defined as a random multiphase material in which the permittivity of a particular cell is statistically independent of that of any other cell. We shall confine ourselves to the symmetric case where the fluctuation properties of the geometry of all phases are identical. Let $P(\mathbf{r}_1, \mathbf{r}_2)$ be the probability that the two points $\mathbf{r}_1$ and $\mathbf{r}_2$ fall into the same cell and $P(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ be the probability that the three points $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ lie in the same cell. Then the independence hypothesis asserts that

$$\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\rangle = \langle\epsilon'^2\rangle g(\mathbf{r}_{12}) = \langle\epsilon'^2\rangle P(\mathbf{r}_1, \mathbf{r}_2), \tag{3.1}$$

$$\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\epsilon'(\mathbf{r}_3)\rangle = \langle\epsilon'^3\rangle g(\mathbf{r}_{12}, \mathbf{r}_{23}) = \langle\epsilon'^3\rangle P(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3), \tag{3.2}$$

whence

$$A_{ij}^{(2)} = -\frac{1}{2\pi}\int_S d\omega_{12}\frac{x_{12,i}}{r_{12}^2}\frac{\partial P(\mathbf{r}_1, \mathbf{r}_2)}{\partial x_{12,j}}, \tag{3.3}$$

$$A_{ij}^{(3)} = \frac{1}{(2\pi)^2}\int_S d\omega_{12}\int_S d\omega_{23}\frac{x_{12,i}}{r_{12}^2}\frac{x_{23,k}}{r_{23}^2}\frac{\partial^2 P(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)}{\partial x_{12,k}\partial x_{23,j}}. \tag{3.4}$$

We first treat a symmetric cell material composed of cells of uniform shape, size, and orientation. Proceeding in the same way as we did in Sec. 3B of I, we have

$$A_{ij}^{(2)} = -\frac{1}{s}\int_s d\sigma_{10}\frac{\partial^2}{\partial x_{10,i}\partial x_{10,j}}\int_s d\sigma_{20}\frac{\log(1/r_{12})}{2\pi} = \frac{1}{s}\int_s d\sigma_{10}\int_s d\sigma_{20}\frac{\partial^2}{\partial x_{10,i}\partial x_{20,j}}\left(\frac{\log(1/r_{12})}{2\pi}\right), \tag{3.5}$$

where s is the cell area. Thus, the second-order perturbation coefficient $A_{ij}^{(2)}$ is equal to the magnetometric demagnetization tensor of two dimensions. Define the 2D point-function demagnetization tensor by

$$L_{ij}(\mathbf{r}_{10}) = \int_s d\sigma_{20}\frac{\partial^2}{\partial x_{10,i}\partial x_{20,j}}\left(\frac{\log(1/r_{12})}{2\pi}\right). \tag{3.6}$$

Then Eq. (3. 5) is written as

$$A_{ij}^{(2)} = \frac{1}{s}\int_s d\sigma_{10}L_{ij}(\mathbf{r}_{10}). \tag{3.7}$$

Likewise,

$$A_{ij}^{(3)} = \frac{1}{s}\int_s d\sigma_{20}L_{ik}(\mathbf{r}_{20})L_{kj}(\mathbf{r}_{20}). \tag{3.8}$$

Note that $A_{ij}^{(3)}$ as well as $A_{ij}^{(2)}$ depends only on the shape of cells and not on their size.

Next we shall discuss the case where identical cells are oriented at random. In the present case the cell material is statistically isotropic, so that $A_{ij}^{(2)}$ must be $\frac{1}{2}\delta_{ij}$ as shown in Eq. (2. 22). For the third-order coefficient $A^{(3)}$ we obtain

$$A^{(3)} = \tfrac{1}{2}A_{ii}^{(3)} = \frac{1}{2s}\int_s d\sigma_{20}L_{kk}^2(\mathbf{r}_{20}). \tag{3.9}$$

The inequalities corresponding to Eq. (II. 5. 9) are

$$\tfrac{1}{4} \le A^{(3)} \le \tfrac{1}{2}, \tag{3.10}$$

that have already been conjectured by Beran and Silnutzer.[6] Clearly the argument in this paragraph holds true even when the medium consists of cells of varying size. In addition, Eqs. (5. 18) and (5. 19) in II guarantee that the bounding equations (3. 10) apply equally well to an asymmetric cell material.

Finally, let us consider cell materials where elliptic cells are uniformly or randomly oriented. The point-function demagnetization tensor of an ellipse and its degenerate shapes is constant throughout the interior of the body. Consequently, Eqs. (3. 7) and (3. 8) become

$$A_{ij}^{(2)} = L_{ij}, \quad A_{ij}^{(3)} = L_{ik}L_{kj}. \tag{3.11}$$

When axes of elliptic cells are aligned parallel with the coordinate axes,

$$A_i^{(2)} = L_i, \quad A_i^{(3)} = L_{(i)}^2. \tag{3.12}$$

Needless to say, L_i signify the depolarizing or demag-

netizing factors of the ellipse. For randomly oriented ellipses, on the other hand, $A^{(3)}$ is given by

$$A^{(3)} = \tfrac{1}{2}(L_1^2 + L_2^2). \tag{3.13}$$

The depolarizing factors of an ellipse may be determined exactly as in the 3D case of an ellipsoid.[11] In the process of calculation we need to adopt the elliptic coordinate instead of the ellipsoidal coordinate. The result is

$$L_i = \frac{a_1 a_2}{2}\int_0^\infty \frac{dl}{(l+a_i^2)[(l+a_1^2)(l+a_2^2)]^{1/2}} = 1 - [a_i/(a_1+a_2)], \tag{3.14}$$

$$A^{(3)} = \tfrac{1}{2}\{1 - [2a_1a_2/(a_1+a_2)^2]\}. \tag{3.15}$$

Here a_1 and a_2 are the semiaxes of the x_1 and x_2 directions, respectively. For a circular cell where $a_1 = a_2$, we have $L_1 = L_2 = \frac{1}{2}$, so that Eqs. (3.12) and (3.13) yield

$$A_1^{(2)} = A_2^{(2)} = \tfrac{1}{2}, \quad A_1^{(3)} = A_2^{(3)} = \tfrac{1}{4}, \quad A^{(3)} = \tfrac{1}{4}. \tag{3.16}$$

If $a_2 = 0$ or $a_1 \to \infty$, the ellipse degenerates into a parallel lamella (two-dimensional needle), for which

$$A_1^{(2)} = A_1^{(3)} = 0, \quad A_2^{(2)} = A_2^{(3)} = 1, \quad A^{(3)} = \tfrac{1}{2}. \tag{3.17}$$

The bounds presented in Eqs. (2.28), (2.29), (2.32), and (2.33) can be simplified for a symmetric cell material composed of elliptic cells. Since $A_i^{(2)} = B_i^{(2,1)} = L_i$ and $A_i^{(3)} = B_i^{(3,2)} = L_{(i)}^2$, Eqs. (2.28) and (2.29) reduce to

$$\epsilon_i^* \le \langle\epsilon\rangle\left(1 - \frac{L_i(\langle\epsilon'^2\rangle/\langle\epsilon\rangle^2)^2}{\langle\epsilon'^2\rangle/\langle\epsilon\rangle^2 + L_i\langle\epsilon'^3\rangle/\langle\epsilon\rangle^3}\right), \tag{3.18}$$

$$\epsilon_i^* \ge \frac{1}{\langle 1/\epsilon\rangle}\left\{1 - (1-L_i)^2\left(1 - \frac{1}{\langle 1/\epsilon\rangle\langle\epsilon\rangle}\right)^2 \times\left[(1 - 2L_i + L_{(i)}^2)\left(1 - \frac{1}{\langle 1/\epsilon\rangle\langle\epsilon\rangle}\right) + (L_i - L_{(i)}^2) \times\frac{\langle\epsilon'^2\rangle}{\langle\epsilon\rangle^2}\right]^{-1}\right\}^{-1}, \tag{3.19}$$

which are formally equivalent to Eqs. (5.16) and (5.17) of II. Especially for an isotropic circular-cell material,

$$\epsilon^* \le \langle\epsilon\rangle\left(1 - \frac{(\langle\epsilon'^2\rangle/\langle\epsilon\rangle^2)^2}{2\langle\epsilon'^2\rangle/\langle\epsilon\rangle^2 + \langle\epsilon'^3\rangle/\langle\epsilon\rangle^3}\right), \tag{3.20}$$

$$\epsilon^* \ge \frac{1}{\langle 1/\epsilon\rangle}\left\{1 - \left(1 - \frac{1}{\langle 1/\epsilon\rangle\langle\epsilon\rangle}\right)^2\left[1 - \frac{1}{\langle 1/\epsilon\rangle\langle\epsilon\rangle} + \frac{\langle\epsilon'^2\rangle}{\langle\epsilon\rangle^2}\right]^{-1}\right\}^{-1}. \tag{3.21}$$

It is extremely difficult to evaluate the fourth- or higher-order perturbation coefficient. The four-point moment $\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\epsilon'(\mathbf{r}_3)\epsilon'(\mathbf{r}_4)\rangle$ takes the value $\langle\epsilon'^4\rangle$ when four points are in the same cell, $\langle\epsilon'^2\rangle^2$ when two pairs of points are in two different cells, and 0 otherwise. Let $P(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4)$ be the probability that the points $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, $\mathbf{r}_4$ are in the same cell, $P(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}_3, \mathbf{r}_4)$ the probability that the two pairs of points, $(\mathbf{r}_1, \mathbf{r}_2)$ and $(\mathbf{r}_3, \mathbf{r}_4)$, are in two different cells, and so on. It is evident that Eq. (III. 2.18) also holds in our 2D case; namely,

$$\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\epsilon'(\mathbf{r}_3)\epsilon'(\mathbf{r}_4)\rangle = \langle\epsilon'^4\rangle g(\mathbf{r}_{12}, \mathbf{r}_{23}, \mathbf{r}_{34}) = \langle\epsilon'^4\rangle P(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) + \langle\epsilon'^2\rangle^2[P(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}_3, \mathbf{r}_4) + P(\mathbf{r}_1, \mathbf{r}_4; \mathbf{r}_2, \mathbf{r}_3) + P(\mathbf{r}_1, \mathbf{r}_3; \mathbf{r}_2, \mathbf{r}_4)]. \tag{3.22}$$

Hence the fourth-order coefficient $A_{ij}^{(4)}$ is separated into four parts as

$$A_{ij}^{(4)}\langle\epsilon'^4\rangle = A_{1,ij}^{(4)}\langle\epsilon'^4\rangle + (A_{2,ij}^{(4)} + A_{3,ij}^{(4)} + A_{4,ij}^{(4)})\langle\epsilon'^2\rangle^2, \tag{3.23}$$

where

$$A_{1,ij}^{(4)} = -\frac{1}{(2\pi)^3}\int_S d\sigma_{12}\int_S d\sigma_{23}\int_S d\sigma_{34}\frac{x_{12,i}}{r_{12}^2}\frac{x_{23,k}}{r_{23}^2}\frac{x_{34,h}}{r_{34}^2} \times\frac{\partial^3 P(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4)}{\partial x_{12,k}\partial x_{23,h}\partial x_{34,j}}, \tag{3.24}$$

$$A_{2,ij}^{(4)} = -\frac{1}{(2\pi)^3}\int_S d\sigma_{12}\int_S d\sigma_{23}\int_S d\sigma_{34}\frac{x_{12,i}}{r_{12}^2}\frac{x_{23,k}}{r_{23}^2}\frac{x_{34,h}}{r_{34}^2} \times\frac{\partial^3 P(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}_3, \mathbf{r}_4)}{\partial x_{12,k}\partial x_{23,h}\partial x_{34,j}}, \tag{3.25}$$

$$A_{3,ij}^{(4)} = -\frac{1}{(2\pi)^3}\int_S d\sigma_{12}\int_S d\sigma_{23}\int_S d\sigma_{34}\frac{x_{12,i}}{r_{12}^2}\frac{x_{23,k}}{r_{23}^2}\frac{x_{34,h}}{r_{34}^2} \times\frac{\partial^3 P(\mathbf{r}_1, \mathbf{r}_4; \mathbf{r}_2, \mathbf{r}_3)}{\partial x_{12,k}\partial x_{23,h}\partial x_{34,j}}, \tag{3.26}$$

$$A_{4,ij}^{(4)} = -\frac{1}{(2\pi)^3}\int_S d\sigma_{12}\int_S d\sigma_{23}\int_S d\sigma_{34}\frac{x_{12,i}}{r_{12}^2}\frac{x_{23,k}}{r_{23}^2}\frac{x_{34,h}}{r_{34}^2} \times\frac{\partial^3 P(\mathbf{r}_1, \mathbf{r}_3; \mathbf{r}_2, \mathbf{r}_4)}{\partial x_{12,k}\partial x_{23,h}\partial x_{34,j}}. \tag{3.27}$$

Out of these coefficients, $A_{1,ij}^{(4)}$ is independent not only of the size of cells but also of their relative arrangement, because $P(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4)$ concerns the geometry of a single cell. If homothetic ellipses are uniformly or randomly oriented, we find

$$A_{1,ij}^{(4)} = L_{ik}L_{kh}L_{hj}, \tag{3.28}$$

$$A_1^{(4)} = \tfrac{1}{8}, \tag{3.29}$$

whose proof goes in a similar manner to that described in Appendix A of III. On the contrary, it is practically impossible to determine $A_{2,ij}^{(4)}$, $A_{3,ij}^{(4)}$, and $A_{4,ij}^{(4)}$ exactly. In addition, Eqs. (3.28) and (3.29) are generalized to

$$A_{1,ij}^{(n)} = \left(-\frac{1}{2\pi}\right)^{n-1}\int_S d\sigma_{12}\int_S d\sigma_{23}\cdots\int_S d\sigma_{n-1,n} \times\frac{x_{12,i}}{r_{12}^2}\frac{x_{23,k}}{r_{23}^2}\cdots\frac{\partial^{n-1}P(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)}{\partial x_{12,k}\partial x_{23,h}\cdots\partial x_{n-1,n;j}} = L_{ik}L_{kh}\cdots L_{mj}\ [(n-1)\text{-fold product}], \tag{3.30}$$

$$A_1^{(n)} = 1/2^{n-1}. \tag{3.31}$$

4. DETERMINATION OF LOW-ORDER PERTURBATION TERMS FOR COMPLETELY RANDOM MATERIALS

Now, for the purpose of studying the fourth- and fifth-order terms of perturbation expansions, we follow the discussion of III. We treat completely random heterogeneous materials of two dimensions, in which physical constants at different points are statistically independent. The many-point correlation functions $g(\mathbf{r}_{12}, \mathbf{r}_{23}, \ldots, \mathbf{r}_{n-1,n})$ in this case are expressed by means of the appropriate products and sums of such null functions as $\delta_{\mathbf{r}_{12}}$ where

$$\delta_{\mathbf{r}_{12}}=\begin{cases}1 & \text{for } \mathbf{r}_1=\mathbf{r}_2,\\ 0 & \text{for } \mathbf{r}_1\neq\mathbf{r}_2.\end{cases} \tag{4.1}$$

For illustration let us show some of lower-order moments in terms of $\delta_{\mathbf{r}_{12}}$, $\delta_{\mathbf{r}_{23}}$, etc.:

$$\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\rangle=\langle\epsilon'^2\rangle\delta_{\mathbf{r}_{12}}, \tag{4.2}$$

$$\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\epsilon'(\mathbf{r}_3)\rangle=\langle\epsilon'^3\rangle\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}, \tag{4.3}$$

$$\begin{aligned}&\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\epsilon'(\mathbf{r}_3)\epsilon'(\mathbf{r}_4)\rangle\\&=\langle\epsilon'^4\rangle\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{34}}+\langle\epsilon'^2\rangle^2[\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{34}}(1-\delta_{\mathbf{r}_{23}})\\&\quad+\delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}}(1-\delta_{\mathbf{r}_{12}})+\delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{24}}(1-\delta_{\mathbf{r}_{12}})].\end{aligned} \tag{4.4}$$

The factors $(1-\delta_{\mathbf{r}_{12}})$ and $(1-\delta_{\mathbf{r}_{23}})$ arise from the exclusion effect.[12] It should be noticed that a completely random material is statistically isotropic as well as statistically homogeneous.

In order to facilitate the formulation, we regard a completely random material as a limiting case of a symmetric cell material where circular cells of infinitesimal size are distributed at random. Based upon this idea of limiting cell materials, the evaluation of $A^{(n)}$ for an arbitrary n is carried out in accordance with the following conventions:

(i) Express the many-point moment appearing in Eq. (2.11) in terms of $\delta_{\mathbf{r}_{12}}$, etc.;

(ii) contract each product of δ's to the simplest form with respect to $\mathbf{r}_{12}, \mathbf{r}_{23}, \cdots$, and arrange the subscripts in order whenever possible; for example, use $\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{34}}$ instead of $\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}}$, $\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{24}}$, $\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{34}}$, etc.;

(iii) replace $\delta_{\mathbf{r}_{12}}, \delta_{\mathbf{r}_{13}}, \cdots$ by characteristic functions $I_\rho(r_{12}), I_\rho(r_{13}), \cdots$ such that

$$I_\rho(r)=\begin{cases}1 & \text{for } r<\rho,\\ 0 & \text{for } r>\rho.\end{cases} \tag{4.5}$$

(iv) take the limits as $\rho\to+0$.

Let us first calculate the fourth-order perturbation term. Considering the statistical isotropy of a completely random material, we rewrite Eq. (3.23) in the form

$$A^{(4)}\langle\epsilon'^4\rangle=A_1^{(4)}\langle\epsilon'^4\rangle+(A_2^{(4)}+A_3^{(4)}+A_4^{(4)})\langle\epsilon'^2\rangle^2, \tag{4.6}$$

which corresponds to Eq. (4.4). On the introduction of cumulant averages $\langle\epsilon^n\rangle_c$, Eq. (4.6) is rearranged as

$$A^{(4)}\langle\epsilon'^4\rangle=B_1^{(4)}\langle\epsilon^4\rangle_c+(B_2^{(4)}+B_3^{(4)}+B_4^{(4)})\langle\epsilon^2\rangle_c^2, \tag{4.7}$$

where

$$A_1^{(4)}=B_1^{(4)}, \tag{4.8}$$

$$A_2^{(4)}=B_2^{(4)}-B_1^{(4)}, \tag{4.9}$$

$$A_3^{(4)}=B_3^{(4)}-B_1^{(4)}, \tag{4.10}$$

$$A_4^{(4)}=B_4^{(4)}-B_1^{(4)}. \tag{4.11}$$

It follows directly from the definitions of $B_1^{(4)}$ and $B_2^{(4)}$ that

$$B_1^{(4)}=A_1^{(4)}=\tfrac{1}{8}, \tag{4.12}$$

and that

$$B_2^{(4)}=0, \quad A_2^{(4)}=-\tfrac{1}{8}. \tag{4.13}$$

To compute $A_3^{(4)}$ we contract $\delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}}$ as

$$\delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}}=\delta_{\mathbf{r}_{12}+\mathbf{r}_{23}+\mathbf{r}_{34}}\delta_{\mathbf{r}_{23}}=\delta_{\mathbf{r}_{12}+\mathbf{r}_{34}}\delta_{\mathbf{r}_{23}}; \tag{4.14}$$

then

$$\begin{aligned}B_3^{(4)}&=-\frac{1}{(2\pi)^3}\int_S d\sigma_{12}\int_S d\sigma_{23}\int_S d\sigma_{34}\frac{x_{12,i}}{r_{12}^2}\frac{x_{23,k}}{r_{23}^2}\frac{x_{34,h}}{r_{34}^2}\\&\quad\times\frac{\partial^3\delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}}}{\partial x_{12,k}\partial x_{23,h}\partial x_{34,(i)}}\\&=-\lim_{\rho\to+0}\frac{1}{(2\pi)^3}\int_S d\sigma_{23}\frac{x_{23,k}}{r_{23}^2}\frac{\partial I_\rho(r_{23})}{\partial x_{23,h}}\\&\quad\times\int_S d\sigma_{12}\int_S d\sigma_{34}\frac{x_{12,i}}{r_{12}^2}\frac{x_{34,h}}{r_{12}^2}\frac{\partial^2 I_\rho(|\mathbf{r}_{12}+\mathbf{r}_{34}|)}{\partial x_{12,k}\partial x_{34,(i)}}.\end{aligned} \tag{4.15}$$

Similarly to the derivation of Eqs. (2.16),

$$\begin{aligned}&\frac{1}{(2\pi)^2}\int_S d\sigma_{12}\int_S d\sigma_{34}\frac{x_{12,i}}{r_{12}^2}\frac{x_{34,h}}{r_{34}^2}\frac{\partial^2 I_\rho(|\mathbf{r}_{12}+\mathbf{r}_{34}|)}{\partial x_{12,k}\partial x_{34,i}}\\&=\frac{1}{(2\pi)^2}\int_S d\sigma_{12}\int_S d\sigma_{34}\frac{x_{12,i}}{r_{12}^2}\frac{x_{34,i}}{r_{34}^2}\frac{\partial^2 I_\rho(|\mathbf{r}_{12}+\mathbf{r}_{34}|)}{\partial x_{12,k}\partial x_{34,h}}\\&=-\frac{1}{2\pi}\int_S d\sigma\frac{x_k}{r^2}\frac{\partial I_\rho(r)}{\partial x_h}.\end{aligned} \tag{4.16}$$

Accordingly,

$$B_3^{(4)}=A^{(2)}\cdot A^{(2)}=\tfrac{1}{4}, \quad A_3^{(4)}=\tfrac{1}{8}. \tag{4.17}$$

We turn to the numerical estimation of $A_4^{(4)}$. By definition,

$$\begin{aligned}B_4^{(4)}&=-\lim_{\rho\to+0}\frac{1}{(2\pi)^3}\int_S d\sigma_{12}\int_S d\sigma_{23}\int_S d\sigma_{34}\frac{x_{12,i}}{r_{12}^2}\frac{x_{23,k}}{r_{23}^2}\frac{x_{34,h}}{r_{34}^2}\\&\quad\times\frac{\partial^3 I_\rho(r_{13})I_\rho(r_{24})}{\partial x_{12,k}\partial x_{23,h}\partial x_{34,(i)}}\\&=-\lim_{\rho\to+0}\frac{1}{2\pi}\int_S d\sigma_{23}\frac{x_{23,k}}{r_{23}^2}\frac{\partial J_{\rho,ik}(\mathbf{r}_{23})J_{\rho,h(i)}(\mathbf{r}_{23})}{\partial x_{23,h}}.\end{aligned} \tag{4.18}$$

In Appendix B we prove that

$$\begin{aligned}J_{\rho,ik}(\mathbf{r}_{23})&=-\frac{1}{2\pi}\int_S d\sigma_{12}\frac{x_{12,i}}{r_{12}^2}\frac{\partial I_\rho(r_{13})}{\partial x_{12,k}}\\&=C_\rho(r_{23})\frac{x_{23,i}}{r_{23}}\frac{x_{23,k}}{r_{23}}+D_\rho(r_{23})\delta_{ik},\end{aligned} \tag{4.19}$$

where

$$C_\rho(r_{23})=\begin{cases}0 & \text{for } r_{23}<\rho,\\ -\rho^2/r_{23}^2 & \text{for } r_{23}>\rho,\end{cases} \tag{4.20}$$

$$D_\rho(r_{23})=\begin{cases}\frac{1}{2} & \text{for } r_{23}<\rho,\\ \rho^2/2r_{23}^2 & \text{for } r_{23}>\rho.\end{cases} \tag{4.21}$$

Substitution of Eqs. (4.19)—(4.21) into Eq. (4.18) yields

$$\begin{aligned}B_4^{(4)}&=-\lim_{\rho\to+0}\frac{1}{4\pi}\int_S\frac{d\sigma_{23}}{r_{23}}\left(\frac{d}{dr_{23}}[C_\rho(r_{23})+D_\rho(r_{23})]^2\right.\\&\quad\left.+\frac{C_\rho(r_{23})[C_\rho(r_{23})+2D_\rho(r_{23})]}{r_{23}}\right)\\&=\tfrac{1}{8}-\tfrac{1}{2}\lim_{\rho\to+0}\int_0^\infty\frac{C_\rho(r_{23})[C_\rho(r_{23})+2D_\rho(r_{23})]}{r_{23}}dr_{23},\end{aligned} \tag{4.22}$$

from which we get

$$B_4^{(4)}=\tfrac{1}{8}, \quad A_4^{(4)}=0. \tag{4.23}$$

This implies that $A_4^{(4)}$ vanishes in the 2D case, while $B_4^{(4)}$ vanishes in the 3D case.

On the other hand, the five-point moment becomes

$$\langle\epsilon'(\mathbf{r}_1)\epsilon'(\mathbf{r}_2)\epsilon'(\mathbf{r}_3)\epsilon'(\mathbf{r}_4)\epsilon'(\mathbf{r}_5)\rangle$$
$$=\langle\epsilon'^5\rangle\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{34}}\delta_{\mathbf{r}_{45}}+\langle\epsilon'^3\rangle\langle\epsilon'^2\rangle[\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{45}}(1-\delta_{\mathbf{r}_{34}})$$
$$+\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{34}}\delta_{\mathbf{r}_{45}}(1-\delta_{\mathbf{r}_{23}})+\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{25}}\delta_{\mathbf{r}_{34}}(1-\delta_{\mathbf{r}_{23}})$$
$$+\delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{45}}\delta_{\mathbf{r}_{23}}(1-\delta_{\mathbf{r}_{12}})+\delta_{\mathbf{r}_{15}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{34}}(1-\delta_{\mathbf{r}_{12}})$$
$$+\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{24}}\delta_{\mathbf{r}_{35}}(1-\delta_{\mathbf{r}_{23}})+\delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{24}}\delta_{\mathbf{r}_{45}}(1-\delta_{\mathbf{r}_{12}})$$
$$+\delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{34}}\delta_{\mathbf{r}_{25}}(1-\delta_{\mathbf{r}_{12}})+\delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{35}}(1-\delta_{\mathbf{r}_{12}})$$
$$+\delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{35}}\delta_{\mathbf{r}_{24}}(1-\delta_{\mathbf{r}_{12}})]. \tag{4.24}$$

The eleven coefficients corresponding to the terms on the right-hand side of Eq. (4.24) are

$$A^{(5)}\langle\epsilon'^5\rangle=A_1^{(5)}\langle\epsilon'^5\rangle+\sum_{m=2}^{11}A_m^{(5)}\langle\epsilon'^3\rangle\langle\epsilon'^2\rangle. \tag{4.25}$$

By means of the cumulant averages, Eq. (4.25) is converted into

$$A_4^{(5)}\langle\epsilon'^5\rangle=B_1^{(5)}\langle\epsilon^5\rangle_c+\sum_{m=2}^{11}B_m^{(5)}\langle\epsilon^3\rangle_c\langle\epsilon^2\rangle_c, \tag{4.26}$$

where

$$A_m^{(5)}=\begin{cases}B_m^{(5)} & \text{if } m=1,\\ B_m^{(5)}-B_1^{(5)} & \text{otherwise.}\end{cases} \tag{4.27}$$

With the help of the arguments in Appendix D of III, we have

$$B_1^{(5)}=A^{(2)}\cdot A^{(2)}\cdot A^{(2)}\cdot A^{(2)}=\tfrac{1}{16}, \tag{4.28}$$
$$B_2^{(5)}=B_3^{(5)}=0, \tag{4.29}$$
$$B_4^{(5)}=B_5^{(5)}=B_6^{(5)}=A^{(2)}\cdot B_3^{(4)}=\tfrac{1}{8}, \tag{4.30}$$
$$B_7^{(5)}=B_8^{(5)}=B_9^{(5)}=B_{10}^{(5)}=A^{(2)}\cdot B_4^{(4)}=\tfrac{1}{16}. \tag{4.31}$$

In analogy with $B_7^{(5)}$ to $B_{10}^{(5)}$, we expect that

$$B_{11}^{(5)}=\tfrac{1}{16}, \tag{4.32}$$

although we have not been successful in proving this rigorously. Moreover, Eqs. (4.28) to (4.32) enable us to determine $A_m^{(5)}$ as

$$A_1^{(5)}=\tfrac{1}{16}, \tag{4.33}$$
$$A_2^{(5)}=A_3^{(5)}=-\tfrac{1}{16}, \tag{4.34}$$
$$A_4^{(5)}=A_5^{(5)}=A_6^{(5)}=\tfrac{1}{16}, \tag{4.35}$$
$$A_7^{(5)}=A_8^{(5)}=\cdots=A_{11}^{(5)}=0. \tag{4.36}$$

Therefore, the effective permittivity ϵ^* up to the fifth order is

$$\frac{\epsilon^*}{\langle\epsilon\rangle}=1-\frac{\langle\epsilon^2\rangle_c}{2\langle\epsilon\rangle^2}+\frac{\langle\epsilon^3\rangle_c}{4\langle\epsilon\rangle^3}-\frac{\langle\epsilon^4\rangle_c+3\langle\epsilon^2\rangle_c^2}{8\langle\epsilon\rangle^4}$$
$$+\frac{\langle\epsilon^5\rangle_c+11\langle\epsilon^3\rangle_c\langle\epsilon^2\rangle_c}{16\langle\epsilon\rangle^5}-\cdots, \tag{4.37}$$

or

$$\frac{\epsilon^*}{\langle\epsilon\rangle}=1-\frac{\langle\epsilon'^2\rangle}{2\langle\epsilon\rangle^2}+\frac{\langle\epsilon'^3\rangle}{4\langle\epsilon\rangle^3}-\frac{\langle\epsilon'^4\rangle}{8\langle\epsilon\rangle^4}+\frac{\langle\epsilon'^5\rangle+\langle\epsilon'^3\rangle\langle\epsilon'^2\rangle}{16\langle\epsilon\rangle^5}-\cdots. \tag{4.38}$$

5. DIAGRAMMATIC REPRESENTATION OF THE PERTURBATION EXPANSION

In parallel with the discussion in IV, let us introduce diagrams to represent perturbation terms appearing in Eq. (2.12). Before explaining our diagram method in detail, we will review how the expansion terms of general orders are constructed. As in the 3D case, the effective permittivity of a completely random material is expressed as a sum of infinite perturbation series in the form

$$\epsilon'^*=\epsilon^*-\langle\epsilon\rangle=\sum_{n=2}^{\infty}\epsilon^{(n)}=\sum_{n=2}^{\infty}\frac{\sum_{\{\nu_m\}}B_{\{\nu_m\}}^{(n)}\epsilon(n,\{\nu_m\})}{(-\langle\epsilon\rangle)^{n-1}}. \tag{5.1}$$

Here $\{\nu_m\}$ denotes the way of partitioning n variables $\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n$ into m subsets of at least two elements each and $\sum_{\{\nu_m\}}$ signifies the sum over all possible partitions. The factors $\epsilon(n,\{\nu_m\})$ and $B_{\{\nu_m\}}^{(n)}$ are defined by

$$\epsilon(n,\{\nu_m\})=\langle\epsilon^{\nu_1}\rangle_c\langle\epsilon^{\nu_2}\rangle_c\cdots\langle\epsilon^{\nu_m}\rangle_c, \tag{5.2}$$

$$B_{\{\nu_m\}}^{(n)}=\left(-\frac{1}{2\pi}\right)^{n-1}\int_S d\sigma_{12}\int_S d\sigma_{23}\cdots\int_S d\sigma_{n-1,n}\frac{x_{12,i}}{r_{12}^2}$$
$$\times\frac{x_{23,k}}{r_{23}^2}\cdots\frac{\partial^{n-1}F_n(\{\nu_m\};\mathbf{r}_{12},\mathbf{r}_{23},\ldots,\mathbf{r}_{n-1,n})}{\partial x_{12,k}\partial x_{23,h}\cdots\partial x_{n-1,n;(i)}}$$
$$=(-\langle\epsilon\rangle)^{n-1}\int_S d\sigma_{12}\int_S d\sigma_{23}\cdots\int_S d\sigma_{n-1,n}G_{ik}(\mathbf{r}_{12})$$
$$\times G_{kh}(\mathbf{r}_{23})\cdots F_n(\{\nu_m\};\mathbf{r}_{12},\mathbf{r}_{23},\ldots,\mathbf{r}_{n-1,n}), \tag{5.3}$$

where $F_n(\{\nu_m\};\mathbf{r}_{12},\mathbf{r}_{23},\ldots,\mathbf{r}_{n-1,n})$ stands for the product of δ's corresponding to the partition $\{\nu_m\}$ and

$$G_{ik}(\mathbf{r}_{12})=\frac{1}{\langle\epsilon\rangle}\frac{\partial^2[\log(1/r_{12})/2\pi]}{\partial x_{12,i}\partial x_{12,k}}. \tag{5.4}$$

By the use of quantities $\gamma(n,\{\nu_m\})$ such that

$$\gamma(n,\{\nu_m\})=\langle\epsilon'^{\nu_1}\rangle\langle\epsilon'^{\nu_2}\rangle\cdots\langle\epsilon'^{\nu_m}\rangle, \tag{5.5}$$

Eq. (5.1) is rearranged as

$$\epsilon'^*=\sum_{n=2}^{\infty}\frac{\sum_{\{\nu_m\}}A_{\{\nu_m\}}^{(n)}\gamma(n,\{\nu_m\})}{(-\langle\epsilon\rangle)^{n-1}}. \tag{5.6}$$

In the foregoing sections, we have estimated lower-order perturbation coefficients as follows:

(Second-order term)

(i) $F_2(1;\mathbf{r}_{12})=\delta_{\mathbf{r}_{12}}$ (5.7a)

(ii) $\epsilon(2,1)=\langle\epsilon^2\rangle_c$, $\gamma(2,1)=\langle\epsilon'^2\rangle$, (5.7b)

(iii) $B^{(2)}=A^{(2)}=\frac{1}{2}$, (5.7c)

(iv) $\epsilon^{(2)}=-\dfrac{\langle\epsilon^2\rangle_c}{2\langle\epsilon\rangle}=-\dfrac{\langle\epsilon'^2\rangle}{2\langle\epsilon\rangle}$. (5.7d)

(third-order term)

(i) $F_3(1;\mathbf{r}_{12},\mathbf{r}_{23})=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}$, (5.8a)

(ii) $\epsilon(3,1)=\langle\epsilon^3\rangle_c$, $\gamma(3,1)=\langle\epsilon'^3\rangle$, (5.8b)

(iii) $B^{(3)}=A^{(3)}=\frac{1}{4}$, (5.8c)

(iv) $\epsilon^{(3)}=\dfrac{\langle\epsilon^3\rangle_c}{4\langle\epsilon\rangle^2}=\dfrac{\langle\epsilon'^3\rangle}{4\langle\epsilon\rangle^2}$. (5.8d)

(fourth-order term)

(i) $F_4(1;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34})=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{34}}$,
$F_4(2;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34})=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{34}}$,
$F_4(3;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34})=\delta_{\mathbf{r}_{14}}\delta_{\mathbf{r}_{23}}$,
$F_4(4;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34})=\delta_{\mathbf{r}_{13}}\delta_{\mathbf{r}_{24}}$; (5.9a)

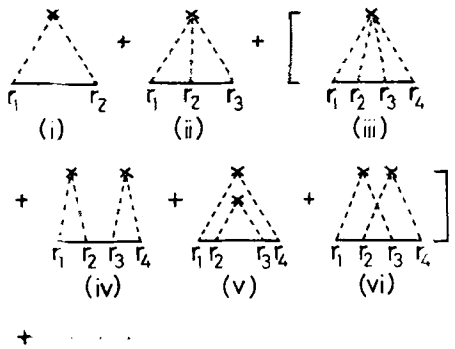


FIG. 1. Diagrammatic representation of Eq. (5.1).

(ii) $\epsilon(4,1)=\langle\epsilon^4\rangle_c$, $\gamma(4,1)=\langle\epsilon'^4\rangle$,

$$\epsilon(4,2)=\epsilon(4,3)=\epsilon(4,4)=\langle\epsilon^2\rangle_c^2, \qquad (5.9b)$$

$$\gamma(4,2)=\gamma(4,3)=\gamma(4,4)=\langle\epsilon'^2\rangle^2;$$

(iii) $B_1^{(4)}=A_1^{(4)}=\frac{1}{8}$,

$$B_2^{(4)}=0, \quad A_2^{(4)}=-\tfrac{1}{8}, \qquad (5.9c)$$

$$B_3^{(4)}=\tfrac{1}{4}, \quad A_3^{(4)}=\tfrac{1}{8},$$

$$B_4^{(4)}=\tfrac{1}{8}, \quad A_4^{(4)}=0;$$

(iv) $$\epsilon^{(4)}=-\frac{\langle\epsilon^4\rangle_c}{8\langle\epsilon\rangle^3}-\frac{3\langle\epsilon^2\rangle_c^2}{8\langle\epsilon\rangle^3}=-\frac{\langle\epsilon'^4\rangle}{8\langle\epsilon\rangle^3}; \qquad (5.9d)$$

(fifth-order term)

(i) $$F_5(1;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34},\mathbf{r}_{45})=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{34}}\delta_{\mathbf{r}_{45}},$$

$$F_5(2;\mathbf{r}_{12},\mathbf{r}_{23},\mathbf{r}_{34},\mathbf{r}_{45})=\delta_{\mathbf{r}_{12}}\delta_{\mathbf{r}_{23}}\delta_{\mathbf{r}_{45}}, \text{ etc.}; \qquad (5.10a)$$

(ii) $\epsilon(5,1)=\langle\epsilon^5\rangle_c$, $\gamma(5,1)=\langle\epsilon'^5\rangle$,

$$\epsilon(5,2)=\epsilon(5,3)=\cdots=\epsilon(5,11)=\langle\epsilon^3\rangle_c\langle\epsilon^2\rangle_c \qquad (5.10b)$$

$$\gamma(5,2)=\gamma(5,3)=\cdots=\gamma(5,11)=\langle\epsilon'^3\rangle\langle\epsilon'^2\rangle;$$

(iii) $B_1^{(5)}=A_1^{(5)}=\frac{1}{16}$,

$$B_2^{(5)}=B_3^{(5)}=0, \quad A_2^{(5)}=A_3^{(5)}=-\tfrac{1}{16},$$

$$B_4^{(5)}=B_5^{(5)}=B_6^{(5)}=\tfrac{1}{8}, \quad A_4^{(5)}=A_5^{(5)}=A_6^{(5)}=\tfrac{1}{16}, \qquad (5.10c)$$

$$B_7^{(5)}=B_8^{(5)}=B_9^{(5)}=B_{10}^{(5)}=\tfrac{1}{16},$$

$$A_7^{(5)}=A_8^{(5)}=A_9^{(5)}=A_{10}^{(5)}=0;$$

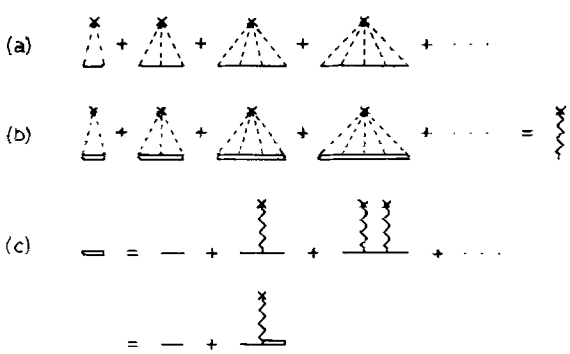


FIG. 2. Process for modifying single-site diagrams. (a) Single-site diagrams with free propagators. (b) Single-site diagrams with modified propagators. (c) Definition of a modified propagator.

(iii') $B_{11}^{(5)}=\frac{1}{16}$, $A_{11}^{(5)}=0$ (unproved); (5.10c')

(iv) $$\epsilon^{(5)}=\frac{\langle\epsilon^5\rangle_c}{16\langle\epsilon\rangle^4}+\frac{11\langle\epsilon^3\rangle_c\langle\epsilon^2\rangle_c}{16\langle\epsilon\rangle^4}$$

$$=\frac{\langle\epsilon'^5\rangle}{16\langle\epsilon\rangle^4}+\frac{\langle\epsilon'^3\rangle\langle\epsilon'^2\rangle}{16\langle\epsilon\rangle^4}. \qquad (5.10d)$$

In order to represent the perturbation series in Eq. (5.1), we adopt cumulant diagrams proposed by Yonezawa and Matsubara.[13] As mentioned in IV, these diagrams are composed of cross vertices, dashed vertical lines, and solid horizontal lines. According to usual nomenclature in the quantum-mechanical diagram methods, we shall name a vertical line an interaction line and a horizontal line a propagator. The perturbation expansion of ϵ'^* up to fourth order is expressed diagrammatically by Fig. 1. With the knowledge of lower-order terms obtained in the preceding paragraph, we make the rules for counting contributions from cumulant diagrams:

(i) Identify unconnected or improper diagrams with zero. In the quantum-mechanical problems we do not encounter the situation that all the improper diagrams contribute zero.

(ii) Assign ϵ' to a dashed interaction line and take the cumulant of the product of all ϵ''s that correspond to interaction lines starting from a cross vertex. That is to say, allot the νth-order cumulant $\langle\epsilon'^\nu\rangle_c$ to ν dashed interaction lines which start from a cross vertex. Note that $\langle\epsilon'^\nu\rangle_c=\langle\epsilon^\nu\rangle_c$ for $\nu\geq 2$.

(iii) Assign $-1/\langle\epsilon\rangle$ to each propagator.

(iv) Assign $\frac{1}{2}$ to each independent propagator. More generally, use the reciprocal of the dimensionality of the medium. In the Green's function formalism and diagrams in the momentum space, this independent propagator corresponds to the independent integral with respect to an inner propagator k. For detail see Refs. 12 and 14.

(v) Calculate the product of all factors thus determined.

The above prescriptions (ii) to (v) suffice to determine the contributions of nested diagrams as well as those of one-vertex diagrams. Although nested diagrams are concerned with two or more vertices, they are reducible to single-site diagrams with modified propagators. The process of modifying propagators is illus-

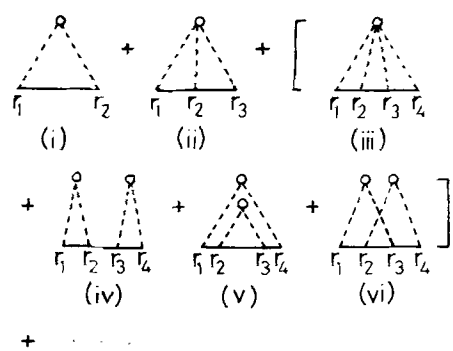


FIG. 3. Diagrammatic representation of Eq. (5.6).

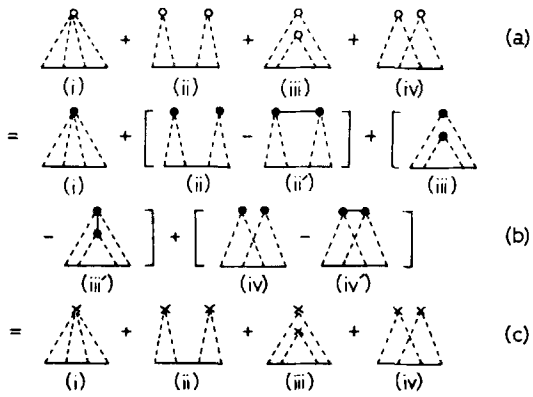


FIG. 4. Process for transforming restricted moment diagrams into cumulant diagrams. The diagram (a-i) corresponds to $\langle\epsilon'^4\rangle\delta_{r_{12}}\delta_{r_{23}}\delta_{r_{34}}$, (b-i) to $\langle\epsilon'^4\rangle\delta_{r_{12}}\delta_{r_{23}}\delta_{r_{34}}$, (c-i) to $\langle\epsilon^4\rangle_c\delta_{r_{12}}\delta_{r_{23}}\delta_{r_{34}}$, (a-ii) to $\langle\epsilon'^2\rangle^2\delta_{r_{12}}\delta_{r_{34}}(1-\delta_{r_{23}})$, (b-ii) to $\langle\epsilon'^2\rangle^2\delta_{r_{12}}\delta_{r_{34}}$, (b-ii) to $\langle\epsilon'^2\rangle^2\delta_{r_{12}}\delta_{r_{34}}\delta_{r_{23}}$, (c-ii) to $\langle\epsilon^2\rangle_c^2\delta_{r_{12}}\delta_{r_{34}}$, and so on.

trated in Fig. 2. Figure 2(a) represents the sum of all one-vertex diagrams. In Fig. 2(b), an inner free propagator expressed by a solid horizontal line is replaced by a modified propagator expressed by a double line which in turn is defined by Fig. 2(c). It is obvious that all nested diagrams are included in Fig. 2(b). Therefore, nested diagrams belong to single-site diagrams in a wide sense. The procedure to evaluate such single-site diagrams is similar to that described in Appendix A of IV.

Besides cumulant diagrams having cross vertices, we employ two other types of diagrams due to Yonezawa[12] which contain vertices indicated by open or closed circles. For a diagram with closed-circle vertices, called an unrestricted moment diagram, the νth-order moment $\langle\epsilon'^\nu\rangle$ instead of the νth-order cumulant $\langle\epsilon'^\nu\rangle_c$ is allotted to ν interaction lines starting from the same vertex. The prescriptions (i), (iii), (iv), and (v) hold good for a diagram of this sort. On the other hand, diagrams with open-circle vertices (which are called restricted moment diagrams) are introduced to represent perturbation terms in Eq. (5.6) rather than in Eq. (5.1). The diagram expression of Eq. (5.6) up to

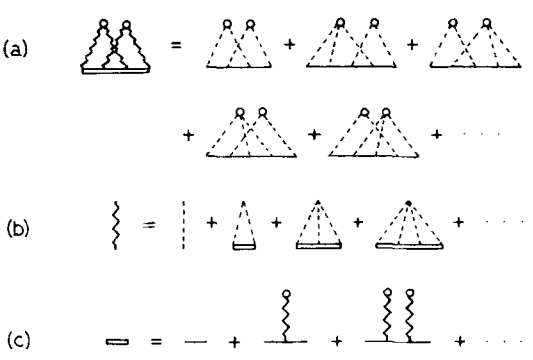


FIG. 5. Crossed diagrams with open-circle vertices whose contribution is zero. (a) Sum of vanishing crossed diagrams. (b) Definition of a wavy interaction line. (c) Definition of a double-line propagator.

fourth order is presented in Fig. 3. By way of illustration let us compare the fourth-order terms in Figs. 1 and 3. The process of transforming restricted moment diagrams into cumulant diagrams is shown in Fig. 4.

The contributions from crossed diagrams are quite different in our 2D case from the 3D case. Recall that $A_4^{(4)}=0$ and $A_7^{(5)}=A_8^{(5)}=A_9^{(5)}=A_{10}^{(5)}=0$ for a 2D material while $B_4^{(4)}=0$ and $B_7^{(5)}=B_8^{(5)}=B_9^{(5)}=B_{10}^{(5)}=0$ for a 3D material. In the 3D formulation, zero contributions are due to crossed diagrams with cross vertices. In the 2D problem, however, crossed diagrams with open-circle vertices are proved or expected to vanish. Thus, the following rule concerning crossed diagrams with open-circle vertices is framed:

(vi) Associate zero with crossed diagrams which are comprised as addends in Fig. 5.

The proof of the prescription (vi) proceeds exactly as in Appendix B of IV.

Strictly speaking, the above six rules are not comprehensive in the sense that all crossed diagrams do not belong to the category defined by Fig. 5. In fact, we are not able to prove $A_{11}^{(5)}=0$ and a number of other crossed diagrams of higher orders are undetermined. As we shall see later, we disregard all crossed diagrams with open-circle vertices in the course of calculating partial summation of the perturbation series. It is expected, however, that the neglect of these complicated diagrams is not serious partly because zero contributions are demonstrated for important higher-order diagrams included in Fig. 5, and partly because there are not wanting indications that some other crossed diagrams vanish.

6. APPROXIMATE PERTURBATION SOLUTIONS AND NUMERICAL RESULTS

In view of finding the best estimate for the effective permittivity in 2D systems, we calculate various approximate solutions to the perturbation series (5.1) or (5.6). First, Kröner's approximation[15-17] corresponds to summing only the first terms of the nth-order coefficients in Eq. (5.6). Consequently,

$$\epsilon_K^*=\langle\epsilon\rangle+\sum_{n=2}^{\infty}A_1^{(n)}\frac{\langle\epsilon'^n\rangle}{(-\langle\epsilon\rangle)^{n-1}}\,. \tag{6.1}$$

Considering that $A_1^{(n)}=B_1^{(n)}=1/2^{n-1}$, we have

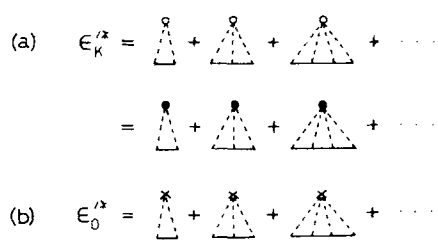


FIG. 6. Diagram equations for evaluating ϵ_K^* and ϵ_0^*. (a) Kröner's approximation. (b) Non-self-consistent cumulant approximation.

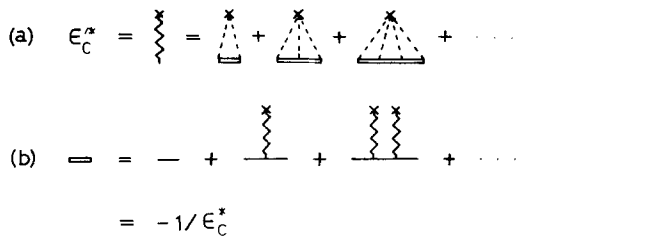


FIG. 7. Diagram equations for evaluating ϵ_C^*. (a) Self-consistent cumulant approximation. (b) Definition of a modified propagator.

$$\epsilon_K^* = \langle\epsilon\rangle - \frac{\langle\epsilon'^2\rangle}{2\langle\epsilon\rangle} + \frac{\langle\epsilon'^3\rangle}{4\langle\epsilon\rangle^2} - \frac{\langle\epsilon'^4\rangle}{8\langle\epsilon\rangle^3} + \frac{\langle\epsilon'^5\rangle}{16\langle\epsilon\rangle^4} - \cdots$$
$$= \langle\epsilon\rangle - \left\langle \frac{\epsilon'^2}{2\langle\epsilon\rangle + \epsilon'} \right\rangle. \tag{6.2}$$

In the language of diagrams, Eq. (6.2) is identical with the sum of restricted or unrestricted moment diagrams as shown in Fig. 6(a).

As discussed in IV, Kröner's theory neglects the exclusion effect; the ordinary moment averages must be replaced by the appropriate cumulant averages when the exclusion effect is properly taken into account. Diagrammatically, this reduces to the employment of cross vertices instead of circular vertices. The single-site diagrams along this line are given in Fig. 6(b). As a result, we obtain an infinite series of the form

$$\epsilon_0^* = \langle\epsilon\rangle - \sum_{n=2}^{\infty} \left(\frac{1}{2}\right)^{n-1} \frac{\langle\epsilon^n\rangle_c}{(-\langle\epsilon\rangle)^{n-1}}$$
$$= \langle\epsilon\rangle - \frac{\langle\epsilon^2\rangle_c}{2\langle\epsilon\rangle} + \frac{\langle\epsilon^3\rangle_c}{4\langle\epsilon\rangle^2} - \frac{\langle\epsilon^4\rangle_c}{8\langle\epsilon\rangle^3} + \frac{\langle\epsilon^5\rangle_c}{16\langle\epsilon\rangle^4} - \cdots$$
$$= \langle\epsilon\rangle - \frac{\langle\epsilon'^2\rangle}{2\langle\epsilon\rangle} + \frac{\langle\epsilon'^3\rangle}{4\langle\epsilon\rangle^2} - \frac{\langle\epsilon'^4\rangle - 3\langle\epsilon'^2\rangle^2}{8\langle\epsilon\rangle^3}$$
$$+ \frac{\langle\epsilon'^5\rangle - 10\langle\epsilon'^3\rangle\langle\epsilon'^2\rangle}{16\langle\epsilon\rangle} - \cdots . \tag{6.3}$$

The partial summation in Eq. (6.3) is equivalent to counting $B_1^{(n)}$ correctly and discarding all $B_m^{(n)}$ for $m \geq 2$. It has been pointed out in IV that Eq. (6.3) is brought into a compact form as

$$\epsilon_0^* = \int_0^1 \langle \epsilon z^{\epsilon/2\langle\epsilon\rangle}\rangle / \langle z^{\epsilon/2\langle\epsilon\rangle}\rangle \, dz. \tag{6.4}$$

The cumulant expansion solution (6.4) is usually regarded as non-self-consistent in the sense that the corresponding diagrams are expressed by free propagators. We shall make the approximation self-consistent so that all of nested diagrams discarded in the non-self-consistent approximation are included. The self-consistent cumulant solution is derived by replacing inner free propagators with modified propagators illustrated in Fig. 2(c). Namely, we use the diagram equation in Fig. 7, where a wavy interaction line is identified with $\epsilon_C'^*$. Then, it is easy to see that a modified propagator is calculated as

$$\frac{-1}{\langle\epsilon\rangle} + \frac{-1}{\langle\epsilon\rangle}\epsilon_C'^*\frac{-1}{\langle\epsilon\rangle} + \frac{-1}{\langle\epsilon\rangle}\epsilon_C'^*\frac{-1}{\langle\epsilon\rangle}\epsilon_C'^*\frac{-1}{\langle\epsilon\rangle} + \cdots$$
$$= \frac{-1}{\langle\epsilon\rangle}\frac{1}{1+\epsilon_C'^*/\langle\epsilon\rangle} = -\frac{1}{\epsilon_C^*}. \tag{6.5}$$

The resultant effective permittivity satisfies a self-consistent equation

$$\epsilon_C^* = \int_0^1 \langle \epsilon z^{\epsilon/2\epsilon_C^*}\rangle / \langle z^{\epsilon/2\epsilon_C^*}\rangle. \tag{6.6}$$

The series expansion of ϵ_C^* up to fifth order is

$$\epsilon_C^* = \langle\epsilon\rangle - \frac{\langle\epsilon^2\rangle_c}{2\langle\epsilon\rangle} + \frac{\langle\epsilon^3\rangle_c}{4\langle\epsilon\rangle^2} - \frac{\langle\epsilon^4\rangle_c + 2\langle\epsilon^2\rangle_c^2}{8\langle\epsilon\rangle^3}$$
$$+ \frac{\langle\epsilon^5\rangle_c + 6\langle\epsilon^3\rangle_c\langle\epsilon^2\rangle_c}{16\langle\epsilon\rangle^4} - \cdots$$
$$= \langle\epsilon\rangle - \frac{\langle\epsilon'^2\rangle}{2\langle\epsilon\rangle} + \frac{\langle\epsilon'^3\rangle}{4\langle\epsilon\rangle^2} - \frac{\langle\epsilon'^4\rangle - \langle\epsilon'^2\rangle^2}{8\langle\epsilon\rangle^3}$$
$$+ \frac{\langle\epsilon'^5\rangle - 4\langle\epsilon'^3\rangle\langle\epsilon'^2\rangle}{16\langle\epsilon\rangle^4} - \cdots . \tag{6.7}$$

It has been argued in Ref. 12 that, for the purpose of obtaining both mathematically correct and physically reasonable results, the corrections from the exclusion effect and the degree of approximations must be self-contained. The self-consistent cumulant solution is not self-contained since the theory counts the corrections from all higher-order diagrams, some of which are neglected in the approximation. The procedure of making the single-site approximation self-contained is given in Fig. 8, where dotted lines denote renormalized interaction lines. This treatment leads to the coherent-potential approximation (CPA) for the average Green's function of noninteracting electrons in disordered binary alloys.[12,14] In IV it has been proved that the CPA is identical with the effective-medium (EM) theory for the problem of classical systems.[18] Moreover, we remark that the self-contained single-site approximation or the EM theory essentially ignores crossed diagrams with open-circle vertices, while the self-consistent cumulant approximation disregards crossed diagrams with cross vertices.

According to Fig. 8, the effective permittivity determined by the self-contained single-site approximation is written as

$$\epsilon_{EM}^* = \langle\epsilon\rangle - \frac{\langle\epsilon'(\epsilon - \epsilon_{EM}^*)\rangle}{2\epsilon_{EM}^*} + \frac{\langle\epsilon'(\epsilon - \epsilon_{EM}^*)^2\rangle}{4\epsilon_{EM}^{*2}} - \cdots$$
$$= \langle\epsilon\rangle - \left\langle \frac{\epsilon'(\epsilon - \epsilon_{EM}^*)}{\epsilon + \epsilon_{EM}^*} \right\rangle, \tag{6.8}$$

which yields the relation

$$\left\langle \frac{\epsilon - \epsilon_{EM}^*}{\epsilon + \epsilon_{EM}^*} \right\rangle = 0. \tag{6.9}$$

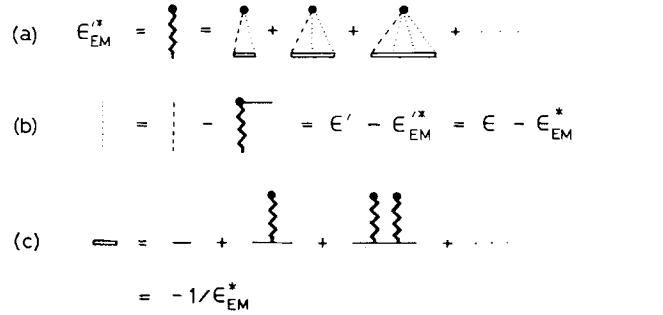


FIG. 8. Diagram equations for evaluating ϵ_{EM}^*. (a) Self-contained single-site approximation. (b) Definition of a renormalized interaction line. (c) Definition of a modified propagator.

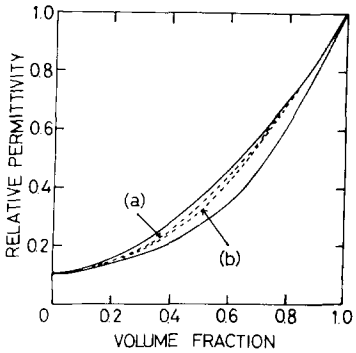


FIG. 9. Plots of the relative permittivity ϵ_C^*/ϵ_1 and $\epsilon_{EM}^*/\epsilon_1$ against the volume fraction x for the choice of $\epsilon_2/\epsilon_1 = 0.1$. The dashed lines (a) and (b) represent ϵ_C^* and ϵ_{EM}^*, respectively, and the solid lines indicate the bounds on ϵ^*.

Equation (6.9) can also be obtained from the EM theory in an analogous manner to the derivation of Eq. (IV.4.3). The explicit expansion series of ϵ_{EM}^* up to fifth order becomes

$$\epsilon_{EM}^* = \langle\epsilon\rangle - \frac{\langle\epsilon^2\rangle_c}{2\langle\epsilon\rangle} + \frac{\langle\epsilon^3\rangle_c}{4\langle\epsilon\rangle^2} - \frac{\langle\epsilon^4\rangle_c + 3\langle\epsilon^2\rangle_c^2}{8\langle\epsilon\rangle^3} + \frac{\langle\epsilon^5\rangle_c + 11\langle\epsilon^3\rangle_c\langle\epsilon^2\rangle_c}{16\langle\epsilon\rangle^4} - \cdots$$
$$= \langle\epsilon\rangle - \frac{\langle\epsilon'^2\rangle}{2\langle\epsilon\rangle} + \frac{\langle\epsilon'^3\rangle}{4\langle\epsilon\rangle^2} - \frac{\langle\epsilon'^4\rangle}{8\langle\epsilon\rangle^3} + \frac{\langle\epsilon'^5\rangle + \langle\epsilon'^3\rangle\langle\epsilon'^2\rangle}{16\langle\epsilon\rangle^4} - \cdots . \tag{6.10}$$

Now we consider 2D systems with binary disorder. Let the two constituent phases have permittivities ϵ_1 and ϵ_2 and occupy fractions of the total volume $v_1 = x$ and $v_2 = 1 - x$. Then it follows that

$$\epsilon_K^* = [\epsilon_2 + (\epsilon_1 - \epsilon_2)x]\left(1 - 2(\epsilon_1 - \epsilon_2)^2 \times \frac{x(1-x)}{[\epsilon_1 + \epsilon_2 + (\epsilon_1 - \epsilon_2)x][2\epsilon_2 + (\epsilon_1 - \epsilon_2)x]}\right), \tag{6.11}$$

$$\epsilon_0^* = \epsilon_2 + (\epsilon_1 - \epsilon_2)x \int_0^1 dz\, [x + (1-x)z^{-(\epsilon_1-\epsilon_2)/2[\epsilon_2+(\epsilon_1-\epsilon_2)x]}]^{-1}, \tag{6.12}$$

$$\epsilon_C^* = \epsilon_2 + (\epsilon_1 - \epsilon_2)x \int_0^1 dz [x + (1-x)z^{-(\epsilon_1-\epsilon_2)/2\epsilon_C^*}]^{-1}, \tag{6.13}$$

$$\epsilon_{EM}^{*2} + (\epsilon_1 - \epsilon_2)(1-2x)\epsilon_{EM}^* - \epsilon_1\epsilon_2 = 0. \tag{6.14}$$

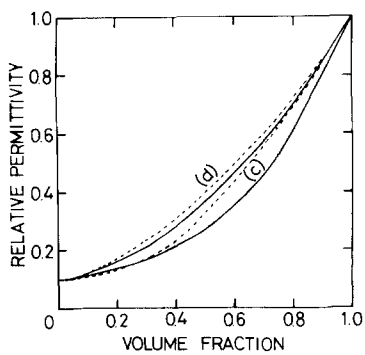


FIG. 10. Plots of the relative permittivity ϵ_K^*/ϵ_1 and ϵ_0^*/ϵ_1 against the volume fraction x for the choice of $\epsilon_2/\epsilon_1 = 0.1$. The dashed lines (c) and (d) represent ϵ_K^* and ϵ_0^*, respectively, and the solid lines indicate the bounds on ϵ^*.

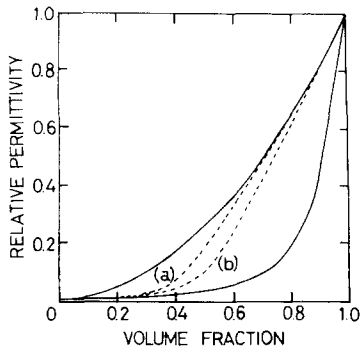


FIG. 11. Plots of the relative permittivity ϵ_C^*/ϵ_1 and $\epsilon_{EM}^*/\epsilon_1$ against the volume fraction x for the choice of $\epsilon_2/\epsilon_1 = 0.01$. Labeling conventions are the same as in Fig. 9.

The root of the quadratic equation (6.14) is

$$\epsilon_{EM}^* = \tfrac{1}{2}(-(\epsilon_1 - \epsilon_2)(1-2x) + [(\epsilon_1+\epsilon_2)^2 - 4(\epsilon_1-\epsilon_2)^2x + 4(\epsilon_1-\epsilon_2)^2x^2]^{1/2}). \tag{6.14'}$$

The bounding inequalities (3.20) and (3.21) reduce to

$$\epsilon^* \le \epsilon_2 + \frac{2\epsilon_2(\epsilon_1-\epsilon_2)}{\epsilon_1+\epsilon_2}x + \frac{(\epsilon_1-\epsilon_2)^2}{\epsilon_1+\epsilon_2}x^2, \tag{6.15}$$

$$\epsilon^* \ge \frac{\epsilon_1\epsilon_2}{\epsilon_1 - [2\epsilon_1(\epsilon_1-\epsilon_2)/(\epsilon_1+\epsilon_2)]x + [(\epsilon_1-\epsilon_2)^2/(\epsilon_1+\epsilon_2)]x^2}. \tag{6.16}$$

In Figs. 9 to 14 we plot Eqs. (6.11) to (6.14') together with the bounds (6.15) and (6.16). The relative values of the effective permittivity, ϵ^*/ϵ_1, for the three cases of $\epsilon_2/\epsilon_1 = 0.1$, 0.01, 0 are exhibited as functions of the volume fraction x. As in the case of 3D materials, ϵ_K^* sometimes violates the lower bound and even takes negative values. If $\epsilon_2/\epsilon_1 = 0$ we find from Eq. (6.11)

$$\epsilon_K^*/\epsilon_1 = -[x(1-3x)/(1+x)], \tag{6.17}$$

whence $\epsilon_K^* < 0$ for $0 < x < \frac{1}{3}$. The non-self-consistent cumulant approximation systematically gives ϵ_0^* larger than the upper bound. This is owing to the fact that a number of important diagrams are assumed to be zero and the corrections from the exclusion effect are over-

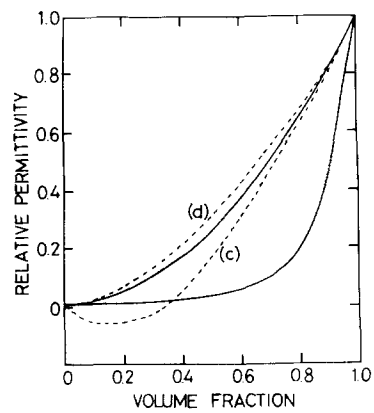


FIG. 12. Plots of the relative permittivity ϵ_K^*/ϵ_1 and ϵ_0^*/ϵ_1 against the volume fraction x for the choice of $\epsilon_2/\epsilon_1 = 0.01$. Labeling conventions are the same as in Fig. 10.

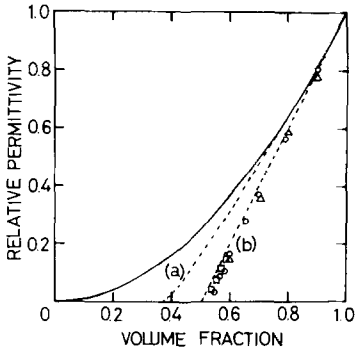


FIG. 13. Plots of the relative permittivity ϵ_C^*/ϵ_1 and $\epsilon_{EM}^*/\epsilon_1$ against the volume fraction x for the choice of $\epsilon_2/\epsilon_1 = 0$. Data points represent the results of the computer simulation carried out by Kirkpatrick.[8,21,22] Other labeling conventions are the same as in Fig. 9.

counted. In both the self-consistent cumulant theory and the EM approximation, however, the estimates of ϵ_C^* and ϵ_{EM}^* lie always within the bounds.

To discuss the comparative merits of ϵ_C^* and ϵ_{EM}^* we investigate the behavior of effective permittivities in the vicinity of $\epsilon_2/\epsilon_1 = 0$. The difference between ϵ_C^* and ϵ_{EM}^* becomes serious as ϵ_2/ϵ_1 approaches zero. From Eq. (6.14′) we find

$$\frac{\epsilon_{EM}^*}{\epsilon_1} = \begin{cases} 0 & \text{for } x < \frac{1}{2}, \\ -1 + 2x & \text{for } x > \frac{1}{2}. \end{cases} \tag{6.18}$$

The critical concentration $\tilde{x}_c = \frac{1}{2}$ reached by Eq. (6.18) is in agreement with the critical probability for the bond percolation on a square lattice.[7,19,20] On the contrary, the self-consistent cumulant method fails to give a correct critical percolation concentration. Actually, substitution of $\epsilon_2 = 0$ into Eq. (6.13) yields

$$\frac{\epsilon_C^*}{\epsilon_1} = x \int_0^1 \frac{z^{\epsilon/2\epsilon_C^*}}{1 - x + xz^{\epsilon/2\epsilon_C^*}} dz. \tag{6.19}$$

By reference to Appendix C in IV, the critical concentration is estimated at

$$\bar{x}_c = 1 - e^{-1/2} = 0.393\cdots. \tag{6.20}$$

The numerical data by Kirkpatrick,[8,21,22] who carried out computer simulation on 2D square networks with 50 ×50 sites, also fit Eq. (6.18) rather than Eq. (6.19) (see Fig. 13).

The above situation is different from the 3D case in which ϵ_C^* serves as a better approximation than ϵ_{EM}^*. The difference is clearly explained by means of diagrams. In the 3D problem, contributions of crossed diagrams with cross vertices are proved or expected to vanish. Zero contributions from crossed diagrams with cross vertices guarantee the validity of the self-consistent cumulant solution, because in this theory the corrections coming from the exclusion effect of crossed diagrams are wholly taken into account. Although the crossed diagrams themselves contribute nothing, their exclusion corrections are not zero and these terms are exactly counted in the cumulant approximation. Now that crossed diagrams are concerned with cluster effects, ϵ_C^* corresponds to an improved approximation which includes cluster effects. As stated in IV, this is the reason why ϵ_C^* gives a better result than ϵ_{EM}^* near the percolation threshold.

In our case of 2D systems, we have demonstrated that crossed diagrams with open-circle vertices as expressed by Fig. 5 vanish. That is, the contributions from a crossed cumulant diagram and the corrections due to the same diagram are equal to zero only all together but not independently. In this sense, the effective-medium theory is regarded as an approximation in which crossed cumulant diagrams and exclusion-effect corrections from them are both taken into consideration in a self-contained manner, and therefore some important cluster effects are included. In contrast to ϵ_{EM}^*, the cumulant solution ϵ_C^* contains corrections alone and apparently violates the self-containedness requirement. Here we recall that Eq. (6.10) is exact up to fourth order but Eq. (6.7) is not true in fourth order. This reasoning accounts for the numerical behavior of ϵ_{EM}^* and ϵ_C^* in 2D materials.

7. ONE-DIMENSIONAL SYSTEMS

We conclude this paper with a discussion of the mathematically trivial but nevertheless interesting 1D problem. In the 1D case the governing equation (2.3) will be

$$\frac{d\epsilon(x_1)E_1(x_1)}{dx_1} = 0, \tag{7.1}$$

or

$$\epsilon(x_1)E_1(x_1) = \text{const} = \langle \epsilon(x_1)E_1(x_1) \rangle. \tag{7.2}$$

Since the effective permittivity ϵ^* is defined by

$$\langle \epsilon(x_1)E_1(x_1) \rangle = \epsilon^* \langle E_1 \rangle, \tag{7.3}$$

we arrive at an exact result

$$\epsilon^* = 1/\langle 1/\epsilon \rangle. \tag{7.4}$$

The perturbation series for ϵ^* is[23]

$$\epsilon^* = \langle \epsilon \rangle - \frac{\langle \epsilon'^2 \rangle}{\langle \epsilon \rangle} + \frac{\langle \epsilon'^3 \rangle}{\langle \epsilon \rangle^2} - \frac{\langle \epsilon'^4 \rangle - \langle \epsilon'^2 \rangle^2}{\langle \epsilon \rangle^3} + \frac{\langle \epsilon'^5 \rangle - 2\langle \epsilon'^3 \rangle \langle \epsilon'^2 \rangle}{\langle \epsilon \rangle^4} - \cdots$$

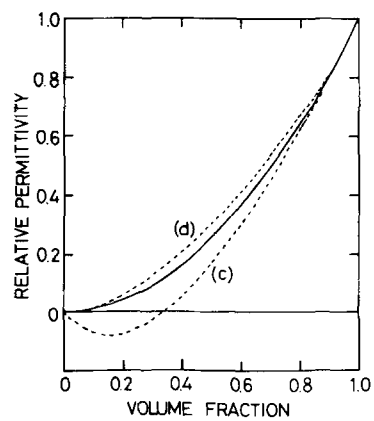


FIG. 14. Plots of the relative permittivity ϵ_K^*/ϵ_1 and $\epsilon_\delta^*/\epsilon_1$ against the volume fraction x for the choice of $\epsilon_2/\epsilon_1 = 0$. Labeling conventions are the same as in Fig. 10.

$$= \langle\epsilon\rangle - \frac{\langle\epsilon^2\rangle_c}{\langle\epsilon\rangle} + \frac{\langle\epsilon^3\rangle_c}{\langle\epsilon\rangle^2} - \frac{\langle\epsilon^4\rangle_c + 2\langle\epsilon^2\rangle_c}{\langle\epsilon\rangle^3} + \frac{\langle\epsilon^5\rangle_c + 8\langle\epsilon^3\rangle_c\langle\epsilon^2\rangle_c}{\langle\epsilon\rangle^4} - \cdots . \tag{7.5}$$

It has already been pointed out that $A_1^{(n)} = 1/3^{n-1}$ for a 3D system and $A_1^{(n)} = 1/2^{n-1}$ for a 2D system [see Eqs. (III. 2. 27) and (3. 31)]. The corresponding expression in the present case is

$$A_1^{(n)} = 1/1^{n-1} = 1, \tag{7.6}$$

so that Kröner's approximation gives

$$\epsilon_K^* = \langle\epsilon\rangle + \sum_{n=2}^{\infty} \frac{\langle\epsilon'^n\rangle}{(-\langle\epsilon\rangle)^{n-1}}$$
$$= \langle\epsilon\rangle - \frac{\langle\epsilon'^2\rangle}{\langle\epsilon\rangle} + \frac{\langle\epsilon'^3\rangle}{\langle\epsilon\rangle^2} - \frac{\langle\epsilon'^4\rangle}{\langle\epsilon\rangle^3} + \frac{\langle\epsilon'^5\rangle}{\langle\epsilon\rangle^4} - \cdots$$
$$= \langle\epsilon\rangle - \left\langle \frac{\epsilon'^2}{\langle\epsilon\rangle + \epsilon'} \right\rangle = \langle\epsilon\rangle \left(2 - \langle\epsilon\rangle \left\langle \frac{1}{\epsilon} \right\rangle \right). \tag{7.7}$$

As the non-self-consistent cumulant solution we get

$$\epsilon_0^* = \langle\epsilon\rangle + \sum_{n=2}^{\infty} \frac{\langle\epsilon^n\rangle_c}{(-\langle\epsilon\rangle)^{n-1}}$$
$$= \langle\epsilon\rangle - \frac{\langle\epsilon^2\rangle_c}{\langle\epsilon\rangle} + \frac{\langle\epsilon^3\rangle_c}{\langle\epsilon\rangle^2} - \frac{\langle\epsilon^4\rangle_c}{\langle\epsilon\rangle^3} + \frac{\langle\epsilon^5\rangle_c}{\langle\epsilon\rangle^4} - \cdots$$
$$= \langle\epsilon\rangle - \frac{\langle\epsilon'^2\rangle}{\langle\epsilon\rangle} + \frac{\langle\epsilon'^3\rangle}{\langle\epsilon\rangle^2} - \frac{\langle\epsilon'^4\rangle - 3\langle\epsilon'^2\rangle^2}{\langle\epsilon\rangle^3} + \frac{\langle\epsilon'^5\rangle - 10\langle\epsilon'^2\rangle^2}{\langle\epsilon\rangle^4} - \cdots$$
$$= \int_0^1 \langle\epsilon z^{\epsilon/\langle\epsilon\rangle}\rangle / \langle z^{\epsilon/\langle\epsilon\rangle}\rangle \, dz. \tag{7.8}$$

Similarly, for the self-consistent cumulant solution,

$$\epsilon_C^* = \langle\epsilon\rangle - \frac{\langle\epsilon^2\rangle_c}{\langle\epsilon\rangle} + \frac{\langle\epsilon^3\rangle_c}{\langle\epsilon\rangle^2} - \frac{\langle\epsilon^4\rangle_c + \langle\epsilon^2\rangle_c^2}{\langle\epsilon\rangle^3} + \frac{\langle\epsilon^5\rangle_c + 3\langle\epsilon^3\rangle_c\langle\epsilon^2\rangle_c}{\langle\epsilon\rangle^4} - \cdots$$
$$= \langle\epsilon\rangle - \frac{\langle\epsilon'^2\rangle}{\langle\epsilon\rangle} + \frac{\langle\epsilon'^3\rangle}{\langle\epsilon\rangle^2} - \frac{\langle\epsilon'^4\rangle - 2\langle\epsilon'^2\rangle^2}{\langle\epsilon\rangle^3} + \frac{\langle\epsilon'^5\rangle - 7\langle\epsilon'^3\rangle\langle\epsilon'^2\rangle}{\langle\epsilon\rangle^4} - \cdots$$
$$= \int_0^1 \langle\epsilon z^{\epsilon/\epsilon_C^*}\rangle / \langle z^{\epsilon/\epsilon_C^*}\rangle \, dz. \tag{7.9}$$

It is important to note that the EM approximation for a 1D materiai coincides with the exact solution. The EM theory is exact in one dimension because the replacement of the surrounding medium of a particular element by the effective medium does not affect the electric displacement in that element. Actually, the diagram equation shown in Fig. 8 yields

$$\epsilon_{EM}^* = \langle\epsilon\rangle - \frac{\langle\epsilon'(\epsilon - \epsilon_{EM}^*)\rangle}{\epsilon_{EM}^*} + \frac{\langle\epsilon'(\epsilon - \epsilon_{EM}^*)^2\rangle}{\epsilon_{EM}^{*2}} - \cdots$$
$$= \langle\epsilon\rangle - \left\langle \frac{\epsilon'(\epsilon - \epsilon_{EM}^*)}{\epsilon} \right\rangle, \tag{7.10}$$

which is equivalent to

$$\epsilon_{EM}^* = 1/\langle 1/\epsilon \rangle. \tag{7.11}$$

To sum up, the EM theory provides an exact result in 1D systems and serves as a better approximation than the cumulant expansion method in 2D systems.

APPENDIX A

Integrating by parts and applying Gauss' theorem of two dimensions, we recast Eq. (2. 13) as

$$A_{ij}^{(2)} = \frac{1}{2\pi} \int_S d\sigma_{12} \frac{\partial \log(1/r_{12})}{\partial x_{12,i}} \frac{\partial g(\mathbf{r}_{12})}{\partial x_{12,j}}$$
$$= -\frac{1}{2\pi} \int_C \frac{d\mu_{12}}{r_{12}} n_{12,j} \frac{x_{12,i}}{r_{12}} g(\mathbf{r}_{12}) - \frac{1}{2\pi} \int_S d\sigma_{12} \times \frac{\partial^2 \log(1/r_{12})}{\partial x_{12,i} \partial x_{12,j}} g(\mathbf{r}_{12}). \tag{A1}$$

Here C is the bounding curve of the medium S, $d\mu_{12}$ a line element of the boundary C, and $n_{12,j}$ the x_j component of an outward unit normal to the line element $d\mu_{12}$. Since the boundary condition (2. 14) guarantees that the curvilinear integral on the right approaches zero as $S \to \infty$, the proof of Eq. (2. 15) is completed.

To demonstrate Eq. (2. 16) we transform the right-hand side into

$$\frac{1}{(2\pi)^2} \int_S d\sigma_1 \int_S d\sigma_2 \frac{x_{1,k}}{r_1^2} \frac{x_{2,k}}{r_2^2} \frac{\partial^2 g(\mathbf{r}_{12})}{\partial x_{1,j} \partial x_{2,i}}$$
$$= -\frac{1}{(2\pi)^2} \int_S d\sigma_1 \int_S d\sigma_2 \frac{x_{1,k}}{r_1^2} \frac{\partial}{\partial x_{2,k}} \left(\frac{x_{2,i}}{r_2^2} \right) \frac{\partial g(\mathbf{r}_{12})}{\partial x_{1,j}}$$
$$= -\frac{1}{(2\pi)^2} \int_S d\sigma_{12} \frac{\partial g(\mathbf{r}_{12})}{\partial x_{12,j}} \int_S d\sigma_1 \frac{x_{1,k}}{r_1^2} \frac{\partial}{\partial x_{1,k}} \left(\frac{x_{2,i}}{r_2^2} \right). \tag{A2}$$

In polar coordinates we have

$$\int_S d\sigma_1 \frac{x_{1,k}}{r_1^2} \frac{\partial}{\partial x_{1,k}} \left(\frac{x_{2,i}}{r_2^2} \right) = \int_0^{2\pi} d\theta_1 \int_0^{\infty} dr_1 \frac{\partial}{\partial r_1} \left(\frac{x_{2,i}}{r_2^2} \right)$$
$$= -2\pi \frac{x_{12,i}}{r_{12}^2}. \tag{A3}$$

Insertion of Eq. (A3) into Eq. (A2) yields

$$\frac{1}{(2\pi)^2} \int_S d\sigma_1 \int_S d\sigma_2 \frac{x_{1,k}}{r_1^2} \frac{x_{2,k}}{r_2^2} \frac{\partial^2 g(\mathbf{r}_{12})}{\partial x_{1,i} \partial x_{2,j}}$$
$$= -\frac{1}{2\pi} \int_S d\sigma_{12} \frac{x_{12,i}}{r_{12}^2} \frac{\partial g(\mathbf{r}_{12})}{\partial x_{12,j}} = A_{ij}^{(2)}, \tag{A4}$$

which accomplishes the proof.

APPENDIX B

The quantity $J_{\rho,ik}(\mathbf{r})$ entering in Eq. (4. 18) is an isotropic tensor and has the form

$$J_{\rho,ik}(\mathbf{r}_{23}) = -\frac{1}{2\pi} \int_S d\sigma_{12} \frac{x_{12,i}}{r_{12}^2} \frac{\partial I_\rho(r_{13})}{\partial x_{12,k}}$$
$$= C_\rho(r_{23}) \frac{x_{23,i}}{r_{23}} \frac{x_{23,k}}{r_{23}} + D_\rho(r_{23}) \delta_{ik}. \tag{B1}$$

It follows that

$$J_{\rho,ii}(\mathbf{r}_{23}) = -\frac{1}{2\pi} \int_S d\sigma_{12} \frac{x_{12,i}}{r_{12}^2} \frac{\partial I_\rho(r_{13})}{\partial x_{12,i}}$$
$$= -\frac{1}{2\pi} \int_0^{2\pi} d\theta_{12} \int_0^{\infty} dr_{12} \frac{\partial I_\rho(r_{13})}{\partial r_{12}}$$
$$= I_\rho(r_{23}) = C_\rho(r_{23}) + 2D_\rho(r_{23}). \tag{B2}$$

For evaluating $C_\rho(r_{23})$ and $D_\rho(r_{23})$ choose $\mathbf{r}_{23}$ to lie along the x_1 axis; then

$$\frac{\partial r_{13}}{\partial x_{13,1}} = \frac{x_{13,1}}{r_{13}} = \cos\theta_{13}. \tag{B3}$$

From Eqs. (B1) and (B3) we obtain

$$\begin{aligned} J_{\rho,11}(\mathbf{r}_{23}) &= -\frac{1}{2\pi}\int_S d\sigma_{13}\frac{x_{12,1}}{r_{12}^2}\frac{\partial I_\rho(r_{13})}{\partial x_{13,1}} \\ &= -\frac{1}{2\pi}\int_0^{2\pi} d\theta_{13}\int_0^\infty dr_{13} \\ &\quad\times\frac{r_{13}(r_{13}\cos\theta_{13}-r_{23})\cos\theta_{13}}{r_{13}^2+r_{23}^2-2r_{23}r_{13}\cos\theta_{13}}\frac{dI_\rho(r_{13})}{dr_{13}}. \end{aligned} \tag{B4}$$

Because

$$\begin{aligned} &-\frac{1}{2\pi}\int_0^{2\pi} d\theta_{13}\frac{(r_{13}\cos\theta_{13}-r_{23})\cos\theta_{13}}{r_{13}^2+r_{23}^2-2r_{23}r_{13}\cos\theta_{13}} \\ &= \begin{cases} r_{13}/2r_{23}^2 & \text{for } r_{13}<r_{23}, \\ -1/2r_{13} & \text{for } r_{13}>r_{23}, \end{cases} \end{aligned} \tag{B5}$$

Eq. (B4) becomes

$$\begin{aligned} J_{\rho,11}(\mathbf{r}_{23}) &= \frac{1}{2r_{23}^2}\int_0^{r_{23}} r_{13}^2\,dI_\rho(r_{13}) + \tfrac{1}{2}I_\rho(r_{23}) \\ &= C_\rho(r_{23}) + D_\rho(r_{23}). \end{aligned} \tag{B6}$$

The solutions of Eqs. (B2) and (B6) are

$$\begin{aligned} C_\rho(r_{23}) &= \frac{1}{r_{23}^2}\int_0^{r_{23}} r_{13}^2\,dI_\rho(r_{13}) \\ &= \begin{cases} 0 & \text{for } r_{23}<\rho, \\ -\rho^2/r_{23}^2 & \text{for } r_{23}>\rho, \end{cases} \end{aligned} \tag{B7}$$

$$\begin{aligned} D_\rho(r_{23}) &= \frac{I_\rho(r_{23})}{2} - \frac{1}{2r_{23}^2}\int_0^{r_{23}} r_{13}^2\,dI_\rho(r_{13}) \\ &= \begin{cases} \frac{1}{2} & \text{for } r_{23}<\rho, \\ \rho^2/2r_{23}^2 & \text{for } r_{23}>\rho. \end{cases} \end{aligned} \tag{B8}$$

It is seen that

$$\begin{aligned} &J_{\rho,ik}(\mathbf{r}_{23})J_{\rho,hl}(\mathbf{r}_{23}) \\ &= C_\rho(r_{23})[C_\rho(r_{23})+2D_\rho(r_{23})]\frac{x_{23,k}}{r_{23}}\frac{x_{23,h}}{r_{23}} + [D_\rho(r_{23})]^2\delta_{kh}, \end{aligned} \tag{B9}$$

$$\begin{aligned} \frac{\partial J_{\rho,ik}(\mathbf{r}_{23})J_{\rho,hl}(\mathbf{r}_{23})}{\partial x_{23,h}} &= \frac{x_{23,k}}{r_{23}}\left(\frac{d[C_\rho(r_{23})+D_\rho(r_{23})]^2}{dr_{23}}\right. \\ &\quad\left.+\frac{C_\rho(r_{23})[C_\rho(r_{23})+2D_\rho(r_{23})]}{r_{23}}\right). \end{aligned} \tag{B10}$$

Accordingly, combination of Eqs. (4.18), (B7), (B8), and (B10) leads to the conclusion that

$$\begin{aligned} B_4^{(4)} &= \tfrac{1}{8} - \tfrac{1}{2}\lim_{\rho\to+0}\int_0^\infty \frac{C_\rho(r_{23})[C_\rho(r_{23})+2D_\rho(r_{23})]}{r_{23}}\,dr_{23} \\ &= \tfrac{1}{8}, \end{aligned} \tag{B11}$$

as asserted in Eq. (4.23).

[1] M. Hori, J. Math. Phys. **14**, 514 (1973).
[2] M. Hori, J. Math. Phys. **14**, 1942 (1973).
[3] M. Hori and F. Yonezawa, J. Math. Phys. **15**, 2177 (1974).
[4] M. Hori and F. Yonezawa, J. Math. Phys. **16**, 352 (1975) (preceeding article).
[5] Z. Hashin, in *Mechanics of Composite Materials*, edited by F.W. Wendt, H. Liebowitz, and N. Perrone (Pergamon, Oxford, 1970), p. 201.
[6] M.J. Beran and N.R. Silnutzer, J. Compos. Mater. **5**, 246 (1971).
[7] V.K.S. Shante and S. Kirkpatrick, Adv. Phys. **20**, 325 (1971).
[8] S. Kirkpatrick, Rev. Mod. Phys. **45**, 574 (1973).
[9] M.N. Miller, J. Math. Phys. **10**, 1988 (1969).
[10] M.N. Miller, J. Math. Phys. **10**, 2005 (1969); **12**, 1057 (1971).
[11] See, for example, O.D. Kellogg, *Foundations of Potential Theory* (Springer, Berlin, 1967), p. 192.
[12] F. Yonezawa, Prog. Theor. Phys. **40**, 734 (1968).
[13] F. Yonezawa and T. Matsubara, Prog. Theor. Phys. **35**, 357, 759 (1966).
[14] F. Yonezawa and K. Morigaki, Prog. Theor. Phys. Suppl. **53**, 1 (1973).
[15] E. Kröner, J. Mech. Phys. Solids **15**, 319 (1967).
[16] E. Kröner, Phys. Status Solidi **41**, 499 (1970).
[17] E. Kröner, Phys. Status Solidi B **49**, K23 (1972).
[18] A comprehensive list of references on the EM theory is found in Ref. 4.
[19] M.F. Sykes and J.W. Essam, J. Math. Phys. **5**, 1117 (1964).
[20] J.W. Essam, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M.S. Green (Academic, New York, 1972), Vol. 2, p. 197.
[21] S. Kirkpatrick, Phys. Rev. Lett. **27**, 1722 (1971).
[22] S. Kirkpatrick, in *The Properties of Liquid Metals*, edited by S. Takeuchi (Taylor and Francis, London, 1973), p. 351.
[23] M. Beran and J. Molyneux, Nuovo Cimento **30**, 1406 (1963).

Simple expression for the linear boson transformation coefficients. II

M. A. Rashid

International Centre for Theoretical Physics, Trieste, Italy
(Received 30 October 1973)

We obtain a simple expression for the coefficients which connect a Fock state containing an arbitrary number of quasiparticles with its transformed state under a boson Bogolubov transformation.

I. INTRODUCTION

On account of the importance of the Bogolubov transformation in the theoretical understanding of superconductivity, [1] there has been interest in obtaining an explicit expression for the transformation coefficients which connect a Fock state containing an arbitrary number of quasiparticles with its transform under a boson Bogolubov transformation. In fact, Tanabe[2] in a recent attempt obtained an expression for these coefficients. However, his answer is extremely complicated on account of the presence of four explicit and two implicit summations.[3] Moreover, his calculations are not at all straightforward and involve a lot of computations. It seems that his *indirect* method of using the eigenfunctions of a linear harmonic oscillator to effect the calculations makes the calculations unnecessarily involved when one attempts to obtain a simple analytic expression for these coefficients.

We attempt the problem *directly* by writing difference equations for the coefficients in question. A slight manipulation of these equations gives us a form which can be used to solve the problem using the technique of generating functions. Finally we can compute the explicit answer by expanding the generating function.

The scheme of this paper is as follows. In Sec. II we give the calculations for the zero momentum case. In Sec. III we do the same calculation for the momentum-dependent case wherein we arrive at the expression mentioned above. To make comparison with Tanabe's work easier, we shall be using his notations throughout.

II. EXPLICIT DETERMINATION OF THE COEFFICIENTS FOR THE ZERO-MOMENTUM BOSONS

In this case, the Bogolubov transformation is e^S with

$$S = -\tfrac{1}{2}x(a^\dagger a^\dagger - aa) = -S^\dagger, \tag{1}$$

where the creation and the annihilation operators a and $a^\dagger$ satisfy the commutation relation

$$[a, a^\dagger] = 1. \tag{2}$$

The operators b and $b^\dagger$ which are the transforms of the operators a and $a^\dagger$ are given by

$$\begin{aligned} b &= ua + va^\dagger = e^S a e^{-S}, \\ b^\dagger &= ua^\dagger + va = e^S a^\dagger e^{-S}, \end{aligned} \tag{3}$$

where $u = \cosh x$ and $v = \sinh x$. The inverse transformation e^{-S} gives the answers obtained from Eqs. (3) by replacing $v = \sinh x$ by $-v$.

We must compute

$$G_{k;l}(x) = \frac{1}{\sqrt{k!l!}}\,{}_a\langle 0 \,|\, a^k e(-\tfrac{1}{2}x(a^\dagger a^\dagger - aa))(a^\dagger)^l \,|\, 0\rangle_a \tag{4}$$

$$= \sqrt{k!l!}\,H_{k;l}(x). \tag{5}$$

Using Eqs. (2)–(5), we obtain

$$kH_{k;l}(x) = -(l+1)\sinh x\, H_{k-1;l+1}(x) + \cosh x\, H_{k-1;l-1}(x) \tag{6}$$

$$lH_{k;l}(x) = (k+1)\sinh x\, H_{k+1;l-1}(x) + \cosh x\, H_{k-1;l-1}(x). \tag{7}$$

The above recursion relations have the indices k and l both increasing and decreasing. In order to evolve a situation where k and l are both nonincreasing, we substitute, in Eq. (6), the expression obtained from Eq. (7) for $H_{k-1;l+1}(x)$. This gives

$$k\cosh x\, H_{k;l}(x) = -\sinh x\, H_{k-2;l}(x) + H_{k-1;l-1}(x). \tag{8}$$

Similarly

$$l\cosh x\, H_{k;l}(x) = \sinh x\, H_{k;l-2}(x) + H_{k-1;l-1}(x). \tag{9}$$

Now from the definition of $H_{k;l}(x)$ it is evident that $H_{k;l}(x) = 0$ unless $k \pm l =$ even, i.e., $H_{k;l}(x) \neq 0$ only if k and l are both even or both odd. Together with the above recursion relations this gives

$$H_{-n;l}(x) = H_{k;-n}(x) = 0,$$

where $n > 0$, as should also be evident physically. Thus our recursion relations should, *in principle*, give us all the $H_{k;l}(x)$ in terms of a *single* starting $H_{0;0}(x)$. Before we show how this is explicitly achieved, we should like to mention a special case. Putting $k = l = 1$ in Eqs. (8) or (9), we obtain

$$H_{1;1}(x) = (1/\cosh x)H_{0;0}(x). \tag{10}$$

By defining the generating function

$$H(a, b; x) = \sum_{k,l=0}^{\infty} H_{k;l}(x)a^k b^l, \tag{11}$$

Eq. (8) becomes

$$\cosh x \frac{\partial}{\partial a} H(a, b; x) = \left(-a\tanh x + \frac{b}{\cosh x}\right) H(a, b; x), \tag{12}$$

which has the solution

$$H(a, b; x) = C(b; x)\exp(-\tfrac{1}{2}a^2\tanh x + ab/\cosh x), \tag{13}$$

where

$$C(b; x) = H(0, b; x) = \sum_{l=0}^{\infty} H_{0;l}(x)b^l = \sum_{l=0}^{\infty} H_{0;2l}(x)b^{2l}, \tag{14}$$

using Eq. (11) and $H_{0;2l+1}(x) = 0$.

To obtain $H_{0;2l}(x)$ for $l \geq 0$, let us put $k=0$ in Eq. (9). This leads to

$$lH_{0;l}(x) = \tanh x\, H_{0;l-2}(x), \tag{15}$$

i.e.

$$H_{0;2l}(x) = (\tfrac{1}{2}\tanh x)^l \cdot (1/l!) H_{0;0}(x), \tag{16}$$

which gives

$$C(b;x) = H_{0;0}(x)\exp(\tfrac{1}{2}b^2\tanh x). \tag{17}$$

Thus Eq. (13) takes the form

$$H(a, b; x) = H_{0;0}(x)\exp[ab/\cosh x - \tfrac{1}{2}(a^2 - b^2)\tanh x]. \tag{18}$$

Finally we have to compute $H_{0;0}(x)$. For this purpose, we differentiate the definition of $H_{0;0}(x)$ with regard to x to arrive at

$$\frac{d}{dx}H_{0;0}(x) = {}_{2;0}(x). \tag{19}$$

But from Eq. (8)

$$H_{2;0}(x) = -\tfrac{1}{2}H_{0;0}(x)\tanh x. \tag{20}$$

The above two equations result in

$$H_{0;0}(x) = (\cosh x)^{-1/2}, \tag{21}$$

since $H_{0;0}(0) = 1$ by definition.

Substituting this value of $H_{0;0}(x)$ in Eq. (18) and expanding the generating function in powers of a and b, we arrive at[4]

$$\begin{aligned} G_{k;l}(x) &= \sqrt{k!l!}\, H_{k;l}(x) \\ &= (-1)^{k/2}\left(\frac{k!l!}{\cosh x}\right)\left(\frac{\tanh x}{2}\right)^{(k+l)/2} \\ &\quad \times \sum_{\lambda} \frac{(-4/\sinh^2 x)^{\lambda}}{(2\lambda)!\,(\frac{1}{2}k-\lambda)!\,(\frac{1}{2}l-\lambda)!}, \quad \text{for } k, l \text{ even}, \\ &= (-1)(k-1)/2\left(\frac{k!l!}{\cosh^3 x}\right)\left(\frac{\tanh x}{2}\right)^{(k+l)/2-1} \\ &\quad \times \sum_{\lambda} \frac{(-4/\sinh^2 x)^{\lambda}}{(2\lambda+1)!\,(\frac{1}{2}(k-1)-\lambda)!\,(\frac{1}{2}(l-1)-\lambda)!}, \\ &\qquad\qquad \text{for } k, l \text{ odd}, \\ &= 0 \quad \text{otherwise}. \end{aligned} \tag{22}$$

These results are exactly the same as obtained by Tanabe using complicated and lengthy mathematical techniques.

III. EXPLICIT DETERMINATION OF THE COEFFICIENTS FOR NONZERO MOMENTUM BOSONS

In this case, the Bogolubov transformation is e^T with

$$T = -x_{\mathbf{k}}(a_{\mathbf{k}}^{\dagger}a_{-\mathbf{k}}^{\dagger} - a_{\mathbf{k}}a_{-\mathbf{k}}) = -T^{\dagger}, \tag{23}$$

where the momentum $\mathbf{k} \neq 0$ and the operators $a_{\mathbf{k}}$, $a_{-\mathbf{k}}$, $a_{\mathbf{k}}^{\dagger}$, $a_{-\mathbf{k}}^{\dagger}$ satisfy the commutation relations

$$[a_{\mathbf{k}}, a_{\mathbf{k}'}^{\dagger}] = \delta_{\mathbf{k},\mathbf{k}'}. \tag{24}$$

The transforms $b_{\mathbf{k}}$ and $b_{\mathbf{k}}^{\dagger}$ of these operators are given by

$$b_{\mathbf{k}} = e^T a_{\mathbf{k}} e^{-T} = u_{\mathbf{k}}a_{\mathbf{k}} + v_{\mathbf{k}}a_{-\mathbf{k}}^{\dagger},$$

$$b_{\mathbf{k}}^{\dagger} = e^T a_{\mathbf{k}}^{\dagger} e^{-T} = u_{\mathbf{k}}a_{\mathbf{k}}^{\dagger} + v_{\mathbf{k}}a_{-\mathbf{k}}, \tag{25}$$

where

$$u_{\mathbf{k}} = u_{-\mathbf{k}} = \cosh x_{\mathbf{k}} \quad \text{and} \quad v_{\mathbf{k}} = v_{-\mathbf{k}} = \sinh x_{\mathbf{k}}. \tag{26}$$

We have to compute

$$\begin{aligned} G_{p,q;r,s}(x_{\mathbf{k}}) &= \frac{1}{\sqrt{p!q!r!s!}}\,{}_a\langle 0 \,|\, (a_{\mathbf{k}})^p (a_{-\mathbf{k}})^q e^T (a_{\mathbf{k}}^{\dagger})^r (a_{-\mathbf{k}}^{\dagger})^s \,|\, 0\rangle_a \\ &= \sqrt{p!q!r!s!}\, H_{p,q;r,s}(x_{\mathbf{k}}). \end{aligned} \tag{27}$$

Following the techniques used in the previous section, we first obtain *four* recursion relations satisfied by $H_{p,q;r,s}(x_{\mathbf{k}})$ which have indices both increasing and decreasing. Exactly the same manipulations as were used in the previous section lead to the following recursion relations with indices all nonincreasing:

$$\begin{aligned} p\cosh x_{\mathbf{k}} H_{p,q;r,s}(x_{\mathbf{k}}) &= -\sinh x_{\mathbf{k}} H_{p-1,q-1;r,s}(x_{\mathbf{k}}) + H_{p-1,q;r-1,s}(x_{\mathbf{k}}), \end{aligned} \tag{28}$$

$$\begin{aligned} q\cosh x_{\mathbf{k}} H_{p,q;r,s}(x_{\mathbf{k}}) &= -\sinh x_{\mathbf{k}} H_{p-1,q-1;r,s}(x_{\mathbf{k}}) + H_{p,q-1;r,s-1}(x_{\mathbf{k}}), \end{aligned} \tag{29}$$

$$\begin{aligned} r\cosh x_{\mathbf{k}} H_{p,q;r,s}(x_{\mathbf{k}}) &= \sinh x_{\mathbf{k}} H_{p,q;r-1,s-1}(x_{\mathbf{k}}) + H_{p-1,q;r-1,s}(x_{\mathbf{k}}), \end{aligned} \tag{30}$$

$$\begin{aligned} s\cosh x_{\mathbf{k}} H_{p,q;r,s}(x_{\mathbf{k}}) &= \sinh x_{\mathbf{k}} H_{p,q;r-1,s-1}(x_{\mathbf{k}}) + H_{p,q-1;r,s-1}(x_{\mathbf{k}}). \end{aligned} \tag{31}$$

From the definition of H's we immediately conclude

$$H_{p,q;r,s}(x_{\mathbf{k}}) = 0 \quad \text{unless } p - q = r - s. \tag{32}$$

[This result can also be seen as a consequence of momentum conservation. A *special* case of it is $H_{p,q;r,s}(x_{\mathbf{k}}) = 0$ if $p+q+r+s = \text{odd}$.] The above relations now give us the physical result that an H with any of the indices equal to a negative integer is equal to zero. Thus the above recursion relations *should* enable us to compute all the $H_{p,q;r,s}(x_{\mathbf{k}})$ in terms of a *single* starting one, i.e., $H_{0,0;0,0}(x_{\mathbf{k}})$. We give below this explicit calculation.

By defining the generating function

$$H(a, b, c, d; x_{\mathbf{k}}) = \sum_{p,q,r,s=0}^{\infty} H_{p,q;r,s}(x_{\mathbf{k}})a^p b^q c^r d^s, \tag{33}$$

Eq. (28) becomes

$$\frac{\partial}{\partial a}H(q, b, c, d; x_{\mathbf{k}}) = \left(-b\tanh x_{\mathbf{k}} + \frac{c}{\cosh x_{\mathbf{k}}}\right)H(a, b, c, d; x_{\mathbf{k}}), \tag{34}$$

which has the solution

$$H(a, b, c, d; x_{\mathbf{k}}) = C(b, c, d; x_{\mathbf{k}})\exp(-ab\tanh x_{\mathbf{k}} + ac/\cosh x_{\mathbf{k}}), \tag{35}$$

where

$$C(b, c, d; x_{\mathbf{k}}) = H(0, b, c, d; x_{\mathbf{k}}) = \sum_{\substack{q,r,s,=0 \\ s=q+r}}^{\infty} H_{0,q;r,s}(x_{\mathbf{k}})b^q c^r d^s. \tag{36}$$

To obtain $H_{0,q;r,q+r}(x_{\mathbf{k}})$, we put $p=0$ in Eq. (30), which results in

$$H_{0,q;r,q+r}(x_{\mathbf{k}}) = [\tanh x_{\mathbf{k}}/r]H_{0,q;r-1,q+r-1}(x_{\mathbf{k}})$$

or

$$H_{0,q;r,q+r}(x_{\mathbf{k}}) = [(\tanh x_{\mathbf{k}})^r/r!]H_{0,q;0,q}(x_{\mathbf{k}}). \tag{37}$$

Similarly from Eq. (31) putting $p=r=0$ and solving, we obtain

$$H_{0,q;0,q}(x_k)=(1/q!)(\cosh x_k)^q H_{0,0;0,0}(x_k). \tag{38}$$

On using Eqs. (37) and (38), we can now compute the $C(b, c, d; x_k)$ appearing in Eq. (36) as

$$C(b, c, d; x_k)=H_{0,0;0,0}(x_k)\exp(cd\tanh x_k+bd/\cosh x_k). \tag{39}$$

Thus we obtain the generating function in Eq. (35) in terms of the only unknown $H_{0,0;0,0}(x_k)$ as

$$H(a, b, c, d; x_k)=H_{0,0;0,0}(x_k)\times\exp[(ac+bd)/\cosh x_k-(ab-cd)\tanh x_k]. \tag{40}$$

To compute $H_{0,0;0,0}(x_k)$, we differentiate its definition. This gives

$$\frac{d}{dx_k}H_{0,0;0\ 0}(x_k)=H_{1,1;0,0}(x_k). \tag{41}$$

Also from Eq. (28), putting $p=q=1$ and $r=s=0$, we get

$$H_{1,1;0,0}(x_k)=-\tanh x_k H_{0,0;0,0}(x_k). \tag{42}$$

Thus

$$H_{0,0;0,0}(x_k)=(\cosh x_k)^{-1}, \tag{43}$$

since $H_{0,0;0,0}(0)=1$ from its definition. Our generating function is now finally obtained as

$$H(a, b, c, d; x_k)=[1/(\cosh x_k)]\times\exp[(ac+bd)/\cosh x_k-(ab-cd)\tanh x_k]. \tag{44}$$

On expanding this generating function in powers of a, b, c, d, we arrive at

$$G_{p,q;r,s}(x_k)=\sqrt{p!q!r!s!}\,H_{p,q;r,s}(x_k)$$
$$=(-1)^p\sqrt{p!q!r!s!}\,\frac{(\tanh x_k)^{p+r}}{(\cosh x_k)^{q-p+1}}\times\sum_\lambda\frac{(-1/\sinh^2 x_k)^\lambda}{\lambda!(p-\lambda)!(r-\lambda)!(q-p+\lambda)!},$$
when $p-q=r-s$, (45)
$$=0 \text{ otherwise.}$$

This extremely simple result may be compared with Tanabe's extremely complicated result containing four explicit and two implicit summations.[3]

ACKNOWLEDGMENTS

The author wishes to thank Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste.

*On leave of absence from Department of Mathematics, University of Islamabad, Islamabad, Pakistan.

[1]N. N. Bogolubov, Zh. Eksp. Teor. Fiz. **34**, 58 (1958) [Sov. Phys. —JETP 7, 41 (1958)].
[2]K. Tanabe, J. Math. Phys. **14**, 618 (1973).
[3]Eq. (4.4) of Ref. 2.
[4]Eq. (3.7) of Ref. 2.

On the equivalence of the Einstein–Mayer and Einstein–Cartan theories for describing a spinning medium

M. N. Mahanta

Department of Mathematics, Indian Institute of Technology, Delhi, New Delhi-110029, India

A. Sudhakar Rao

Department of Physics, Indian Institute of Technology, Delhi, New Delhi-110029, India

(Received 30 October 1973; revised manuscript received 20 December 1973)

It is shown that the five-dimensional formalism of the Einstein–Mayer theory of 1932 can be given a new interpretation and used to describe a continuous medium with intrinsic angular momentum. The equations of motion of a spinning particle are derived from the generalized field equations and are seen to be essentially the same as the Mathisson–Papapetrou equations. Finally the correspondence of the new approach with the theories based on Cartan's torsion tensor is established.

1. INTRODUCTION

In recent years two possible ways of generalizing the field equations of general relativity for the case of a continuous medium endowed with intrinsic angular momentum (spin) have been developed. The first having its roots in Cartan's original suggestion[1–4] has been worked out in detail by F.W. Hehl,[5–8] A. Trautman,[9–11] and R. L. Clerc.[12,13] Similar results using the tetrad formalism were earlier obtained by D.W. Sciama[14] and T.W.B. Kibble.[15] A second and more recent approach of the first author (M.N.M.)[16] uses a five-dimensional formalism which was originally developed by A. Einstein and W. Mayer[17] in 1932 to describe electromagnetism and gravitation in a unified manner. It was proved that the Einstein—Mayer theory can be reinterpreted to describe a spinning medium and the equations of motion of a spinning point particle derived from it are essentially the Mathisson—Papapetrou equations.[18,19]

The complete equivalence of the two approaches (with a slight modification in the second) was established by the second author (A.S.R.) and a combined account of the works of both is presented below.

2. APPROACH BASED ON EINSTEIN-MAYER FORMALISM

In this formalism[17] at each point of the V_4, side-by-side with the four-dimensional vector space formed from contravariant and covariant vectors there is also assumed a five-dimensional vector space V_5 with a metric tensor $g_{\mu\nu}$ (greek indices run from 1 to 5 and italic indices from 1 to 4). The connection between 5-vectors in V_5 and 4-vectors in V_4 is realized by a mixed tensor γ_σ^k. For example,

$$a^k = \gamma_\sigma^k a^\sigma, \tag{2.1}$$

$$b_\sigma = \gamma_\sigma^k b_k + (b_\tau A^\tau)A_\sigma, \tag{2.2}$$

the nonnull vector A^σ being defined by

$$\gamma_\sigma^k A^\sigma = 0 \tag{2.3}$$

and normalized by the condition

$$g_{\rho\sigma}A^\rho A^\sigma = 1. \tag{2.4}$$

Further we identify

$$g_{ik} = g_{\rho\sigma}\gamma_i^\rho \gamma_k^\sigma \tag{2.5}$$

as the metric tensor of V_4. Then it is easily shown that

$$\gamma_\rho^i \gamma_k^\rho = \delta_k^i, \tag{2.6}$$

$$\gamma_\sigma^\rho \gamma_\rho^\tau = \delta_\sigma^\tau - A_\sigma A^\tau. \tag{2.7}$$

From (2.1) and (2.7) we get

$$a^\tau = \gamma_k^\tau a^k + (A_\sigma a^\sigma)A^\tau. \tag{2.8}$$

We shall define covariant derivatives of vectors and tensors in V_5 with respect to x^q by analogy with the corresponding definition in V_4. For example,

$$a^\sigma_{;q} = \partial_q a^\sigma + \Gamma^\sigma_{q\pi} a^\pi \tag{2.9}$$

the mixed Γ symbols being functions of x^i, subject to the condition

$$g_{\sigma\rho;q} \equiv \partial_q g_{\sigma\rho} - g_{\sigma\pi}\Gamma^\pi_{q\rho} - g_{\rho\pi}\Gamma^\pi_{q\sigma} = 0.$$

The conditions for integrability of the conditions for parallel transport of the vector a^σ, viz.,

$$\delta a^\sigma = a^\sigma_{;q}\, dx^q = 0 \tag{2.10}$$

are given by

$$P_{q\rho\lambda}{}^\sigma a^\lambda = 0, \tag{2.11}$$

where $P_{pq\lambda}{}^\sigma$ is a Riemann—Christoffel-like tensor defined by

$$P_{q\rho\lambda}{}^\sigma \equiv \partial_q \Gamma^\sigma_{p\lambda} - \partial_p \Gamma^\sigma_{q\lambda} + \Gamma^\sigma_{q\pi}\Gamma^\pi_{p\lambda} - \Gamma^\sigma_{p\pi}\Gamma^\pi_{q\lambda}. \tag{2.12}$$

We may also prove the following identities:

$$\gamma_{\sigma k;q;p} - \gamma_{\sigma k;p;q} = \gamma_{\sigma l} R_{qpk}{}^l(\{\}) + \gamma_{\rho k} P_{qp\sigma}{}^\rho, \tag{2.13}$$

$$A_{\sigma;p;q} - A_{\sigma;q;p} = P_{pq\sigma}{}^\lambda A_\lambda, \tag{2.14}$$

$R_{qpk}{}^l(\{\})$ being the Riemann—Christoffel tensor for V_4 with Christoffel brackets $\{\}$.

In order to make the above formalism suitable for the description of a spinning medium we introduce the following hypotheses:

(1) $$\gamma_{\sigma k;q} = \gamma_\sigma^r V_{rkq}, \tag{2.15}$$

V_{rkq} being a covariant tensor of rank 3 in V_4;

(2) $$A^\sigma_{;q} = 0. \tag{2.16}$$

The assumption (2.15) together with

$$g_{lk;q}=0 \tag{2.17}$$

implies

$$V_{rkq}=-V_{krq}. \tag{2.18}$$

Equation (2.18) together with (2.16) ensures that

$$g^{\rho\sigma}_{;q}\equiv(A^\rho A^\sigma+\gamma^\rho_k\gamma^{\sigma k})_{;q}=0.$$

The identities (2.13) and (2.14) together with the hypotheses (2.15) and (2.16) lead to an asymmetric Einstein tensor G^{ij}, where

$$G^{ij}\equiv P^{ij}-\tfrac{1}{2}Pg^{ij}, \tag{2.19}$$

$$\begin{aligned}P^{ij}&\equiv P^{i\ \ \rho}_{r\ \ \sigma}\gamma^r_\rho\gamma^{\sigma j}\\&=R^{ij}+V^{jti}(g^{kl}V_{tkl})+V^{tjk}V_{tk}{}^{i}\\&\quad+(g^{it}g^{kl}V^{j}{}_{kl})_{;t}+V^{lji}_{;l},\end{aligned} \tag{2.20}$$

$$P=g_{ij}P^{ij}. \tag{2.21}$$

For physical application of the above formalism we consider an incoherent spinning fluid of the Weyssenhoff–Raabe type[20] with energy–momentum tensor given by p^iu^j, p^i being the 4-momentum density and u^j the 4-velocity of the fluid. The spin density of the medium is represented by the skew symmetric tensor s^{ij} whose components are subject to the restriction

$$s^{ij}u_j=0. \tag{2.22}$$

By analogy with coventional general relativity and the Weyssenhoff–Raabe theory of spinning fluid for the special relativistic case, we propose the following system of field equations:

$$G^{ij}=Kp^iu^j \tag{2.23}$$

and

$$V^{lji}-V^{lij}=Ks^{ij}u^l, \tag{2.24}$$

K being a universal constant. Writing (2.24) in the equivalent form, viz.,

$$V^{lij}=-\frac{K}{2}(u^ls^{ij}+u^is^{jl}-u^js^{li}). \tag{2.24'}$$

We may write (2.22) equivalently as

$$g^{lj}V_{lij}=g^{lj}V_{ilj}=0. \tag{2.22'}$$

In view of (2.22') the skew symmetric part of Eq. (2.23) yields

$$(s^{ij}u^k)_{;k}=p^iu^j-p^ju^i. \tag{2.25}$$

Also Eq. (2.23) itself simplifies to

$$\begin{aligned}G^{ij}&\equiv R^{ij}+(K^2/4)(s_{lm}s^{lm})u^iu^j+V^{lji}_{;l}\\&\quad-\tfrac{1}{2}g^{ij}[R+(K^2/4)(s_{lm}s^{lm})(u_ru^r)]\\&=Kp^iu^j.\end{aligned} \tag{2.26}$$

For $s^{lm}=0$, Eqs. (2.25) and (2.26) together reduce to the Einstein field equations for a spinless incoherent fluid, viz.,

$$G^{ij}\equiv R^{ij}-\tfrac{1}{2}g^{ij}R=K\rho_0u^iu^j, \tag{2.27}$$

where $p^i=\rho_0u^i$, ρ_0 being the invariant expressing the proportionality of p^i and u^i.

To derive the equations of motion of a spinning point particle we first get, taking the divergence of (2.26) w.r.t. j,

$$\begin{aligned}&K(p^iu^j)_{;j}-(K^2/4)\{(s_{lm}s^{lm})u^iu^j\}_{;j}-V^{lji}_{;l;j}\\&\quad+(K^2/8)g^{ij}\{(s_{lm}s^{lm})(-c^2)\}_{;j}=0.\end{aligned} \tag{2.28}$$

Integration of Eqs. (2.22), (2.25), and (2.28) over the small proper volume V_0 of the particle leads to the following system of equations of motion for the spinning particle:

$$S^{ij}u_j=0, \tag{2.29}$$

$$\dot S^{ij}=P^iu^j-P^ju^i, \tag{2.30}$$

$$\dot P^i-(KS_0/4)\dot u^i+\tfrac{1}{2}u^lS^{mn}R_{lmn}{}^{i}=0, \tag{2.31}$$

where the dot represents differentiation w.r.t the proper time element $d\tau$ of the particle and

$$S^{ij}=\int_{V_0}s^{ij}\,dV_0=\text{total spin of the particle},$$

$$P^i=\int_{V_0}p^i\,dV_0=\text{total momentum of the particle},$$

$$S_0=\int_{V_0}s_{lm}s^{lm}\,dV_0,$$

which can be shown to be constant over the world line of the particle. In deriving (2.31) we have made use of the relation (which is easily proved)

$$V^{lji}_{;l;j}=-\frac{K}{2}u^ls^{mn}R_{lmn}{}^{i}. \tag{2.32}$$

Equations (2.30) and (2.31) essentially represent the Mathisson–Papapetrou equations for a spinning point particle.

3. EQUIVALENCE WITH THE EINSTEIN-CARTAN THEORY

We shall now establish the correspondence of the above theory with that of the Einstein–Cartan theory as developed in F.W. Hehl's papers.[5–8]

From the connection (used by Hehl)

$$\begin{aligned}\Gamma^k_{ij}&=\left\{{k\atop ij}\right\}+S_{ij}{}^{k}-S_j{}^{k}{}_{i}+S^k{}_{ij}\\&=\left\{{k\atop ij}\right\}-K_{ij}{}^{k},\end{aligned} \tag{3.1}$$

where

$$S_{ij}{}^{k}=\tfrac{1}{2}(\Gamma^k_{ij}-\Gamma^k_{ji}) \tag{3.2}$$

is the Cartan's torsion tensor and $K_{ij}{}^{k}$ is the contortion tensor, we may form the Ricci tensor $R^{ij}(\Gamma)$ given by

$$R^{ij}(\Gamma)=R^{ij}(\{\})-K^{ijl}_{;l}-(g^{il}K_r{}^{rj})_{;l}-K_i{}^{l}{}_mK^{ijm}-K^i{}_m{}^iK_l{}^{jm}. \tag{3.3}$$

A comparison of (3.3) with (2.20) shows that if we put

$$V^{lji}=-K^{ijl}, \tag{3.4}$$

Then the two expressions are identical and hence the Einstein tensors G^{ij} in the two cases become identical. Now (3.4) leads to

$$V^{lji}-V^{lij}=2S^{ijl}. \tag{3.5}$$

In the preceding section we set $V^{lji}-V^{lij}$ equal to $Ks^{ij}u^l$ and made the further assumption (2.22). Now we deviate from this line and identify

$$T^{ijl}=\frac{K}{2}s^{ij}u^l, \tag{3.6}$$

where

$$T^{ijk} = S^{ijk} + g^{ik}S^{j}{}_{l}{}^{l} - g^{jk}S^{i}{}_{l}{}^{l}. \tag{3.7}$$

From (3.5) and (3.6) we get

$$(V^{lji} - V^{lij}) + g^{il}V^{j}{}_{k}{}^{k} - g^{jl}V^{i}{}_{k}{}^{k} = Ks^{ij}u^{l}. \tag{3.8}$$

A further contraction of suffixes leads to

$$V^{i}{}_{k}{}^{k} = 2S^{i}{}_{k}{}^{k} = -\frac{K}{2}s^{ij}u_{j}. \tag{3.9}$$

Now (2.22) implies $S^{i}{}_{k}{}^{k} = 0$, in which case T^{ijk} and S^{ijk} become identical and we get

$$T^{ijl} = S^{ijl} = \frac{K}{2}s^{ij}u^{l} = \tfrac{1}{2}(V^{lji} - V^{lij})$$

which is the same as (2.24).

It only remains to replace $\frac{1}{2}s^{ij}u^{k}$ with the general spin tensor τ^{ijk} and the asymmetric e—m tensor $p^{i}u^{j}$ with $\sum^{ij}$, appearing in Hehl's works. Hence by making the identifications (3.4) and (3.6) and not assuming (2.22), we can have complete identity of all relations in the two approaches after converting the operators ∇_k, $\overset{*}{\nabla}_k$ and $\overset{+}{\nabla}_k$ introduced by Hehl, into covariant derivatives for Christoffel symbols.

A deeper understanding of the correspondence between the two approaches can be gained from the following considerations:

Let a^{σ} be a 5-vector in V_5 and a^{k} the corresponding 4-vector in V_4 where

$$a^{k} = \gamma_{\sigma}^{k}a^{\sigma}.$$

Now we may consider two covariant derivatives: (1) First we transform a^{σ} to a^{k} and then take the covariant derivative. This is nothing but the covariant derivative of a^{k} w.r.t. to Christoffel bracket $\{{}^{k}_{ij}\}$. (2) We may first take the covariant derivative of a^{σ} as defined by (2.9) and then transform this tensor to V_4 with help of γ_{σ}^{k}.

The second procedure is very interesting and we may ask: What meaning is to be attached to the tensor so obtained? It is very tempting to consider it also to be a covariant derivative of a^{k}. But if it is taken to be the covariant derivative of a^{k} w.r.t. $\{{}^{k}_{ij}\}$, then it means that the procedures (1) and (2) lead to the same tensor and this would imply that the covariant derivative of γ_{σ}^{k} is zero. Because

$$\begin{aligned}\gamma_{\sigma}^{k}(a^{\sigma}_{;q}) &= (\gamma_{\sigma}^{k}a^{\sigma})_{;q} - a^{\sigma}\gamma^{k}_{\sigma;q} \\ &= a^{k}_{;q} - a^{\sigma}\gamma^{k}_{\sigma;q} \overset{\text{def}}{=} a^{k}_{;q}\end{aligned}$$

for all a^{σ}. Hence

$$\gamma^{k}_{\sigma;q} = 0. \tag{3.10}$$

This will not lead to any new results at all and we might as well work throughout with V_4.

But we may consider the second procedure to lead to a covariant derivative of a^{k} but w.r.t. a connection different from $\{{}^{k}_{ij}\}$ which we denote by Γ^{k}_{ij}, i.e.,

$$\gamma_{\sigma}^{k}(a^{\sigma}_{;q}) = \nabla_{q}a^{k} = \partial_{q}a^{k} + \Gamma^{k}_{ql}a^{l}. \tag{3.11}$$

Then it can be shown that

$$\Gamma^{k}_{ql} = \{{}^{k}_{ql}\} - \gamma_{l}^{\sigma}\gamma^{k}_{\sigma;q}; \tag{3.12}$$

writing out $\gamma^{k}_{\sigma;q}$ explicitly in (3.12), we get

$$\Gamma^{k}_{ql} = \gamma_{\sigma}^{k}\partial_{q}\gamma_{l}^{\sigma} + \gamma_{l}^{\sigma}\gamma_{\tau}^{k}\Gamma^{\tau}_{q\sigma}. \tag{3.13}$$

Using this expression for Γ^{k}_{ql} we may compute the corresponding Riemann—Christoffel tensor $R_{ijk}{}^{l}(\Gamma)$ and after performing the calculations [using (2.3), (2.7), and (2.16)] we may show that

$$R_{ijk}{}^{l}(\Gamma) = P_{ijk}{}^{l}, \tag{3.14}$$

where

$$P_{ijk}{}^{l} = P_{ij\sigma}{}^{\tau}\gamma_{\tau}^{l}\gamma_{k}^{\sigma}, \tag{3.15}$$

$P_{ij\sigma}{}^{\tau}$ being defined by (2.12). Equation (3.14) shows why the Einstein tensor defined by (2.19) and $G^{ij}(\Gamma)$ were identical. It should be noted the explicit expressions for the Ricci tensors in the two cases involving $K_{ij}{}^{k}$ and $V_{ij}{}^{k}$ have not been used.

We may denote

$$\gamma_{l}^{\sigma}\gamma^{k}_{\sigma;q} = -V^{k}{}_{lq}$$

which immediately leads to the results (2.15), (2.18), and (3.4).

ACKNOWLEDGMENT

The authors express their gratitude to Professor O. Costa de Beauregard and Dr. R. L. Clerc for their kind interest in the work and for having made a number of critical remarks.

[1]E. Cartan, C.R. Acad. Sc. (Paris) **174**, 593 (1922).
[2]E. Cartan, Ann-Ec. Norm. Sup. **40**, 325 (1923).
[3]E. Cartan, Ann-Ec. Norm. Sup. **41**, 1 (1924).
[4]E. Cartan, Ann-Ec. Norm. Sup. **42**, 17 (1925).
[5]F.W. Hehl, "Habilitationsschrift," Technical University of Clausthal (August 1970).
[6]F.W. Hehl, Phys. Lett. **36A**, 225 (1971).
[7]F.W. Hehl, "Spin and torsion in GR I" (to be published); "Spin and torsion in GR II" (unpublished).
[8]F.W. Hehl and Paul Von Der Hyde, "Spin and the structure of space—time" (to be published).
[9]A. Trautman, Bull. Acad. Polon. Sci. **XX**, No. 2 (1972).
[10]A. Trautman, Bull. Acad. Polon. Sci. **XX**, No. 6 (1972).
[11]A. Trautman, Bull. Acad. Polon. Sci. **XX**, No. 10 (1972).
[12]R. L. Clerc, Thesis, Toulouse (1972).
[13]R. L. Clerc, Ann. Inst. Henri Poincaré **XVII**, No. 3, 227, 259 (1972).
[14]D.W. Sciama, *Recent Developments in General Relativity* (Pergamon, London, 1962), p. 415.
[15]T.W.B. Kibble, J. Math. Phys. **2**, 212 (1961).
[16]M.N. Mahanta, "On a possible generalization of the field equations of general relativity for the case of a spinning medium," seminar at Institut Henri Poincaré (Paris) (1972).
[17]A. Einstein and W. Mayer, Sitzungsber Preuss. Akad. Wiss. Phys. Math. Kl. 541 (1931); 130 (1932).
[18]M. Mathisson, Acta Phys. Pol. **6**. 218 (1937).
[19]A. Papapetrou, Proc. Roy. Soc. A **209**, 248 (1951).
[20]J. Weyssenhoff and A. Raabe, Acta Phys. Pol. **9**, 7 (1947).

Exact propagator for a time-dependent harmonic oscillator with and without a singular perturbation

D. C. Khandekar and S. V. Lawande

Reactor Engineering Division, Bhabha Atomic Research Centre, Trombay, Bombay 400 085, India
(Received 21 November 1973)

By using Feynman's definition of a path integral, exact propagators for a time-dependent harmonic oscillator with and without an inverse quadratic potential have been evaluated. It is shown that these propagators depend only on the solutions of the classical unperturbed oscillator. The relations between these propagators, the invariants, and the Schrödinger equation are also discussed.

1. INTRODUCTION

The path integral formulation of Feynman[1] provides an alternative approach towards solving dynamical problems in quantum mechanics. In this approach the usual Schrödinger equation is replaced by the integral equation.

$$\psi(x'',t'')=\int K(x'',t'';x',t')\psi(x',t')\,dx' \quad (t''>t') \tag{1}$$

with the initial condition $\psi(x'',t')=\psi(x',t')$; the propagator or the kernel K is defined by a path integral

$$K(x'',t'';x',t')=\int \exp[(i/\hbar)S(x'',t'';x',t')]Dx(t). \tag{2}$$

Here, the integrations are over all possible paths starting at $x'=x(t')$ and terminating at $x''=x(t'')$. The function $S(x'',t'';x',t')$ in the integrand is the classical action

$$S(x'',t'';x',t')=\int_{t'}^{t''} L(\dot{x},x,t)\,dt, \tag{3}$$

$L(\dot{x},x,t)$ being the Lagrangian of the system considered.

Despite its intuitive appeal the applicability of this approach has been limited because of analytical difficulties and explicit expressions for path integrals are available only for a few cases.[2-5] Furthermore, systems for which the Lagrangian has explicit time dependence have received very little attention. In this context, an exact path integral solution is hitherto available for the lone case of a forced harmonic oscillator.[2,6]

In this paper we have obtained exact propagators for two systems where the Lagrangian is explicitly time-dependent. These are: (i) a time-dependent harmonic oscillator[7-10] in one dimension (harmonic oscillator with a frequency which in a function of time) and (ii) a time-dependent harmonic oscillator perturbed by a constant inversely quadratic potential.[11] We find that propagators for both these cases depend only on the solutions of the classical unperturbed oscillator. Further, it is found that the propagators admit expansions in terms of certain solutions of the corresponding time-dependent Schrödinger equation. These solutions of the Schrödinger equation (apart from time-dependent phase factors) are also the eigenfunctions of an invariant operator associated with the problem. In Sec. 2 we give the formulation and evaluation of the propagators, while in Sec. 3, we discuss the relations between the Schrödinger equation, invariants, and the propagators. Brief derivations of some mathematical results necessary for Sec 2 are given in an appendix.

2. FORMULATION AND EVALUATION OF THE PROPAGATOR

The path integral in Eq. (2) is usually defined by

$$K(x'',t'';x',t')=\lim_{N\to\infty}A_N\int\cdots\int \exp\left(\frac{i}{\hbar}\sum_{j=1}^{N}S_j(x_j,x_{j-1})\right)\prod_{j=1}^{N-1}dx_j, \tag{4}$$

where $x_j=x(t_j)$, $x_0=x'$, $x_N=x''$, $t_j-t_{j-1}=(t''-t')/N=\epsilon$, and A_N is the normalization factor in the N-th approximation. The action $S_j(x_j,x_{j-1})$ over an infinitesimally small time interval $t_j-t_{j-1}=\epsilon$ may be approximated by

$$S_j(x_j,x_{j-1})=\epsilon L(x_j,x_{j-1},\epsilon). \tag{5}$$

Time-dependent harmonic oscillator

This is characterized by a Lagrangian

$$L(\dot{x},x,t)=\tfrac{1}{2}m\dot{x}^2-\tfrac{1}{2}m\omega^2(t)x^2, \tag{6}$$

where $\omega(t)$ is assumed to be a regular function of t. From Eq. (5) we have

$$S_j(x_j,x_{j-1})=(m/2\epsilon)(x_j-x_{j-1})^2-(\epsilon/2)m\omega_j^2x_j^2, \tag{7}$$

with $\omega_j=\omega(t_j)$ and we may write

$$K(x'',t'';x',t')=\lim_{N\to\infty}A_N\int\prod_{j=1}^{N}\exp\left[\frac{im}{2\hbar}\left(\frac{(x_j-x_{j-1})^2}{\epsilon}-\epsilon\omega_j^2x_j^2\right)\right]\prod_{j=1}^{N-1}dx_j. \tag{8}$$

Since K has to be unitary A_N is given by

$$A_N=(m/2\pi i\hbar\epsilon)^{N/2}. \tag{9}$$

Equation (8) may be cast in the following form:

$$\begin{aligned}K(x'',t'';x',t')&=\lim_{N\to\infty}\left(\frac{-i\beta}{2\pi}\right)^{N/2}\exp\left(\frac{i\beta}{2}(x'^2+x''^2)\right)\\&\times\int\exp[i(\alpha_1x_1^2+\alpha_2x_2^2+\cdots+\alpha_{N-1}x_{N-1}^2)]\\&\times\exp(-i\beta x_0x_1)\cdots\exp(-i\beta x_{N-1}x_N)\prod_{j=1}^{N-1}dx_j,\end{aligned} \tag{10}$$

where

$$\beta=m/\hbar\epsilon,\quad \alpha_j=\beta(1-\tfrac{1}{2}\omega_j^2\epsilon^2),\quad j=1,2,\ldots,N-1. \tag{11}$$

After performing the integrations by repeated use of the formula

$$\begin{aligned}\int_{-\infty}^{\infty}\exp(i\alpha x^2)\exp(-iax)\exp(-ibx)\,dx&=\sqrt{i\pi/\alpha}\exp[-i(a^2+b^2)/4\alpha]\exp(-iab/2\alpha),\end{aligned} \tag{12}$$

we obtain

$$K(x'',t'';x',t')$$

$$= \lim_{N\to\infty} \sqrt{a_N/2\pi}\exp(ip_N x'^2 + iq_N x''^2)\exp(-ia_N x'' x'), \quad (13)$$

where

$$a_N = \beta \prod_{j=1}^{N-1} \frac{\beta}{2\gamma_j}, \quad (14)$$

$$p_N = \frac{\beta}{2} - \sum_{j=1}^{N-1} \frac{\beta_j^2}{4\gamma_j}, \quad (15)$$

$$q_N = \frac{\beta}{2} - \frac{\beta^2}{4\gamma_{N-1}}, \quad (16)$$

$$\gamma_1 = \alpha_1, \quad \gamma_j = \alpha_j - \beta^2/4\gamma_{j-1}, \quad (17)$$

$$\beta_1 = \beta, \quad \beta_j = \beta \prod_{k=1}^{j-1} \frac{\beta}{2\gamma_k}. \quad (18)$$

It is shown in the Appendix that the limiting values of a_N, p_N, and q_N are given by

$$\lim_{N\to\infty} a_N = \frac{m}{\hbar}\left(\frac{\dot\gamma(t'')\dot\gamma(t')}{\sin\phi(t'',t')}\right)^{1/2}, \quad (19)$$

$$\lim_{N\to\infty} p_N = \frac{m}{2\hbar}\left(\frac{-\dot s(t')}{s(t')} + \dot\gamma(t')\cot\phi(t'',t')\right), \quad (20)$$

$$\lim_{N\to\infty} q_N = \frac{m}{2\hbar}\left(\frac{\dot s(t'')}{s(t'')} + \dot\gamma(t'')\cot\phi(t'',t')\right), \quad (21)$$

where the dots over s and γ denote differentations with respect to time,

$$\phi(t'',t') = \gamma(t'') - \gamma(t'), \quad (22)$$

and $s(t)$ and $\gamma(t)$ obey the differential equations

$$\ddot s - C^2 s^{-3} + \omega^2(t)s = 0, \quad (23)$$

$$\dot\gamma s^2 = C, \quad (24)$$

where C is an orbitrary constant. In fact, $s(t)$ and $\gamma(t)$ represent, respectively, the amplitude and phase of a classical time-dependent oscillator [with real $\omega(t)$]:

$$\ddot\eta(t) + \omega^2(t)\eta(t) = 0, \quad (25)$$

$$\eta(t) = s(t)\exp[i\gamma(t)]. \quad (26)$$

The propagator K thus takes the form

$$K(x'',t'';x',t') = \left(\frac{m\sqrt{\dot\gamma''\dot\gamma'}}{2\pi i\hbar\sin\phi(t'',t')}\right)^{1/2}\exp\left[\frac{im}{2\hbar}\left(\frac{\dot s''}{s''}x''^2 - \frac{\dot s'}{s'}x'^2\right)\right] \times\exp\left(\frac{im}{2\hbar\sin\phi(t'',t')}\left[(\dot\gamma'' x''^2 + \dot\gamma' x'^2)\cos\phi(t'',t') - 2\sqrt{\dot\gamma''\dot\gamma'}x''x'\right]\right), \quad (27)$$

where the prime and the double-prime denote the quantities evaluated at time t' and t'', respectively. It may be easily verified that for the case when $\omega(t)$ is a real positive constant, we have $s(t) = \sqrt{m/\omega}$, $\gamma(t) = \omega t$, and the propagator of Eq. (27) reduces to the usual expression for a harmonic oscillator with constant frequency.[2]

Harmonic oscillator with a inverse quadratic potential

We now consider a particle of mass m moving in one dimension under an external quadratic time-dependent potential and a constant inversely quadratic potential. The Lagrangian is given by

$$L(\dot x, x, t) = (m/2)\dot x^2 - (m/2)\omega^2(t)x^2 - (g/x^2), \quad (28)$$

where $\omega(t)$ is a regular function of time t and $g > -\hbar^2/8m$ to avoid "the fall to the center."[12] We note that the singular nature of the perturbation term prevents any transition between the states for $x > 0$ and $x < 0$. We may therefore restrict ourselves to the region $x \geq 0$; the solution in this region can be extended to the $x < 0$ region without any further conditions on continuity.

According to Eq. (5), we may write

$$S_j(x_j, x_{j-1}) = \frac{m}{2\epsilon}(x_j - x_{j-1})^2 - \frac{\epsilon}{2}m\omega_j^2 x_j^2 - \frac{\epsilon g}{x_j x_{j-1}} = \frac{m}{2}[(x_j^2 + x_{j-1}^2)/\epsilon - \epsilon\omega_j^2 x_j^2] - \left(\frac{m x_j x_{j-1}}{\epsilon} + \frac{\hbar^2(a^2 - \frac{1}{4})\epsilon}{2m x_j x_{j-1}}\right), \quad (29)$$

where

$$a = \tfrac{1}{2}(1 + 8mg/\hbar^2)^{1/2}. \quad (30)$$

Thus the integrand of Eq. (4) becomes

$$\exp\left(\frac{i}{\hbar}\sum_{j=1}^{N} S_j(x_j, x_{j-1})\right) = \prod_{j=1}^{N}\exp\left[\frac{im}{2\hbar}\left(\frac{x_j^2 + x_{j-1}^2}{\epsilon} - \epsilon\omega_j^2 x_j^2\right)\right] \times\exp\left(\frac{m x_j x_{j-1}}{i\hbar\epsilon} - \frac{i\hbar(a^2 - \frac{1}{4})\epsilon}{2m x_j x_{j-1}}\right). \quad (31)$$

Noting that the asymptotic form of $I_a(u/\epsilon)$, the modified Bessel function, for small ϵ, is given by

$$I_a\left(\frac{u}{\epsilon}\right) \approx \left(\frac{\epsilon}{2\pi u}\right)^{1/2}\exp\left(\frac{u}{\epsilon} - \tfrac{1}{2}(a^2 - \tfrac{1}{4})\frac{\epsilon}{u} + O(\epsilon^2)\right), \quad (32)$$

we may replace the last exponential in Eq. (31) by

$$\left(\frac{2\pi m x_j x_{j-1}}{i\hbar\epsilon}\right)^{1/2} I_a\left(\frac{m x_j x_{j-1}}{i\hbar\epsilon}\right). \quad (33)$$

Hence the path integral becomes

$$K(x'',t'';x',t') = \lim_{N\to\infty} A_N \int \prod_{j=1}^{N} R(x_j, x_{j-1}) \prod_{j=1}^{N} dx_j, \quad (34)$$

where

$$R(x_j, x_{j-1}) = \left(\frac{2\pi m x_j x_{j-1}}{i\hbar\epsilon}\right)^{1/2}\exp\left[\frac{im}{2\hbar}\left(\frac{(x_j^2 + x_{j-1}^2)}{\epsilon} - \epsilon\omega_j^2 x_j^2\right)\right] \times I_a\left(\frac{m x_j x_{j-1}}{i\hbar\epsilon}\right), \quad (35)$$

and A_N is given by Eq. (9).

Steps involved in the evaluation propagator of Eq. (34) are essentially the same as before. We first write Eq. (34) as

$$K(x'',t'';x',t') = \sqrt{x''x'}\lim_{N\to\infty}(-i\beta)^N\exp\left(\frac{i\beta}{2}(x''^2 + x'^2)\right) \times\int\exp[i(\alpha_1 x_1^2 + \alpha_2 x_2^2 + \cdots + \alpha_{N-1}x_{N-1}^2)] \times I_a(-i\beta x_0 x_1)\cdots I_a(-i\beta x_{N-1}x_N)\prod_{j=1}^{N-1} x_j\, dx_j, \quad (36)$$

where α_j and β are defined in Eq. (11). Next, we carry out the integrations in Eq. (36) by repeatedly using the formula[3]

$$\int_0^\infty \exp(i\alpha x^2) I_\nu(-ibx) I_\nu(-icx) x\,dx = \left(\frac{i}{2\alpha}\right) \exp\left(\frac{-i(b^2+c^2)}{4\alpha}\right) I_\nu\left(\frac{-ibc}{2\alpha}\right), \quad (\mathrm{Re}\alpha > 0,\ \mathrm{Re}\nu > -1) \tag{37}$$

and finally obtain

$$K(x'',t'';x',t') = -i\sqrt{x''x'} \lim_{N\to\infty} a_N \exp(ip_N x'^2 + iq_N x''^2) I_a(-ia_N x'x''), \tag{38}$$

where a_N, p_N, q_N are determined from Eqs. (14)—(24). Thus, the propagator K may be written explicitly as

$$K(x'',t'';x',t') = \frac{m\sqrt{x''\dot\gamma''x'\dot\gamma'}}{i\hbar \sin\phi(t'',t')} \exp\left[\frac{im}{2\hbar}\left(\frac{\dot s''}{s''}x''^2 - \frac{\dot s'}{s'}x'^2\right)\right] \times \exp\left(\frac{im}{2\hbar}(\dot\gamma'' x''^2 + \dot\gamma' x'^2)\cot\phi(t'',t')\right) \times I_a\left(\frac{m}{i\hbar}\frac{\sqrt{\dot\gamma''\dot\gamma'}x''x'}{\sin\phi(t'',t')}\right), \tag{39}$$

where ϕ, s, and γ are as defined by Eqs. (22)—(24).

3. SCHRÖDINGER EQUATION, INVARIANTS, AND PROPAGATOR

We now consider the expansions of the propagators of Eqs. (27) and (39) which we denote by K^0 and K^p, respectively. We first consider K^0, the propagator corresponding to the unperturbed oscillator, which may be rewritten as

$$K^0(x'',t'';x',t') = \left(\frac{m}{\pi\hbar}\right)^{1/2} (\dot\gamma''\dot\gamma')^{1/4} \frac{\exp(-i\phi/2)}{[1-\exp(-2i\phi)]^{1/2}} \times \exp\left[\frac{im}{2\hbar}\left(\frac{\dot s''}{s''}x''^2 - \frac{\dot s'}{s'}x'^2\right)\right] \exp\left(\frac{m}{2\hbar}(\dot\gamma'' x''^2 + \dot\gamma' x'^2)\right) \times \exp\left(\frac{-m}{\hbar[1-\exp(-2i\phi)]}[\dot\gamma'' x''^2 + \dot\gamma' x'^2 - 2\sqrt{\dot\gamma'\dot\gamma''}x''x' \times \exp(-i\phi)]\right). \tag{40}$$

If we now use the Mehler's formula[13]

$$\frac{\exp[-(x^2+y^2-2xyz)/(1-z^2)]}{\sqrt{1-z^2}} = \exp[-(x^2+y^2)] \sum_{n=0}^\infty \frac{z^n}{2^n n!} H_n(x) H_n(y), \tag{41}$$

and let $z = \exp(-i\phi)$ $[\phi = \gamma(t'') - \gamma(t')]$, $x = \sqrt{m\dot\gamma''/\hbar}\,x''$, $y = \sqrt{m\dot\gamma'/\hbar}\,x'$, we obtain

$$K^0(x'',t'';x',t') = \sum_{n=0}^\infty \psi_n^{0*}(x',t')\psi_n^0(x'',t''), \tag{42}$$

where

$$\psi_n^0(x,t) = \left(\frac{1}{2^n n!}\left(\frac{m\dot\gamma}{\pi\hbar}\right)^{1/2}\right)^{1/2} \exp[-i(n+\tfrac12)\gamma(t)] \times \exp\left[\frac{im}{2\hbar}\left(\frac{\dot s}{s} + i\dot\gamma\right)x^2\right] H_n(\sqrt{m\dot\gamma/\hbar}\,x). \tag{43}$$

Using Eqs. (23)—(24) defining s and γ, it is easily seen that $\psi_n^0(x,t)$ are normalized solutions of the time-dependent Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = H^0(t)\psi, \tag{44}$$

where

$$H^0 \equiv -\frac{\hbar^2}{2m}\frac{\partial^2}{2x^2} + \frac{m}{2}\omega^2(t)x^2. \tag{45}$$

Moreover, we may write

$$\psi_n^0(x,t) = \exp[i\alpha_n^0(t)]\phi_n^0(x,t), \tag{46}$$

where

$$\alpha_n^0(t) = -(n+\tfrac12)\gamma(t), \tag{47}$$

$$\phi_n^0(x,t) = \left(\frac{1}{2^n n!}\left(\frac{m\dot\gamma}{\hbar}\right)^{1/2}\right)^{1/2} \exp\left[\frac{im}{2\hbar}\left(\frac{\dot s}{s} + i\dot\gamma\right)x^2\right] H_N\left(\left(\frac{m\dot\gamma}{\hbar}\right)^{1/2}x\right) \tag{48}$$

$\phi_n^0(x,t)$ are then found to be the normalized eigenfunctions of a hermitian invariant operator[14] associated with the time-dependent harmonic oscillator. According to Ref. 8 this invariant operator is given by

$$I^0 = \frac{1}{2m}[m^2 s^{-2}x^2 + (sp - m\dot s x)^2], \tag{49}$$

where $p = -i\hbar\partial/\partial x$ and $s(t)$ is a solution of the equation

$$\ddot s + \omega^2(t)s - s^{-3} = 0. \tag{50}$$

It may be pointed out here that Eq. (50) is equivalent to Eqs. (23)—(24) because the arbitrary constant C occurring in the latter equations may be obsorbed in s, (by letting $s/\sqrt{C} \to s$ so that $\dot\gamma \to 1/s^2$).

Thus, $\phi_n^0(x,t)$ are found to obey the equation.

$$I^0\phi_n^0(x,t) = \hbar(n+\tfrac12)\phi_n^0(x,t), \tag{51}$$

while the phases $\alpha_n^0(t)$ satisfy the equation

$$\hbar\frac{d\alpha_n^0(t)}{dt} = \left\langle \phi_n^0 \middle| i\hbar\frac{\partial}{\partial t} - H^0 \middle| \phi_n^0 \right\rangle, \tag{52}$$

as is required by the discussion of Ref. 9.

Next, consider the case of the perturbed oscillator. Here the propagator K^p may be rewritten as

$$K^p(x'',t'';x',t') = \frac{2m\sqrt{x''\dot\gamma''x'\dot\gamma'}}{\hbar} \frac{\exp(-i\phi/2)}{[1-\exp(-2i\phi)]} \exp\left[\frac{im}{2\hbar}\left(\frac{\dot s''}{s''}x''^2 - \frac{\dot s'}{s'}x'^2\right)\right] \times \exp\left(\frac{-m}{\hbar}(\dot\gamma'' x''^2 + \dot\gamma' x'^2)\frac{\exp(-2i\phi)}{[1-\exp(-2i\phi)]}\right) \times I_a\left(\frac{2m\sqrt{\dot\gamma''\dot\gamma'}\exp(-i\phi)x'x''}{\hbar[1-\exp(-2i\phi)]}\right). \tag{53}$$

We now use the Hille-Hardy formula[15]

$$\frac{1}{(1-z)}\exp\left(-\frac{(x+y)z}{(1-z)}\right) I_a\left(\frac{2\sqrt{xyz}}{(1-z)}\right) = \sum_{n=0}^\infty \frac{n!}{\Gamma(n+a+1)} z^n (xyz)^{a/2} L_n^a(x) L_n^a(y) \quad (|z|<1) \tag{54}$$

and set $z = \exp(-2i\phi)$ $[\phi = \gamma(t'') - \gamma(t')]$, $x = m\dot\gamma'' x''^2/\hbar$, $y = m\dot\gamma' x'^2/\hbar$ to obtain

$$K^p(x'',t'';x',t') = \sum_{n=0}^\infty \psi_n^{p*}(x',t')\psi_n^p(x'',t''), \tag{55}$$

where

$$\psi_n^p(x,t) = \left(\frac{2(n!)}{\Gamma(n+a+1)}\right)^{1/2}\left(\frac{m\dot\gamma}{\hbar}\right)^{(a+1)/2} x^{a+1/2}\exp[-i(2n+a+1)\gamma(t)] \times\exp\left[\frac{im}{2\hbar}\left(\frac{\dot s}{s}+i\dot\gamma\right)x^2\right]L_n^a\left(\frac{m\dot\gamma}{\hbar}x^2\right). \tag{56}$$

As before, $\psi_n^p(x,t)$ are found to be normalized solutions of the time-dependent Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = H^p\psi, \tag{57}$$

with

$$H^p \equiv -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{m}{2}\omega^2(t)x^2 + \frac{g}{x^2}. \tag{58}$$

The invariant operator I^p corresponding to this problem is given by

$$I^p = \frac{1}{2m}\left(m^2s^{-2}x^2 + (sp - m\dot s x)^2 + \frac{2gms^2}{x^2}\right), \tag{59}$$

where $p = -i\hbar\partial/\partial x$ and $s(t)$ obeys Eq. (50). Nothing that $2gm = \hbar^2(a^2 - \frac{1}{4})$ [according to Eq. (30)] and writing

$$\psi_n^p(x,t) = \exp[i\alpha_n^p(t)]\phi_n^p(x,t), \tag{60}$$

where

$$\alpha_n^p(t) = -(2n+a+1)\gamma(t), \tag{61}$$

and

$$\phi_n^p(x,t) = \left(\frac{2(n!)}{\Gamma(n+a+1)}\right)^{1/2}\left(\frac{m\dot\gamma}{\hbar}\right)^{(a+1)/2} x^{a+1/2} \times\exp\left[\frac{im}{2\hbar}\left(\frac{\dot s}{s}+i\dot\gamma\right)x^2\right]L_n^a\left(\frac{m\dot\gamma}{\hbar}x^2\right), \tag{62}$$

we may easily verify that $\phi_n^p(x,t)$ are eigenfunctions of I^p, and that

$$I^p\phi_n^p(x,t) = \hbar(2n+a+1)\phi_n^p(x,t), \tag{63}$$

while the phases $\alpha_n^p(t)$ are determined by

$$\hbar\frac{d\alpha_n^p(t)}{dt} = \left\langle\phi_n^p\middle|i\hbar\frac{\partial}{\partial t} - H^p\middle|\phi_n^p\right\rangle. \tag{64}$$

These examples bring out an interesting observation. Consider a system with a time dependent Hamiltonian $H(t)$ and assume that there exists a Hermitian invariant operator $I(t)$ which does not involve time differentiation. Assume that $\{\phi_n(x,t)\}$ form a complete set of normalized eigenfunctions of $I(t)$ with eigenvalues $\{\lambda_n\}$. Then the propagator $K(x'',t'';x',t')$ for the system admits an expansion

$$K(x'',t'';x',t') = \sum_n \exp\{i[\alpha_n(t'') - \alpha_n(t')]\}\phi_n^*(x',t')\phi_n(x'',t''), \tag{65}$$

where the phases $\alpha_n(t)$ are determined from the equation

$$\hbar\frac{d\alpha_n}{dt} = \left\langle\phi_n\middle|i\hbar\frac{\partial}{\partial t} - H(t)\middle|\phi_n\right\rangle, \tag{66}$$

and

$$I(t)\phi_n(x,t) = \lambda_n\phi_n(x,t). \tag{67}$$

λ_n are time-dependent since the operator $I(t)$ does not involve time differentiation. In particular, if the Hamiltonian $H(t)$ is a time-independent operator H, the invariant operator I may be taken as H itself and both ϕ_n and λ_n are time-independent. In fact, in this case λ_n are the energy eigenvalues E_n and the phases α_n, as determined by Eq. (66), are given by

$$\alpha_n = -E_nt/\hbar, \tag{68}$$

within an arbitrary constant; and hence the propagator K assumes the usual form of expansion[2]

$$K(x'',t'';x',t') = \sum_n \exp[-(i/\hbar)E_n(t''-t')]\phi_n^*(x')\phi_n(x''). \tag{69}$$

APPENDIX: DETERMINATION OF THE COEFFICIENTS a_N, p_N, AND q_N, AS $N\to\infty$

We first define

$$\lambda_j = 2\gamma_j/\beta, \tag{A1}$$

$$\Lambda_k = \prod_{j=1}^{k} 1/\lambda_j, \tag{A2}$$

and express the coefficients a_N, p_N, q_N in terms of β and Λ_k:

$$a_N = \beta\Lambda_{N-1}, \tag{A3}$$

$$p_N = \tfrac{1}{2}\beta\left(1 - \sum_{j=1}^{N-1}\Lambda_j\Lambda_{j-1}\right), \tag{A4}$$

$$q_N = \tfrac{1}{2}\beta(1 - \Lambda_{N-1}/\Lambda_{N-2}). \tag{A5}$$

Now from Eqs. (11) and (17) it follows that λ_j satisfy the recurrence relation

$$\lambda_j = 2(1 - \tfrac{1}{2}\omega_j^2\epsilon^2) - 1/\lambda_{j-1}. \tag{A6}$$

Let us write

$$\lambda_j = Q_{j+1}/Q_j, \tag{A7}$$

so that Q_j satisfy the recurrence relation

$$Q_{j+1} = 2(1 - \tfrac{1}{2}\omega_j^2\epsilon^2)Q_j - Q_{j-1}, \tag{A8}$$

which may be written as

$$(Q_{j+1} - 2Q_j + Q_{j-1})/\epsilon^2 = -\omega_j^2Q_j, \tag{A9}$$

which in the limit as $\epsilon\to 0$ reduces to the differential equation

$$\ddot Q(t) + \omega^2(t)Q(t) = 0 \tag{A10}$$

with

$$Q_0 = Q(t') = 0. \tag{A11}$$

The latter condition follows from Eqs. (A7) and (A8) when j is set equal to 1.

A solution of Eq. (A10) satisfying Eq. (A11) is given by

$$Q(t) = s(t)\sin[\gamma(t) - \gamma(t')], \tag{A12}$$

where $s(t)$ and $\gamma(t)$ obey the Eqs. (22) and (23), respectively. We may now proceed to determine the limiting values of a_N, p_N, q_N as $N\to\infty$. First, all these coefficients are expressed in terms of β and Q_k:

$$a_N = \beta\Lambda_{N-1} = \beta(Q_1/Q_N), \tag{A13}$$

$$p_N = \tfrac{1}{2}\beta\left(1 - \sum_{j=1}^{N-1}\frac{Q_1}{Q_{j+1}}\frac{Q_1}{Q_j}\right), \tag{A14}$$

$$q_N = \tfrac{1}{2}\beta(Q_N - Q_{N-1})/Q_N. \tag{A15}$$

Noting that $\beta = \hbar/m\epsilon$, we find that in the limit $N \to \infty$ $(\epsilon \to 0)$,

$$a_N \to \frac{m}{\hbar}\frac{\dot{Q}(t')}{Q(t'')} = \frac{m}{\hbar}\frac{\sqrt{\dot\gamma(t'')\dot\gamma(t')}}{\sin\phi(t'',t')}\ , \tag{A16}$$

$$p_N \to \frac{m}{2\hbar}\lim_{\epsilon\to 0}\left(1 - \epsilon\int_{t''+\epsilon}^{t''}\frac{Q^2(t')}{Q^2(t)}\,dt\right)$$

$$= \frac{m}{2\hbar}\left(-\frac{\dot s(t')}{s(t')} + \dot\gamma(t')\cot\phi(t'',t')\right), \tag{A17}$$

and

$$q_N \to \frac{m}{2\hbar}\frac{\dot Q(t'')}{Q(t'')} = \frac{m}{2\hbar}\left(\frac{\dot s(t'')}{s(t'')} + \dot\gamma(t'')\cot\phi(t'',t')\right). \tag{A18}$$

[1] R. P. Feynman, Rev. Mod. Phys. 20, 367 (1948).
[2] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
[3] D. Peak and A. Inomata, J. Math. Phys. 10, 1422 (1969).
[4] D. C. Khandekar and S. V. Lawande, J. Phys. A 5, 812 (1972).
[5] D. C. Khandekar and S. V. Lawande, J. Phys. A 5, 157 (1972).
[6] A. V. Jones and G. J. Papadopoulos, J. Phys. A 4, 186 (1971).
[7] H. R. Lewis, Phys. Rev. Lett. 18, 510 (1967).
[8] H. R. Lewis, J. Math. Phys. 9, 1976 (1968).
[9] H. R. Lewis and W. B. Reisenfeld, J. Math. Phys. 10, 1458 (1969).
[10] An exhaustive list of references on "time-dependent harmonic oscillator" is given at the end of Ref. 11 in this text.
[11] P. Camiz, A. Gerardi, C. Marchioro, E. Presutti, and E. Scacciatelli, J. Math. Phys. 12, 2040 (1971).
[12] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, New York, 1958), Sec. 35.
[13] *Higher Transcendental Functions*, edited by A. Erdelyi (McGraw-Hill, New York, 1953), Vol. 2, p. 194.
[14] An invariant operator $I(t)$ is defined by $dI/dt = \partial I/\partial t + (1/i\hbar)[I, H] = 0$.
[15] *Higher Transcendental Functions*, edited by A. Erdelyi (McGraw-Hill, New York, 1953), Vol. 2, p. 189.

Calculation of moments in the harmonic oscillator basis

P. C. Joshi

Nuclear Physics Section, Department of Physics, Banaras Hindu University, Varanasi 221005, India
(Received 27 November 1973)

The matrix element of r^k between the eigenstates of harmonic oscillator potential is generally calculated by using the value of an integral listed in available literature for certain specific conditions. Recently, while calculating the matrix elements of deformed spin–orbit potential in the harmonic oscillator basis for nonspherical nuclei, we found that the evaluation of this integral is needed under more general conditions than discussed previously. We discuss the results of the evaluation of this integral for these additional cases.

The three-dimensional harmonic oscillator potential has been extensively used in practically all the branches of physics. The eigenfunctions describing this potential contain Laguerre functions (or confluent hypergeometric functions). The radial part of these wavefunctions is written as[1]

$$|Nl\rangle = N_0 r^l e^{-r^2/2} L_n^{l+1/2}(r^2). \tag{1}$$

Generally one comes across the problem of calculating the matrix element of r^k for integer k (or the moments) between these wavefunctions, which are

$$\langle N'l' | r^k | Nl\rangle = C\int Z^p e^{-Z} L_m^{p-\mu}(Z) L_n^{p-\nu}(Z) dZ, \tag{2}$$

where

$$\begin{aligned} &Z = r^2, \quad p = \tfrac{1}{2}(l' + l + k + 1), \\ &N = 2n + l, \quad N' = 2m + l', \\ &p - \mu = l' + \tfrac{1}{2}, \quad p - \nu = l + \tfrac{1}{2}. \end{aligned} \tag{3}$$

The value of integral in Eq. (2) is given by[2]

$$I = (-1)^{m+n}\Gamma(p+m-\mu+1)\Gamma(p+n-\nu+1)\mu!\,\nu! \times \sum_\sigma \frac{\Gamma(p+\sigma+1)}{\sigma!(m-\sigma)!(n-\sigma)!(\sigma+\mu-m)!(\sigma+\nu-n)!}, \tag{4}$$

where m, n, μ, ν are integers or zero and σ takes on integral values such that

$$\begin{matrix} m \\ n \end{matrix} \Big\rangle\!\Big\rangle\ \sigma\ \Big\rangle\!\Big\rangle \begin{matrix} m-\mu \\ n-\nu \end{matrix}. \tag{5}$$

It is clear that condition (5) is not satisfied for negative values of μ and/or ν; the evaluation of this integral corresponding to such values has not been reported so far. Recently, we came across a physical problem where the value of μ and/or ν becomes negative. We have evaluated the integral (2) for these cases, starting with the definition of generating functions of Laguerre functions and the results are given below.

Case I: If one of the parameters, say μ is negative, the value of the integral is as follows:

$$I = \frac{\Gamma(p+m-\mu+1)\Gamma(p+n-\nu+1)}{\Gamma(-\mu)}\nu! \times \sum_\sigma (-1)^{n-\sigma}\frac{\Gamma(m-\sigma-\mu)\Gamma(p+\sigma+1)}{(m-\sigma)!(n-\sigma)!(\nu+\sigma-n)!\sigma!} \tag{6}$$

where

$$\begin{matrix} n \\ m \end{matrix} \Big\rangle\!\Big\rangle\ \sigma\ \Big\rangle\!\Big\rangle\ n-\nu. \tag{7}$$

Case II: If both μ and ν are negative integers, then we obtain the value

$$I = \frac{\Gamma(p+m-\mu+1)\Gamma(p+n-\nu+1)}{\Gamma(-\mu)\Gamma(-\nu)} \times \sum_\sigma \frac{\Gamma(p+\sigma+1)\Gamma(m-\sigma-\mu)\Gamma(n-\sigma-\nu)}{\sigma!(n-\sigma)!(m-\sigma)!} \tag{8}$$

where σ can have values from 1 to m or n (whichever is smaller).

The physical problem investigated in our case concerns the introduction of deformations[3] in the spin-orbit coupling term in the Nilsson model[4] Hamiltonian for the case of deformed nuclei. Assuming the nucleons to be Dirac particles one follows the analogy with the electrons and writes the spin orbit term as

$$V_{s.o.} = \lambda(\nabla U \times \mathbf{p}\cdot\sigma). \tag{9}$$

while calculating the matrix elements of this operator in the oscillator basis we came across a term of the form

$$\begin{aligned} J\delta_{l',l\text{ or }l\pm2} = N\Big(&\int Z^{(l'+l+3)/2} e^{-Z} L_{n'}^{l'+1/2}(Z) L_n^{l+1/2}(Z) dZ \\ &- (l+1)\int Z^{(l+l'+1)/2} e^{-Z} L_{n'}^{l'+1/2}(Z) L_n^{l+1/2}(Z)\, dZ \\ &+ 2(n+1)\int Z^{(l+l'+1)/2} e^{-Z} L_{n'}^{l'+1/2}(Z) \\ &\times L_{n+1}^{l-1/2}(Z) dZ\Big)\delta_{l',l\text{ or }l\pm2} \\ = J_1 + J_2 + J_3. \end{aligned} \tag{10}$$

Thus the integrals (10) are to be evaluated for these three cases, i.e., $l' = l'$, $l' = l \pm 2$. These integrals are of the form (2) and (6). On establishing identification of labels p, μ, and ν, we find the values of these

TABLE I. The values of μ and ν for possible values of l' in the integrals (10).

l'	n'	For J_1		For J_2		For J_3	
		μ	ν	μ	ν	μ	ν
$l-2$	$n+1$	2	0	1	−1	1	0
l	n	1	1	0	0	0	1
$l+2$	$n-1$	0	2	−1	1	−1	2

parameters as given in Table I. We see that three of the integrals in Eq. (10) are to be evaluated using Eq. (6) and the rest can be evaluated using Eq. (4). Thus, the value of integral J (using the proper normalization constants) is found to be:

$$J = -\tfrac{3}{2}\delta_{l',l} + (l+1)\left(\frac{n+1}{n+l+1/2}\right)^{1/2} \delta_{l',l-2=1} - \left(\frac{n}{n+l+3/2}\right)^{1/2} \delta_{l',l+2}. \tag{11}$$

ACKNOWLEDGMENT

The author wishes to thank Professor P.C. Sood and Dr. A.N. Mantri for helpful discussions and suggestions. The award of a fellowship from the Council of Scientific and Industrial Research (India) is gratefully acknowledged.

[1] P.M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), p. 1662.
[2] Ref. 1, p. 785.
[3] B.L. Andersen, Nucl. Phys. A **162**, 208 (1971).
[4] S.G. Nilsson, Mat. Fys. Medd. Dan. Vid. Selsk. **29**, No. 16 (1955).

Disproof of a conjecture of D. Muhlerin and I. Zinnes

John D. Dollard

Department of Mathematics, University of Texas, Austin, Texas 78712
(Received 28 January 1974)

It is shown that the operator $U_c^{(+)}(t)$ defined by Muhlerin and Zinnes[1] in their account of Coulomb potential scattering theory is not an isometry when the Coulomb potential is weak enough. It is opined that for this reason their account of Coulomb potential scattering theory does not have a straightforward physical interpretation.

INTRODUCTION

Notation: In this paper, the boldface letters $\mathbf{x}$ and $\mathbf{k}$ denote elements of three-dimensional Euclidean space $\mathbf{R}^3$, while x and k denote the magnitudes of $\mathbf{x}$ and $\mathbf{k}$. All integrals occurring are over all of $\mathbf{R}^3$ unless otherwise specified. We denote by H the Coulomb Hamiltonian given by

$$H = -\frac{\Delta}{2m} + \frac{\lambda}{m}\frac{1}{x}, \tag{1}$$

where Δ is the Laplacean operator, m is the mass of the particle we shall consider, and λ is a constant representing the strength of the Coulomb potential. Finally, in this introductory section of the paper, the notations Φ_{0f} and f are used interchangeably for identically the same function in $L^2(\mathbf{R}^3)$. This is done because Muhlerin and Zinnes[1] do it, and the author wishes to make his equations easily comparable with theirs. After the introductory section, this practice is discontinued, except in the summary of results occurring in the Conclusion of the paper.

The work of Muhlerin and Zinnes: Muhlerin and Zinnes[1] have given an account of time-dependent Coulomb potential scattering theory in which a new kind of asymptotic state is introduced. In their notation, let Φ_{0f} (this function is also interchangeably denoted by f) be a normalized function in $L^2(\mathbf{R}^3)$ with Fourier transform $\tilde{f}$, so that

$$\Phi_{0f}(\mathbf{x}) = f(\mathbf{x}) = \lim \frac{1}{(2\pi)^{3/2}} \int \exp(i\mathbf{k}\cdot\mathbf{x})\tilde{f}(\mathbf{k})\,d\mathbf{k}. \tag{2}$$

Then the asymptotic states $\Phi_{cf}^{(\pm)}$ are defined by

$$\Phi_{cf}^{(\pm)}(\mathbf{x}) = (2\pi)^{-3/2}\int \phi_c^{(\pm)}(\mathbf{k},\mathbf{x})\tilde{f}(\mathbf{k})\,d\mathbf{k}, \tag{3}$$

where

$$\phi_c^{(\pm)}(\mathbf{k},\mathbf{x}) = \exp\{i[\mathbf{k}\cdot\mathbf{x} \mp (\lambda/k)\ln(kx \pm \mathbf{k}\cdot\mathbf{x})]\}. \tag{4}$$

Because $\phi_c^{(\pm)}$ is a bounded function, it is clear that the integral in (3) exists at least for $\tilde{f} \in L^1(\mathbf{R}^3)$, except at the point $\mathbf{x}=0$, where it is undefined. Muhlerin and Zinnes discuss definition (3) for functions f belonging to a certain set $/\!/\!/$ dense in $L^2(\mathbf{R}^3)$, and such that $f \in /\!/\!/$ implies $\tilde{f} \in L^1(\mathbf{R}^3)$. They define the "free Coulomb propagator" $U_c^{(\pm)}(t)$ by

$$[U_c^{(\pm)}(t)\Phi_{0f}](\mathbf{x}) = (2\pi)^{-3/2}\int \phi_c^{(\pm)}(\mathbf{k},\mathbf{x})\exp(-ik^2t/2m)\tilde{f}(\mathbf{k})\,d\mathbf{k} \tag{5}$$

and prove that $U_c^{(\pm)}(t)\Phi_{0f}$ converges strongly for $f \in /\!/\!/$ as $t \to \pm\infty$ to solutions $\exp(-iHt)\psi_{cf}^{(\pm)}$ of the Schrödinger equation with interaction. That is,

$$\lim_{t\to\pm\infty} \| \exp(-iHt)\psi_{cf}^{(\pm)} - U_c^{(\pm)}(t)\Phi_{0f}\| = 0 \quad \text{for } f \in /\!/\!/. \tag{6}$$

The Møller wave matrices $\Omega_c^{\pm}$ can then be found in the usual way:

$$\Omega_c^{\pm} = \lim_{t\to\pm\infty} \exp(iHt)U_c^{(\pm)}(t) \quad \text{on } /\!/\!/\,. \tag{7}$$

It is not stated explicitly in the paper that $U_c^{(\pm)}(t)\Phi_{0f}$ is to be thought of as representing the state of the particle at large positive and negative times. However, the reader familiar with the subject cannot help but wonder whether or not this is a correct view. If not, then while (6) is an interesting mathematical fact, it is not capable of the interpretation that is usual in scattering theory, namely that it represents the asymptotic agreement of two descriptions of a moving particle. If $U_c^{(\pm)}(t)\Phi_{0f}$ is to be thought of as representing the state of the particle, then we must have

$$\|U_c^{(\pm)}(t)\Phi_{0f}\| = \|\Phi_{0f}\| = \|f\|, \quad \text{for all } t, \tag{8}$$

an equation which represents the conservation of probability. By taking into account the definitions of $U_c^{(\pm)}(t)$ and $\Phi_{cf}^{(\pm)} = U_c^{(\pm)}(0)\Phi_{0f}$, it is easy to show that (8) will hold for all $f \in L^2(\mathbf{R}^3)$ if and only if we have

$$\|\Phi_{cf}^{(\pm)}\| = \|f\| \tag{9}$$

for all $f \in L^2(\mathbf{R}^3)$. [The same statement holds with $L^2(\mathbf{R}^3)$ replaced by $/\!/\!/$.] Muhlerin and Zinnes in fact conjecture that (9) holds for all $f \in L^2(\mathbf{R}^3)$. However, they are able to prove only

$$\|\Phi_{cf}^{(\pm)}\| < \infty \quad \text{for } f \in /\!/\!/. \tag{10}$$

It is the purpose of the present note to prove that (9) is false, at least for a certain class of functions f for which $\Phi_{cf}^{(\pm)}$ can be defined (the functions in question belong to the class $/\!/\!/$), and for small enough (but nonzero) values of the constant λ representing the strength of the Coulomb potential. The present author believes that there is then no reason to expect that (9) should hold for large values of λ. Muhlerin and Zinnes remark: "Whether one can give the asymptotic states... $(\Phi_{cf}^{(\pm)})$... objective significance depends to some extent on the validity of the conjecture ... ((9) above)...." In the present author's view, the situation is this: $U_c^{(\pm)}(t)\Phi_{0f}$ does not represent the state of a particle. It is part of a mathematical prescription which does not have a straightforward physical interpretation but whose end result is a rederivation of the known Coulomb Møller wave matrix.[2,3]

I. RESULTS, AND THE IDEA OF THE PROOF

To economize on notation, we use new symbols. We define for all $f \in L^1(\mathbf{R}^3)$ the function $T_\lambda f$ by the equation

$$(T_\lambda f)(\mathbf{x}) = (2\pi)^{-3/2}\int \phi_\lambda(\mathbf{k},\mathbf{x})\,f(\mathbf{k})\,d\mathbf{k}, \tag{11}$$

with

$$\phi_\lambda(\mathbf{k},\mathbf{x}) = \exp\{i[\mathbf{k}\cdot\mathbf{x} + (\lambda/k)\ln(kx - \mathbf{k}\cdot\mathbf{x})]\}. \tag{12}$$

$T_\lambda f$ is the function Φ_{cf}^- of the last section. (We will deal only with Φ_{cf}^- in order to avoid clutter. As the reader should be able to see, a corresponding discussion for Φ_{cf}^+ would not look much different and would lead to the same sort of results.) We wish to decide whether or not the equation

$$\|T_\lambda f\| = \|f\| \tag{13}$$

holds. In other words, we are asking whether or not T_λ is an isometry. Now T_0 is the identity operator, which is isometric. We shall show that, for small $\lambda \neq 0$, T_λ is not isometric. The idea of the proof is as follows:

(a) Find a function f (actually a one-parameter family of such functions) for which the function $T_\lambda f$ is a strongly differentiable function of λ. Then $\|T_\lambda f\|^2$ will be a differentiable function of λ.

(b) Prove that

$$\frac{d}{d\lambda}\|T_\lambda f\|^2 \neq 0 \quad \text{for } \lambda = 0 \tag{14}$$

(c) Conclusion: For small enough $\lambda \neq 0$ we have

$$\|T_\lambda f\|^2 \neq \|T_0 f\|^2 = \|f\|^2. \tag{15}$$

This apparently roundabout approach is adopted because of the formidable difficulty of a direct attack, something which can only be appreciated by attempting such a direct attack. The execution of (a) and (b) above requires a straightforward, though grim, application of the theory of integration. This execution is presented in outline in Sec. II. In order to preserve the reader's sanity, we do not give full details. The author will be pleased to supply them on request.

II. OUTLINE OF THE PROOF THAT T_λ IS NOT ISOMETRIC

We consider a function f whose Fourier transform $\tilde{f}$ has the following properties:

(i) $\tilde{f}(\mathbf{k})$ depends only on the magnitude k of $\mathbf{k}$. So we shall write $\tilde{f}(k)$ instead of $\tilde{f}(\mathbf{k})$.

(ii) $\tilde{f}(k)$ is three times continuously differentiable as a function of its argument (this condition makes our work correspond more closely with that of Zinnes and Muhlerin. Two derivatives would be enough for our proof.)

(iii) $\tilde{f}(k)$ vanishes in a neighborhood of $k = 0$.

(iv) $\tilde{f}(k)$ and its derivatives to order three die off faster than any reciprocal power of k as $k \to \infty$.

Eventually, we will specialize our considerations to the case

$$\tilde{f}(k) = C_\delta(k)\exp(-\mu k), \tag{16}$$

where μ is any positive real number, and $C_\delta(k)$ is a three times continuously differentiable function vanishing in a small sphere about $k = 0$ and equaling 1 outside a slightly larger sphere about $k = 0$.

If $\tilde{f}$ satisfies conditions (i)—(iv), then $\tilde{f}$ is in $L^1(\mathbf{R}^3)$, and so the definition in (11) certainly makes sense for all $\mathbf{x} \neq 0$. We will now outline the proof that $T_\lambda f$ is strongly differentiable with respect to λ. For this purpose, first write the integration variable $\mathbf{k}$ in parabolic coordinates ξ, η, ψ, with the z direction in k-space chosen along the vector $\mathbf{x}$ ($\mathbf{x} \neq 0$):

$$\begin{aligned} &\xi = (kx + \mathbf{k}\cdot\mathbf{x})/x, \quad k = (\xi+\eta)/2, \\ &\eta = (kx - \mathbf{k}\cdot\mathbf{x})/x, \quad \mathbf{k}\cdot\mathbf{x} = (\xi-\eta)/2, \\ &\psi = \tan^{-1}(k_1/k_2), \end{aligned} \tag{17}$$

where k_1 and k_2 are the projections of $\mathbf{k}$ on a pair of orthogonal unit vectors perpendicular to $\mathbf{x}$. Then for $\mathbf{x} \neq 0$ we have

$$\begin{aligned} &(T_\lambda f)(\mathbf{x}) \\ &= \frac{2\pi}{(2\pi)^{3/2}}\int_0^\infty\int_0^\infty \exp\left\{i\left[\left(\frac{\xi-\eta}{2}\right)x + \frac{2\lambda}{\xi+\eta}\ln\eta x\right]\right\} \\ &\quad\times \tilde{f}\left(\frac{\xi+\eta}{2}\right)\frac{\xi+\eta}{2}\,d\xi\,d\eta, \end{aligned} \tag{18}$$

where the factor of 2π results from doing the trivial ψ integration and the factor $(\xi+\eta)/2$ in the integrand is part of the volume element in parabolic coordinates. The derivative of the integrand with respect to λ is (for $\eta \neq 0$)

$$i\ln\eta x \exp\left\{i\left[\left(\frac{\xi-\eta}{2}\right)x + \frac{2\lambda}{\xi+\eta}\ln\eta x\right]\right\}\tilde{f}\left(\frac{\xi+\eta}{2}\right). \tag{19}$$

(The point $\xi + \eta = 0$ causes no trouble because $\tilde{f}$ vanishes in a neighborhood of this point.) For each $\eta \neq 0$ the absolute value of the function in (19) is bounded as follows:

$$\left|\ln(\eta x)\tilde{f}\left(\frac{\xi+\eta}{2}\right)\right| \leq (|\ln\eta| + |\ln x|)\left|\tilde{f}\left(\frac{\xi+\eta}{2}\right)\right|. \tag{20}$$

Thus for $\eta \neq 0$ the differentiated integrand (19) is bounded by a function independent of λ, and easy estimates using the properties of $\tilde{f}$ show that this function is integrable over the integration region of (18). It follows[4] that for $\mathbf{x} \neq 0$ the derivative $(d/d\lambda)(T_\lambda f)(\mathbf{x})$ exists *at each point* $\mathbf{x}$, and equals $(2\pi)^{-1/2}$ times the integral over the integration region indicated in (18) of the function in (19). We have also for $\mathbf{x} \neq 0$

$$\frac{d}{d\lambda}(T_\lambda f)(\mathbf{x}) = \lim_{h\to 0}\Delta_h(\mathbf{x}), \tag{21}$$

where

$$\Delta_h(\mathbf{x}) = \frac{(T_{\lambda+h}f)(\mathbf{x}) - (T_\lambda f)(\mathbf{x})}{h} \tag{22}$$

We wish to show that actually Δ_h converges strongly to $(d/d\lambda)T_\lambda f$ as $h \to 0$. To do this, using Lebesgue's dominated convergence theorem and some elementary manipulations, it suffices to show that $|\Delta_h(\mathbf{x})|$ is bounded almost everywhere by a fixed square-integrable function of $\mathbf{x}$. We proceed as follows to show this: We have

$$\begin{aligned} &(1+x^2)\Delta_h(\mathbf{x}) \\ &= \frac{(1+x^2)}{(2\pi)^{1/2}}\int_0^\infty\int_0^\infty \phi_\lambda(\mathbf{k},\mathbf{x})\left[\exp\left(\frac{2ih}{\xi+\eta}\ln\eta x\right) - 1\right] \\ &\quad\times \tilde{f}\left(\frac{\xi+\eta}{2}\right)\frac{\xi+\eta}{2}\,d\xi\,d\eta. \end{aligned} \tag{23}$$

We now remark that $\phi_\lambda(\mathbf{k},\mathbf{x})$ contains the factor $\exp\{i[(\xi-\eta)/2]x\}$ [see (18)] and that

$$x^2\exp\{i[(\xi-\eta)/2]x\} = -4(\partial^2/\partial\xi^2)\exp\{i[(\xi-\eta)/2]x\} \tag{24}$$

and use this fact in (23). Two integrations by parts and some technical arguments concerning boundary terms, combined with estimates of the type made in (20), lead to the result

$$(1+x^2)\left|\Delta_\lambda(\mathbf{x})\right| \leq A + B\left|\ln x\right| + C\left|\ln x\right|^2 + D\left|\ln x\right|^4 \quad (25)$$

for all $\mathbf{x} \neq 0$, where A, B, C, and D are constants. Equation (25) shows that $|\Delta_\lambda(\mathbf{x})|$ is bounded almost everywhere by a fixed square-integrable function of $\mathbf{x}$. Thus the strong derivative of $T_\lambda f$ exists and is given for $\mathbf{x} \neq 0$ by $(2\pi)^{-1/2}$ times the integral of the function in (19). We denote the value of this strong derivative at $\lambda = 0$ by Kf. Thus

$$(Kf)(\mathbf{x}) = \left(\frac{d}{d\lambda} T_\lambda f\right)(\mathbf{x})\Big|_{\lambda=0} = \frac{i}{(2\pi)^{1/2}} \int_0^\infty \int_0^\infty \ln \eta x \exp\left\{i\left[\left(\frac{\xi-\eta}{2}\right)x\right]\right\} \times \tilde{f}\left(\frac{\xi+\eta}{2}\right) d\xi\, d\eta. \quad (26)$$

Because the strong derivative of $T_\lambda f$ exists, it follows that $\|T_\lambda f\|^2$ is differentiable as a function of λ, and

$$\frac{d}{d\lambda}\|T_\lambda f\|^2 = \frac{d}{d\lambda}(T_\lambda f, T_\lambda f) = 2\,\mathrm{Re}\left(\frac{d}{d\lambda} T_\lambda f, T_\lambda f\right) \quad (27)$$

In particular, we have

$$\frac{d}{d\lambda}\|T_\lambda f\|^2\Big|_{\lambda=0} = 2\,\mathrm{Re}(Kf, T_0 f) = 2\,\mathrm{Re}(Kf, f). \quad (28)$$

We now choose f to be a function f_δ whose Fourier transform $\tilde{f}_\delta$ has the form

$$\tilde{f}_\delta(k) = C_\delta(k)\exp(-\mu k), \quad (29)$$

where $\mu > 0$, $0 < \delta < 1$, and $C_\delta(k)$ is a function satisfying the conditions:

$$\begin{aligned} &C_\delta(k) \text{ is three times continuously differentiable,} \\ &C_\delta(k) = 0, \quad \text{for } 0 \leq k \leq \delta, \\ &C_\delta(k) = 1, \quad \text{for } k \geq \delta^{1/2}. \end{aligned} \quad (30)$$

For $\delta < k < \delta^{1/2}$, $C_\delta(k)$ is a polynomial of degree 7.

These conditions specify $C_\delta(k)$ uniquely. The last condition is a convenience intended to simplify calculations. It would be easy to write down the coefficients of the polynomial mentioned in this condition explicitly, but there is no point in doing so. It should be clear that with this choice of $C_\delta(k)$, the function $\tilde{f}_\delta(k)$ of (29) satisfies conditions (i)—(iv) at the beginning of this section, so that $T_\lambda f_\delta$ is strongly differentiable with respect to λ.

Because $\tilde{f}_\delta(k)$ converges pointwise almost everywhere as $\delta \to 0$ to the function $\exp(-\mu k)$, it is plausible to think that the function $Kf_\delta = (d/d\lambda)T_\lambda f_\delta|_{\lambda=0}$ defined as in (26) converges as $\delta \to 0$ to the function g defined by

$$g(\mathbf{x}) = \frac{i}{(2\pi)^{1/2}} \int_0^\infty \int_0^\infty \ln(\eta x) \exp\left\{i\left[\left(\frac{\xi-\eta}{2}\right)x\right]\right\} \times \exp[-\mu(\xi+\eta)/2]\, d\xi\, d\eta. \quad (31)$$

This is in fact true, and furthermore the convergence is strong convergence, as a detailed estimate shows. Since the function $f_\delta(k)$ also converges strongly as $\delta \to 0$ to the function $f(k) = \exp(-\mu k)$, it follows that

$$\lim_{\delta \to 0}(Kf_\delta, f_\delta) = (g, f). \quad (32)$$

The rest of our argument consists in showing that

$$2\,\mathrm{Re}(g, f) \neq 0. \quad (33)$$

In fact, a rather involved calculation shows that

$$2\,\mathrm{Re}(g, f) = \frac{16\mu^4}{\pi^2} \int \frac{\tan^{-1}(x/\mu)}{(\mu^2+x^2)^3}\, d\mathbf{x} > 0. \quad (34)$$

We can now assemble the pieces: Because of (32) and (34) we have

$$2\,\mathrm{Re}(Kf_\delta, f_\delta) > 0 \quad \text{for small enough } \delta > 0. \quad (35)$$

Thus for small enough $\delta > 0$ we have, using (28),

$$\frac{d}{d\lambda}\|T_\lambda f_\delta\|^2\Big|_{\lambda=0} = 2\,\mathrm{Re}(Kf_\delta, f_\delta) > 0. \quad (36)$$

and hence for small enough $\delta > 0$ and small enough $\lambda \neq 0$ we have

$$\begin{aligned} &\|T_\lambda f_\delta\|^2 > \|T_0 f_\delta\|^2 = \|f_\delta\|^2 \quad \text{if } \lambda > 0, \\ &\|T_\lambda f_\delta\|^2 < \|T_0 f_\delta\|^2 = \|f_\delta\|^2 \quad \text{if } \lambda < 0. \end{aligned} \quad (37)$$

Thus T_λ is norm-changing for small enough $\lambda \neq 0$ on the class of functions considered. This class is rather large because of the arbitrariness of μ. As a special application, it follows that T_λ cannot be isometric on the set $/\!/$ of Muhlerin and Zinnes, because an elementary argument involving a partition of unity shows that each of the functions f_δ considered, for any value of μ, belongs to $/\!/$. (Rather than give the definition of $/\!/$ here, we refer the interested reader to the paper of Muhlerin and Zinnes.)

CONCLUSION

We have shown that for an appropriate class of functions f in $/\!/$ and small enough nonzero values of the parameter λ representing the strength of the Coulomb potential we have

$$\|\Phi_{cf}^{(\pm)}\| \neq \|f\|. \quad (38)$$

Thus $U_c^{(\pm)}(t)$ is not an isometry, and Muhlerin and Zinnes' expression $U_c^{(\pm)}(t)\Phi_{0f}$ cannot in general represent the state of a quantum-mechanical particle. In the view of the author, their prescription for calculating the Møller wave matrix therefore does not have a straightforward physical interpretation. It might perhaps be hoped that an interpretation could be salvaged if at least $U_c^{(\pm)}(t)^*$ were an isometry. However, it is not.

[1] D. Muhlerin and I. Zinnes, J. Math. Phys. 11, 1402 (1970).
[2] W. Gordon, Z. Phys. 48, 180 (1928).
[3] J. D. Dollard, J. Math. Phys. 5, 729 (1964).
[4] R. G. Bartle, *The Elements of Integration* (Wiley, New York, 1966), Corollary 5.9, p. 46.

Semisimple subgroups of linear semisimple Lie groups

J. F. Cornwell and J. M. Ekins*

Department of Theoretical Physics, University of St. Andrews, St. Andrews, Fife, Scotland
(Received 26 December 1973)

The problem of embedding a semisimple Lie group in a linear semisimple Lie group [where one (or both) may be noncompact] is investigated in detail. The analysis is based on a previous set of papers dealing with the corresponding problem for real Lie algebras.

I. INTRODUCTION

In a previous series of papers (Cornwell,[1-3] Ekins and Cornwell[4,5]), hereafter referred to as I to V, a detailed investigation was carried out on the problem of embedding a semisimple real Lie algebra $\mathcal{L}'$ in another semisimple real Lie algebra $\mathcal{L}$, one or both of these algebras being noncompact. In the present paper this study is complemented and completed by examining the corresponding problem for semisimple Lie *groups*.

The previous analysis was based on the work of Cartan[6,7] and Gantmacher[8,9] in which noncompact semisimple real Lie algebras are generated from their compact forms by involutive automorphisms of the compact forms. This construction again plays a vital role in the present problem.

In Sec. II the relevant part of the theory of linear semisimple Lie groups is reviewed and extended, the essential details for each of the real groups being summarized in the Appendix. In Sec. III the embedding problem for linear connected semisimple Lie groups is examined in detail, and an explicit method of construction is given. An example illustrating this construction is described in Sec. IV.

II. LINEAR SEMISIMPLE LIE GROUPS

It is well known that all the locally isomorphic connected Lie groups $\mathcal{G}$ corresponding to the same real Lie algebra $\mathcal{L}$ can be obtained from a single unique group $\tilde{\mathcal{G}}$, the *universal covering group* of $\mathcal{L}$, by taking factor groups of $\tilde{\mathcal{G}}$ with respect to all possible discrete central normal subgroups $\mathcal{N}$ of $\tilde{\mathcal{G}}$; that is, $\mathcal{G} \approx \tilde{\mathcal{G}}/\mathcal{N}$. If $\mathcal{G}$ possesses at least one faithful finite linear representation, then $\mathcal{G}$ is said to be a *linear Lie group*. There do exist semisimple noncompact Lie groups that are not linear, but these have not yet found application in physical problems. Therefore attention will be concentrated on the linear groups.

The following basic theorem on linear Lie groups is due to Mal'cev[10]: "Among the locally isomorphic connected linear semisimple Lie groups having real Lie algebra $\mathcal{L}$ there is one, denoted by $\hat{\mathcal{G}}$ and called the *universal linear group* of $\mathcal{L}$, which is unique up to isomorphism, such that all the others are isomorphic to factor groups of $\hat{\mathcal{G}}$. That is, for every connected linear semisimple Lie group $\mathcal{G}$ there exists a subgroup $\mathcal{N}$ of $Z(\hat{\mathcal{G}})$, the center of $\hat{\mathcal{G}}$, such that $\mathcal{G} \approx \hat{\mathcal{G}}/\mathcal{N}$. If $\hat{\mathcal{G}}_1$ and $\hat{\mathcal{G}}_2$ are the universal linear groups corresponding to the real Lie algebras $\mathcal{L}_1$ and $\mathcal{L}_2$, then the universal linear group corresponding to the direct sum $\mathcal{L}_1 \oplus \mathcal{L}_2$ is isomorphic to the direct product $\hat{\mathcal{G}}_1 \otimes \hat{\mathcal{G}}_2$." Clearly this theorem shows that a complete knowledge of the connected linear semisimple Lie groups can be obtained from the simple universal linear groups $\hat{\mathcal{G}}$ and their centers $Z(\hat{\mathcal{G}})$.

Dynkin and Oniscik[11] have studied the universal linear groups and their centers for the *compact* simple Lie groups $\mathcal{G}_C$, $\hat{\mathcal{G}}_C$ and $\hat{\mathcal{G}}_C$ being isomorphic in this case. The generalization to the *non-compact* groups was made by Sirota[12,13] and Solodovnikov,[14] and reviewed in a survey article by them.[15] The present section is devoted to summarizing the results on universal linear groups that are essential for dealing with the subgroup problem. Some of these results are given explicity in the papers just mentioned, others merely appear implicitly, while some are new. The notations and conventions will be taken to be those defined in papers I to V.[1-5] In particular $\mathcal{G}$ will denote a linear connected simple Lie group, $\mathcal{L}$ its real Lie algebra, $\tilde{\mathcal{L}}$ the complexification of $\mathcal{L}$, H the Cartan subalgebra of $\tilde{\mathcal{L}}$, $\mathcal{L}_C$ the compact real form of $\tilde{\mathcal{L}}$, and $\mathcal{G}_C$ a connected compact Lie group having $\mathcal{L}_C$ as its Lie algebra.

There are four cases to be considered:

(i) $\mathcal{L}$ compact,

(ii) $\mathcal{L}$ noncompact and generated from $\mathcal{L}_C$ by an inner involutive automorphism (cf. I),

(iii) $\mathcal{L}$ noncompact and generated from $\mathcal{L}_C$ by an outer involutive automorphism, with $\tilde{\mathcal{L}}$ simple (cf. II, III, V),

(iv) $\mathcal{L}$ noncompact and generated from $\mathcal{L}_C$ by an outer involutive automorphism, but $\tilde{\mathcal{L}} = \tilde{\mathcal{L}}_1 \oplus \tilde{\mathcal{L}}_2$, where $\tilde{\mathcal{L}}_1$ and $\tilde{\mathcal{L}}_2$ are isomorphic simple complex Lie algebras (cf. IV).

The groups corresponding to cases (i), (ii), and (iii) are often called *real* Lie groups, while those corresponding to (iv) are referred to as *complex* Lie groups.

As $Z(\hat{\mathcal{G}})$ is necessarily finite and Abelian, $Z(\hat{\mathcal{G}})$ must be a finite cyclic group or a direct product of a finite number of finite cyclic groups. The cyclic group of order n generated by $\exp h$ will be denoted by $Z_n(\exp h)$. Using the results of Dynkin and Oniscik[11] and Sirota and Solodovnikov,[15] the generators $\exp h$ of $Z(\hat{\mathcal{G}})$ may easily be found by the following argument. Suppose that l is the rank of $\tilde{\mathcal{L}}$, that $h_1, h_2, \ldots, h_l$ form the canonical basis of H (cf. Jacobson[16]), and that $\mathcal{L}$ is obtained from $\mathcal{L}_C$ by the chief involutive automorphism S.[1,2,8] Then S maps H into H, so that if the basis of $\mathcal{L}_C$ is chosen so that the first l elements are $ih_1, ih_2, \ldots, ih_l$, then the elements $h_1', h_2', \ldots, h_l'$ provide a basis for $H \cap \mathcal{L}$, where

$$h_j' = \sum_{k=1}^{l} (\sqrt{S})_{jk}\, ih_k, \tag{1}$$

and

$$\sqrt{S} = \tfrac{1}{2}(1-i)S + \tfrac{1}{2}(1+i)I, \tag{2}$$

I being the identity transformation. Clearly, if $\exp h$ is a generator of $Z(\hat{\mathcal{G}})$, then $h \in \mathcal{H} \cap \mathcal{L}$, so $h = \sum_{j=1}^{l} \kappa_j h_j'$, where the κ_j are all real. [In particular, for cases (i) and (ii), $S_{jk} = \delta_{jk}$, $j,k = 1,2,\ldots,l$, so that $h_j' = ih_j, j = 1,2,\ldots,l.$] Using the explicit matrix representations of $\mathcal{L}$ in which h_j is represented by a matrix h_j (as given in Appendix A of I for the classical Lie algebras and Sec. 2 of V for the exceptional Lie algebras) then

$$\exp\left(\sum_{j=1}^{l} \kappa_j \mathrm{h}_j'\right) = \gamma \mathrm{I} \tag{3}$$

for some γ, where I is the unit matrix. Condition (3) allows the determination of all the possible values of κ_j and hence all the possible generators of $Z(\hat{\mathcal{G}})$ (even when the representation of $\mathcal{G}$ obtained from the h_j by exponentiation is not a faithful representation of $\mathcal{G}$). For later use it may be noted that when $\mathcal{G}$ is of type (i) or (ii), condition (3) becomes

$$\exp\left(\sum_{j=1}^{l} i\kappa_j \mathrm{h}_j\right) = \gamma \mathrm{I}. \tag{4}$$

In the subsequent sections a vital role is played by the representations of $\mathcal{G}$ that are both single-valued and faithful, so that a criterion is needed to distinguish these representations. Suppose that $\mathbf{\Gamma}$ is a representation of $\hat{\mathcal{G}}$, and $\mathrm{Ker}\mathbf{\Gamma}$ is defined as the kernel of $\mathbf{\Gamma}$ in $\hat{\mathcal{G}}$. i.e., $\exp h \in \mathrm{Ker}\mathbf{\Gamma}$ if and only if $\exp\mathbf{\Gamma}(h)$ is the unit matrix. Obviously $\mathrm{Ker}\mathbf{\Gamma}$ is a subgroup of $Z(\hat{\mathcal{G}})$. Then it is clear that the required criterion is

$\mathbf{\Gamma}$ *provides a single-valued and faithful representation of* $\mathcal{G}$,

where (A)

$\mathcal{G} \approx \hat{\mathcal{G}}/\mathcal{N}$, *if and only if* $\mathcal{N} = \mathrm{Ker}\mathbf{\Gamma}$.

This implies that each representation of $\hat{\mathcal{G}}$ (and hence each representation obtained by exponentiation from $\mathcal{L}$) provides a single-valued and faithful representation of one and only one group $\mathcal{G}$ that is a factor of $\hat{\mathcal{G}}$. If $\mathbf{\Gamma}$ is reducible, and $\mathbf{\Gamma}^1$, $\mathbf{\Gamma}^2$, ... are its irreducible constituents, then

$$\mathrm{Ker}\mathbf{\Gamma} = \mathrm{Ker}\mathbf{\Gamma}^1 \cap \mathrm{Ker}\mathbf{\Gamma}^2 \cap \ldots,$$

so that the only problem remaining is that of determining $\mathrm{Ker}\mathbf{\Gamma}$ when $\mathbf{\Gamma}$ is irreducible.

Suppose that $\exp h \in Z(\hat{\mathcal{G}})$, where $h = \sum_{j=1}^{l} \kappa_j h_j'$ (as above). Then $\exp\mathbf{\Gamma}(h)$ must be a multiple of the unit matrix I, and as the diagonal element of $\exp\mathbf{\Gamma}(h)$ corresponding to the highest weight $\Lambda(h)$ of the IR $\mathbf{\Gamma}$ in $\tilde{\mathcal{L}}$ is $\exp\{\Lambda(h)\}$, it follows that

$$\exp\mathbf{\Gamma}(h) = [\exp\{\Lambda(h)\}]\mathrm{I}. \tag{5}$$

Then the element $\exp h$ of $Z(\hat{\mathcal{G}})$ is a member of $\mathrm{Ker}\mathbf{\Gamma}$ if and only if $\exp\{\Lambda(h)\} = 1$. If $\lambda_1(h), \lambda_2(h) \ldots, \lambda_l(h)$ are the fundamental weights of $\tilde{\mathcal{L}}$, and

$$\Lambda(h) = \sum_{j=1}^{l} m_j \lambda_j(h),$$

where $(m_1, m_2, \ldots, m_l)$ is a set of nonnegative integers, then, as $\lambda_j(h_k) = \delta_{jk}$, it follows that

$$\Lambda(h) = i\sum_{j=1}^{l}\sum_{k=1}^{l} m_j \kappa_k (\sqrt{S})_{kj}.$$

Thus $\exp h \in \mathrm{Ker}\mathbf{\Gamma}$ if and only if

$$\sum_{j=1}^{l}\sum_{k=1}^{l} m_j \kappa_k (\sqrt{S})_{kj} = 0 \pmod{2\pi}. \tag{6}$$

For cases (i) and (ii), (6) simplifies to give

$$\sum_{j=1}^{l} m_j \kappa_j = 0 \pmod{2\pi}. \tag{7}$$

By virtue of condition (A), $\mathbf{\Gamma}$ provides a single-valued and faithful representation of $\hat{\mathcal{G}}$ if and only if $\mathrm{Ker}\mathbf{\Gamma}$ consists only of the identity. Of course such a representation provides an explicit realization of $\hat{\mathcal{G}}$. As shown in the Appendix, for all $\hat{\mathcal{G}}$ [except those corresponding to $\mathcal{G} = SO(4n-2p, 2p), p = 0,1,\ldots,[n]$, to $\mathcal{G} = ND_{4n}$ and to $\mathcal{G} = SO(4n, C)$], there exists an *irreducible* single-valued faithful representation of $\hat{\mathcal{G}}$. For the exceptional cases the realizations of $\hat{\mathcal{G}}$ are reducible.

In case (ii) Sirota and Solodovnikov[15] have shown that $Z(\hat{\mathcal{G}}) = Z(\hat{\mathcal{G}}_C)$. This follows from the fact that the conditions (4) are identical for $\hat{\mathcal{G}}$ and $\hat{\mathcal{G}}_C$. The same is true of the conditions (7), so that if $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$ are the representations of $\hat{\mathcal{G}}$ and $\hat{\mathcal{G}}_C$ that are obtained by exponentiation of representations of $\mathcal{L}$ and $\mathcal{L}_C$ which coincide when complexified to $\mathcal{L}$, then $\mathrm{Ker}\mathbf{\Gamma}_1 = \mathrm{Ker}\mathbf{\Gamma}_2$.

In case (iii), consider a set of nonisomorphic real Lie algebras $\mathcal{L}$ generated by chief outer automorphisms $S = Z_0 \exp(\mathrm{ad}h'')$ from the same $\mathcal{L}_C$ and with the same Z_0, but with different h''. [The set of Lie algebras corresponding to the groups $SO(2l-2p-1, 2p+1)$, $p = 0,1,\ldots,[\frac{1}{2}l]$, provide an example (cf. III)]. As $\exp(\mathrm{ad}h'')$ acts as the identity on $\mathcal{H}$, for all such Lie algebras $(\sqrt{S})_{jk} = (\sqrt{Z_0})_{jk}$, $j,k = 1,2,\ldots,l$, *independently* of h''. Hence, by (3), all the $\hat{\mathcal{G}}$ belonging to real Lie algebras $\mathcal{L}$ of this set have identical centers. Moreover, by (6), if $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$ are representations of two universal linear groups $\hat{\mathcal{G}}_1$ and $\hat{\mathcal{G}}_2$ of this set that are obtained by exponentiation of representations of real Lie algebras that coincide when complexified to $\tilde{\mathcal{L}}$, then $\mathrm{Ker}\mathbf{\Gamma}_1 = \mathrm{Ker}\mathbf{\Gamma}_2$.

The universal linear groups $\hat{\mathcal{G}}$, their centers $Z(\hat{\mathcal{G}})$, and the kernels $\mathrm{Ker}\mathbf{\Gamma}$ of irreducible representations are listed in the Appendix for all the connected linear simple *real* Lie groups.

For the *complex* simple Lie groups $\mathcal{G}$ [i.e., those of type (iv)] some further analysis is required. Mal'cev[10] showed that for these groups $\hat{\mathcal{G}} \approx \tilde{\mathcal{G}}$. It will now be demonstrated that $Z(\hat{\mathcal{G}})$ and $\mathrm{Ker}\mathbf{\Gamma}$ are essentially determined by the corresponding quantities belonging to the maximal compact simple subgroup of $\mathcal{G}$.

Suppose then that $\mathcal{G}$ is a connected complex simple Lie group whose real Lie algebra $\mathcal{L}$ is generated by a chief outer involutive automorphism S from the semi-simple compact Lie algebra $\mathcal{L}_C = \mathcal{L}_{1C} \oplus \mathcal{L}_{2C}$, where $\mathcal{L}_{1C}$ and $\mathcal{L}_{2C}$ are isomorphic simple compact Lie algebras, and that $\mathcal{K}$ is a simple compact Lie group having $\mathcal{L}_{1C}$ as its Lie algebra. (Then $\mathcal{K}$ is homomorphic to the maximal compact subgroup of $\mathcal{G}$). It is convenient to change the previous convention and assume that $\mathcal{L}$ has rank $2l$ (instead

of l as previously), so that the complexification $\tilde{L}_1$ of L_{1C} has rank l. [As an example, for $G=SL(l+1,C)$, $L=A_l\oplus A_l$, and $K=SU(l+1)$].

Let $ih_1, ih_2, \ldots, ih_l$ be the basis of $H\cap L_{1C}$ and $ih_{l+1}, ih_{l+2}\ldots, ih_{2l}$ be the basis of $H\cap L_{2C}$. As shown in IV,

$$Sh_j=h_{j+l},\quad Sh_{j+l}=h_j,\quad j=1,2,\ldots,l,$$

so that the generators $h'_j (j=1,2,\ldots,2l)$ of $H\cap L$ are given according to (1) and (2) by

$$\left.\begin{aligned} h'_j &= \tfrac{1}{2}(1+i)ih_j+\tfrac{1}{2}(1-i)ih_{j+l}\\ h_{j+l} &= \tfrac{1}{2}(1-i)ih_j+\tfrac{1}{2}(1+i)ih_{j+l}\end{aligned}\right\},\quad j=1,2,\ldots,l. \tag{8}$$

Consider the representation Γ of $\tilde{L}=\tilde{L}_1\oplus\tilde{L}_2$ that corresponds to the direct product representation $\Gamma_1\otimes\Gamma_2$ of the direct product Lie groups corresponding to $L_C=L_{1C}\oplus L_{2C}$. In this representation

$$\Gamma(h_j)=\begin{cases}\Gamma_1(h_j)\otimes \mathrm{I}_2, & j=1,2,\ldots,l,\\ \mathrm{I}_1\otimes\Gamma_2(h_{j-l}), & j=l+1,\ldots,2l,\end{cases} \tag{9}$$

where I_1, I_2 are unit matrices of the dimensions of Γ_1 and Γ_2 respectively. An explicit matrix representation of L of the same dimension as that given for $\tilde{L}_1$ in I and V can now be obtained as follows. In the special case of (9) in which Γ_2 is the trivial one-dimensional representation in which $\Gamma_2(h_{j-l})=(0)$ and in which $\Gamma_1(h_j)=\mathrm{h}_j$, (8) and (9) imply that

$$\Gamma(h'_j)=\begin{cases}\tfrac{1}{2}(1+i)i\mathrm{h}_j, & j=1,2,\ldots,l,\\ \tfrac{1}{2}(1-i)i\mathrm{h}_{j-l}, & j=l+1,\ldots,2l.\end{cases} \tag{10}$$

Every element of $Z(\hat{G})$ must have the form $\exp(\sum_{j=1}^{2l}\kappa_j h'_j)$, where $\kappa_1,\kappa_2,\ldots,\kappa_{2l}$ are all real. Clearly $\exp(\sum_{j=1}^{2l}\kappa_j h'_j)\in Z(\hat{G})$ if and only if $\exp(\sum_{j=1}^{2l}\kappa_j\Gamma(h'_j))$ is a multiple of the unit matrix.

From (10),

$$\sum_{j=1}^{2l}\kappa_j\Gamma(h'_j)=\sum_{j=1}^{l}i\kappa_j^C\mathrm{h}_j,$$

where

$$\kappa_j^C=\tfrac{1}{2}(\kappa_j+\kappa_{j+l})+\tfrac{1}{2}i(\kappa_j-\kappa_{j+l}),$$

so the condition is that $\exp(\sum_{j=1}^{l}i\kappa_j^C\mathrm{h}_j)$ must be a multiple of the unit matrix. However, this is precisely the condition (4) for $\exp(\sum_{j=1}^{l}\kappa_j^C h_j)$ to be a member of $Z(\hat{K})$, where $\hat{K}$ is the universal linear group of K. Moreover, (4) only has solutions for κ_j^C real, so

$$\kappa_j=\kappa_{j+l}=\kappa_j^C,$$

and consequently

$$\sum_{j=1}^{2l}\kappa_j h'_j=\sum_{j=1}^{l}i\kappa_j^C(h_j+h_{j+l}).$$

Thus we have proved that $Z(\hat{G})$ is isomorphic to $Z(\hat{K})$, and the elements of $Z(\hat{G})$ are all of the form $\exp[\sum_{j=1}^{l}i\kappa_j^C(h_j+h_{j+l})]$, where the elements of $Z(\hat{K})$ are of the form $\exp(\sum_{j=1}^{l}i\kappa_j^C h_j)$. These latter elements are listed explicitly in the Appendix.

As an example consider the connected complex simple Lie group $G=SL(l+1,C)$. Then $K=SU(l+1)$, and, as shown in the Appendix, $Z(\hat{K})=Z_{l+1}(\exp\sum_{j=1}^{l}2\pi ijh_j/(l+1))$. Thus $Z(\hat{G})=Z_{l+1}(\exp\sum_{j=1}^{l}2\pi ij(h_j+h_{j+l})/(l+1))$. Moreover, it is easily verified directly that $Z(G)=Z(\hat{G})$, and so $G\approx\hat{G}$.

Every irreducible representation of a connected simple complex group G can be obtained from (8) and (9) by exponentiation. Suppose that Γ_1 and Γ_2 are irreducible representations of L_1 with highest weights $\Lambda_1(h)$ and $\Lambda_2(h)$, respectively, and that

$$\Lambda_1(h)=\sum_{j=1}^{l}m_{1j}\lambda_j(h),\quad \Lambda_2(h)=\sum_{j=1}^{l}m_{2j}\lambda_j(h),$$

where $\lambda_j(h)$ are the fundamental weights of $\tilde{L}_1$ and $(m_{11},\ldots,m_{1l})$ and $(m_{21},\ldots,m_{2l})$ are two sets of nonnegative integers. Then

$$\exp\left[\Gamma\left(\sum_{j=1}^{l}i\kappa_j^C(h_j+h_{j+l})\right)\right]$$
$$=\exp\left(\sum_{j=1}^{l}i\kappa_j^C[\Gamma_1(h_j)\otimes\mathrm{I}_2+\mathrm{I}_1\otimes\Gamma_2(h_j)]\right.$$

But as $\exp(\sum_{j=1}^{l}i\kappa_j^C h_j)\in Z(\hat{K})$, then, as in (5),

$$\exp\left(\sum_{j=1}^{l}i\kappa_j^C\Gamma_k(h_j)\right)$$
$$=\left[\exp\left(\sum_{j=1}^{l}i\kappa_j^C\Lambda_k(h_j)\right)\right]\mathrm{I}_k,\quad k=1,2,$$

and hence

$$\exp\left[\Gamma\left(\sum_{j=1}^{l}i\kappa_j^C(h_j+h_{j+l})\right)\right]$$
$$=\left[\exp\left(\sum_{j=1}^{l}i\kappa_j^C[\Lambda_1(h_j)+\Lambda_2(h_j)]\right)\right]\mathrm{I}_{12},$$

where I_{12} is the unit matrix of the dimension of Γ. *Thus the element* $\exp[\sum_{j=1}^{l}i\kappa_j^C(h_j+h_{j+1})]$ *of* $Z(\hat{G})$ is a *member of* KerΓ *if and only if*

$$\sum_{j=1}^{l}\kappa_j^C(m_{1j}+m_{2j})=0\ (\mathrm{mod}\,2\pi).$$

This should be compared with the condition (7) for KerΓ in $\hat{K}$.

III. CONSTRUCTION OF SUBGROUPS

Suppose that G and G' are two connected linear semisimple Lie groups, that L and L' are their real Lie algebras, and that L and $\tilde{L}'$ are the complexifications of L and L'. Clearly a *necessary* condition for G' to be a subgroup of G is that L' is a subalgebra of L. The conditions for L' to be a subalgebra of L have been investigated in great detail in Papers I to V. In I the general criterion was derived, and theory was developed for the case in which both L and L' were simple classical Lie algebras generated from their compact forms by *inner* involutive automorphisms. In Paper II the theory was extended to the case in which L or L' was a real form of A_l that was generated by an *outer* involutive automorphism. The corresponding case with A_l replaced by D_l was considered in Paper III, while in IV the situation was investigated in which L or L' was semisimple (but not simple). The study was concluded in Paper V by an examination of the exceptional Lie algebras.

Obviously the condition that $\mathcal{L}'$ be a subalgebra of $\mathcal{L}$ is not a *sufficient* condition on its own. However, it can be complemented by the following construction:

(1) Let $\mathcal{G}$ be *any* connected linear semisimple Lie group having $\mathcal{L}$ as its real Lie algebra, and suppose that $\mathcal{N}$ is a subgroup of $Z(\hat{\mathcal{G}})$ such that $\mathcal{G} \approx \hat{\mathcal{G}}/\mathcal{N}$. Choose a representation $\mathbf{\Gamma}$ of $\hat{\mathcal{G}}$ such that $\mathrm{Ker}\mathbf{\Gamma} = \mathcal{N}$.

(2) Construct the representation Γ of $\mathcal{L}$ such that $\mathbf{\Gamma}(\exp a) = \exp\Gamma(a)$ for all $a \in \mathcal{L}$. (There will be no confusion in denoting this representation of $\mathcal{L}$ by the same symbol as that used for the corresponding representation of $\hat{\mathcal{G}}$).

(3) Assuming $\mathcal{L}'$ to be a subalgebra of $\mathcal{L}$, construct the representation $\mathbf{\Gamma}'$ of $\mathcal{L}'$ such that $\mathbf{\Gamma}'(a') = \mathbf{\Gamma}(a')$ for all $a' \in \mathcal{L}'$.

(4) Construct the representation $\mathbf{\Gamma}'$ of the universal linear group $\hat{\mathcal{G}}'$ of $\mathcal{L}'$ such that $\mathbf{\Gamma}'(\exp a') = \exp\mathbf{\Gamma}'(a')$ for all $a' \in \mathcal{L}'$.

(5) *Then*

$$\mathcal{G}' \approx \hat{\mathcal{G}}'/\mathrm{Ker}\mathbf{\Gamma}'$$

is a subgroup of $\mathcal{G}$ *that corresponds to the embedding of* $\mathcal{L}'$ *in* $\mathcal{L}$. *For a set of conjugate embeddings of* $\mathcal{L}'$ in $\mathcal{L}$ *the subgroup* $\mathcal{G}'$ *is unique* (*up to conjugacy*).

{The proofs of the statements in (5) are straightforward. They depend on the fact that the representation Γ of $\hat{\mathcal{G}}$ constructed in stage (1) provides [by condition (A) of Sec. II] a faithful and single-valued representation of $\mathcal{G}$.

Suppose first that $\mathcal{G}'$ *is* a subgroup of $\mathcal{G}$. Then the representation $\Gamma_{\mathcal{G}'}$ of $\mathcal{G}'$ defined by $\Gamma_{\mathcal{G}'}(u') = \mathbf{\Gamma}(u')$ for all $u' \in \mathcal{G}'$ is a faithful and single-valued representation of $\mathcal{G}'$. Define $\mathbf{\Gamma}_{\mathcal{G}'}$ on $\mathcal{L}'$ by $\mathbf{\Gamma}_{\mathcal{G}'}(\exp a') = \exp\mathbf{\Gamma}_{\mathcal{G}'}(a')$ for all $a' \in \mathcal{L}'$. Then $\mathbf{\Gamma}_{\mathcal{G}'}$ is identical to $\mathbf{\Gamma}'$ on $\mathcal{L}'$. However, by condition (A) of section 2, $\mathbf{\Gamma}'$ provides a faithful and single-valued representation of only one group having Lie algebra $\mathcal{L}'$, namely $\mathcal{G}'/\mathrm{Ker}\mathbf{\Gamma}'$. Thus $\mathcal{G}' = \hat{\mathcal{G}}'/\mathrm{Ker}\Gamma'$.

Conversely, if $\mathcal{L}'$ is a subalgebra of $\mathcal{L}$ and $\mathcal{G}' \approx \hat{\mathcal{G}}'/\mathrm{Ker}\mathbf{\Gamma}'$, then $\mathbf{\Gamma}'$ is a faithful and single-valued representation of $\mathcal{G}'$. By construction the matrices of the representation $\mathbf{\Gamma}'$ of $\mathcal{G}'$ are a *subset* of the matrices of the representation $\mathbf{\Gamma}$ of $\mathcal{G}$. As the former group of matrices is isomorphic to $\mathcal{G}'$ and the latter group of matrices is isomorphic to $\mathcal{G}$, $\mathcal{G}'$ must be a subgroup of $\mathcal{G}$.}

For a given set of conjugate embeddings of $\mathcal{L}'$ in $\mathcal{L}$, this process can be repeated for every group $\mathcal{G}$ having $\mathcal{L}$ as its real Lie algebra. A different choice of $\mathcal{G}$ will correspond to a different $\mathcal{N}$, which may then lead to a different $\mathbf{\Gamma}'$ and thence to a different $\mathcal{G}' \approx \hat{\mathcal{G}}'/\mathrm{Ker}\mathbf{\Gamma}'$.

If $\mathcal{L}'$ has more than one set of conjugate embeddings in $\mathcal{L}$, then the above process can be repeated for each such set. Different sets of subgroups may emerge because the reduction in stage (3) may not be the same for different sets of conjugate embeddings. The simplest example of this behavior is provided by the embedding of compact groups corresponding to A_1 in compact groups belonging to A_2. There are two non-conjugate embeddings of A_1 in A_2, one of which leads to the embedding of $SU(2)$ in $SU(3)$ and the other to the embedding of $SO(3)$ ($\approx SU(2)/Z(SU(2))$ in $SU(3)$.

In stage (3) of the above construction it is easily demonstrated that if the representations $\mathbf{\Gamma}$ and $\mathbf{\Gamma}'$ are extended from the real Lie algebras $\mathcal{L}$ and $\mathcal{L}'$ to the complex Lie algebras $\tilde{\mathcal{L}}$ and $\tilde{\mathcal{L}}'$ in the usual way, then $\mathbf{\Gamma}'(a') = \mathbf{\Gamma}(a')$ for all $a' \in \mathcal{L}'$ if and only if $\mathbf{\Gamma}'(b') = \mathbf{\Gamma}(b')$ for all $b' \in \tilde{\mathcal{L}}'$. Thus stage (3) reduces immediately to an application of the branching rules for irreducible representations of simple *complex* Lie algebras, which have been very extensively studied, particularly by Strauman,[17] Whippman,[18] Navon and Patera,[19] Delaney and Gruber,[20] Devi,[21] Stone,[22] Wong,[23] and Patera and Sankoff.[24] The other nontrivial stages (1) and (5) simply involve the application of the information given in the appendix.

It should be noted that every subgroup $\mathcal{G}'$ of $\mathcal{G}$ constructed by the above process is *topologically closed* in $\mathcal{G}$. This follows from the theorem of Sirota and Solodovnikov[15] (p. 93) that an arbitrary connected semisimple subgroup $\mathcal{G}'$ of a Lie group $\mathcal{G}$ is closed in $\mathcal{G}$ if the center $Z(\mathcal{G})$ is finite, which it is if $\mathcal{G}$ is linear.

IV. AN EXAMPLE: $\tilde{\mathcal{L}}' = C_2$, $\tilde{\mathcal{L}} = A_3$

This simple but nontrivial example is included merely to give an explicit demonstration of the construction of Sec. III, some of the results in this example being very obvious from other considerations. To avoid confusion, the real Lie algebras corresponding to the groups $Sp(2)$, NSp_4^2, $Sp(2,R)$, $SU(4)$, $SU(2,2)$, Q_2, and $SL(4,R)$ will be denoted by $sp(2)$, nsp_4^2, $sp(2,R)$, $su(4)$, $su(2,2)$, q_2, and $sl(4,R)$ respectively. As shown in Paper I, Sec. 7.3, and Paper II, Sec. 6.3, the *only* real Lie algebra embeddings corresponding to $C_2 \subset A_3$ are $sp(2) \subset su(4)$, $nsp_4^2 \subset su(2,2)$, $sp(2,R) \subset su(2,2)$, $sp(2) \subset q_2$, $nsp_4^2 \subset q_2$, and $sp(2,R) \subset sl(4,R)$, there being only one set of conjugate embeddings in each case.

A. Group embeddings corresponding to $sp(2) \subset su(4)$, $nsp_4^2 \subset su(2,2)$, and $sp(2,R) \subset su(2,2)$

Consider first embeddings in $\mathcal{G} = \hat{\mathcal{G}}/\{e\}$. Application of the formula given in Sec. A(1) of the Appendix shows that for the IR $(0,0,1)$ $\mathrm{Ker}\mathbf{\Gamma} = \{e\}$, so in stage (1) $\mathbf{\Gamma}$ may be taken to be $(0,0,1)$. Then (cf. Patera and Sankoff[24]) $\mathbf{\Gamma}'$ is $(1,0)$, so by Sec. C of the Appendix $\mathrm{Ker}\mathbf{\Gamma}' = \{e\}$. Thus the embeddings in $\mathcal{G} = \hat{\mathcal{G}}/\{e\}$ are

$$Sp(2) \subset SU(4), \quad NSp_4^2 \subset SU(2,2), \quad Sp(2,R) \subset SU(2,2).$$

Now consider embeddings in $\mathcal{G} = \hat{\mathcal{G}}/Z_2(\exp\pi i(h_1 + 2h_2 + 3h_3))$. As $\mathrm{Ker}\mathbf{\Gamma} = Z_2(\exp\pi i(h_1 + 2h_2 + 3h_3))$ for the IR $(0,1,0)$, this is the appropriate choice in stage (1). Then $\mathbf{\Gamma}'$ is $(0,1) \oplus (0,0)$, so that $\mathrm{Ker}\mathbf{\Gamma}' = Z_2(\exp\pi i h_1)$, and the corresponding embeddings are

$$SO(5) \approx Sp(2)/Z_2(\exp\pi i h_1) \subset SU(4)/Z_2(\exp\pi i(h_1 + 2h_2 + 3h_3)) \approx SO(6),$$

$$SO(4,1) \approx NSp_4^2/Z_2(\exp\pi i h_1) \subset SU(2,2)/Z_2(\exp\pi i(h_1 + 2h_2 + 3h_3)) \approx SO(4,2),$$

$$SO(3,2) \approx Sp(2,R)/Z_2(\exp\pi i h_1) \subset SU(2,2)/Z_2(\exp\pi i(h_1 + 2h_2 + 3h_3)) \approx SO(4,2).$$

Finally, consider embeddings in $\mathcal{G}=\hat{\mathcal{G}}/Z_4(\exp\frac{1}{2}\pi i(h_1+2h_2+3h_3))$. As $\mathrm{Ker}\Gamma=Z_4(\exp\frac{1}{2}\pi i(h_1+2h_2+3h_3))$ for the IR $(1,0,1)$, this is the appropriate choice in stage (1). Then Γ' is $(2,0)\oplus(0,1)$, so that $\mathrm{Ker}\Gamma'=Z_2(\exp\pi i h_1)$, and the corresponding embeddings are

$$SO(5)\approx Sp(2)/Z_2(\exp\pi i h_1)$$
$$\subset SU(4)/Z_4(\exp\tfrac{1}{2}\pi i(h_1+2h_2+3h_3)),$$
$$SO(4,1)\approx NSp_4^2/Z_2(\exp\pi i h_1)$$
$$\subset SU(2,2)/Z_4(\exp\tfrac{1}{2}\pi i(h_1+2h_2+3h_3)),$$
$$SO(3,2)\approx Sp(2,R)/Z_2(\exp\pi i h_1)$$
$$\subset SU(2,2)/Z_4(\exp\tfrac{1}{2}\pi i(h_1+2h_2+3h_3)).$$

B. Group embeddings corresponding to $sp(2)\subset q_2$, $nsp_4^2\subset q_2$, and $sp(2,R)\subset s\ell(4,R)$

First consider embeddings in $\mathcal{G}=\hat{\mathcal{G}}/\{e\}$. Application of the formula given in Sec. A(3) of the Appendix shows that $\mathrm{Ker}\Gamma=\{e\}$ for the IR$(0,0,1)$, so this is again an appropriate choice for Γ. Then, as above, Γ' is $(0,1)$ and $\mathrm{Ker}\Gamma'=\{e\}$, so that the embeddings in $\hat{\mathcal{G}}$ are

$$Sp(2)\subset Q_2,\quad NSp_4^2\subset Q_2,\quad Sp(2,R)\subset SL(4,R).$$

For embeddings in $\mathcal{G}=\hat{\mathcal{G}}/Z_2(\exp\pi i(h_1+h_3))$, as $\mathrm{Ker}\Gamma=Z_2(\exp\pi i(h_1+h_3))$ for the IR $(0,1,0)$, and as Γ' $(0,1)\oplus(0,0)$, so that $\mathrm{Ker}\Gamma'=Z_2(\exp\pi i h_1)$, the corresponding embeddings are

$$SO(5)\approx Sp(2)/Z_2(\exp\pi i h_1)$$
$$\subset Q_2/Z_2(\exp\pi i(h_1+h_3))\approx SO(5,1),$$
$$SO(4,1)\approx NSp_4^2/Z_2(\exp\pi i h_1)$$
$$\subset Q_2/Z_2(\exp\pi i(h_1+h_3))\approx SO(5,1),$$
$$SO(3,2)\approx Sp(2,R)/Z_2(\exp\pi i h_1)$$
$$\subset SL(4,R)/Z_2(\exp\pi i(h_1+h_3)\approx SO(3,3).$$

APPENDIX: THE UNIVERSAL LINEAR GROUPS $\hat{\mathcal{G}}$, THEIR CENTERS $Z(\hat{\mathcal{G}})$, AND THE KERNELS $\mathrm{Ker}\Gamma$ OF IRREDUCIBLE REPRESENTATIONS Γ IN $\hat{\mathcal{G}}$ FOR THE CONNECTED LINEAR SIMPLE REAL LIE GROUPS

The notations for the simple real Lie groups $\mathcal{G}$ are those given in Sec. 4 of Paper I[1] and in Secs. 3 and 4 of Paper V.[5] In every case only the *connected* part of the group is being considered. The conventions for the elements h_j of $\mathcal{H}$ are as given in Appendix A of Paper I for the classical Lie algebras and in Sec. 2 and the Appendix of Paper V for the exceptional Lie algebras. The trivial group consisting only of the identity will be denoted by $\{e\}$. The expression $[a]$ denotes the largest integer not greater than a.

A. Groups for which $\tilde{\mathcal{L}}=A_l$

(1) $\mathcal{G}=SU(l+1)$ or $SU(l+1-p,p)$, $p=1,2,\ldots,[\frac{1}{2}(l+1)]$:

$$Z(\hat{\mathcal{G}})=Z_{l+1}\left(\exp\sum_{j=1}^{l}2\pi ijh_j/(l+1)\right)\approx Z(\mathcal{G}),$$

so that $\mathcal{G}\approx\hat{\mathcal{G}}$. The element $\exp[s\sum_{j=1}^{l}2\pi ijh_j/(l+1)]$ is a member of $\mathrm{Ker}\Gamma$ if and only if $s\sum_{j=1}^{l}jm_j/(l+1)$ is an integer.

(2) $\mathcal{G}=SL(l+1,R)$ for l even:

$$Z(\hat{\mathcal{G}})=\{e\}\approx Z(\mathcal{G}),$$
so that $\mathcal{G}\approx\hat{\mathcal{G}}$ and $\mathrm{Ker}\Gamma=\{e\}$.

(3) $\mathcal{G}=SL(l+1,R)$ or $Q_{(l+1)/2}$ for l odd:

$$Z(\hat{\mathcal{G}})=Z_2\left(\exp\sum_{j\,\mathrm{odd}}\pi ih_j\right)\approx Z(\mathcal{G}),\quad\text{so that } \mathcal{G}\approx\hat{\mathcal{G}}.$$

Then

$$\mathrm{Ker}\Gamma=\begin{cases}\{e\}, & \text{if } \sum_{j\,\mathrm{odd}}m_j \text{ is odd.}\\ Z(\hat{\mathcal{G}}), & \text{if } \sum_{j\,\mathrm{odd}}m_j \text{ is even.}\end{cases}$$

B. Groups for which $\tilde{\mathcal{L}}=B_l$

$\mathcal{G}=SO(2l+1)$ or $SO(2l+1-2p,2p)$, $p=1,2,\ldots,l$:

$$Z(\hat{\mathcal{G}})=Z_2(\exp\pi ih_l)\text{ while } Z(\mathcal{G})=\{e\},$$
so that $\mathcal{G}\approx\hat{\mathcal{G}}/Z(\hat{\mathcal{G}})$.

Then

$$\mathrm{Ker}\Gamma=\begin{cases}\{e\}, & \text{if } m_l \text{ is odd,}\\ Z(\hat{\mathcal{G}}), & \text{if } m_l \text{ is even,}\end{cases}$$

so that an explicit realization of $\hat{\mathcal{G}}$ is provided by the IR $(0,0,0,\ldots,0,1)$.

C. Groups for which $\tilde{\mathcal{L}}=C_l$

$\mathcal{G}=Sp(l)$, NSp_{2l}^{2p}, $p=1,2,\ldots,[\frac{1}{2}l]$, or $Sp(l,R)$:

$$Z(\hat{\mathcal{G}})=Z_2\left(\exp\sum_{j\,\mathrm{odd}}\pi ih_j\right)\approx Z(\mathcal{G}),\quad\text{so that } \mathcal{G}\approx\hat{\mathcal{G}}.$$

Then

$$\mathrm{Ker}\Gamma=\begin{cases}\{e\}, & \text{if } \sum_{j\,\mathrm{odd}}m_j \text{ is odd,}\\ Z(\hat{\mathcal{G}}), & \text{if } \sum_{j\,\mathrm{odd}}m_j \text{ is even.}\end{cases}$$

D. Groups for which $\tilde{\mathcal{L}}=D_l$, $l\geqslant 2$

(1) $\mathcal{G}=SO(2l)$, $SO(2l-2p,2p)$, $p=1,2,\ldots,[\frac{1}{2}l]$, or ND_{2l}: There are two cases:

(i) l odd:

$Z(\hat{\mathcal{G}})=Z_4(\exp h)$ and $\mathcal{G}\approx\hat{\mathcal{G}}/Z_2(\exp 2h)$, where $h=\pi i(h_1+h_3+h_5+\cdots+h_{l-2}+\frac{1}{2}h_{l-1}-\frac{1}{2}h_l)$. Also $Z(\mathcal{G})=Z_2(-\mathrm{I})$. Then

$$\mathrm{Ker}\Gamma=\begin{cases}\{e\}, & \text{if } \phi(m)\neq\text{integer},\\ Z_2(\exp 2h), & \text{if } \phi(m) \text{ is odd},\\ Z_4(\exp h), & \text{if } \phi(m) \text{ is even},\end{cases}$$

where

$$\phi(m)=m_1+m_3+m_5+\cdots+m_{l-2}+\tfrac{1}{2}m_{l-1}-\tfrac{1}{2}m_l.$$

Thus the IR $(0,0,\ldots,0,1)$ provides an explicit realization of $\hat{\mathcal{G}}$.

(ii) l even:

$$Z(\hat{\mathcal{G}})=Z_2(\exp h)\otimes Z_2(\exp h')\text{ and }\mathcal{G}\approx\hat{\mathcal{G}}/Z_2(\exp h'),$$

where $h=\pi i\sum_{j\,\mathrm{odd}}h_j$ and $h'=\pi i(h_{l-1}+h_l)$. Also $Z(\mathcal{G})=Z_2(-\mathrm{I})$.

Then

$$\mathrm{Ker}\Gamma=\begin{cases} Z_2(\exp h), & \text{if } \sum_{j\,\mathrm{odd}} m_j \text{ is even and } (m_{l-1}+m_l) \text{ is odd},\\ Z_2(\exp h'), & \text{if } \sum_{j\,\mathrm{odd}} m_j \text{ is odd and } (m_{l-1}+m_l) \text{ is even},\\ Z(\hat{\mathcal{G}}) & \text{if } \sum_{j\,\mathrm{odd}} m_j \text{ is even and } (m_{l-1}+m_l) \text{ is even},\\ Z_2(\exp(h+h')), & \text{if } \sum_{j\,\mathrm{odd}} m_j \text{ is odd and } (m_{l-1}+m_l) \text{ is odd}. \end{cases}$$

As $\mathrm{Ker}\Gamma\neq\{e\}$ for any irreducible representation, no irreducible representation provides a realization of $\hat{\mathcal{G}}$. However the reducible representation that is the direct sum of the IR's $(0,0,\ldots,0,1)$ and $(1,0,\ldots,0)$ provides a realization.

(2) $\mathcal{G}=SO(2l-2p-1,2p+1)$, $p=0,1,\ldots,[\frac{1}{2}l]$:

$Z(\hat{\mathcal{G}})=Z_2(\exp\pi i(h_{l-1}+h_l))$, while $Z(\mathcal{G})=\{e\}$ (as $-\mathbf{I}$ is not in the *connected* group $\mathcal{G}$). Thus $\mathcal{G}\approx\hat{\mathcal{G}}/Z(\hat{\mathcal{G}})$, and

$$\mathrm{Ker}\Gamma=\begin{cases}\{e\}, & \text{if } m_{l-1}+m_l \text{ is odd},\\ Z(\hat{\mathcal{G}}), & \text{if } m_{l-1}+m_l \text{ is even}.\end{cases}$$

Thus the IR $(0,0,\ldots,0,1)$ provides an explicit realization of $\hat{\mathcal{G}}$.

E. Groups for which $\tilde{L}=G_2$

$\mathcal{G}=CG_2$ or NG_2:

$Z(\hat{\mathcal{G}})=\{e\}\approx Z(\mathcal{G})$,

so that

$\mathcal{G}\approx\hat{\mathcal{G}}$ and $\mathrm{Ker}\Gamma=\{e\}$.

F. Groups for which $\tilde{L}=F_4$

$\mathcal{G}=CF_4$, NF_4^1, or NF_4^2:

$Z(\hat{\mathcal{G}})=\{e\}\approx Z(\mathcal{G})$,

so that

$\mathcal{G}=\hat{\mathcal{G}}$ and $\mathrm{Ker}\Gamma=\{e\}$.

G. Groups for which $\tilde{L}=E_6$

(1) $\mathcal{G}=CE_6$, NE_6', or NE_6^2:

$Z(\hat{\mathcal{G}})=Z_3(\exp\frac{2}{3}\pi i(h_1+2h_2+4h_4-h_5))\approx Z(\mathcal{G})$,

so that

$\mathcal{G}\approx\hat{\mathcal{G}}$.

Thus

$$\mathrm{Ker}\Gamma=\begin{cases}\{e\}, & \text{if } m_1+2m_2+4m_4-m_5\neq 0 \pmod 3,\\ Z(\hat{\mathcal{G}}), & \text{if } m_1+2m_2+4m_4-m_5=0 \pmod 3.\end{cases}$$

(2) $\mathcal{G}=NE_6^3$ or NE_6^4:

$Z(\hat{\mathcal{G}})=\{e\}\approx Z(\mathcal{G})$,

so that

$\mathcal{G}\approx\hat{\mathcal{G}}$ and $\mathrm{Ker}\Gamma=\{e\}$.

H. Groups for which $\tilde{L}=E_7$

$\mathcal{G}=CE_7$, NE_7^1, NE_7^2, or NE_7^3:

$Z(\hat{\mathcal{G}})=Z_2(\exp\pi i(h_4+h_6+h_7))\approx Z(\mathcal{G})$,

so that

$\mathcal{G}\approx\hat{\mathcal{G}}$.

Thus

$$\mathrm{Ker}\Gamma=\begin{cases}\{e\}, & \text{if } m_4+m_6+m_7 \text{ is odd},\\ Z(\mathcal{G}), & \text{if } m_4+m_6+m_7 \text{ is even}.\end{cases}$$

I. Groups for which $\tilde{L}=E_8$

$\mathcal{G}=CE_8$, NE_8^1, or NE_8^2:

$Z(\hat{\mathcal{G}})=\{e\}\approx Z(\mathcal{G})$,

so that

$\mathcal{G}\approx\hat{\mathcal{G}}$ and $\mathrm{Ker}\Gamma=\{e\}$.

*Present address: Faculty of Mathematics, The Open University, Milton Keynes, Bucks., England.

[1] J.F. Cornwell, Repts. Math. Phys. 2, 239 (1971).
[2] J.F. Cornwell, Repts. Math. Phys. 2, 289 (1971).
[3] J.F. Cornwell, Repts. Math. Phys. 3, 91 (1972).
[4] J.M. Ekins and J.F. Cornwell, Repts. Math. Phys. 5, 17 (1974).
[5] J.M. Ekins and J.F. Cornwell (to be published).
[6] M.E. Cartan, Ann. l'Ecole Norm. Sup. 3-eme Sér. 31, 263 (1914).
[7] M.E. Cartan, J. Math. Pures Appl. 8, 1 (1929).
[8] F. Gantmacher, Rec. Math. (Mat. Sbornik) N.S. 5 (47), 101 (1939).
[9] F. Gantmacher, Rec. Math. (Mat. Sbornik) N.S. 5 (47), 217 (1939).
[10] A.I. Mal'cev, Dokl. Akad. Nauk. SSSR 40, 108 (1943).
[11] E.B. Dynkin and A.L. Oniscik, Usp. Mat. Nauk 10, 3 (1953) [Amer. Math. Soc. Trans., Ser. 2, 21, 119 (1962)]
[12] A.I. Sirota, Dokl. Akad. Nauk SSSR 134, 44 (1960) [Sov. Math. Dokl. 1, 1021 (1960)].
[13] A.I. Sirota, Dokl. Adad. Nauk SSSR 134, 540 (1960) [Sov. Math. Dokl. 1, 1118 (1960)].
[14] A.S. Solodovnikov, Dokl. Akad. Nauk SSSR 129, 272 (1959).
[15] A.I. Sirota and A.S. Solodovnikov, Usp. Mat. Nauk 18, 85 (1963) [Russ. Math. Surveys 18, 87 (1963)].
[16] N. Jacobson, *Lie Algebras* (Interscience, New York, 1962), p. 121.
[17] N. Straumann, Helv. Acta Phys. 38, 481 (1965).
[18] M. L. Whippman, J. Math. Phys. 6, 1534 (1965).
[19] A. Navon and J. Patera, J. Math. Phys. 8, 489 (1967).
[20] R.M. Delaney and B. Gruber, J. Math. Phys. 10, 252 (1969).
[21] V.S. Devi, Nuovo Cimento A 68, 1 (1970).
[22] A.D. Stone, J. Math. Phys. 11, 29 (1970).
[23] M.K.F. Wong, J. Math. Phys. 11, 1489 (1970).
[24] J. Patera and D. Sankoff, *Tables of Branching Rules for Representations of Simple Lie Algebras* (Presses Université de Montréal, 1973).

The representation of the $SO(4,1)$ group in four-dimensional Euclidean and spinor space

R. Bogdanović and M. A. Whitehead

Department of Chemistry, McGill University, Montreal, Canada H3C 3G1
(Received 27 September 1974)

An explicit representation of the $SO(4,1)$ group in both $SO(4)$ spinor space and four-dimensional Euclidean space has been found. The two spaces are related by linearly transforming the variables. It has been shown that the four-dimensional hyperspherical harmonics in both four-dimensional polar coordinate systems transform in accordance with the $W = 0$, $Q = 2$ representation of the $SO(4,1)$ group. A set of new recursion relations is derived for the $SO(3)$ group reduced rotational matrices, together with a set of standard recursion relations for the Legendre and Gegenbauer polynomials, all of which are obtained by transforming both forms of the $SO(4)$ group basis states in spinor space into four-dimensional space. The matrix elements of the noncompact $SO(4,1)$ group generators are given in the (j,m) basis.

1. INTRODUCTION

The de Sitter $SO(4,1)$ group[1] is an important subgroup of the full dynamical group of the Coulomb potential $O(4,2)$. The importance of the $SO(4,1)$ group lies in the fact that when a basis is chosen for which the $SO(4)$ subgroup is diagonal, then two classes of its unitary irreducible representation (hereafter called the representation) contain the bound states of the Coulomb potential. These bound states belong to the special class of the $SO(4)$ group representations called "square",[2] which are characterized by $j^+ = j^- = (\frac{1}{2})n$, where $n = 0, 1, 2, \cdots$, and $n+1$ is the energy quantum number of the Coulomb potential bound states; the standard notation for these representations is $[0, n+1]$.

In three-dimensional coordinate space, the bound states are the solutions to the nonrelativistic Coulomb Hamiltonian, the hydrogenic wavefunctions. The solutions to the same Hamiltonian in three-dimensional momentum space are related to the hyperspherical harmonics in four-dimensional space through Fock's transformations.[3] These hyperspherical harmonics are basis states for the square representations of the $SO(4)$ group.

The main concern of this paper is the relation of the hyperspherical harmonics to the representation of the $SO(4,1)$ group in four-dimensional space. Since the noncompact generators of this group act as ladder operators for n the energy quantum number, this representation of the $SO(4,1)$ group has two important applications:

(i) It can be used to study the transformation properties of the radial parts of the hydrogenic wavefunctions, which are still not fully understood.[4]

(ii) The representation is relevant to a group theoretical description of an atom in an external magnetic field, when the $SO(4)$ symmetry is broken. The group which leaves the magnetic quantum number, m, unchanged is the $SO(2,1)\times SO(2)$ group in the $SO(4,1)$ group framework; if the full dynamical group $SO(4,2)$ is considered, then the required subgroup is $SO(2,2)\times O(2)$. Thus instead of the usual reduction for the field-free case, $SO(4,1)\supset SO(4)$, the reduction $SO(4,1)\supset SO(2,1)\times O(2)$ is more appropriate when a field is applied. The representation of the $SO(4,1)$ group established in this paper will facilitate this reduction. Interest in (i) and (ii) motivated this research.

A. Preliminaries

The unitary irreducible representations of the $SO(4,1)$ group were derived by Thomas.[5] The representations are labelled by two invariants (p, q) which are related to the Casimir invariants (Q, W). Thomas also calculated the matrix elements of the generators.[5] The list of representations was later corrected by Newton who labelled the representations with (Q, W).[6] Additional representations have been added by Dixmier.[7]

The square representations of the $SO(4)$ group are contained in two unitary irreducible representations of the $SO(4,1)$ group, Table I. The minimum number of dimensions of space to carry all irreducible representations of the $SO(4,1)$ and $SO(4)$ groups are six and four, respectively.[8] Therefore, a representation of $SO(4,1)$ group in four-dimensional space is admissable only if *both* Casimir invariants are specified; four-dimensional space has only three independent dimensions with respect to the $SO(4)$ group, since the four-coordinates are linked through an invariant Hermitian form, and consequently one Casimir invariant has to be specified; this is usually the second invariant, q or W, and is set to zero. Therefore, only square representations occur.

TABLE I. Comparison of the notations for two classes of the $SO(4,1)$ group representations. The Eigenvalue spectrum of the two Casimir invariants for two classes of $SO(4,1)$ group representations relevant to the Coulomb problem, are shown in the notations of Thomas[5] and Newton[6]. Representation 1 contains all bound states; representation 2 contains only states for $n \geq s$.

Representation	Thomas	Newton	$SO(4)$ Content
1	$q=0$ $-\frac{1}{2}<p<1$ $p+\frac{1}{2}=i\sigma$ σ real	class I $W=0$ $0<Q<\frac{9}{4}$ $\frac{9}{4}\leq Q<\infty$	$n\geq 0$
2	$q=0$ $p=s$	class II $W=0$ $Q=-(s-1)(s+2)$ $s=1, 2, 3, \cdots$	$n\geq s$

There are two ways to label basis states in an irreducible representation of the $SO(4)$ group; each corresponds to a different linear combination of the generators. One uses the homomorphism of the $SO(4)$ group to the group $SO(2)\times SO(2)$: The basis states and relevant set of generators are then in the (m^+, m^-) scheme. The other uses the physical $SO(3)$ subgroup, and the basis states and the corresponding generators are in the (j, m) scheme. Both schemes are used.

The basis for the square representations of the $SO(4)$ group in three-dimensional space are hydrogenic wavefunctions; in addition, there are two other bases for both $SO(4)$ schemes: one in spinor space[2] and one in four-dimensional Euclidean space.[9]

B. Method and results

The notation is established in Sec. 2. The matrix elements of the generators of Thomas,[4] which are in the (m^+, m^-) scheme, are reformulated in terms of their tensor properties and then are reformulated in the (j, m) scheme (Sec. 3).

In three-dimensional space the representation of the $SO(4)$ group generators has been found in spherical coordinates, the (j, m) scheme,[10] and in parabolic coordinates, the (m^+, m^-) scheme.[11]

The spinor space and four-dimensional space are explicitly connected through a linear transformation of the variables. In addition to the already known compact generators of the $SO(4,1)$ group, the relative simplicity of spinor space permits four-noncompact $SO(4,1)$ group generators to be constructed. These generators together form a representation of the $SO(4,1)$ group in spinor space. The basis states in this spinor space transform according to representation 1 of Table I, with

$$q=0 \text{ and } p=0 \tag{1.1}$$

or

$$W=0 \text{ and } Q=2.$$

These basis states and generators will be transformed into four-dimensional Euclidean space, when the states become hyperspherical harmonics in the (j, m) scheme, and functions related to the $SO(3)$ reduced rotational matrices in the (m^+, m^-) scheme. The transformed states are the bases for the square representation of the $SO(4)$ group in four-dimensional space.[9] The transformation properties (1.1) remain the same. The noncompact $SO(4,1)$ group generators assume comparatively simple forms. A similar but different method of using spinor space has been employed to study certain representations of the Lorenz group and the $U(2,2)$ group.[12,13]

This representation, with the matrix elements of the $SO(4,1)$ group generators from Sec. 3, permit derivation of the set of standard recursion relations for the Legendre and Gengenbauer polynomials and a set of new recursion relations for the $SO(3)$ reduced rotational matrices, and prove the validity of (1.1) and of the representation of the $SO(4,1)$ group based on (1.1), Sec. 4.

C. Remark

While preparing the manuscript, our attention was drawn to the paper of Han[14] where the representation of the $SO(4,1)$ group generators in four-dimensional space is found which corresponds to the representation 1 in Table I with

$$q=0 \quad \text{and} \quad p=-1/2$$

or (1.2)

$$W=0 \quad \text{and} \quad Q=9/4$$

which differs from the representation in the present work which uses (1.1). With Han's generators, the hyperspherical harmonics in both the (j, m) and (m^+, m^-) schemes do not transform correctly, but their linear combination should. This further indicates that the choice of (1.1) for the values of the $SO(4,1)$ group Casimir invariants is natural for the hyperspherical harmonics and consequently for the hydrogenic wave functions. The relation betweeen Han's generators and the ones used herein, is given in Appendix A.

2. THE GENERATORS OF THE $SO(4,1)$ GROUP

As an attempt to promote a unified notation for continuous groups, the notation for the $O(4,2)$ group[15] has been truncated to give the notation for the $SO(4,1)$ group. The compact axes are labelled 1, 2, 3, and 5, and the noncompact axis is labelled 0.

The generators of the $SO(4,1)$ group commute between themselves

$$[L_{jk}, L_{ls}] = i(g_{jl}L_{ks} - g_{js}L_{kl} - g_{kl}L_{js} + g_{ks}L_{jl}),$$

where

$$\begin{aligned} &i=\sqrt{-1}, \\ &j,k,l,s=0,1,2,3, \text{ or } 5, \\ &g_{11}=g_{22}=g_{33}=g_{55}=1, \quad g_{00}=-1, \end{aligned} \tag{2.1}$$

and zero otherwise.

The compact generators can be represented by

$$L_{lj} = -i(x_l\partial_{x_j} - x_j\partial_{x_l}) \text{ with } l,j=1,2,3, \text{ or } 5 \tag{2.2}$$

and the noncompact generators by

$$L_{0j} = x_0'\partial_{x_j} + x_j\partial_{x_0'} \text{ with } j=1,2,3, \text{ or } 5, \tag{2.3}$$

where

$x_0 = ix_0'$ and x_0' is a real variable.

A. The generators for the (m^+, m^-) scheme

In the (m^+, m^-) scheme the generators are given by,

$$\begin{aligned} &j_\pm^+ = \tfrac{1}{2}[(L_{23} \pm iL_{31}) + (L_{15} \pm iL_{25})], \quad j_0^+ = \tfrac{1}{2}(L_{12} + L_{35}), \\ &j_\pm^- = \tfrac{1}{2}[(L_{23} \pm iL_{31}) - (L_{15} \pm iL_{25})], \quad j_0^- = \tfrac{1}{2}(L_{12} - L_{35}), \\ &P_\pm = \frac{1}{\sqrt{2}}(L_{03} \pm iL_{05}), \quad S_\pm = \pm\frac{1}{\sqrt{2}}(L_{01} \pm iL_{02}). \end{aligned} \tag{2.4}$$

The $(j_\pm^+, j_0^+)$ and $(j_\pm^-, j_0^-)$ constitute two mutually commuting $SO(3)$ subgroups with respect to which $P_\pm$ and $S_\pm$ trans-

forms like $\frac{1}{2}$ tensor; namely,

$$
\begin{aligned}
&[j_+^+, S_-] = P_+, \quad [j_+^+, P_-] = S_+,\\
&[j_-^+, P_+] = S_-, \quad [j_-^+, S_+] = P_-,\\
&[j_0^+, S_\pm] = \pm\tfrac{1}{2}S_\pm, \quad [j_0^+, P_\pm] = \pm\tfrac{1}{2}P_\pm,\\
&[j_+^+, P_+] = [j_-^+, P_-] = [j_+^+, S_+] = [j_-^+, S_-] = 0,\\
&[j_+^-, S_-] = P_-, \quad [j_+^-, P_+] = S_+,\\
&[j_-^-, P_-] = S_-, \quad [j_-^-, S_+] = P_+,\\
&[j_0^-, S_\pm] = \pm\tfrac{1}{2}S_\pm, \quad [j_0^-, P_\pm] = \pm\tfrac{1}{2}P_\pm,\\
&[j_0^-, P_-] = [j_-^-, P_+] = [j_+^-, S_+] = [j_-^-, S_-] = 0\,.
\end{aligned}
\tag{2.5}
$$

From these commutational relations it is seen that the pairs (P_+, S_-) and (S_+, P_-) form the 1/2 and $-1/2$ components of a 1/2 tensor with respect to the $\mathbf{j}^+$ subgroup, and that the pairs (P_-, S_-) and (S_+, P_+) form the 1/2 and $-1/2$ components of a 1/2 tensor with respect to the $\mathbf{j}^-$ subgroup.

B. The generators for the (j,m) scheme

In the (j, m) scheme the generators assume the standard form,

$$
\begin{aligned}
&L_\pm = L_{23} \pm iL_{31}, \quad L_0 = L_{12},\\
&K_\pm = L_{15} \pm iL_{25}, \quad K_0 = L_{35},\\
&L_{03} = \frac{1}{\sqrt{2}}(P_+ + P_-),\\
&iL_{05} = \frac{1}{\sqrt{2}}(P_+ - P_-),
\end{aligned}
$$

and $S_\pm$.

The L_+, L_-, and L_0 are generators of the $SO(3)$ subgroup with respect to which the generators K_+, K_-, and K_0 form a vector, the noncompact generators S_+, S_-, and L_{03} form a spherical vector and L_{05} is a scalar.

3. THE MATRIX ELEMENTS OF THE GENERATORS

A. The matrix elements in the (m^+, m^-) scheme

The matrix elements of the generators have been calculated by Thomas[5] in the (m^+, m^-) scheme for the whole range of eigenvalues of two Casimir operators for unitary representations of the $SO(4,1)$ group.[16]

To make the notation consistent with the tensor properties of $P_\pm$ and $S_\pm$, and for future reference, the matrix elements are reformulated as

$$
\begin{aligned}
&\left\langle \begin{matrix} j_b^+ & j_b^- \\ m_b^+ & m_b^- \end{matrix} \middle| S_\pm \middle| \begin{matrix} j_a^+ & j_a^- \\ m_a^+ & m_a^- \end{matrix} \right\rangle\\
&= \left\langle \begin{matrix} j_a^+ & \frac{1}{2} \\ m_a^+ \pm \frac{1}{2} & \end{matrix} \middle| \begin{matrix} j_b^+ \\ m_b^+ \end{matrix} \right\rangle \left\langle \begin{matrix} j_a^- & \frac{1}{2} \\ m_a^- \pm \frac{1}{2} & \end{matrix} \middle| \begin{matrix} j_b^- \\ m_b^- \end{matrix} \right\rangle \langle j_b^+ j_b^- pq \| S_\pm \| j_a^+ j_a^- pq \rangle .
\end{aligned}
\tag{3.1}
$$

Similarly, for $P_\pm$,

$$
\left\langle \begin{matrix} j_b^+ & j_b^- \\ m_b^+ & m_b^- \end{matrix} \middle| P_\pm \middle| \begin{matrix} j_a^+ & j_a^- \\ m_a^+ & m_a^- \end{matrix} \right\rangle
\tag{3.2}
$$

$$
= \left\langle \begin{matrix} j_a^+ & \frac{1}{2} \\ m_a^+ \pm \frac{1}{2} & \end{matrix} \middle| \begin{matrix} j_b^+ \\ m_b^+ \end{matrix} \right\rangle \left\langle \begin{matrix} j_a^- & \frac{1}{2} \\ m_a^- \pm \frac{1}{2} & \end{matrix} \middle| \begin{matrix} j_b^- \\ m_b^- \end{matrix} \right\rangle \langle j_b^+ j_b^- pq \| P_\pm \| j_a^+ j_a^- pq \rangle .
$$

The p and q are the eigenvalues of two Casimir invariants which specify an irreducible representation of the $SO(4,1)$ group and $\langle \circ\circ | \circ \rangle$ are Clebsh–Gordon coefficients and $\langle \cdot \| \| \circ \rangle$ are reduced matrix elements. The matrix elements are different from zero only in four cases:

$$
j_b^+ = j_a^+ \pm \tfrac{1}{2} \quad \text{and} \quad j_b^- = j_a^- \pm \tfrac{1}{2}.
\tag{3.3}
$$

B. The reduced matrix elements

For these nonzero cases in the Clebsh–Gordon coefficients, the reduced matrix elements are

$$
\begin{aligned}
&\langle j^+ + \tfrac{1}{2} j^- + \tfrac{1}{2} pq \| S_\pm \| j^+ j^- pq \rangle = \langle j^+ + \tfrac{1}{2} j^- + \tfrac{1}{2} pq \| P_\pm \| j^+ j^- pq \rangle\\
&\quad = \left[\frac{(j^+ + j^- - p + 1)(j^+ + j^- + p + 2)(j^+ + j^- + q + 2)}{2(2j^+ + 2)(2j^- + 2)} \right]^{1/2},\\
&\langle j^+ + \tfrac{1}{2} j^- - \tfrac{1}{2} pq \| S_\pm \| j^+ j^- pq \rangle = \langle j^+ + \tfrac{1}{2} j^- - \tfrac{1}{2} pq \| P_\pm \| j^+ j^- pq \rangle\\
&\quad = \left[\frac{(j^+ - j^- - p)(j^+ - j^- + p + 1)(j^+ - j^- - q)(j^+ - j^- + q + 1)}{2(2j^+ + 2)(2j^- + 2)} \right]^{1/2},\\
&\langle j^+ - \tfrac{1}{2} j^- + \tfrac{1}{2} pq \| S_\pm \| j^+ j^- pq \rangle = \langle j^+ - \tfrac{1}{2} j^- + \tfrac{1}{2} pq \| P_\pm \| j^+ j^- pq \rangle\\
&\quad = -\left[\frac{(j^+ - j^- - p - 1)(j^+ - j^- + p)(j^+ - j^- - q - 1)(j^+ - j^- + q)}{2(2j^+)(2j^-)} \right]^{1/2},\\
&\langle j^+ - \tfrac{1}{2} j^- - \tfrac{1}{2} pq \| S_\pm \| j^+ j^- pq \rangle = \langle j^+ - \tfrac{1}{2} j^- - \tfrac{1}{2} pq \| P_\pm \| j^+ j^- pq \rangle\\
&\quad = -\left[\frac{(j^+ + j^- - p)(j^+ + j^- + p + 1)(j^+ + j^- + q + 1)}{2(2j^+)(2j^-)} \right]^{1/2}.
\end{aligned}
\tag{3.4}
$$

Because the representations are unitary, the generators (L_{ij}) are Hermitian, therefore $(P_+$ and $P_-)$ are Hermitian conjugates of each other, and $(S_+$ and $S_-)$ are an anti-Hermitian conjugate pair.

C. The matrix elements in the (j, m) scheme

In the (j, m) scheme, the reduced matrix elements (3.4) are unchanged. Applying then the Wigner–Eckart theorem to the tensors (2.6), the matrix elements are

$$
\begin{aligned}
&\left\langle \begin{matrix} j_b^+ & j_b^- \\ j_b & m_b \end{matrix} \middle| S_\pm \middle| \begin{matrix} j_a^+ & j_a^- \\ j_a & m_a \end{matrix} \right\rangle = \left\langle \begin{matrix} \frac{1}{2} & \frac{1}{2} \\ \pm\frac{1}{2} \pm \frac{1}{2} & \end{matrix} \middle| \begin{matrix} 1 \\ \pm 1 \end{matrix} \right\rangle\\
&\times [3(2j_a + 1)(2j_b^+ + 1)(2j_b^- + 1)]^{1/2} \left\langle \begin{matrix} j_a & 1 \\ m_a & \pm 1 \end{matrix} \middle| \begin{matrix} j_b \\ m_b \end{matrix} \right\rangle \begin{Bmatrix} j_b & j_a & 1 \\ j_b^+ & j_a^+ & \frac{1}{2} \\ j_b^- & j_a^- & \frac{1}{2} \end{Bmatrix}\\
&\times \langle j_b^+ j_b^- pq \| S_\pm \| j_a^+ j_a^- pp \rangle,
\end{aligned}
\tag{3.5}
$$

$$
\begin{aligned}
&\left\langle \begin{matrix} j_b^+ & j_b^- \\ j_b & m_b \end{matrix} \middle| L_{03} \middle| \begin{matrix} j_a^+ & j_a^- \\ j_a & m_a \end{matrix} \right\rangle = \delta_{m_a m_b} \sqrt{2} \left\langle \begin{matrix} j_a & 1 \\ m_a & 0 \end{matrix} \middle| \begin{matrix} j_b \\ m_a \end{matrix} \right\rangle\\
&\times \left\langle \begin{matrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{matrix} \middle| \begin{matrix} 1 \\ 0 \end{matrix} \right\rangle \times [3(2j_a + 1)(2j_b^+ + 1)(2j_b^- + 1)]^{1/2} \begin{Bmatrix} j_b & j_a & 1 \\ j_b^+ & j_a^+ & \frac{1}{2} \\ j_b^- & j_a^- & \frac{1}{2} \end{Bmatrix}
\end{aligned}
$$

$$\times\langle j_b^+ j_b^- pq\|L_{03}\|j_a^+ j_a^- pq\rangle, \tag{3.6}$$

$$\left\langle \begin{matrix} j_b^+ & j_b^- \\ j_b & m_b \end{matrix} \middle| L_{05} \middle| \begin{matrix} j_a^+ & j_a^- \\ j_a & m_a \end{matrix} \right\rangle = \delta_{m_a m_b}\delta_{j_a j_b}\sqrt{2}\left\langle \begin{matrix} j_a & 0 \\ m_a & 0 \end{matrix} \middle| \begin{matrix} j_b \\ m_a \end{matrix} \right\rangle$$

$$\times\left\langle \begin{matrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{matrix} \middle| \begin{matrix} 0 \\ 0 \end{matrix} \right\rangle \frac{(-1)^{1/2+j_a^- -j_b^- -j_a^-}}{\sqrt{2}}[(2j_b^+ +1)(2j_b^- +1)]^{1/2} \tag{3.7}$$

$$\times W(j_b^+ j_a^+ j_b^- j_a^-;\ \tfrac{1}{2}j_a)\langle j_b^+ j_b^- pq\|iL_{05}\|j_a^+ j_a^- pq\rangle,$$

where

$$\langle j_b^+ j_b^- pq\|L_{03}\|j_a^+ j_a^- pq\rangle = \langle j_b^+ j_b^- pq\|iL_{05}\|j_a^+ j_a^- pq\rangle = \langle j_b^+ j_b^- pq\|P_\pm\|j_a^+ j_a^- pq\rangle$$

and $W(\cdots)$ is a Racah coefficient. The results can be verified directly by inserting the expansion

$$\left|\begin{matrix} j^+ & j^- \\ m^+ & m^- \end{matrix}\right\rangle = \sum_{j,m}\left|\begin{matrix} j^+ & j^- \\ m^+ & m^- \end{matrix}\right\rangle\left\langle\begin{matrix} j^+ & j^- \\ m^+ & m^- \end{matrix}\middle|\begin{matrix} j \\ m \end{matrix}\right\rangle \tag{3.8}$$

into Eq. (3.1).

The generators L_{03} and L_{05} leave m unchanged, and together with L_{35} they constitute the $SO(2,1)$ subgroup which commutes with L_{12}. This subgroup is of interest in calculations involving and external magnetic field. The generator L_{05} also leaves j unchanged.

4. THE REPRESENTATION OF THE $SO(4,1)$ GROUP

In this section the known forms of the square representations of the $SO(4)$ group are extended to include the $SO(4,1)$ group representation.

A. The representation of the $SO(4,1)$ group in the (m^+,m^-) scheme, and a method of constructing noncompact $SO(4,1)$ group generators

The basis states for the square representation of the $SO(4)$ group in the (m^+, m^-) scheme have been obtained by Sharp by superimposing two orthogonalized $SO(3)$ states in spinor variables (μ, θ) and (ξ, ζ).[2] For a square representation $j^+ = j^- = \frac{1}{2}n$, the basis states are

$$\left|\begin{matrix} j^+ & j^- \\ m^+ & m^- \end{matrix}\right\rangle' = \left|\begin{matrix} n \\ m^+ & m^- \end{matrix}\right\rangle' = \left[\frac{(\frac{1}{2}n+m^+)!\,(\frac{1}{2}n-m^+)!\,(\frac{1}{2}n+m^-)!\,(\frac{1}{2}n-m^-)!}{n!}\right]^{1/2}$$

$$\times\sum_{m'}\frac{\eta^{n/4+m^+/2+m'}\theta^{n/4+m^+/2-m'}\xi^{n/4-m^+/2-m'}\zeta^{n/4-m^+/2+m'}}{(\frac{1}{4}n+\frac{1}{2}m^+ +m')!\,(\frac{1}{4}n+\frac{1}{2}m^+ -m')!\,(\frac{1}{4}n-\frac{1}{2}m^+ -m')!\,(\frac{1}{4}n-\frac{1}{2}m^+ -m^- +m')!} \tag{4.1}$$

where $-j^\pm \le m^\pm \le j^\pm$, and m' is restricted to nonnegative factorials.

The square representations of the $SO(4)$ group appear in the representations of the $SO(4,1)$ group only with $q=0$ (Table I), and only two matrix elements of the noncompact generators differ from zero, (3.1) and (3.3) for $n_b = n_a \pm 1$. For the noncompact generators to commute with the compact ones according to (2.5) they have to contain at least first derivatives with respect to the spinor variables, (4.1). In addition, the noncompact generators change the n of the operand states, (4.1) to $n\pm1$. Therefore, the following form of the noncompact generators is proposed:

$$G^t = \sum_{ijk} A^t(i,j,k)\gamma_i\gamma_j\partial_{\gamma_k} + \sum_l B^t(l)\partial_{\gamma_l}, \tag{4.2}$$

where i,j,k and $l=1,2,3$ or 4, γ_i is the spinor variable, μ, θ, ξ, or ζ, and G^t is a noncompact generator. The A^t and B^t are coefficients.

From the identity,

$$G^t\left|\begin{matrix} n \\ m^+ & m^- \end{matrix}\right\rangle = C^t\left|\begin{matrix} n+1 \\ m_t^+ & m_t^- \end{matrix}\right\rangle + D^t\left|\begin{matrix} n-1 \\ m_t^+ & m_t^- \end{matrix}\right\rangle, \tag{4.3}$$

where m_t^+ is related to $m^\pm$ by the operation of a particular generator G^t, (3.1). With a noncompact generator in the form (4.2), and the states (4.1), the solutions for the A^t and B^t, which do not depend on n, can be found, as well as the solutions for C^t and D^t. The results of this lengthy straightforward calculation for the coefficients is

$$C^t = n!\left\langle\begin{matrix} n+1 \\ m_t^+ & m_t^- \end{matrix}\middle|G^t\middle|\begin{matrix} n \\ m^+ & m^- \end{matrix}\right\rangle\Big|_{q,\,p=0},$$

$$D^t = n!\left\langle\begin{matrix} n-1 \\ m_t^+ & m_t^- \end{matrix}\middle|G^t\middle|\begin{matrix} n \\ m^+ & m^- \end{matrix}\right\rangle\Big|_{q,\,p=0}. \tag{4.4}$$

The noncompact generators appropriate for the (m^+, m^-) scheme, (2.3) or (2.4) are

$$S_+ = \frac{1}{\sqrt{2}}\eta(1+\eta\partial_\eta+\theta\partial_\theta+\xi\partial_\xi)+\frac{1}{\sqrt{2}}(\theta\xi-1)\partial_\zeta,$$

$$S_- = \frac{1}{\sqrt{2}}\zeta(1+\zeta\partial_\zeta+\theta\partial_\theta+\xi\partial_\xi)+\frac{1}{\sqrt{2}}(\theta\xi-1)\partial_\eta, \tag{4.5}$$

$$P_+ = \frac{1}{\sqrt{2}}\theta(1+\zeta\partial_\zeta+\theta\partial_\theta+\eta\partial_\eta)+\frac{1}{\sqrt{2}}(\eta\zeta+1)\partial_\xi,$$

and

$$P_- = \frac{1}{\sqrt{2}}\xi(1+\xi\partial_\xi+\zeta\partial_\zeta+\eta\partial_\eta)+\frac{1}{\sqrt{2}}(\eta\zeta+1)\partial_\theta.$$

For completeness, the set of six compact $SO(4)$ generators[2] is

$$j_+^+ = \eta\partial_\xi+\theta\partial_\zeta,\quad j_-^+ = \xi\partial_\eta+\zeta\partial_\theta,\quad j_0^+ = \tfrac{1}{2}(\eta\partial_\eta+\theta\partial_\theta-\xi\partial_\xi-\zeta\partial_\zeta),$$
$$j_-^- = \zeta\partial_\xi+\theta\partial_\eta,\quad j_+^- = \xi\partial_\zeta+\eta\partial_\theta,\quad \text{and}\quad j_0^- = \tfrac{1}{2}(\xi\partial_\xi+\eta\partial_\eta-\zeta\partial_\zeta-\theta\partial_\theta). \tag{4.6}$$

Direct calculation verifies their commutational laws (2.1) and (2.5).

The appearance of $n!$ in (4.4) is because the four-dimensional space is too small to carry all unitary irreducible representations of the $SO(4,1)$ group. This can be seen from the following considerations:

The norm of the basis states with respect to n can be determined in two ways:

(i) the norm of the next highest state can be generated from the norm of the next lowest state using the matrix elements of the noncompact generator; or
(ii) the norm can be determined by dividing the state with the square root of its length. But the form of a generator has to depend on p, q and n in order to produce the correct matrix elements of the noncompact generator.

The simplest solution is to adopt the norm for the states as determined from (i), when the form of a generator will remain free of the representation indices, and the basis is orthogonal, but the length of the basis states is $(n!)^{1/2}$ instead of 1. Therefore, the set of generators (4.5) and (4.6), together with the basis states

$$\left|\begin{matrix} n \\ m^+\ m^- \end{matrix}\right\rangle = (n!)^{1/2}\left|\begin{matrix} n \\ m^+\ m^- \end{matrix}\right\rangle', \tag{4.7}$$

constitute the representations of the $SO(4,1)$ group in the (m^+, m^-) scheme in spinor space, with p and q equal to zero (Table I).

This representation is now transformed into four-dimensional Euclidean space by a change of variables effected by a matrix T_1:

$$T_1\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_5 \end{pmatrix} = \begin{pmatrix} \eta \\ \xi \\ \theta \\ \xi \end{pmatrix},$$

where (4.8)

$$T_1 = \begin{pmatrix} -1 & -i & 0 & 0 \\ 1 & -i & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & 1 & -i \end{pmatrix} \text{and } T_1^{-1} = \tfrac{1}{2}T_1^{\dagger} = \tfrac{1}{2}\begin{pmatrix} -1 & 1 & 0 & 0 \\ i & i & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -i & i \end{pmatrix}.$$

For example, by this transformation, the compact generators (2.4), when expressed as differential operators (2.2), are brought to the form (4.6).

To arrive at a convenient form of the representation of the $SO(4,1)$ group in four-dimensional space in the (m^+, m^-) scheme, it is necessary to express the basis states, and the generators in polar variables defined by

$$\begin{aligned} x_1 &= \exp(-\rho)\cos\omega\cos\alpha, \\ x_2 &= \exp(-\rho)\cos\omega\sin\alpha, \quad \text{(System A)} \\ x_3 &= \exp(-\rho)\sin\omega\cos\beta, \end{aligned} \tag{4.9}$$

and

$$x_5 = \exp(-\rho)\sin\omega\sin\beta.$$

The volume element of this system is g_A defined by

$$g_A = (1/2)\exp(-4\rho)\sin 2\omega.$$

The spinor variables are linked with these polar variables through Eq. (4.8). The noncompact generators (4.5) become in these polar variables

$$S_{\pm} = \frac{1}{\sqrt{2}}\exp(\pm i\alpha)\left\{\pm\cosh\rho\cos\omega\partial_{\rho} \pm \sinh\rho\sin\omega\partial_{\omega} - i\frac{\sinh\rho}{\cos\omega}\partial_{\alpha} \pm \exp(-\rho)\cos\omega\right\}$$

and (4.10)

$$P_{\pm} = \frac{1}{\sqrt{2}}\exp(\pm i\beta)\left\{-\cosh\rho\sin\omega\partial_{\rho} + \sinh\rho\cos\omega\partial_{\omega} \pm i\frac{\sinh\rho}{\sin\omega}\partial_{\beta} + \exp(-\rho)\sin\omega\right\}$$

The compact generators assume the familiar form[17]

$$\begin{aligned} j_{\pm}^{+} &= \tfrac{1}{2}\exp[\pm i(\alpha+\beta)]\{\pm\partial_{\omega} - i\cot\omega\partial_{\beta} + i\tan\omega\partial_{\alpha}\}, \\ j_0^{+} &= -\tfrac{1}{2}\{\partial_{\alpha} + \partial_{\beta}\}, \\ j_{\pm}^{-} &= \tfrac{1}{2}\exp[\pm i(\alpha-\beta)]\{\pm\partial_{\omega} + i\cot\omega\partial_{\beta} + i\tan\omega\partial_{\alpha}\}, \end{aligned} \tag{4.11}$$

and

$$j_0^{-} = -\tfrac{1}{2}\{\partial_{\alpha} - \partial_{\beta}\}.$$

The basis states (4.7) become after a little manipulation:

$$\left|\begin{matrix} n \\ m^+\ m^- \end{matrix}\right\rangle = (-1)^{n/2+m^+}\exp(-n\rho)[\exp(i\alpha)]^{m^+ + m^-} \times [\exp(i\beta)]^{m^+ - m^-} d^{n/2}_{m^+ m^-}(2\omega), \tag{4.12}$$

where $d^{\circ}_{\circ}(2\omega)$ is an $SO(3)$ reduced rotational matrix as defined in Ref. 18. The range of parameters is

$$0 < \rho < \infty, \quad 0 < \omega < \pi/2, \quad 0 < \alpha \text{ and } \beta < 2\pi.$$

The length of these states in system A, (4.9) is

$$\left\langle \begin{matrix} n \\ m^+\ m^- \end{matrix} \middle| \begin{matrix} n \\ m^+\ m^- \end{matrix} \right\rangle^{1/2} = \frac{\pi}{[2(n+1)(n+2)]^{1/2}}. \tag{4.13}$$

Since the ρ appear explicitly in the noncompact generators (4.9) and in the basis states (4.2), the integrals involving these states will diverge unless the exponential radial dependence, $\exp\{-\rho\}$, is used in system A, (4.9), instead of the usual linear dependence on ρ.

Applying the noncompact generators (4.10) to the basis states (4.12) gives a set of eight new recursion relations for the $SO(3)$ group reduced rotational matrices. Reducing the general matrix elements of the noncompact generators (3.1) and (3.2) to the case $j^+ = j^- = \frac{1}{2}n$, and replacing ω by $\frac{1}{2}\omega$ so that the reduced $SO(3)$ group rotational matrices depend on a single argument,

$S_{\pm}$ gives

$$[\tfrac{1}{2}n(1+\cos\omega) \pm (m^+ + m^-)]\, d^{n/2}_{m^+ m^-}(\omega) - \sin\omega\partial_{\omega} d^{n/2}_{m^+ m^-}(\omega) = 2[(\tfrac{1}{2}n \pm m^+)(\tfrac{1}{2}n \pm m^-)]^{1/2}\cos\frac{\omega}{2} d^{n/2-1/2}_{m^+ \pm m^-/2 \pm 1/2}(\omega),$$

$P_{\pm}$ gives

$$[\tfrac{1}{2}n(1-\cos\omega) \pm (m^+ - m^-)]\, d^{n/2}_{m^+ m^-}(\omega) + \sin\omega\partial_{\omega} d^{n/2}_{m^+ m^-}(\omega) = \pm 2[(\tfrac{1}{2}n \pm m^+)(\tfrac{1}{2}n \pm m^-)]^{1/2}\sin\frac{\omega}{2} d^{n/2-1}_{m^+ \pm m^-/2 \pm 1/2}(\omega), \tag{4.14}$$

S_+ gives $[\frac{1}{2}(n+2)(1+\cos\omega)\pm(m^+ + m^-)]^{1/2} d^{n/2}_{m^+ m^-}(\omega)$

$$+\sin\omega\partial_\omega d^{n/2}_{m^+ m^-}(\omega)$$

$$=2[(\tfrac{1}{2}n\pm m^+ +1)(\tfrac{1}{2}n\pm m^- +1)]^{1/2}\cos\frac{\omega}{2} d^{n/2+1/2}_{m^+\pm m^-/2\pm 1/2}(\omega);$$

and

$P_\pm$ gives $[\frac{1}{2}(n+2)(1-\cos\omega)\pm(m^+ - m^-)\, d^{n/2}_{m^+ m^-}(\omega)$

$$-\sin\omega\partial_\omega d^{n/2}_{m^+ m^-}(\omega)$$

$$=\pm 2[(\tfrac{1}{2}n\pm m^+ +1)(\tfrac{1}{2}n\pm m^- +1)]^{1/2}\sin\frac{\omega}{2} d^{n/2+1/2}_{m^+\pm m^-/2\pm 1/2}(\omega),$$

where $S_\pm$ and $P_\pm$ are the particular generators which lead to the recurrence relations.

B. The representation of the $SO(4,1)$ group in the (j,m) scheme

The basis states and the set of generators for the (j,m) scheme are obtained from those for the (m^+, m^-) scheme by substitution:

$$\alpha=(1/\sqrt{2})(\theta-\xi)$$
$$\beta=(1/\sqrt{2})(\theta+\xi) \tag{4.15}$$

The basis states become

$$\left|\begin{matrix} n \\ l\ m \end{matrix}\right\rangle=\left(\frac{(2l+1)!(n-l)!(n+l+1)!}{n!l!2^{n-l}}\right)^{1/2}$$
$$\times\sum_s \frac{(l+s)!\alpha^{n-l-2s}(\beta^2-2\eta\xi)^s}{s!(n-l-2s)!(2l+2s+1)!}\left|\begin{matrix} l \\ m \end{matrix}\right\rangle, \tag{4.16}$$

where

$$\left|\begin{matrix} l \\ m \end{matrix}\right\rangle=\left(\frac{l!(l+m)!(l-m)!2^{l-m}}{(2l)!}\right)^{1/2}\sum_t \frac{\eta^{m+t}\beta^{l-m-2t}\xi^t}{t!(m+t)!(l-m-2t)!2^t},$$
$$-l\le m\le l,\quad 0\le l\le n$$

where l replaces j, which is always an integer in a square representation. The set of noncompact generators (2.6) in the (j,m) scheme become, after substitution of (4.15) into (4.5),

$$S_+=\frac{1}{\sqrt{2}}\eta(1+\eta\partial_\eta+\alpha\partial_\alpha+\beta\partial_\beta)-\frac{1}{2\sqrt{2}}(\alpha^2-\beta^2+2)\partial_\xi,$$
$$S_-=\frac{1}{\sqrt{2}}\xi(1+\xi\partial_\xi+\alpha\partial_\alpha+\beta\partial_\beta)-\frac{1}{2\sqrt{2}}(\alpha^2-\beta^2+2)\partial_\eta,$$
$$L_{03}=\frac{1}{\sqrt{2}}\beta(1+\xi\partial_\xi+\eta\partial_\eta)+\frac{1}{2\sqrt{2}}(\alpha^2+\beta^2+2\eta\xi+2)\partial_\alpha$$
$$+\frac{1}{\sqrt{2}}\alpha\beta\partial_\beta, \tag{4.17}$$

and

$$iL_{05}=\frac{1}{\sqrt{2}}\alpha(1+\xi\partial_\xi+\eta\partial_\eta)+\frac{1}{2\sqrt{2}}(\alpha^2+\beta^2-2\eta\xi-2)\partial_\alpha$$
$$+\frac{1}{\sqrt{2}}\alpha\beta\partial_\beta$$

and the compact generators become

$$L_+=\sqrt{2}(\beta\partial_\xi+\eta\partial_\beta),\quad L_-=\sqrt{2}(\beta\partial_\eta+\xi\partial_\beta),\quad L_0=\eta\partial_\eta-\xi\partial_\xi,$$
$$K_+=\sqrt{2}(\alpha\partial_\xi-\eta\partial_\alpha),\quad K_-=\sqrt{2}(\xi\partial_\alpha-\alpha\partial_\eta)\quad K_0=\alpha\partial_\beta+\beta\partial_\alpha. \tag{4.18}$$

By analogy with Sec. 4A the representation is transformed into four-dimensional Euclidean space

$$T_2\begin{pmatrix}x_1\\x_2\\x_3\\x_5\end{pmatrix}=\begin{pmatrix}\eta\\\xi\\\beta\\\alpha\end{pmatrix},$$

where (4.19)

$$T_2=2\begin{pmatrix}-\frac{1}{\sqrt{2}}&-\frac{i}{\sqrt{2}}&0&0\\-\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}&0&0\\0&0&1&0\\0&0&0&i\end{pmatrix},\quad T_2^{-1}=\tfrac{1}{4}T_2^\dagger=\tfrac{1}{2}\begin{pmatrix}-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0&0\\\frac{i}{\sqrt{2}}&\frac{i}{\sqrt{2}}&0&0\\0&0&1&0\\0&0&0&-i\end{pmatrix}$$

and the basis states and the generators are expressed in polar variables appropriate to the (j,m) scheme:

$$x_1=\exp(-\rho)\sin\chi\sin\theta\cos\phi,$$
$$x_2=\exp(-\rho)\sin\chi\sin\theta\sin\phi,$$
$$x_3=\exp(-\rho)\sin\chi\cos\theta, \qquad \text{(System B)} \tag{4.20}$$

and

$$x_5=\exp(-\rho)\cos\chi.$$

The volume element of this system is g_B

$$g_B=\exp(-4\rho)\sin^2\chi\sin\theta.$$

The range of parameters is

$$0\le\rho<+\infty,\quad 0\le\chi\le\pi,$$

and

$$0\le\theta\le\pi,\quad 0\le\phi\le 2\pi.$$

The noncompact generators (4.17) become

$$S_\pm=\tfrac{1}{4}\exp(\pm i\phi)\{\pm(e^\rho+2e^{-\rho})\sin\chi\sin\theta\partial_\rho\pm(e^\rho-2e^{-\rho})$$
$$\times\cos\chi\sin\theta\partial_\chi\pm(e^\rho-2e^{-\rho})\frac{\cos\theta}{\sin\chi}\partial_\theta$$
$$-i(e^\rho-2e^{-\rho})\frac{1}{\sin\chi\sin\theta}\partial_\phi\pm 4e^{-\rho}\sin\chi\sin\theta\}, \tag{4.21}$$
$$L_{03}=\frac{1}{2\sqrt{2}}\{-(e^\rho+2e^{-\rho})\sin\chi\cos\theta\partial_\rho+(e^\rho-2e^{-\rho})\cos\chi\cos\theta\partial_\chi$$
$$-(e^\rho-2e^{-\rho})\frac{\sin\theta}{\sin\chi}\partial_\theta+4e^{-\rho}\sin\chi\cos\theta\},$$
$$L_{05}=\frac{1}{2\sqrt{2}}\{-(e^\rho+2e^{-\rho})\cos\chi\partial_\rho-(e^\rho-2e^{-\rho})\sin\chi\partial_\chi$$
$$+4e^{-\rho}\cos\chi\}.$$

The compact generators (4.18) take the usual form,[17]

$$L_\pm=\exp(\pm i\phi)(\pm\partial_\theta+i\cot\theta\partial_\phi),\quad L_0=-i\partial_\phi,$$

$$K_{\pm}=\exp(\pm i\phi)\left(i\sin\theta\partial_{\chi}+i\cot\chi\cos\theta\partial_{\theta}\pm\frac{\cot\chi}{\sin\theta}\partial_{\phi}\right), \tag{4.22}$$

and

$$K_0=i(\cos\theta\partial_{\chi}-\cot\chi\sin\theta\partial_{\theta}).$$

The basis states (4.16) become, after manipulation (Appendix B), the hyperspherical harmonics,

$$\left|\begin{matrix}n\\ l\ m\end{matrix}\right\rangle=(i)^{n-l}\left(\frac{2^{n+2l}(n-l)!(l!)^2}{(n+l+1)!}(2l+1)\frac{(l-m)!}{(l+m)!}\right)^{1/2} \tag{4.23}$$

$$\exp(-n\rho)(\sin\chi)^l C_{n-l}^{l+1}(\cos\chi)P_l^m(\cos\theta)\exp(im\phi),$$

where $C_{n-l}^{l+1}(\cos\chi)$ is a Gegenbauer polynomial, and $P_l^m(\cos\theta)$ is an associated Legendre polynomial as defined in Ref. 19. The same state, apart from normalization, has been derived on group theoretical grounds by Sharp.[2]

The length, the norm of the states (4.23) is

$$\left\langle\begin{matrix}n\\ l\ m\end{matrix}\middle|\begin{matrix}n\\ l\ m\end{matrix}\right\rangle^{1/2}=(-1)^{n-l}\left(\frac{\pi^2 2^n}{(n+1)(n+2)}\right)^{1/2}.$$

Similarly, as in Sec. 4A, the known matrix elements in the (j,m) scheme (3.4), the set of noncompact operators (4.10) and the states (4.23), allow the recursion relations for the hyperspherical harmonics to be derived. All of them will not be written explicitly, because the functions of two variables are involved; instead, several examples will be considered in which only one function appears. The following recursion relations are all known; nevertheless, they will be derived because they establish the representation given in (4.10) for noncompact generators and the basis states (4.23).

The operation of the generators $S_{\pm}$ on a basis state for which $m=l$ is

$$\begin{aligned}S_{\pm}\left|\begin{matrix}l\\ l\ m\end{matrix}\right\rangle=&\left\langle\begin{matrix}l+1\\ l+1\ m\pm1\end{matrix}\middle|S_{\pm}\middle|\begin{matrix}l\\ l\ m\end{matrix}\right\rangle\left|\begin{matrix}l+1\\ l+1\ m\pm1\end{matrix}\right\rangle\\&+\left\langle\begin{matrix}l+1\\ l-1\ m\pm1\end{matrix}\middle|S_{\pm}\middle|\begin{matrix}l\\ l\ m\end{matrix}\right\rangle\left|\begin{matrix}l+1\\ l-1\ m\pm1\end{matrix}\right\rangle\\&+\left\langle\begin{matrix}l-1\\ l-1\ m\pm1\end{matrix}\middle|S_{\pm}\middle|\begin{matrix}l\\ l\ m\end{matrix}\right\rangle\left|\begin{matrix}l-1\\ l-1\ m\pm1\end{matrix}\right\rangle.\end{aligned} \tag{4.24}$$

The identity (4.24), with the representation of the noncompact generators (4.21) and the states (4.23), splits into two equations, one for each exponent $\exp[-(n\pm1)\rho]$, after the matrix elements have been calculated from (3.4) and (3.5) for the case $p=0$, $q=0$, and $j^+=j^-=\frac{1}{2}n$.

This procedure gives four standard recursion relations for the associated Legendre polynomials of a real variable (five are actually found, but two are identical),

$$(2l+1)\sin\theta P_l^m(\cos\theta)=P_{l+1}^{m+1}(\cos\theta),$$

$$\begin{aligned}&l\sin\theta P_l^m(\cos\theta)+\cos\theta\partial_{\theta}P_l^m(\cos\theta)-\frac{m}{\sin\theta}P_l^m(\cos\theta)\\&\quad=P_{l-1}^{m+1}(\cos\theta),\end{aligned} \tag{4.25}$$

$$(2l+1)\sin\theta P_l^m(\cos\theta)=(1-m+1)(1-m+2)P_{l+1}^{m-1}(\cos\theta)$$

$$-(1+m)(l+m-1)P_{l-1}^{m-1}(\cos\theta),$$

and

$$\begin{aligned}&-l\sin\theta P_l^m(\cos\theta)-\cos\theta\partial_{\theta}P_l^m(\cos\theta)-\frac{m}{\sin\theta}P_l^m(\cos\theta)\\&\quad=(1+m)(l+m-1)P_{l-1}^{m-1}(\cos\theta).\end{aligned}$$

Another two recursion relations can be derived from the identity

$$\begin{aligned}L_{03}\left|\begin{matrix}l\\ l\ m\end{matrix}\right\rangle=&\left\langle\begin{matrix}l+1\\ l+1\ m\end{matrix}\middle|L_{03}\middle|\begin{matrix}l\\ l\ m\end{matrix}\right\rangle\left|\begin{matrix}l+1\\ l+1\ m\end{matrix}\right\rangle\\&+\left\langle\begin{matrix}l+1\\ l-1\ m\end{matrix}\middle|L_{03}\middle|\begin{matrix}l\\ l\ m\end{matrix}\right\rangle\left|\begin{matrix}l+1\\ l-1\ m\end{matrix}\right\rangle\\&+\left\langle\begin{matrix}l-1\\ l-1\ m\end{matrix}\middle|L_{03}\middle|\begin{matrix}l\\ l\ m\end{matrix}\right\rangle\left|\begin{matrix}l-1\\ l-1\ m\end{matrix}\right\rangle.\end{aligned} \tag{4.26}$$

This gives

$$\begin{aligned}&(2l+1)\cos\theta P_l^m(\cos\theta)=(1-m+1)P_{l+1}^m(\cos\theta)\\&\quad+(l+m)P_{l-1}^m(\cos\theta),\\&l\cos\theta P_l^m(\cos\theta)-\sin\theta\partial_{\theta}P_l^m(\cos\theta)=(l+m)P_{l-1}^m(\cos\theta).\end{aligned} \tag{4.27}$$

Thus six out of eight possible[19] recursion relations appear. The remaining two recursion relations can be derived by applying the compact $SO(4,1)$ group generators (4.22) to the states (4.23). In particular, from the identity

$$L_{\pm}\left|\begin{matrix}l\\ l\ m\end{matrix}\right\rangle=\left\langle\begin{matrix}l\\ l\ m\pm1\end{matrix}\middle|L_{\pm}\middle|\begin{matrix}l\\ l\ m\end{matrix}\right\rangle\left|\begin{matrix}l\\ l\ m\pm1\end{matrix}\right\rangle, \tag{4.28}$$

the two remaining recursion relations follows:

$$\partial_{\theta}P_l^m(\cos\theta)-m\cot\theta P_l^m(\cos\theta)=P_l^{m+1}(\cos\theta),$$

and (4.29)

$$\begin{aligned}&\partial_{\theta}P_l^m(\cos\theta)+m\cot\theta P_l^m(\cos\theta)=\\&\quad-(l+m)(l-m+1)P_l^{m-1}(\cos\theta).\end{aligned}$$

The $K_{\pm}$ gives the second recursion relation of (4.25); the K_0 gives the second recursion relation of (4.27); L_0 gives a trivial result. The recursion relations (4.25), (4.27), and (4.29) are in a form which corresponds to that of the generators. The relations easily reduce to their standard form as in Ref. 19, by eliminating the derivatives with respect to θ.

Thus to obtain all the resursion relations group theoretically, the $SO(4,1)$ group representation has to be used. It is of interest to note that, if it is assumed that all the recursion relations for the Legendre polynomials are known, and contained in (4.25), (4.27), and (4.29), then consideration of the full dynamical group of the Coulomb potential $O(4,2)$ *cannot* give new recursion relations. A representation of the $SO(4,2)$ group in four-dimensional space, with the hyperspherical harmonics as basis states, has to coincide with the representation of the $SO(4,1)$ group. In this sense, it is sufficient to use the $SO(4,1)$ group as the dynamical group for this system, instead of the more complex $SO(4,2)$ group.

The corresponding recursion relations for the Gegenbauer polynomial is most easily derived from the operation of the noncompact generators on the basis state with $m=l$. From

$$S_+\left|\begin{matrix}n\\ l\ l\end{matrix}\right\rangle=\left\langle\begin{matrix}n+1\\ l+1\ l+1\end{matrix}\right|S_+\left|\begin{matrix}n\\ l\ l\end{matrix}\right\rangle\left|\begin{matrix}n+1\\ l+1\ l+1\end{matrix}\right\rangle + \left\langle\begin{matrix}n-1\\ l+1\ l+1\end{matrix}\right|S_+\left|\begin{matrix}n\\ l\ l\end{matrix}\right\rangle, \tag{4.30}$$

two standard recursion relations are obtained, after changing the indices

$$\nu=l+1, \quad k=n-l; \tag{4.31}$$

namely,

$$(k+2\nu)C_k^\nu(\cos\chi)-\cot\chi\,\partial_\chi C_k^\nu(\cos\chi)=2\nu C_k^{\nu+1}(\cos\chi)$$

and (4.32)

$$kC_k^\nu(\cos\chi)+\cot\chi\,\partial_\chi C_k^\nu(\cos\chi)=-2\nu C_{k-2}^{\nu+1}(\cos\chi).$$

The operation of the generators S_- and L_{03} does not lead to a different recursion relation in this case. As iL_{05} does not change either l or m, it can be applied to an arbitrary state,

$$iL_{05}\left|\begin{matrix}n\\ l\ m\end{matrix}\right\rangle=\left\langle\begin{matrix}n+1\\ l\ m\end{matrix}\right|iL_{05}\left|\begin{matrix}n\\ l\ m\end{matrix}\right\rangle\left|\begin{matrix}n+1\\ l\ m\end{matrix}\right\rangle + \left\langle\begin{matrix}n-1\\ l\ m\end{matrix}\right|iL_{05}\left|\begin{matrix}n\\ l\ m\end{matrix}\right\rangle\left|\begin{matrix}n-1\\ l\ m\end{matrix}\right\rangle, \tag{4.33}$$

to give two more recursion relations:

$$k\cos\chi\,C_k^\nu(\cos\chi)-\sin\chi\,\partial_\chi C_k^\nu(\cos\chi)=(k+2\nu-1)C_{k-1}^\nu(\cos\chi) \tag{4.34}$$

and

$$(k+2\nu)\cos\chi\,C_k^\nu(\cos\chi)+\sin\chi\,\partial_\chi C_k^\nu(\cos\chi)=(k+1)\,C_{k+1}^\nu(\cos\chi).$$

Equations (4.32) and (4.34) give four recursion relations for the Gegenbauer polynomials. The remaining recursion relation can be derived by applying any of the generators K_+, K_-, and K_0 to the basis states (4.23) with $m=1$, $m=-l$, and $m=1$, respectively. The result is

$$-\partial_\chi C_k^\nu(\cos\chi)=2\nu\sin\chi\,C_{k-1}^{\nu+1}(\cos\chi). \tag{4.35}$$

The last recursion relation is the only one which can be derived by use of the compact generators. Similarly, as in the case of the Legendre polynomials, the full set of recursion relations for the Gegenbauer polynomials is contained in a representation of the $SO(4,1)$ group.

5. CONCLUDING REMARKS

Two points of interest emerge.

First, the connection between the space of the $O(4)$ spinors and four-dimensional Euclidean space; apart from the theoretical interest, it is an advantage in practical calculations, since various operations are far more easily performed in spinor space than in Euclidean space; by the coordinate transformations (4.8) and (4.19), one can always arrive at the corresponding set of operations expressed in the appropriate coordinates in four-dimensional space from those in spinor space.

Second, the explicit representation of the $SO(4,1)$ generators found, should facilitate reduction of the $SO(4,1)$ group with respect to its different subgroups.

As mentioned in the introduction, it is known that the radial part of the hydrogenic wavefunction is linked with the basis for a unitary irreducible representation of the $SO(2,1)$ group.[4] The radial wavefunctions depend on n and l. The wavefunction dependence on m can be eliminated by considering the $SO(4,1)$ group generators which leave this quantum number m unchanged, namely, the L_{35}, L_{05}, and L_{03} which constitute the Lie algebra of the $SO(2,1)$ group.

This suggests that in the reduction of $SO(4,1)\supset SO(2,1)\times SO(2)$ in which L_{12} generates the $SO(2)$ group, the radial wavefunctions are contained in a basis in which the subgroup $SO(2,1)$ is diagonal. The same reduction is of interest in considering the group theoretical description of an atom in a magnetic field where $SO(4,1)\supset SO(2,1)\times SO(2)$.

The noncompact generators (4.10) and (4.21) can be transformed into three-dimensional "physical" space by reversing the procedure in Ref. 3. It will be of interest to compare such transformed generators with the existing representations of the $SO(4,1)$ group in three-dimensional space,[12,13] or with the representations of the $SO(4,2)$ group in three-dimensional space.[20]

The representations of the $SO(4,1)$ group found in this paper for $q=0$ and $p=0$, representation 1 of Table I, fall within the chain of reduction of certain important classes of the $O(4,2)$ group representations $O(4,2)\supset SO(4,1)\supset SO(4)$, the subclass (1) of Ref. 1.

In this paper the representation of the generators of the $SO(4,1)$ group has been found for which the hyperspherical harmonics are basis states, belonging to the representation 1 of Table I with $p=0$ and $q=0$.

The recursion relations of the $SO(3)$ reduced rotational matrices, the Legendre polynomials and the Gegenbauer polynomials have been derived by group theoretical arguments alone, consequently they are of intrinsic interest in the study of the connections between these functions and the group representations.

ACKNOWLEDGMENT

R.B. is indebted to Professor R.T. Sharp of McGill University for many helpful and inspiring discussions during the course of this work and both of the authors are grateful to Dr. R. Shaw of Oxford University for bringing their attention to Han's work. This research was supported by the National Research Council (Canada to whom M.A.W. is grateful for a Travel Fellowship 1971—1974. M.A.W. is indebted to the Theoretical Chemistry Department, Oxford, and the late Professor Coulson for hospitality and to the Science Research Council (U.K.) for a Visiting Scientist Award 1973—1974.

APPENDIX A

Han has obtained a representation of the $SO(4,1)$ group in four-dimensional space by constructing an E_4 algebra relevant to the problem.[14] His noncompact generators

are expressed in terms of a 4-vector P^j, called a "Shirokov operator" which transforms irreducibly with respect to the $SO(4)$ group. The relation between Han's noncompact generators $(\mathbf{B}, S)$ and those found in this paper $(\mathbf{B}', L_{05})$ ($\mathbf{B}'$ has components L_{01}, L_{02}, L_{03}) in the notation of this paper is

$$i\mathbf{B} = \sqrt{2}\,\mathbf{B}' + 2(\mathbf{L}\times\mathbf{x}) + (9/2)\mathbf{x} - (\mathbf{x}\cdot\mathbf{p})\mathbf{x} - (1/2)\mathbf{p},$$
$$iS = -\sqrt{2}\,L_{05} + (\mathbf{x}\cdot\mathbf{p})x_5 - (5/2)x_5 + p_5. \tag{A1}$$

The two vectors $\mathbf{x}$ and $\mathbf{p}$ have components (x_1, x_2, x_3) and $(\partial_{x_1}, \partial_{x_2}, \partial_{x_3})$ and $p_5 \to \partial_{x_5}$.

The difference arises from the arbitrariness in determining the 4-vector P^j; i.e., any vector which transforms irreducibly with respect to the $SO(4)$ group can be identified with four noncompact generators of the $SO(4,1)$ group. Han's choice is $\mathbf{P} \to \mathbf{x}$; compared to that in this paper which is given in (A1).

When expressed in spinor coordinates using (4.8) for example, Han's generators do not contain the term which lowers n to $n-1$. Therefore, the hyperspherical harmonics in their normal form (4.23) cannot be the basis states for Han's generators (although a linear combination with respect to l and m may possibly transform correctly).

APPENDIX B

After the spinor variables in the basis states in (4.16) have been replaced by the spherical ones of (4.20) via (4.18), the two sums in the expression for the basis states of (4.16) take the following form, without the normalization factors,

$$S_s = \sum_s \frac{(-1)^s (l+s)!\,(\cos\chi)^{n-l-2s}(\sin\chi)^{2s}}{s!\,(n-l-2s)!\,(2l+2s+1)!} \tag{B1}$$

and

$$S_t = \sum_t \frac{(-1)^t (\sin\theta)^{m+2t}(\cos\theta)^{l-m-2t}}{t!\,(m+t)!\,(l-m-2t)!\,2^{2t}}. \tag{B2}$$

By developing

$$(\sin^2\chi)^s = (1-\cos^2\chi)^s$$

into the binomial series we find for S_s

$$S_s = \sum_{p=0} \left\{ \sum_{s=p} \frac{(l+s)!}{(n-l-2s)!\,(2l+2s+1)!\,(s-p)!} \right\} \times \frac{(-1)^p}{p!}(\cos\chi)^{n-2p}. \tag{B3}$$

The coefficient in curly brackets may be calculated as follows: First, the index s is replaced:

$$s = l + p.$$

Then, two identities[21] for the gamma functions are used:

$$(n-l-2p-2t)! = \frac{(n-l-2p)!}{2^{2t}[(-n+l)/2+p]_t\,[(-n+l+1)/2+p]_t}$$

and

$$(2l+2s+1)! = \Gamma(2l+2s+2) = \frac{(2)^{2l+2s+1}}{\sqrt{\pi}}\,\Gamma(l+s+1)\Gamma(l+s+\tfrac{3}{2}). \tag{B4}$$

After the last replacements, the coefficient in (B3) becomes

$$\{\cdot\} = \frac{\sqrt{\pi}}{(2)^{2l+2s+1}} \frac{1}{(n-l-2p)!\,\Gamma(l+p+\frac{3}{2})} \times {}_2F_1\left(\frac{-n+l}{2}+p,\ \frac{-n+l+1}{2}\,l+p+\tfrac{3}{2};\ 1\right). \tag{B5}$$

From the well-known formula for the value of the hyperspherical function, ${}_2F_1(\cdots)$, for the argument 1, and using the duplication formula of the gamma functions once more,

$$\sum_{s=p} \frac{(l+s)!}{(n-l-2s)!\,(2l+2s+1)!\,(s-p)!} = \frac{(2)^{n-l}}{(2)^{2p}} \frac{(n-p)!}{(n-l-2p)!\,(n+l+1)!} \tag{B6}$$

and

$$S_s = \frac{1}{(n+l+1)!} \sum_p \frac{(-1)^p (n-p)!}{p!\,(n-l-2p)!} (2\cos\chi)^{n-l-2p} = \frac{l!}{(n+l+1)!} C_{n-l}^{l+1}(\cos\chi), \tag{B7}$$

where $C_{n-l}^{l+1}(\cos\chi)$ is a Gegenbauer polynomial, and the sum is restricted to the nonnegative factorials.

Similarly, the sum (B2) leads to

$$S_t = \frac{(2^m)}{(l+m)!}(-1)^m P_l^m(\cos\theta). \tag{B8}$$

Substitution of these sums, (B7) and (B8), into (4.16) together with the normalization factors give the basis states (4.23).

[1] The notation is that of A. O. Barut and A. Bohm, J. Math. Phys. **11**, 2938 (1970) for the continuous groups relevant to this paper. The $SO(4,1)$ group is a group of all pseudo-rotations excluding inversion.

[2] The notation for the $SO(4)$ group is that of R. T. Sharp, J. Math. and Phys. **47**, 359 (1968). The irreducible representations are labelled with a pair of indices (j^+, j^-) which are linked to the Casimir invariants

$$\tfrac{1}{2}C_1 = (j^+)^2 + (j^-)^2,$$
$$\tfrac{1}{2}C_2 = (j^+)^2 - (j^-)^2.$$

[3] M. Bander and C. Itzykson, Rev. Mod. Phys. **38**, 346 (1966).

[4] The radial part of the hydrogenic wavefunction has been considered in relation to the representations of the $O(2,1)$ and $O(3,2)$ group by L. Armstrong Jr., Phys. Rev. A**3**, 1546 (1971) and D. Herrick and O. Sinanoğlu, Phys. Rev. A**5**, 2309 (1972), respectively. Within the $O(4,2)$ group scheme, the problem has been treated by M. J. Cunningham, J. Math. Phys. **13**, 33, 1108 (1972).

[5] L. H. Thomas, Ann. Math. **42**, 113 (1941).

[6] T. D. Newton, Ann. Math. **51**, 730 (1950) where the connection between p, q and Q, W is given.

[7] J. Dixmier, Bull. Soc. Math. (France) **89**, 9 (1961).

[8] A. Khilberg, Arkiv Fysik **30**, 121 (1965).

[9] A. P. Stone, Proc. Camb. Phil. Soc. **52**, 424 (1956).

[10] R. Musto, Phys. Rev. **148**, 1274 (1966).

[11] R. H. Pratt and T. F. Jordan, Phys. Rev. **148**, 1276 (1966).

[12] D. Tz. Stoyanov and I. T. Todorov, J. Math. Phys. **9**, 2146 (1968).

[13] G. Mack and I. T. Todorov, J. Math. Phys. **11**, 2078 (1969).

[14] M. Y. Han, Nuovo Cimento **42B**, 367 (1966).

[15]T. Yao, J. Math. Phys. 8, 1931 (1967); 9, 1615 (1968); 12, 315 (1971).
[16]The noncompact generators used in this paper are related by $S_\pm = \pm (1/\sqrt{2})(U + iV)$ and $P_\pm = \pm (1/\sqrt{2})(T + iW)$ to those of Thomas, Ref. 5
[17]M.J. Englefield, *Group Theory and the Coulomb Problem* (Wiley-Interscience, New York, 1972), Appendix C.
[18]D.M. Brink and G.R. Satchler, *Angular Momentum* (Clarendon, Oxford, 1968), second edition.
[19]*Higher Transcendental Functions*, The Bateman Manuscript Project, edited by A. Erdelyi (McGraw-Hill, New York, 1953), Vol. 1, Chap. III.
[20]A. O. Barut and H. Kleinert, Phys. Rev. 156, 1541 (1966).
[21]Ref. 19, p. 66.

J-matrix method: Extensions to arbitrary angular momentum and to Coulomb scattering

Hashim A. Yamani*†

Department of Physics, Harvard University, Cambridge, Massachusetts 02138

Louis Fishman

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138
(Received 3 September 1974)

The J-matrix method introduced previously for s-wave scattering is extended to treat the lth partial wave kinetic energy and Coulomb Hamiltonians within the context of square integrable (L^2), Laguerre (Slater), and oscillator (Gaussian) basis sets. The determination of the expansion coefficients of the continuum eigenfunctions in terms of the L^2 basis set is shown to be equivalent to the solution of a linear second order differential equation with appropriate boundary conditions, and complete solutions are presented. Physical scattering problems are approximated by a well-defined model which is then solved exactly. In this manner, the generalization presented here treats the scattering of particles by neutral and charged systems. The appropriate formalism for treating many channel problems where target states of differing angular momentum are coupled is spelled out in detail. The method involves the evaluation of only L^2 matrix elements and finite matrix operations, yielding elastic and inelastic scattering information over a continuous range of energies.

1. INTRODUCTION

In two previous publications[1,2] (referred to as I and II) the J-matrix (Jacobi matrix) method was introduced as a new approach for solution of quantum scattering problems. As discussed in I, the principal characteristics of the method are its use of only square integrable (L^2) basis functions and its ability to yield an exact solution to a model scattering Hamiltonian, which, in a well-defined and systematically improvable manner, approximates the actual scattering Hamiltonian. The method is numerically highly efficient as scattering information is obtained over a continuous range of energies from a single matrix diagonalization.

The development of the J-matrix method as presented in I is based primarily upon the observation that the s wave kinetic energy,

$$H_0 = -\frac{1}{2}\frac{d^2}{dr^2} \tag{1.1}$$

can be analytically diagonalized in the Laguerre (Slater) basis:

$$\phi_n(\lambda r) = (\lambda r)\exp(-\lambda r/2)L_n^1(\lambda r), \quad n = 0, 1, \ldots, \infty, \tag{1.2}$$

where λ is a scaling parameter. This follows from the fact that the infinite matrix representation of $(H_0 - k^2/2)$ in the above basis is tridiagonal (i. e., J or Jacobi matrix) and that the resulting three-term recursion scheme can be analytically solved yielding the expansion coefficients of both a "sine-like" $\tilde{S}(r)$ and a "cosine-like" $\tilde{C}(r)$ function. The J-matrix solutions $\tilde{S}(r)$ and $\tilde{C}(r)$ are used to obtain the exact solution of the model scattering problem defined by approximating the potential V by its projection V^N onto the finite subspace spanned by the first N basis functions. That is, the exact solution Ψ of the scattering problem,

$$(H_0 + V^N - k^2/2)\Psi = 0, \tag{1.3}$$

is obtained by determining its expansion coefficients in terms of the basis set $\{\phi_n\}$ subject to the asymptotic boundary condition,

$$\Psi \to \tilde{S}(r) + \tan\delta\tilde{C}(r), \tag{1.4}$$

where δ is the phase shift due to the potential V^N.

This paper is intended to generalize the formalism developed in I in three areas. First, the results of I are extended to all partial waves, in which case the uncoupled Hamiltonian becomes the lth partial wave kinetic energy operator,

$$H_0 = -\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2}, \tag{1.5}$$

which has a Jacobi representation in the Laguerre basis,

$$\phi_n(\lambda r) = (\lambda r)^{l+1}\exp(-\lambda r/2)L_n^{2l+1}(\lambda r), \quad n = 0, 1, \ldots, \infty. \tag{1.6}$$

Secondly, a similar analysis of the Hamiltonian of Eq. (1.5) is presented within the context of the oscillator (Gaussian) basis,

$$\phi_n(\lambda r) = (\lambda r)^{l+1}\exp(-\lambda^2 r^2/2)L_n^{l+1/2}(\lambda^2 r^2), \quad n = 0, 1, \ldots, \infty, \tag{1.7}$$

which preserves the Jacobi representation and is also analytically soluble. The third generalization involves the analysis of the lth partial wave Coulomb Hamiltonian,

$$H_0 = -\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} + \frac{z}{r} \tag{1.8}$$

in the Laguerre function space of Eq. (1.6), which again yields a Jacobi form and is subsequently analytically soluble. It is noted that the analysis of the Coulomb Hamiltonian in the oscillator set of Eq. (1.7) does not lead to a Jacobi form.

For the solution of these problems, a general technique is developed which reduces the solution of the infinite recurrence problem for the asymptotically "sine-like" J-matrix eigenfunction to the solution of a linear second order differential equation with appropriate boundary conditions. An asymptotically "cosine-like"

solution which obeys the same differential equation with different boundary conditions is then constructed. The fact that both J-matrix solutions obey the same recurrence scheme is essential to the success of the method as an efficient technique for solving scattering problems.[1]

The program of the paper is as follows: In Sec. 2 A, the generalized H_0 problem is considered and a general procedure for obtaining the expansion coefficients of the sine-like and cosine-like functions in terms of the basis sets is outlined. In Sec. 2 B, the general method is illustrated in detail for the case of the radial kinetic energy in a Laguerre basis. The analogous results for the oscillator basis and for the Coulomb problem are outlined in Secs. 2 C and 2 D, respectively. The details of the Coulomb derivation are given in the Appendix. Section 3 contains the application of the results thus obtained to potential scattering problems. This section presents a formula which allows for the computation of phase shifts. Section 4 presents the natural generalization of the J-matrix method to multichannel scattering. Finally, Sec. 5 contains a brief discussion of the overall results and suggestions for applications and areas of further theoretical interest.

2. THE H_0 PROBLEM

The problem examined in this section is the "solution" of the equation,

$$(H_0 - k^2/2)\Psi = 0 \tag{2.1}$$

within the framework of the L^2 function space $\{\Phi_n\}$, in such a manner as to obtain both an asymptotically sine-like and asymptotically cosine-like function. The two J-matrix solutions, $\tilde{S}(r)$ and $\tilde{C}(r)$, form the basis for the asymptotic representation of the scattering wavefunction associated with the full problem. It will also be required that the expansion coefficients of both $\tilde{S}(r)$ and $\tilde{C}(r)$ satisfy the same three-term recursion scheme.

A. Generalized H_0 problem

The basic differential equation

$$(H_0 - k^2/2)\Psi^0 = 0 \tag{2.2}$$

possesses both a regular and an irregular solution which behave near the origin as

$$\Psi^0_{\text{reg}} \underset{r\to 0}{\sim} r^{l+1}, \tag{2.3a}$$

$$\Psi^0_{\text{irreg}} \underset{r\to 0}{\sim} r^{-l} \tag{2.3b}$$

and asymptotically as

$$\Psi^0_{\text{reg}} \underset{r\to\infty}{\sim} \sin\xi, \tag{2.4a}$$

$$\Psi^0_{\text{irreg}} \underset{r\to\infty}{\sim} \cos\xi, \tag{2.4b}$$

where $\xi = kr - \pi l/2$ in the free particle case and $\xi = kr + t\ln 2kr - \pi l/2 + \sigma_l$ in the Coulomb case. In the above, the definitions $t = -z/k$ and $\sigma_l = \arg\Gamma(l+1-it)$ have been used.[3]

Since the basis set $\{\phi_n\}$ is complete for functions regular at the origin, Ψ^0_{reg} can be expanded as

$$\Psi^0_{\text{reg}} \equiv \tilde{S}(r) = \sum_{n=0}^{\infty} s_n \phi_n(\lambda r) \tag{2.5}$$

with the expansion coefficients s_n being formally given by[4]

$$s_n = \langle \bar{\phi}_n(\lambda r) \,|\, \tilde{S}(r) \rangle \tag{2.6}$$

with $\bar{\phi}_n$ satisfying $\langle \bar{\phi}_n \,|\, \phi_m \rangle = \delta_{nm}$. A differential equation satisfied by the set of coefficients $\{s_n\}$ can be constructed in the following manner. Since the basis set $\{\phi_n\}$ tridiagonalizes the operator $(H_0 - k^2/2)$, the $\{s_n\}$ satisfy a three-term recursion relation of the form,

$$[a_{1n} + a_{2n}g(\eta)]s_n + a_{3n}s_{n-1} + a_{4n}s_{n+1} = 0, \quad n > 0, \tag{2.7a}$$

$$[a_{10} + a_{20}g(\eta)]s_0 + a_{40}s_1 = 0, \quad n = 0, \tag{2.7b}$$

where η is the energy variable defined by $\eta = k/\lambda$ and $g(\eta)$ is a function dependent upon the particular choice of $\{\phi_n\}$ and H_0. Differentiating Eq. (2.6) with respect to x, where x is a function of the energy variable η appropriate to the particular case, leads to a differential difference equation of the form

$$\eta \frac{dx}{d\eta}\frac{ds_n}{dx} = b_{1n}s_{n+1} + b_{2n}s_n + b_{3n}s_{n-1}, \quad n > 0 \tag{2.8a}$$

$$\eta \frac{dx}{d\eta}\frac{ds_0}{dx} = b_{10}s_1 + b_{20}s_0, \quad n = 0. \tag{2.8b}$$

For the case of the Laguerre function space, $b_{2n} = 0$, while for the oscillator function space the general form of Eq. (2.8) is appropriate. Combining Eqs. (2.7) and (2.8) yields a linear second order differential equation of the form

$$A(x)\frac{d^2 s_n}{dx^2} + B(x)\frac{ds_n}{dx} + D(x)s_n = 0 \tag{2.9}$$

with two linearly independent solutions χ_1 and χ_2 and a general solution of the form

$$s_n = \alpha_1\chi_1 + \alpha_2\chi_2. \tag{2.10}$$

Equation (2.7b) determines s_n to within a normalization constant. The advantage of the differential equation approach is that a cosine-like solution $\tilde{C}(r)$, whose expansion coefficients will also satisfy the differential equation (2.9), can be readily constructed.

The cosine-like J-matrix solution,

$$\tilde{C}(r) = \sum_{n=0}^{\infty} c_n \phi_n(\lambda r), \tag{2.11}$$

is constructed to be (1) regular at the origin like Ψ^0_{reg}, so as to be expandable in the basis set $\{\phi_n\}$, (2) behave asymptotically as Ψ^0_{irreg}, and (3) to have its expansion coefficients $\{c_n\}$ satisfy Eq. (2.7a). This immediately means that $\tilde{C}(r)$ cannot satisfy the homogeneous differential equation (2.2). By choosing $\tilde{C}(r)$ to satisfy the inhomogeneous differential equation

$$(H_0 - k^2/2)\tilde{C}(r) = \beta\bar{\phi}_0(\lambda r), \tag{2.12}$$

the Green's function[5]

$$G = 2\Psi^0_{\text{reg}}(r_<)\Psi^0_{\text{irreg}}(r_>)/W(\Psi^0_{\text{reg}}, \Psi^0_{\text{irreg}}) \tag{2.13}$$

may be used to obtain the solution[5]

$$\tilde{C}(r) = \frac{-2\beta}{W}\left(\Psi^0_{\text{irreg}}\int_0^r dr'\,\Psi^0_{\text{reg}}\bar{\phi}_0(\lambda r') + \Psi^0_{\text{reg}}\int_r^{\infty} dr'\,\Psi^0_{\text{irreg}}\bar{\phi}_0(\lambda r')\right), \tag{2.14}$$

where $W(\Psi^0_{\text{reg}}, \Psi^0_{\text{irreg}})$ is the Wronskian of the two indepen-

dent solutions $\Psi^0_{\rm reg}$ and $\Psi^0_{\rm irreg}$ and is independent of r and β is a free parameter.[6] $\tilde{C}(r)$ as given by Eq. (2.14) is regular at the origin and with the choice

$$\beta = -W/2s_0 \tag{2.15}$$

goes asymptotically as $\Psi^0_{\rm irreg}$. The fact that the inhomogeneity of Eq. (2.12) is orthogonal to the set $\{\phi_n\}$, for $n = 1, 2, \ldots, \infty$, implies that, for $n > 0$, the $\{c_n\}$ satisfy the same three-term recursion relation as the $\{s_n\}$, Eq. (2.7a). The $n = 0$ case has the form

$$(a_{10} + a_{20}g(\eta))c_0 + a_{40}c_1 = l(\eta) \neq 0, \tag{2.16}$$

where $l(\eta)$ is a function which depends upon the form of β and upon any terms that were divided out in the derivation of the homogeneous recursion relation, Eq. (2.7). Equation (2.16) is to be contrasted with the homogeneous initial condition

$$(a_{10} + a_{20}g(\eta))s_0 + a_{40}s_1 = 0, \tag{2.7b}$$

which occurs in the sine-like J-matrix solution. It may be shown from Eqs. (2.14) and (2.15) that the set $\{c_n\}$ satisfies a differential difference equation analogous to Eq. (2.8), which when combined with Eqs. (2.7a) and (2.16) leads to the differential equation (2.9). The application of the inhomogeneous initial condition given by Eq. (2.16), and an additional boundary condition specific to the case being considered, determine the two integration constants γ_1 and γ_2 in the solution

$$c_n = \gamma_1\chi_1 + \gamma_2\chi_2. \tag{2.17}$$

B. Radial kinetic energy: Laguerre basis

For the case of the radial kinetic energy and a Laguerre basis, the detailed construction of the J-matrix solutions is given following the general technique outlined in Sec. 2 A. The Hamiltonian is

$$H^0_l = -\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2}, \tag{2.18}$$

while the L^2 expansion set is given by

$$\phi_n(\lambda r) = (\lambda r)^{l+1}\exp(-\lambda r/2)L_n^{2l+1}(\lambda r), \quad n = 0, 1, \ldots, \infty. \tag{2.19}$$

In essence the infinite matrix problems

$$\langle\phi_m|(H^0_l - k^2/2)|\tilde{S}(r)\rangle = 0, \quad m = 0, 1, 2, \ldots, \infty, \tag{2.20a}$$

and

$$\langle\phi_m|(H^0_l - k^2/2)|\tilde{C}(r)\rangle = 0, \quad m = 1, 2, 3, \ldots, \infty, \tag{2.20b}$$

are solved where

$$\tilde{S}(r) = \sum_{n=0}^{\infty} s_n\phi_n(\lambda r), \tag{2.5}$$

$$\tilde{C}(r) = \sum_{n=0}^{\infty} c_n\phi_n(\lambda r), \tag{2.11}$$

subject to the asymptotic boundary conditions

$$\tilde{S}(r) \underset{r\to\infty}{\sim} \sin(kr - l\pi/2) \tag{2.21a}$$

and

$$\tilde{C}(r) \underset{r\to\infty}{\sim} \cos(kr - l\pi/2). \tag{2.21b}$$

These solutions have the appropriate asymptotic forms to allow for the formulation of partial wave scattering problems for potentials falling off faster than $1/r^2$ at infinity, where the solution of the scattering problem will have the asymptotic form[3]

$$\Psi \underset{r\to\infty}{\sim} \sin(kr - l\pi/2) + \tan\delta\cos(kr - l\pi/2), \tag{2.22}$$

δ being the scattering phase shift.

From the boundary conditions of Eq. (2.21), $\tilde{S}(r)$ is designated the "sine-like" J-matrix solution, and $\tilde{C}(r)$ the "cosine-like" J-matrix solution. The sine-like solution is discussed in Sec. 2 B 1, where the recurrence relation for the coefficients $\{s_n\}$ is solved explicitly, giving closed form expressions. The discussion of $\tilde{C}(r)$ is somewhat more complex: In Sec. 2 B 2, a function $\tilde{C}(r)$ with the appropriate cosine-like behavior is constructed such that, for $n > 0$, the expansion coefficients $\{c_n\}$ obey the same recursion scheme as the set $\{s_n\}$; a fact that is an essential ingredient of the J-matrix method as will be seen in Secs. 3 and 4 and has been discussed in I and II.

1. Sine-like solution

One of the linearly independent eigenfunctions of the radial kinetic energy,

$$H^0_l = -\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} \tag{2.18}$$

may be taken to be regular at $r = 0$ and sine-like asymptotically, that is,

$$\Psi_{\rm reg}(r) \underset{r\to 0}{\sim} r^{l+1}, \tag{2.23a}$$

$$\Psi_{\rm reg}(r) \underset{r\to\infty}{\sim} \sin(kr - l\pi/2), \tag{2.23b}$$

where the eigenfunction satisfying Eq. (2.23) is referred to as the regular solution.[3] A J-matrix solution,

$$\tilde{S}(r) \equiv \Psi_{\rm reg}(r) = \sum_{n=0}^{\infty} s_n\phi_n(\lambda r), \tag{2.5'}$$

satisfying the boundary conditions of Eq. (2.23), is easily found within the context of the Laguerre set of Eq. (2.19).

The matrix

$$\langle\phi_n|(H^0_l - k^2/2)|\phi_m\rangle = J_{nm} = \int_0^{\infty} dr\,\phi_n(\lambda r)\left(-\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} - \frac{k^2}{2}\right)\phi_m(\lambda r) \tag{2.24}$$

may, upon application of the orthogonality and recursion properties of the Laguerre functions $\phi_n(r)$,[4] be reduced to the Jacobi (J-matrix) form

$$J_{nm} = -\frac{\lambda}{2}\frac{\Gamma(n+2l+2)\eta}{n!\sin\theta} \times [2x(n+l+1)\delta_{mn} - n\delta_{m,n-1} - (n+2l+2)\delta_{m,n+1}], \tag{2.25}$$

where

$$x = \cos\theta = (\eta^2 - \tfrac{1}{4})/(\eta^2 + \tfrac{1}{4}) \tag{2.26a}$$

and

$$\eta = k/\lambda = \tfrac{1}{2}\cot(\theta/2). \tag{2.26b}$$

The expansion coefficients $\{s_n\}$ which satisfy the matrix equation $J \cdot s = 0$ may be determined by the solution of the three-term recursion relation

$$2x(n+l+1)u_n(x) - (n+2l+1)u_{n-1}(x) - (n+1)u_{n+1}(x) = 0 \tag{2.27a}$$

with the initial condition

$$2x(l+1)u_0(x) - u_1(x) = 0, \tag{2.27b}$$

where

$$s_n(x) = [n!/\Gamma(n+2l+2)]u_n(x). \tag{2.28}$$

Formally, of course, the $\{u_n\}$ are given by the Fourier projection[4]

$$u_n(x) = \int_0^\infty dr\, \Psi_{\text{reg}}(\eta r)\phi_n(r)/r. \tag{2.29}$$

Rather than obtaining the $\{u_n\}$ by the direct evaluation of the integral, Eq. (2.29), a linear second order differential equation for the $\{u_n\}$ is derived. This differential equation formulation will be utilized for the construction of the cosine-like solution $\tilde{C}(r)$, where the analog of the projection of Eq. (2.29) does not exist.

Differentiation of Eq. (2.29) with respect to x, utilizing the fact that Ψ_{reg} is a function of (ηr), gives, after application of the chain rule, integration by parts, and application of the Laguerre recursion relations,[4]

$$(x^2-1)\frac{d}{dx}u_n(x) = \frac{n+1}{2}u_{n+1}(x) - \frac{(n+2l+1)}{2}u_{n-1}(x) \tag{2.30a}$$

with the initial condition

$$(x^2-1)\frac{d}{dx}u_0(x) = \tfrac{1}{2}u_1(x). \tag{2.30b}$$

Combining Eqs. (2.27) and (2.30) gives the differential equation

$$(1-x^2)u_n''(x) - xu_n'(x) - [x^2 l(l+1)/(1-x^2) - (n^2+2nl+2n+l+1)]u_n(x) = 0, \tag{2.31}$$

where the differentiation is with respect to x.

Equation (2.31) is easily solved. Letting $u_n(x) = (1-x^2)^{(l+1)/2}v_n(x)$ gives

$$(1-x^2)v_n''(x) - (2l+3)xv_n'(x) + [n(n+2l+2)]v_n(x) = 0, \tag{2.32}$$

which is the differential equation satisfied by the Gegenbauer polynomial $C_n^{l+1}(x)$.[6] The general solution of Eq. (2.31) is then[6]

$$u_n^{\text{gen}}(x) = A_{nl}\chi_1(x) + B_{nl}\chi_2(x), \tag{2.33}$$

where

$$\chi_1(x) = (\sin\theta)^{l+1}C_n^{l+1}(\cos\theta), \tag{2.34a}$$

$$\chi_2(x) = [(\cos\theta/2)^{l+1}/(\sin\theta/2)^l] \times {}_2F_1(n+l+\tfrac{3}{2}, -n-l-\tfrac{1}{2}; \tfrac{1}{2}-l; \sin^2\theta/2), \tag{2.34b}$$

where again $x = \cos\theta$ and ${}_2F_1(a,b;c;z)$ is the Gauss hypergeometric function.[7] The coefficients A_{nl} and B_{nl} can be determined to within an l dependent factor by substitution into Eq. (2.27a), resulting in

$$u_n^{\text{gen}}(x) = a_l\chi_1(x) + b_l\chi_2(x). \tag{2.35}$$

The form of $u_n^{\text{gen}}(x)$ appropriate to the initial condition

$$2x(l+1)u_0(x) - u_1(x) = 0 \tag{2.27b}$$

may be determined from Eq. (2.34) as

$$u_n(x) = a_l\chi_1(x) \tag{2.36}$$

since[4]

$$2x(l+1)C_0^{l+1}(x) - C_1^{l+1}(x) = 0 \tag{2.37}$$

while $\chi_2(x)$ does not satisfy Eq. (2.27b).[7] Substitution of Eq. (2.36) into Eq. (2.5) gives

$$\tilde{S}(r) = a_l(\sin\theta)^{l+1}\sum_{n=0}^{\infty}\frac{n!}{\Gamma(n+2l+2)}C_n^{l+1}(\cos\theta)\phi_n(\lambda r), \tag{2.38}$$

where the requirement that

$$\lim_{\substack{r\to 0\\ k\to 0}}[\tilde{S}(kr) - \Psi_{\text{reg}}(kr)] = 0 \tag{2.39}$$

determines a_l as $2^l\Gamma(l+1)$ and thus that

$$s_n(x) = [2^l\Gamma(l+1)n!/\Gamma(n+2l+2)](\sin\theta)^{l+1}C_n^{l+1}(\cos\theta). \tag{2.40}$$

The coefficients $\{s_n\}$ of the regular, sine-like, eigenfunction of the radial kinetic energy have now been determined by the solution of a linear second order differential equation followed by the imposition of the appropriate boundary conditions.

2. Cosine-like solution

The cosine-like eigenfunction of the radial kinetic energy, which is irregular at the origin and defined by the conditions

$$\Psi_{\text{irreg}}(r) \underset{r\to 0}{\sim} r^{-l}, \tag{2.41a}$$

$$\Psi_{\text{irreg}}(r) \underset{r\to\infty}{\sim} \cos(kr - \pi l/2), \tag{2.41b}$$

will be referred to as the irregular solution.[3] For the construction of a cosine-like J-matrix solution

$$\tilde{C}(r) = \sum_{n=0}^{\infty} c_n\phi_n(\lambda r) \tag{2.11}$$

with the asymptotic boundary condition,

$$\tilde{C}(r) \underset{r\to\infty}{\sim} \cos(kr - \pi l/2) \tag{2.42}$$

it is seen that $\tilde{C}(r) \neq \Psi_{\text{irreg}}(r)$ since $\tilde{C}(r) \underset{r\to 0}{\sim} r^{l+1}$ as follows from Eq. (2.11). Thus, the expansion coefficients $\{c_n\}$ cannot be written as a Fourier projection of the form

$$\int_0^\infty dr\, \Psi_{\text{irreg}}(\eta r)\phi_n(r)/r \tag{2.43}$$

in analogy to Eq. (2.29). A cosine-like J-matrix solution $\tilde{C}(r)$ must thus be constructed with the following requirements: (1) $\tilde{C}(r)$ should have a cosine-like asymptotic form; (2) $\tilde{C}(r)$ should be regular at the origin; and (3) the coefficients $\{c_n\}$ should satisfy the same three-term recursion relation as the set $\{s_n\}$ for $n > 0$. Actually, from I it is seen that the most general condition in requirement (3) is $n \geq N+1$, where N is the number of functions in the subspace onto which the potential V is projected in the formulation of the model problem. For the purposes of the J-matrix method, however, it is sufficient to consider the condition $n > 0$. It is immediately seen that $\tilde{C}(r)$ will not satisfy $\langle \phi_m | (H_l^0 - k^2/2) | \times \tilde{C}(r) \rangle = 0$, $m = 0, 1, 2, \ldots, \infty$, as the cosine-like eigenfunction of H_l^0 which is linearly independent of Ψ_{reg} is

Ψ_{irreg}, which is not regular at the origin.

The function $\tilde{C}(r)$ satisfying the above conditions is given by the solution of the equation,

$$(H_l^0 - k^2/2)\tilde{C}(r) = \beta\bar{\phi}_0(\lambda r) \tag{2.44}$$

subject to the boundary conditions

$$\tilde{C}(r) \underset{r\to 0}{\sim} r^{l+1}, \tag{2.45a}$$

$$\tilde{C}(r) \underset{r\to\infty}{\sim} \cos(kr - \pi l/2), \tag{2.45b}$$

where $\bar{\phi}_0(\lambda r) = \phi_0(\lambda r)/r\Gamma(2l+2)$.[4] It is noted that the particular choice of the inhomogeneity in Eq. (2.44) gives the infinite matrix problem

$$\langle\phi_m|(H_l^0 - k^2/2)|\tilde{C}(r)\rangle = \beta\delta_{m0}, \quad m = 0, 1, 2, \ldots, \infty, \tag{2.46}$$

immediately implying that, for $n > 0$, the $\{c_n\}$ and $\{s_n\}$ satisfy the same three-term pure recurrence relation. The parameter β is determined from a Green's function construction of the solution to Eq. (2.44) building in the boundary conditions of Eq. (2.45). Using Ψ_{reg} and Ψ_{irreg} as the two linearly independent solutions of the homogeneous equation $(H_l^0 - k^2/2)\tilde{C}(r) = 0$, the solution of the inhomogeneous problem of Eq. (2.44) may be written as[5]

$$\tilde{C}(r) = \frac{-2\beta}{W}\left(\Psi_{\text{irreg}}\int_0^r dr'\,\Psi_{\text{reg}}\bar{\phi}_0(\lambda r') + \Psi_{\text{reg}}\int_r^\infty dr'\,\Psi_{\text{irreg}}\bar{\phi}_0(\lambda r')\right), \tag{2.47}$$

where W is the Wronskian and is equal to[3] $(-k)$ for the radial kinetic energy case. By taking the $r\to\infty$ limit, β is determined as

$$\beta = \frac{-W}{2s_0} = \frac{2^l k\Gamma(l+\frac{3}{2})}{\sqrt{\pi}(\sin\theta)^{l+1}}, \tag{2.48}$$

where s_0 is the zeroth expansion coefficient of $\tilde{S}(r)$.

To demonstrate that the $\{c_n\}$ satisfy the same differential equation as the $\{s_n\}$, the equation

$$u_n(x) = \frac{\Gamma(n+2l+2)}{n!}c_n(x) = \int_0^\infty dr\,\tilde{C}(r/\lambda)\frac{\phi_n(r)}{r}, \tag{2.49}$$

where $\tilde{C}(r)$ is given by Eq. (2.47) with β given by Eq. (2.48), is differentiated with respect to x. Applying the same procedures that were used in going from Eq. (2.29) to (2.30) yields the differential difference equation of Eq. (2.30a), which, when combined with the previously derived recursion relation, Eq. (2.27a), gives the differential equation (2.31). The solution $u_n(x)$ satisfying the initial condition

$$2x(l+1)u_0(x) - u_1(x) = -2^{l+1}\Gamma(l+\tfrac{3}{2})/\sqrt{\pi}(\sin\theta)^l \tag{2.50}$$

is given by

$$c_n(x) = a_l\chi_1(x) - \frac{\Gamma(l+\frac{1}{2})}{\sqrt{\pi}}\frac{n!}{\Gamma(n+2l+2)}\chi_2(x), \tag{2.51}$$

where $\chi_1(x)$ and $\chi_2(x)$ are given by Eq. (2.34). From the explicit expressions for $\tilde{S}(r)$ and $\tilde{C}(r)$ in terms of Ψ_{reg} and Ψ_{irreg}, it follows that

$$c_n(-\theta) = (-)^l c_n(\theta), \tag{2.52a}$$

$$s_n(-\theta) = (-)^{l+1}s_n(\theta), \tag{2.52b}$$

immediately establishing that $a_l = 0$ and giving

$$c_n(x) = \frac{-\Gamma(l+\frac{1}{2})n!}{\sqrt{\pi}\Gamma(n+2l+2)}\frac{(\cos\theta/2)^{l+1}}{(\sin\theta/2)^l} \times {}_2F_1(n+l+\tfrac{3}{2}, -n-l-\tfrac{1}{2}; \tfrac{1}{2}-l; \sin^2\theta/2). \tag{2.53}$$

A useful alternative form of $c_n(x)$ is given by[7]

$$c_n(x) = \frac{-2^l\Gamma(l+\frac{1}{2})n!}{\sqrt{\pi}\Gamma(n+2l+2)}\frac{1}{(\sin\theta)^l} \times {}_2F_1(-n-2l-1, n+1; \tfrac{1}{2}-l; \sin^2\theta/2), \tag{2.54}$$

which is a finite polynomial in $(\sin^2\theta/2)$.

C. Radial kinetic energy: Oscillator basis

Within the framework of the oscillator basis

$$\phi_n(\lambda r) = (\lambda r)^{l+1}\exp(-\lambda^2 r^2/2)\,L_n^{l+1/2}(\lambda^2 r^2), \quad n = 0, 1, \ldots, \infty, \tag{2.55}$$

the J-matrix defined by

$$J_{nm} = \langle\phi_n|(H_l^0 - k^2/2)|\phi_m\rangle = \int_0^\infty dr\,\phi_n(\lambda r)\left(-\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} - \frac{k^2}{2}\right)\phi_m(\lambda r) \tag{2.56}$$

is a tridiagonal (Jacobi) matrix leading to the fundamental recursion relation

$$-(2n+l+\tfrac{3}{2}-\eta^2)u_n(\eta^2) + (n+l+\tfrac{1}{2})u_{n-1}(\eta^2) + (n+1)u_{n+1}(\eta^2) = 0, \quad n > 0, \tag{2.57}$$

where again $\eta = k/\lambda$, for the solution of the infinite matrix problems

$$\langle\phi_m|(H_l^0 - k^2/2)|\tilde{S}(r)\rangle = 0, \quad m = 1, 2, \ldots, \infty, \tag{2.58a}$$

$$\langle\phi_m|(H_l^0 - k^2/2)|\tilde{C}(r)\rangle = 0, \quad m = 1, 2, \ldots, \infty, \tag{2.58b}$$

for the sine-like and cosine-like J-matrix solutions.

The set $\{u_n\}$ satisfies the differential equation

$$u_n''(\eta^2) + \left((4n+2l+3) - \eta^2 - \frac{l(l+1)}{\eta^2}\right)u_n(\eta^2) = 0, \tag{2.59}$$

where the differentiation is with respect to η. Equation (2.59) has the general solution[6]

$$u_n^{\text{gen}}(\eta) = A_{nl}\exp(-\eta^2/2)\eta^{l+1}L_n^{l+1/2}(\eta^2) + B_{nl}\exp(-\eta^2/2)\eta^{-l} \times {}_1F_1(-n-l-\tfrac{1}{2}, \tfrac{1}{2}-l, \eta^2), \tag{2.60}$$

where ${}_1F_1(a, c, z)$ is the confluent hypergeometric function.[7] The sine-like solution is deduced from Eq. (2.60) by substitution into Eq. (2.57), imposition of the initial condition

$$(l+\tfrac{3}{2}-\eta^2)u_0(\eta^2) - u_1(\eta^2) = 0, \tag{2.61}$$

and normalization in the manner of Eq. (2.39), giving

$$s_n(\eta^2) = \frac{(-)^n n!}{\Gamma(n+l+\frac{3}{2})}u_n(\eta^2) = \frac{\sqrt{2\pi}(-)^n n!}{\Gamma(n+l+\frac{3}{2})}\exp(-\eta^2/2)\eta^{l+1}L_n^{l+1/2}(\eta^2), \tag{2.62}$$

where

$$\tilde{S}(r) = \sum_{n=0}^\infty s_n(\eta^2)\phi_n(\lambda r). \tag{2.63}$$

Substitution into Eq. (2.57), imposition of the initial

condition appropriate to the construction of a cosine-like solution,

$$(l+\tfrac{3}{2}-\eta^2)u_0(\eta^2)-u_1(\eta^2)=\Gamma(l+\tfrac{3}{2})(2/\pi)^{1/2}\exp(\eta^2/2)\eta^{-l}, \tag{2.64}$$

and use of the symmetry conditions given in Eq. (2.52) give

$$c_n(\eta^2)=\left(\frac{2}{\pi}\right)^{1/2}\frac{\Gamma(l+\frac{1}{2})(-)^n n!}{\Gamma(n+l+\frac{3}{2})}\exp(-\eta^2/2)\eta^{-l} \times {}_1F_1(-n-l-\tfrac{1}{2},\tfrac{1}{2}-l,\eta^2), \tag{2.65}$$

where

$$\tilde{C}(r)=\sum_{n=0}^{\infty}c_n(\eta^2)\phi_n(\lambda r). \tag{2.66}$$

$\tilde{S}(r)$ and $\tilde{C}(r)$ are real, regular at the origin, and have sine-like and cosine-like asymptotic forms, respectively.

D. Radial Coulomb Hamiltonian: Laguerre basis

Within the framework of the Laguerre basis, Coulomb J-matrix solutions $\tilde{S}(r)$ and $\tilde{C}(r)$ are constructed, which are regular at the origin and behave asymptotically as[3]

$$\tilde{S}(r)\underset{r\to\infty}{\sim}\sin(kr+t\ln 2kr-\pi l/2+\sigma_l), \tag{2.67a}$$

$$\tilde{C}(r)\underset{r\to\infty}{\sim}\cos(kr+t\ln 2kr-\pi l/2+\sigma_l). \tag{2.67b}$$

The Coulomb J matrix

$$J_{nm}=\langle\phi_n|(H^0_{l,z}-k^2/2)|\phi_m\rangle = \int_0^\infty dr\,\phi_n(\lambda r)\left(-\frac{1}{2}\frac{d^2}{dr^2}+\frac{l(l+1)}{2r^2}+\frac{z}{r}-\frac{k^2}{2}\right)\phi_m(\lambda r) \tag{2.68}$$

is a tridiagonal form, yielding the recurrence relation

$$\left[(n+l+1)\left(\frac{x^2+1}{x}\right)-it\left(\frac{x^2-1}{x}\right)\right]v_n(x) - (n+1)v_{n+1}(x)-(n+2l+1)v_{n-1}(x)=0, \quad n>0, \tag{2.69}$$

where

$$x=\exp(i\theta)=-[(\tfrac{1}{2}-i\eta)/(\tfrac{1}{2}+i\eta)] \tag{2.70}$$

and

$$s_n=\frac{n!\,2^l|\Gamma(l+1-it)|}{\Gamma(n+2l+2)\Gamma(2l+2)\exp(-\pi t/2)}v_n \tag{2.71}$$

for the solution of the infinite matrix equations

$$\langle\phi_m|(H^0_{l,z}-k^2/2)|\tilde{S}(r)\rangle=0, \quad m=1,2,3,\ldots,\infty, \tag{2.72a}$$

$$\langle\phi_m|(H^0_{l,z}-k^2/2)|\tilde{C}(r)\rangle=0, \quad m=1,2,3,\ldots,\infty, \tag{2.72b}$$

for the sine-like and cosine-like J-matrix solutions.

The $\{v_n\}$ satisfy the differential equation

$$(x^2-1)v_n''(x)+\frac{x^2-1}{x}v_n'(x)-\left[\left(\frac{x^2-1}{x^2}\right)(n^2+2nl+2n+l+1-t^2) + \frac{(x^2+1)}{x^2(x^2-1)}[(x^2+1)l(l+1)-2it(n+l+1)(x^2-1)]\right] \times v_n(x)=0 \tag{2.73}$$

where the differentiation is with respect to $x=\exp(i\theta)$ and where $t=-z/k$ is considered to be independent of x. The derivation of this equation is discussed in the Appendix. The general solution of Eq. (2.73) is[7]

$$v_n^{\text{gen}}(x)=A_{nlt}(\sin\theta)^{l+1}\exp(\theta t)\exp(-in\theta) \times {}_2F_1(-n,l+1-it;2l+2;1-\exp(2i\theta))+B_{nlt}(\sin\theta)^{it} \times\exp[-i(n+l+1)\theta] \times {}_2F_1(-l-it,l+1-it;n+l+2-it;1/(1-e^{2i\theta})). \tag{2.74}$$

The sine-like solution is determined from Eq. (2.74) by substitution into Eq. (2.69), imposition of the initial condition

$$\left[(l+1)\left(\frac{x^2+1}{x}\right)-it\left(\frac{x^2-1}{x}\right)\right]v_0(x)-v_1(x)=0, \tag{2.75}$$

and by application of a normalization procedure discussed in the Appendix, giving

$$s_n(\theta)=[2^l n!\,|\Gamma(l+1-it)|/\Gamma(n+2l+2)] \times\exp[(\pi/2+\epsilon\pi)t]\exp(\theta t)(\sin\theta)^{l+1} \times P_n^{l+1}(\cos\theta;2z/\lambda;-2z/\lambda), \tag{2.76}$$

where

$$\epsilon=\begin{cases}(-) & \text{for } \theta[0,\pi]\\ (+) & \text{for } \theta[0,-\pi]\end{cases} \tag{2.77}$$

and $P_n^{\lambda}(z;a;b)$ is the Pollaczek polynomial[4] as discussed in the Appendix. The cosine-like solution is determined by substitution into Eq. (2.69), imposition of the initial condition

$$\left[(l+1)\left(\frac{x^2+1}{x}\right)-it\left(\frac{x^2-1}{x}\right)\right]v_0(x)-v_1(x) = \frac{-[\Gamma(2l+2)]^2}{2^{2l}\exp(\theta t)|\Gamma(l+1-it)|^2(\sin\theta)^l} \tag{2.78}$$

and application of a limiting procedure as discussed in the Appendix, giving

$$c_n+is_n=\frac{-n!\exp(i\sigma_l)\exp(\pi t/2)\exp(-\theta t)\exp[-i(n+1)\theta]}{\Gamma(n+l+2-it)2^l(\sin\theta)^l} \times {}_2F_1(-l-it,n+1;n+l+2-it;\exp(-2i\theta)). \tag{2.79}$$

The functions $\tilde{S}(r)$ and $\tilde{C}(r)$ are real and reduce to the radial kinetic energy results when $z=0$.

3. POTENTIAL SCATTERING

In this section, it will be assumed that the potential V does not couple angular momentum eigenstates; the generalization to the case where coupling occurs is straightforward and will be considered in the multichannel case discussed in Sec. 4. Thus, the good angular momentum quantum number l will be suppressed, it being implicitly assumed that a definite partial wave is under consideration. The aim of this section is then to determine the phase shift caused by the potential V with respect to the uncoupled Hamiltonian H_0, which may be taken to be the lth partial wave kinetic energy or Coulomb Hamiltonian.

As alluded to in the Introduction, and motivated in I, the potential V is approximated by truncating its representation in the function space $\{\phi_n\}$ to a finite, $N\times N$, representation V^N defined by

$$V^N_{n,m} = \begin{cases} \int_0^\infty \phi_n(\lambda r) V(r) \phi_m(\lambda r)\, dr, & n, m \leq N-1, \\ 0, & \text{otherwise.} \end{cases} \tag{3.1}$$

The problem is then to solve

$$\langle \Phi_m | (H_0 + V^N - k^2/2) | \Psi_E \rangle = 0, \quad m = 0, 1, 2, \ldots, \infty, \tag{3.2}$$

where

$$\Psi_E = \sum_{n=0}^{\infty} d_n \phi_n(\lambda r). \tag{3.3}$$

The form of V^N as defined in (3.1) is such, however, that it only couples the first N functions ϕ_m, $m = 0, 1, 2, \ldots, N-1$, in the infinite function space. Thus, outside the space spanned by these N basis functions the generalized sine-like and cosine-like solutions associated with the generalized H_0 problem discussed in Sec. 2 are valid. This leads to the following form for the wavefunction $\Psi_E(r)$,

$$\Psi_E(r) = \Phi(r) + S(r) + tC(r), \tag{3.4}$$

where

$$\Phi(r) = \sum_{n=0}^{N-1} a_n \phi_n(\lambda r), \tag{3.5a}$$

$$S(r) = \sum_{n=N}^{\infty} s_n \phi_n(\lambda r), \tag{3.5b}$$

$$C(r) = \sum_{n=N}^{\infty} c_n \phi_n(\lambda r), \tag{3.5c}$$

and t is the tangent of the phase shift caused by V^N with respect to H_0. Note that the first N terms in the expansions of $\tilde{S}(r)$ and $\tilde{C}(r)$ have been incorporated into the $\Phi(r)$ term, the remainder of the expansions being designated as $S(r)$ and $C(r)$, respectively. The sets of coefficients $\{s_n\}$ and $\{c_n\}$ are just those that were determined in Sec. 2 and are dependent upon the particular H_0 and basis set $\{\phi_n\}$ being considered. The solution given in (3.4) contains $N+1$ unknowns, $(t, \{a_n\}, n = 0, 1, \ldots, N-1)$.

By returning to Eq. (3.2), it is immediately established that it is satisfied for $m \geq N+1$: Since V^N is defined to be zero in this region of the function space and due to the tridiagonal representation of $(H_0 - k^2/2)$,

$$\langle \phi_m | (H_0 + V^N - k^2/2) | \Psi_E \rangle$$

$$= \langle \phi_m | (H_0 - k^2/2) | S + tC \rangle \tag{3.6a}$$

$$= \langle \phi_m | J | S + tC \rangle \tag{3.6b}$$

$$= \sum_{n=0}^{\infty} J_{mn}(s_n + tc_n) \tag{3.6c}$$

$$= J_{m,m-1}(s_{m-1} + tc_{m-1}) + J_{m,m}(s_m + tc_m) + J_{m,m+1}(s_{m+1} + tc_{m+1}) \tag{3.6d}$$

$$= 0. \tag{3.6e}$$

Equation (3.6e) follows from the recursion relation satisfied by both s_n and c_n. The remaining $N+1$ conditions corresponding to the cases $m = 0, 1, \ldots, N$ are sufficient to determine the $N+1$ unknowns, and by following the analysis presented in I the resulting system matrix is obtained:

$$\begin{bmatrix} (J+V)_{0,0} & \cdots & (J+V)_{0,N-2} & (J+V)_{0,N-1} & 0 \\ (J+V)_{1,0} & \cdots & (J+V)_{1,N-2} & (J+V)_{1,N-1} & 0 \\ \vdots & & \vdots & \vdots & \vdots \\ (J+V)_{N-2,0} & \cdots & (J+V)_{N-2,N-2} & (J+V)_{N-2,N-1} & 0 \\ (J+V)_{N-1,0} & \cdots & (J+V)_{N-1,N-2} & (J+V)_{N-1,N-1} & J_{N-1,N}c_N \\ 0 & \cdots & 0 & J_{N,N-1} & -J_{N,N-1}c_{N-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{N-2} \\ a_{N-1} \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -J_{N-1,N}s_N \\ J_{N,N-1}s_{N-1} \end{bmatrix}. \tag{3.7}$$

Equation (3.7) can be immediately solved for t by standard techniques. In particular, an expression for $t = \tan\delta$ can be obtained by prediagonalizing the inner $N \times N$ matrix $(H_0 + V^N - k^2/2)_{mn}$ with the energy independent transformation Γ defined by

$$[\tilde{\Gamma}(H_0 + V^N - k^2/2)\Gamma]_{n,m} = (E_n - E)\delta_{n,m}, \tag{3.8}$$

where $\{E_n\}$ are the Harris eigenvalues. Augmenting Γ to be the $(N+1) \times (N+1)$ matrix,

$$\Gamma_A = \begin{bmatrix} \Gamma & 0 \\ 0 & 1 \end{bmatrix}, \tag{3.9}$$

and applying it to Eq. (3.7) gives

$$t = \tan\delta = -\frac{s_{N-1} + r(E)(J_{N,N-1})s_N}{c_{N-1} + r(E)(J_{N,N-1})c_N}, \tag{3.10}$$

where

$$r(E) = \sum_{m=0}^{N-1} \frac{\Gamma^2_{N-1,m}}{(E_m - E)}. \tag{3.11}$$

At the positive Harris eigenvalues E_n,[8] $\tan\delta$ reduces to

$$\tan\delta(E_n) = s_N(\eta_n)/c_N(\eta_n), \quad \eta_n = k_n/\lambda. \tag{3.12}$$

Also, at the positive energies E_μ, where $r(E_\mu) = 0$, $\tan\delta$ becomes

$$\tan\delta(E_\mu) = -s_{N-1}(\eta_\mu)/c_{N-1}(\eta_\mu), \quad \eta_\mu = k_\mu/\lambda. \tag{3.13}$$

Equations (3.7), (3.10), (3.12), and (3.13) are the generalizations of the corresponding results obtained in I.

4. MULTICHANNEL SCATTERING

The potential scattering results of the previous section will now be generalized to include collisions with targets having internal states. Due to the generalizations developed for solving the H_0 problem, the results of which are given in Secs. 2 B, C, D, collisions with charged, as well as neutral targets, can be considered, while employing the appropriate basis set (Laguerre or

oscillator in the neutral case, and Laguerre in the charged case) to describe the projectile wavefunction. Since the method is formulated in terms of the close-coupling[9] equations, exchange can be treated by the inclusion of the appropriate nonlocal potential followed by its truncation in the J-matrix sense as was done in II.

The target will be described by the collective coordinate ρ, and its dynamics by $H_t(\rho)$. It is assumed that the target possesses a discrete set of L^2 eigenfunctions $R_\gamma(\rho)$ and, further, that the collective quantum number γ includes the total orbital quantum number of the target, L_t, its projection on some specified direction (the z axis), M_t, in addition to the quantum numbers μ that are needed further to completely define the target states. If the target has a dense or continuous spectrum, the method of pseudotarget states may be employed.[2,10] It then follows that

$$H_t(\rho)R_\gamma(\rho) = E_{\mu,L_t}R_\gamma(\rho), \tag{4.1}$$

where $\gamma = \{\mu, L_t, M_t\}$.

The wavefunction Θ describing the projectile-target system satisfies the following Schrödinger equation:

$$(H_t(\rho) + H_0(\mathbf{r}) + V(\mathbf{r},\rho) - E)\Theta(\mathbf{r},\rho) = 0, \tag{4.2}$$

where $\mathbf{r}$ is the projectile coordinate, $H_0(\mathbf{r}) = -\frac{1}{2}\nabla_r^2 + z/r$, its Hamiltonian in atomic units, and $V(\mathbf{r},\rho)$, the interaction between the projectile and the target. In the neutral case z is equal to zero.

For most cases of interest, such as the scattering of electrons by light atoms, the total angular momentum of the system, L, its projection along the z axis, M, as well as the parity, Π, are conserved in the collision process. Therefore, it is more convenient to use a representation which is diagonal in these quantum numbers. Coupling the projectile with the target[9] in a picture with definite total L and total M leads to defining

$$\chi_\Gamma(\hat{r},\rho) = \sum_{m,M_t} C(lL_tL, mM_tM)Y_{lm}(\hat{r})R_\gamma(\rho), \tag{4.3}$$

where $C(lL_tL, mM_tM)$ is the Clebsch–Gordan coefficient, Γ, the channel index $\{\mu L_t l L M\}$, and $Y_{lm}(\hat{r})$, the spherical harmonic functions. It is noted that χ_Γ is an eigenfunction of L and M, satisfying the same equation as R_γ, namely,

$$H_t(\rho)\chi_\Gamma(\hat{r},\rho) = E_{\mu,L_t}\chi_\Gamma(\hat{r},\rho), \tag{4.4}$$

where L satisfies the triangular relations, $|L_t - l| \le L \le L_t + l$. Since the function χ_Γ is composed of L^2 functions and is an eigenfunction of $H_t(\rho)$, the set $\{\chi_\Gamma\}$ may be taken to be orthonormal,

$$\langle\chi_\Gamma|\chi_{\Gamma'}\rangle = \delta_{\Gamma\Gamma'} = \delta_{\mu\mu'}\delta_{L_tL_t'}\delta_{ll'}, \tag{4.5}$$

where $\Gamma = \{\mu L_t l L M\}$ and $\Gamma' = \{\mu' L_t' l' L M\}$.

Within a definite LM picture, the Hilbert space of the system is spanned by the set

$$\{|\chi_\Gamma\rangle|\phi_n^{(l,\Gamma)}\rangle\} \tag{4.6}$$

for all μ, L_t, l, and for $n = 0, 1, \ldots, \infty$. For neutral targets the set $\{\phi_n^{(l,\Gamma)}\}$ can be either the Laguerre functions

$$\phi_n^{(l,\Gamma)}(\lambda_\Gamma r) = (\lambda_\Gamma r)^{l+1}\exp(-\lambda_\Gamma r/2)L_n^{2l+1}(\lambda_\Gamma r),$$

$$n = 0, 1, \ldots, \infty, \tag{4.7}$$

or the oscillator functions

$$\phi_n^{(l,\Gamma)}(\lambda_\Gamma r) = (\lambda_\Gamma r)^{l+1}\exp(-\lambda_\Gamma^2 r^2/2)L_n^{l+1/2}(\lambda_\Gamma^2 r^2),$$

$$n = 0, 1, \ldots, \infty, \tag{4.8}$$

while for charged targets the set $\{\phi_n^{(l,\Gamma)}\}$ can only be the Laguerre functions. Note that the projectile basis can be made channel dependent through the channel dependent scaling parameter λ_Γ.

In general it is not possible to solve Eq. (4.2) for $\Theta(\mathbf{r},\rho)$ exactly in the Hilbert space of Eq. (4.6). By following the procedure in I, the interaction potential V is approximated by $\tilde{V}$, which is defined by the following truncation scheme:

$$\tilde{V}_{nn'}^{\Gamma\Gamma'} = \langle\chi_\Gamma\phi_n^{(l,\Gamma)}|V(\mathbf{r},\rho)|\chi_{\Gamma'}\phi_{n'}^{(l'\Gamma')}\rangle$$

$$\text{for } \Gamma, \Gamma' \le N_c,\ n \le N_{\Gamma-1},\ n' \le N_{\Gamma'-1},$$

$$= 0, \qquad \text{otherwise.} \tag{4.9}$$

where N_Γ is the truncation limit in the channel Γ and N_c is the total number of channels that are allowed to couple through the potential.

It is now proposed to solve the model equation

$$(H_t(\rho) + H_0(\mathbf{r}) + \tilde{V}(\mathbf{r},\rho) - E)\Theta(\mathbf{r},\rho) = 0 \tag{4.10}$$

exactly in the Hilbert space of Eq. (4.6). There will be as many independent solutions Θ_Γ as there are open channels. By following the approach taken in I, Θ_Γ is expanded as

$$\Theta_\Gamma(\mathbf{r},\rho) = \Phi_\Gamma(\mathbf{r},\rho) + \frac{\chi_\Gamma(\hat{r},\rho)S_\Gamma(r)}{\sqrt{k_\Gamma}} + \sum_{\Gamma'}^{N_c}\frac{\chi_{\Gamma'}(\hat{r},\rho)C_{\Gamma'}(r)R_{\Gamma'\Gamma}}{\sqrt{k_{\Gamma'}}},$$

$$\Gamma = 1, 2, \ldots, N_0, \tag{4.11}$$

where N_0 is the number of open channels. The scattering matrix S^{LM} is related to the reactance matrix elements by the relation,

$$S^{LM} = e^{i\sigma}(1 + i[R])(1 - i[R])^{-1}e^{i\sigma}, \tag{4.12a}$$

where $[R]$ is the $N_0 \times N_0$ open-channel part of R and

$$(e^{i\sigma})_{\Gamma\Gamma'} = e^{i\sigma_l}\delta_{\Gamma\Gamma'}. \tag{4.12b}$$

In Eq. (4.11), the quantity k_Γ is the wavenumber of the scattered electron and is given by

$$k_\Gamma = (2|E - E_{\mu,L_t}|)^{1/2}. \tag{4.13}$$

Furthermore, for an open channel Γ,

$$S_\Gamma(r) = \sum_{n=N_\Gamma}^{\infty} s_n(k_\Gamma)\phi_n^{(l,\Gamma)}(\lambda_\Gamma r), \tag{4.14a}$$

$$C_\Gamma(r) = \sum_{n=N_\Gamma}^{\infty} c_n(k_\Gamma)\phi_n^{(l,\Gamma)}(\lambda_\Gamma r), \tag{4.14b}$$

where $\{s_n\}$ and $\{c_n\}$ are given by the results of Sec. 2, while, for a closed channel Γ, the cosine-like term $C_\Gamma(r)$ is replaced by the linear combination

$$C_\Gamma(r) + iS_\Gamma(r) \tag{4.15}$$

and evaluated at $k_\Gamma = i|k_\Gamma|$. The resulting function correctly describes the exponentially decaying closed channel asymptotic behavior. In the neutral case, the function of Eq. (4.15) is related to the spherical Hankel function of the first kind, $h_l^{(1)}(z)$, evaluated at $z = i|k_\Gamma|r$. The internal function, $\Phi_\Gamma(\mathbf{r},\rho)$, which describes the

scattering process at close distances, is written as

$$\Phi_\Gamma(\mathbf{r},\rho)=\sum_{\Gamma'}^{N_c}\sum_{n=0}^{N_{\Gamma'}-1} a_n^{\Gamma'\Gamma}\chi_{\Gamma'}(\hat{r},\rho)\phi_n^{(l',\Gamma')}(\lambda_{\Gamma'}r). \tag{4.16}$$

The remainder of this section briefly demonstrates that the wavefunction Θ is capable of being an exact solution of the model Hamiltonian in the Hilbert space. This is accomplished by uniquely determining the unknowns $\{a_n^{\Gamma'\Gamma}, R_{\Gamma'\Gamma}\}$. It is required that all projections by $\langle\chi_{\Gamma''}\phi_{n''}^{(l'',\Gamma'')}|$ from the left-hand side of Eq. (4.10) vanish:

$$\langle\chi_{\Gamma''}\phi_{n''}^{(l'',\Gamma'')}\left|(H_t(\rho)+H_0(\mathbf{r})+\tilde{V}(\mathbf{r},\rho)-E)\right|\Theta_\Gamma(\mathbf{r},\rho)\rangle=0. \tag{4.17}$$

For the case of $\Gamma=1,2,\ldots,N_0$, $\Gamma''=1,2,\ldots,N_c$, and $n''=0,1,\ldots,N_{\Gamma''}$, $N_0(\sum_{\Gamma=1}^{N_c}(N_{\Gamma''}+1))$ equations are obtained which is equal to the total number of unknowns: $N_0(\sum_{\Gamma'=1}^{N_c}N_{\Gamma'})$ of the a's and N_0N_c of the R's. It only remains now to show that for all other cases Eq. (4.17) is satisfied. The assertion is clear when $\Gamma''>N_c$. Suppose now that $\Gamma''\leq N_c$, but than $n''\geq N_{\Gamma''}+1$. Then $\tilde{V}_{n''n'''}^{\Gamma''\Gamma'''}=0$ by definition and Eq. (4.17) reduces to

$$\left[\langle\phi_{n''}^{(l'',\Gamma'')}\left|\left(-\frac{1}{2}\frac{d^2}{dr^2}+\frac{l''(l''+1)}{2r^2}+\frac{z}{r}-(E-E_{\mu''},{}_{L_t''})\right)\right|\times\sum_{n=0}^{N_{\Gamma''}-1}a_n^{\Gamma'''\Gamma}\phi_n^{(l'',\Gamma'')}+\frac{S_{\Gamma''}}{\sqrt{k_{\Gamma''}}}\delta_{\Gamma\Gamma''}+\frac{C_{\Gamma''}R_{\Gamma''\Gamma}}{\sqrt{k_{\Gamma''}}}\right\rangle=0,\right] \tag{4.18}$$

where $z=0$ in the neutral target case. Since the operator appearing in Eq. (4.18) is tridiagonal in the $\{\phi_n^{(l'',\Gamma'')}\}$ representation, the contribution of the S and C functions vanish since their expansion coefficients $\{s_n\}$ and $\{c_n\}$ satisfy the resulting three-term recursion relation. Since $n''\geq N_{\Gamma''}+1$ while $n\leq N_{\Gamma''}-1$, the contribution of the sum term in Eq. (4.18) vanishes, proving that Eq. (4.18) is identically true.

As in I, the nontrivial equations can be arranged such that the L^2 matrix elements of $(H_t+H_0+\tilde{V})$ appear in an inner block. Additionally, one extra row and column are added to this block for each channel $\Gamma\leq N_c$. The right-hand side driving term and the solution vector, containing the $a_n^{\Gamma\Gamma'}$ and $R_{\Gamma\Gamma'}$, have as many columns as open channels. The R matrix can then be extracted by solving the resulting linear equations by standard techniques. After having done so, the $\mathcal{S}^{LM}$ matrix is constructed via Eq. (4.12), from which physical quantities are then obtained: e.g., the cross section for the transition $(\mu_0L_t^0\rightarrow\mu L_t)$, averaged over M_t^0 states and summed over M_t states, is given by

$$\sigma(\mu_0L_t^0\rightarrow\mu L_t)=\frac{\pi}{(2L_t^0+1)k_{\Gamma 0}^2}\sum_{l_0lL}(2L+1)\left|\delta_{\Gamma 0\Gamma}-\mathcal{S}_{\Gamma 0\Gamma}^{LM}\right|^2. \tag{4.19}$$

Other physical quantities of interest can be similarly constructed.[9]

5. DISCUSSION

The discussion in I compared the J-matrix approach with the R-matrix, separable kernel, and Fredholm methods. The remarks were of such a general nature as to apply to the extensions made in this paper. In particular, there are two points that should be stressed. First, no Kohn-type pseudoresonances are expected to appear in the computed cross sections. This can be demonstrated by bounding these quantities for all energies.[11] Secondly, since H_0 is solved analytically within the function space, it is expected that small basis sets would be adequate to account for the addition of the approximate potential. The resulting physical quantities contain first order errors; however, the analytic nature of the solutions allows for variational correction. This can be accomplished through the application of the Kato correction[12] as discussed in I, and results in reducing the errors to second order.

Presently, work is being done on some of the more mathematical aspects of the J-matrix method. These include the sense of convergence of the expansions for the sine-like and cosine-like functions, possible analytic approaches to the second order Kato correction, and the generality of the solution scheme for the H_0 problem. In addition, the J-matrix method is being applied to the e—He^+ scattering calculation.

ACKNOWLEDGMENTS

The authors are grateful to Professor William P. Reinhardt for encouragement and support of this work along with a critical reading of the manuscript. Additional thanks are expressed to Dr. E. J. Heller for continued interest in this work and a critical reading of the manuscript. Helpful conversations with J. Broad, T. Murtaugh, and J. Winick are acknowledged. This work was supported by grants from the National Science Foundation and the Camille and Henry Dreyfus Foundations.

APPENDIX: THE COULOMB J-MATRIX SOLUTIONS

The Coulomb J matrix of Eq. (2.68) leads to the fundamental recursion relation

$$2[(n+l+1+2z/\lambda)\cos\theta-2z/\lambda]v_n-(n+2l+1)v_{n-1}-(n+1)v_{n+1}=0,\quad n>0, \tag{A1}$$

which may be cast into the form

$$2[(n+l+1)\cos\theta+t\sin\theta]v_n-(n+2l+1)v_{n-1}-(n+1)v_{n+1}=0,\quad n>0, \tag{A2}$$

where $t=-z/k=(-2z/\lambda)\tan\theta/2$ for $-\pi<\theta<\pi$. Since Ψ_{reg} for the Coulomb case has the form[3]

$$\Psi_{\text{reg}}(k,r)=\tfrac{1}{2}(2kr)^{l+1}\exp(ikr)\exp(\pi t/2)\frac{|\Gamma(l+1-it)|}{\Gamma(2l+2)}\times{}_1F_1(l+1-it,2l+2,-2ikr), \tag{A3}$$

the Fourier projection of Eq. (2.29) has the general form

$$v_n=\int_0^\infty dr\,\Phi[\eta r,t(k)]\phi_n(r)/r. \tag{A4}$$

Since the k dependence in Ψ_{reg} appears in both the variable and the order parameter of the confluent hypergeometric function, a differential equation for v_n cannot immediately be constructed as in the radial kinetic energy case.[4] However, since the only requirement is to satisfy a pure recurrence relation and not any differential properties, the procedure of Sec. 2A

can be applied after making one modification. Noting that Eq. (A2) can be viewed as a recursion relation, where t and θ are independent of one another, t is taken to be independent of θ (or k) in Eq. (A4), resulting in the form

$$v_n = \int_0^\infty dr\, \Phi(\eta r)\phi_n(r)/r \tag{A5}$$

and the procedure to obtain a differential difference relation can be carried through as before. Only at the end of the procedure is t set equal to $-z/k$: The differential relations derived will then no longer hold; however, Eq. (A2) will become Eq. (A1), and the original problem will be solved.

It is then straightforward to derive the differential equation

$$(x^2-1)v_n'' + \frac{x^2-1}{x}v_n' - \left[\left(\frac{x^2-1}{x^2}\right)(n^2+2nl+2n+l+1-t^2) + \left(\frac{(x^2+1)}{x^2(x^2-1)}\right) \times [(x^2+1)l(l+1) - 2it(n+l+1)(x^2-1)]\right] v_n = 0, \tag{A6}$$

which can be solved by standard techniques to give[7]

$$v_n = A_{nlt}(\sin\theta)^{l+1}\exp(\theta t)\exp(-in\theta) \times {}_2F_1(-n, l+1-it; 2l+2; 1-\exp(2i\theta)) + B_{nlt}(\sin\theta)^{it}\exp[-i(n+l+1)\theta] \times {}_2F_1(-l-it, l+1-it; n+l+2-it; 1/[1-\exp(2i\theta)]). \tag{A7}$$

Substitution into Eq. (A2) followed by the imposition of the homogeneous initial condition determines s_n to within a factor dependent upon l and t, a_{lt}. Due to the cut structure of the solutions, a_{lt} must be determined in two normalization steps, one corresponding to $k \to 0$ for $\theta[0, \pi]$ and other to $k \to 0$ for $\theta[0, -\pi]$. Applying Eq. (2.39), first in the limit $r \to 0$, $k \to 0^+$ corresponding to $\theta[0, \pi]$, and then in the limit $r \to 0$, $k \to 0^-$ corresponding to $\theta[0, -\pi]$, finally yields

$$s_n = \frac{2^l|\Gamma(l+1-it)|}{\Gamma(2l+2)}\exp[(\pi/2+\epsilon\pi)t]\exp(\theta t)(\sin\theta)^{l+1}\exp(-in\theta) \times {}_2F_1(-n, l+1-it; 2l+2; 1-\exp(2i\theta)). \tag{A8}$$

Letting $t = -z/k$ recovers the original recursion relation and noting the definition of the Pollaczek polynomials,[4]

$$P_n^\lambda(\cos\theta; a; b) = [(2\lambda)_n/n!]\exp(-in\theta)\,{}_2F_1(-n, \lambda - iw; 2\lambda; 1-\exp(2i\theta)), \tag{A9}$$

where $w = (a\cos\theta + b)/\sin\theta$, gives

$$s_n = \frac{2^l n!\,|\Gamma(l+1-it)|}{\Gamma(n+2l+2)}\exp[(\pi/2+\epsilon\pi)t]\exp(\theta t)(\sin\theta)^{l+1} \times P_n^{l+1}(\cos\theta, 2z/\lambda; -2z/\lambda). \tag{A10}$$

Three properties of $\{s_n\}$ will now be verified. Since Ψ_{reg} is real[3] and the basis set is real, $\{s_n\}$ must be real. Application of the linear transformation[7]

$${}_2F_1(a, b; c; z) = (1-z)^{-a}\,{}_2F_1(a, c-b; c; z/z-1) \tag{A11}$$

to Eq. (A8) immediately establishes the reality of $\{s_n\}$. Since Ψ_{reg} transforms as[3]

$$\Psi_{reg}(-k, r) = (-)^{l+1}\exp(-\pi t)\Psi_{reg}(k, r), \tag{A12}$$

the set $\{s_n\}$ should transform as

$$s_n(-k) = (-)^{l+1}\exp(-\pi t)s_n(k). \tag{A13}$$

The transformation $k \to -k$ implies $\theta \to -\theta$, $\epsilon \to -\epsilon$, and $t \to -t$, which when combined with the reality property immediately establishes Eq. (A13). Finally, in the limit as $z \to 0$, the radial kinetic energy in a Laguerre basis results should be recovered. Noting that[4]

$$P_n^\lambda(\cos\theta; 0; 0) = C_n^\lambda(\cos\theta) \tag{A14}$$

immediately establishes the reduction.

Substitution into Eq. (A2) followed by the imposition of the inhomogeneous initial condition, in the derivation of which the result

$$W(k) = -k\exp[(1+\epsilon)\pi t], \tag{A15}$$

as can be derived from the results in Ref. 3, is employed, yields

$$c_n + \alpha(l, t)s_n = -\frac{n!\exp(i\sigma_l)\exp(\pi t/2)\exp(-\theta t)\exp[-i(n+1)\theta]}{2^l\Gamma(n+l+2-it)(\sin\theta)^l} \times {}_2F_1(-l-it, n+1; n+l+2-it; \exp(-2i\theta)). \tag{A16}$$

The coefficient $\alpha(l, t)$ is determined in three steps. Enforcing the reality of $\{c_n\}$ through the application of the linear transformation[7]

$${}_2F_1(a, b; c; z) = [\Gamma(c)\Gamma(b-a)/\Gamma(b)\Gamma(c-a)](-z)^{-a} \times {}_2F_1(a, 1-c+a; 1-b+a; 1/z) + [\Gamma(c)\Gamma(a-b)/\Gamma(a)\Gamma(c-b)](-z)^{-b} \times {}_2F_1(b, 1-c+b; 1-a+b; 1/z) \tag{A17}$$

results in

$$\alpha(l, t) = \rho(l, t) + i, \tag{A18}$$

where $\rho(l, t)$ is a real function. Enforcing the symmetry condition analogous to Eq. (A13),

$$c_n(-k) = (-)^l\exp(-\pi t)c_n(k), \tag{A19}$$

results in the fact that $\rho(l, t)$ is an odd function of t,

$$\rho(l, -t) = -\rho(l, t). \tag{A20}$$

Since $\rho(l, t)$ is independent of n, it is sufficient to consider $n = 0$ in Eq. (A16), which takes the form

$$c_0 + is_0 = [-\exp(i\sigma_l)\exp(\pi t/2)\exp(-\theta t) \times \exp(-i\theta)/2^l\Gamma(l+2-it)(\sin\theta)^l \times {}_2F_1(-l-it, 1; l+2-it; \exp(-2i\theta)) - \rho(l, t)s_0. \tag{A21}$$

The lhs of Eq. (A21) corresponds for $k > 0$ to the zeroth expansion coefficient of the function

$$\tilde{C}(r) + i\tilde{S}(r) = \frac{\exp(\pi t/2)\exp(-\theta t)\lambda}{2^l|\Gamma(l+1-it)|(\sin\theta)^{l+1}} \times \left\{(\Psi_{irreg} + i\Psi_{reg})\int_0^r dr'\,\Psi_{reg}\frac{\phi_0(\lambda r')}{\lambda r'}\right.$$

$$+\Psi_{\text{reg}}\int_r^\infty dr'(\Psi_{\text{irreg}}+i\Psi_{\text{reg}})\frac{\phi_0(\lambda r')}{\lambda r'}\Big\}. \quad \text{(A22)}$$

The coefficient is thus

$$c_0+is_0=\frac{1}{\Gamma(2l+2)}\int_0^\infty dr\,[\tilde{C}(r/\lambda,t)+i\tilde{S}(r/\lambda,t)]\frac{\phi_0(r)}{r} \quad \text{(A23)}$$

Now, letting $k\to ik$ in Eqs. (A22) and (A23), and correspondingly letting $\eta\to i\eta$ in Eq. (A21), and taking the limit $\eta\to\frac{1}{2}-\delta$, gives for Eq. (A21) two terms of order,

$$O(\delta^{l+1-it})-\rho(l,t)O(\delta^{-l-1+it}). \quad \text{(A24)}$$

As $\delta\to 0$, for $(l+1+\mathrm{Im}t)>0$, the $\rho(l,t)$ term is the dominant term in Eq. (A24). The oddness property of $\rho(l,t)$, Eq. (A20), removes the restriction imposed by the inequality. By referring to the expansions given in Eqs. (2.3) and (2.4), it is seen that the integral in Eq. (A23) is convergent in the limit taken above. Thus, in that limit, Eq. (A23) goes as $O(\delta^{l+1-it})$, implying that

$$\rho(l,t)=0 \quad \text{(A25)}$$

and thus that

$$c_n+is_n=\frac{-n!\,\exp(i\sigma_l)\exp(\pi t/2)\exp(-\theta t)\exp[-i(n+1)\theta]}{2^l\Gamma(n+l+2-it)(\sin\theta)^l}$$
$$\times{}_2F_1(-l-it,n+1;n+l+2-it;\exp(-2i\theta)). \quad \text{(A26)}$$

That Eq. (A26) reduces properly in the limit of $z\to 0$ can be established from the theory of Legendre functions.[7] Taking the limit $z\to 0$ in Eq. (A26) and applying formula 3.2 (30) of Ref. 7 establishes the reduction to the radial kinetic energy case.

*Supported by a fellowship from the College of Petroleum and Minerals, Dhahran, Saudi Arabia.
†Present address: Department of Physics, University of Western Ontario, London 72, Ontario, Canada.
[1]E.J. Heller and H.A. Yamani, Phys. Rev. A **9**, 1201 (1974).
[2]E.J. Heller and H.A. Yamani, Phys. Rev. A **9**, 1209 (1974).
[3]A. Messiah, *Quantum Mechanics, Vol. I* (North-Holland, Amsterdam, 1965).
[4]A. Erdélyi, Ed., *Higher Transcendental Functions, Vol. II* (McGraw-Hill, New York, 1953).
[5]R. Courant and D. Hilbert, *Methods of Mathematical Physics, Vol. I* (Interscience, New York, 1966).
[6]P.M. Morse and H. Feschbach, *Methods of Theoretical Physics, Vol. I* (McGraw-Hill, New York, 1953).
[7]A. Erdélyi, Ed., *Higher Transcendental Functions, Vol. I* (McGraw-Hill, New York, 1953).
[8]F.E. Harris, Phys. Rev. Lett. **19**, 173 (1967).
[9]Standard reference: S. Geltman, *Topics in Atomic Collision Theory* (Academic, New York, 1969).
[10]P.G. Burke, D.F. Gallaher, and S. Geltman, J. Phys. B **2**, 1142 (1962).
[11]E.J. Heller, unpublished work.
[12]G. Kato, Prog. Theor. Phys. **6**, 394 (1951).

Acoustic scattering by penetrable homogeneous objects*

R. Kittappa

Mathematics Department, Lincoln University, Oxford, Pennsylvania

R. E. Kleinman

Mathematics Department, University of Delaware, Newark, Delaware
(Received 20 March 1974)

The scattered field is found when a time harmonic acoustic wave is incident upon a finite object characterized by a density and wavenumber different from that of the surrounding medium. The interface is assumed to be either Lyapunoff or piecewise Lyapunoff. The problem is cast as a pair of coupled surface integral equations for the total field and its normal derivative on the interface. The Neumann series obtained by straightforward iteration is proven to be convergent for ranges of density and wavenumber, and specific bounds on these ranges are given. For small enough wavenumbers, the series converges for all values of the interior density. The iteration appears simpler than the usual Born approximation, which involves volume as well as surface integrals. The method is illustrated in the case of a spherical interface.

1. INTRODUCTION

In this paper, we consider the time harmonic three-dimensional diffraction problem for two homogeneous media with a Lyapunoff or piecewise Lyapunoff interface S, which separates all of 3-space into two disjoint regions, one of which is finite in extent. The acoustical properties of each medium are characterized by two real parameters k and ρ, the wavenumber and density, respectively.

The mathematical formulation of this problem involves Helmholtz equations with different wavenumbers for each media and transition conditions at the interface (see Sec. 2 below). Separation of variables can be conveniently used to solve this problem only if S is a sphere. For a general surface, the well-known Born approximation can be applied.[1] However, this method involves the iteration of integral equations not only over the surface but also over the interior volume. This makes the computation of the iterates so involved that seldom is more than the first order iterate ever employed.

Integral equations for the total field not involving a volume integral can be obtained, but these involve the normal derivative as well as the total field, both of which are unknown. The integral equations in the Born approximation involve only the total field. However, an additional surface integral equation for the normal derivative of the field may also be derived resulting in a pair of coupled surface integral equations for the two unknown quantities, the total field and its normal derivative on the surface.

These have been derived by Maue,[2] Koringa,[3] Mitzner,[4] and, for special values of ρ_i and ρ_e, Kupradze.[5] Analogous integral equations for electromagnetics are given by Müller.[6]

Only Mitzner proposes a constructive method of solution of these equations in the same generality as considered here. He derives weakly singular integral equations for the total field and its normal derivative for a "sufficiently smooth" (i. e., Lyapunoff) surface, and he elegantly avoids higher order singularities in his derivation. To solve these equations, Mitzner proposes an iteration method wherein a sequence of boundary value problems have to be solved successively to obtain the iterates. He claims that the iteration must be valid for sufficiently small values of the ratio of the densities of the media (i. e., for greatly different densities). However, his example shows that the convergence of the iterates also depends on the wave numbers.

Ahner and Kleinman[7] derived a direct iterative solution of a surface integral equation formulation of the exterior Neumann problem, and this method was successfully extended to problems in linear elasticity by Ahner and Hsiao.[8] In the present paper, we generalize this method to the acoustic scattering problem for penetrable objects. We derive regularized integral equations which differ from Mitzner's in that the singularity of the integrands is reduced. Moreover, the regularized equations are valid for a wider class of surfaces than the weakly singular formulation. Both the weakly singular and regularized equations are solved directly as Newmann series (convergent for overlapping but not identical ranges of the parameters) without the need to solve associated problems for each iterate. The convergence of the Neumann series is proven for sufficiently small $|\rho_i - \rho_e|$ and $|k_i - k_e|$ thereby establishing the existence of the solution.

2. STATEMENT OF THE PROBLEM

Let S be a simply connected closed piecewise Lyapunoff surface[9] in $\mathbb{R}^3$ (or the union of a finite number of such surfaces) with surface area Σ. In some instances specifically noted, S is further restricted to be Lyapunoff rather than piecewise Lyapunoff. We also assume that S satisfies a modified cone condition.[7] Let V_i be the region interior to $S(\bar{V}_i = V_i \cup S)$ and V_e be the region exterior to $S(\bar{V}_e = V_e \cup S)$. Erect a coordinate system with the origin in V_i and denote by $R(P, P')$ the distance between any two points P and P'.

The unit vectors normal to the surface at P and P' will be denoted by $\hat{n}$ and $\hat{n}'$, respectively, and will always be taken as directed from S into V_e. Differentiation in the direction of $\hat{n}$ or $\hat{n}'$ will be denoted by $\partial/\partial n$ or $\partial/\partial n'$. Furthermore, $u^+(P)$ and $\partial u^+(P)/\partial n$ will denote the limiting values of u and $\partial u/\partial n$ as $P \to P \in S$ from

V_e and similarly $u^-(P)$ and $\partial u^-(P)/\partial n$ will denote limiting values as $P \to P \in S$ from V_i.

By piecewise Lyapunoff surfaces, we mean a surface S, for which

$$S = \sum_{i=1}^{N} S_i$$

where each S_i is a Lyapunoff segment in the sense of Sobolev.[10]

Sobolev lists as one of the requirements of a Lyapunoff surface that there exist some α_1 such that

$$\int \left| \frac{\partial}{\partial n} \frac{1}{R} \right| dS \leq \alpha_1 \leq \infty, \tag{2.1}$$

where the integration may be carried out over any portion of the surface S, including S itself. Mikhlin[9] shows that this is a consequence of the other properties required by both Sobolev and Mikhlin and need not be imposed as a separate condition. This remains true for piecewise Lyapunoff surfaces.[7] Here, however, we add a related but different requirement, namely that

$$\int \left| \frac{\partial}{\partial n'} \frac{1}{R} \right| dS \leq \alpha_2 < \infty, \tag{2.2}$$

where again the integration is over any portion of S. The normal derivative is taken with respect to the field point, not the integration point. This requirement is fulfilled if S is Lyapunoff and can be shown *not* to be fulfilled for specific piecewise Lyapunoff surfaces (e.g., hemisphere). It is not clear if (2.2) is sufficient to guarantee that a piecewise Lyapunoff surface be also Lyapunoff; thus the class of surfaces we treat includes Lyapunoff and possibly some, but not all, piecewise Lyapunoff surfaces.

$C(X)$ and $C^2(X)$ will represent spaces of complex valued functions defined on X which are continuous and twice continuously differentiable respectively and $D'(X)$ is the space of complex valued piecewise continuously differentiable functions on X. Furthermore, we will denote the space of piecewise continuous complex valued functions on S which are also continuous on each Lyapunoff segment by $D(S)$. Similarly, $D'(S)$ denotes piecewise continuously differentiable complex valued functions on S which are also continuously differentiable in each Lyapunoff segment.

Problem A

The problem we consider is that of finding the total field at all points in $\mathbb{R}^3$ given a time harmonic incident field $u^i(P)$ (either a point source in V_e or a plane wave originating in V_e or a distribution of such sources). A time dependence $\exp(-i\omega t)$ is assumed. It is convenient to represent the total field $u(P)$ as a sum of incident and scattered fields for $P \in V_e$, i.e.,

$$u(P) = u^i(P) + u^S(P), \quad P \in V_e. \tag{2.3}$$

The problem then is that of determining $u(P)$ such that

(i) $u(P) \in C^2(V_i) \cap D'(\bar{V}_i) \cap D'(S)$,

(ii) $u^S(P) \in C^2(V_e) \cap D'(\bar{V}_e) \cap D'(S)$,

(iii) $(\nabla^2 + k_i^2)u(P) = 0, \quad P \in V_i$,

(iv) $(\nabla^2 + k_e^2)u^S(P) = 0, \quad P \in V_e$,

(v) $\rho_e u^+(P) = \rho_i u^-(P), \quad P \in S$,

(vi) $\dfrac{\partial u^+(P)}{\partial n} = \dfrac{\partial u^-(P)}{\partial n}, \quad P \in S_i, \ i = 1, 2, \ldots, N$,

(vii) $\lim_{r\to\infty} r\left(\dfrac{\partial u^S}{\partial r} - ik_e u^S\right) = 0$ in all directions.

We note that the prescribed incident field will satisfy $(\nabla^2 + k_e^2)u^i(P) = 0$ almost everywhere in R^3.

This is the basic problem in acoustic diffraction where u is the space part of a time harmonic velocity potential. A similar equation can also be written for the space part of the pressure or density.

Note that if we let

$$\rho_i u(P) = v(P), \quad P \in V_i,$$

and

$$\rho_e u(P) = v(P), \quad P \in V_e,$$

then the boundary conditions become

$$\left.\begin{aligned} &\text{(vii')} \quad v^-(P) = v^+(P) \\ &\text{(viii')} \quad \frac{1}{\rho_i}\frac{\partial v^-(P)}{\partial n} = \frac{1}{\rho_e}\frac{\partial v^+(P)}{\partial n} \end{aligned}\right\} \ P \in S,$$

whereas the other conditions in Problem A remain the same. The transition problem is sometimes posed in this form in current literature.[5,11]

Problem B

We will also consider the corresponding problem for the Laplace equation which results when $k_i = k_e = 0$ and the radiation condition is replaced by a regularity condition in the sense of Kellogg.[12] This problem arises in magnetostatics (Ref. 13, p. 169, and Ref. 14, p. 73).

The solutions of Problems A and B for point sources are essentially the Green's functions for these problems and can, therefore, be employed to yield explicit solutions of a class of inhomogeneous problems.

In the succeeding sections, we derive coupled integral equations of the form

$$u^+ = K_1 u^+ + K_2 \frac{\partial u^+}{\partial n} + u^i,$$

$$\frac{\partial u^+}{\partial n'} = K_3 u^+ + K_4 \frac{\partial u^+}{\partial n} + \frac{\partial u^i}{\partial n'},$$

where K_i are operators. These equations can be written in vector form

$$(I - K)U = U^i,$$

where $U = \begin{pmatrix} u^+ \\ \partial u^+/\partial n \end{pmatrix}$ and U^i is known. The Neumann series solution of this equation,

$$U = \sum_{n=0}^{\infty} K^n U^i,$$

is proven to be convergent for $|\rho_i - \rho_e|$ and $|k_i - k_e|$ sufficiently small.

3. DERIVATION OF INTEGRAL EQUATIONS

A. The total field equations

We begin with the standard integral representation for u^s (Ref. 15)

$$\frac{1}{4\pi}\int_S\left(u^{s+}\frac{\partial}{\partial n}\frac{\exp(ik_eR)}{R}-\frac{\exp(ik_eR)}{R}\frac{\partial u^{s+}}{\partial n}\right)dS=\begin{cases}u^s(P'), & P'\in V_e\\ \frac{\sigma(P')}{4}u^{s+}(P'), & P'\in S\\ 0, & P'\in V_i\end{cases}\tag{3.1}$$

where R stands for $R(P,P')$ with P the integration variable and P' the field point, and $\sigma(P')$ is a measure of solid angle[7] and is equal to 2π when P' lies on a smooth portion of S.

We will always assume that for a piecewise Lyapunoff surface $\int_S f(P)(\partial g/\partial n)\,dS$ stands for $\sum_{i=1}^N\int_{S_i}f(P)(\partial g/\partial n)\,dS$ where each S_i is a Lyapunoff segment.

From Green's identity applied to $\exp(ik_eR)/R$ and u^i in the region V_i

$$\frac{1}{4\pi}\int_S\left(u^i\frac{\partial}{\partial n}\frac{\exp(ik_eR)}{R}-\frac{\exp(ik_eR)}{R}\frac{\partial u^i}{\partial n}\right)dS=\begin{cases}0 & P'\in V_e\\ -\frac{4\pi-\sigma(P')}{4\pi}u^i(P'), & P'\in S\\ -u^i(P'), & P'\in V_i\end{cases}.\tag{3.2}$$

Equations (3.1) and (3.2) together with (2.3) lead to

$$u^i(P')+\frac{1}{4\pi}\int_S\left(u^+\frac{\partial}{\partial n}\frac{\exp(ik_eR)}{R}-\frac{\exp(ik_eR)}{R}\frac{\partial u^+}{\partial n}\right)dS=\begin{cases}u(P'), & P'\in V_e\\ \frac{\sigma(P')}{4\pi}u^+(P'), & P'\in S\\ 0, & P'\in V_i\end{cases}.\tag{3.3}$$

Now, applying Green's identity to $\exp(ik_iR)/R$ and u in the region V_i yields

$$\frac{1}{4\pi}\int_S\left(u^-\frac{\partial}{\partial n}\frac{\exp(ik_iR)}{R}-\frac{\exp(ik_iR)}{R}\frac{\partial u^-}{\partial n}\right)dS=\begin{cases}0, & P'\in V_e\\ -\frac{4\pi-\sigma(P')}{4\pi}u^-(P'), & P'\in S\\ -u(P'), & P'\in V_i\end{cases}.\tag{3.4}$$

Substracting (3.4) from (3.3) and using the boundary conditions, we get

$$u^i(P')+\frac{1}{\rho_i}\frac{1}{4\pi}\int_S u^+\frac{\partial}{\partial n}\frac{\rho_i\exp(ik_eR)-\rho_e\exp(ik_iR)}{R}\,dS-\frac{1}{4\pi}\int_S\frac{\exp(ik_eR)-\exp(ik_iR)}{R}\frac{\partial u^+}{\partial n}\,dS=\begin{cases}u(P'), & P'\notin S\\ \left(\frac{\sigma(P')}{4\pi}\frac{\rho_i-\rho_e}{\rho_i}+\frac{\rho_e}{\rho_i}\right)u^+(P'), & P'\in S\end{cases}.\tag{3.5}$$

We note that it is sufficient to determine the field and its normal derivative on the surface because Eq. (3.5) expresses the field everywhere in terms of these quantities. The representation (3.5) is a weakly singular integral equation for $P'\in S$ and it can be written as

$$\lambda u^i+\lambda Au^+ +\lambda B\frac{\partial u^+}{\partial n}=u^+,\tag{3.6}$$

where

$$\lambda=\left(\frac{\sigma(P')}{4\pi}\frac{\rho_i-\rho_e}{\rho_i}+\frac{\rho_e}{\rho_i}\right)^{-1},\tag{3.7}$$

$$(Af)(P')=\frac{1}{4\pi\rho_i}\int_S f(P)\frac{\partial}{\partial n}\frac{\rho_i\exp(ik_eR)-\rho_e\exp(ik_iR)}{R}\,dS,\tag{3.8}$$

and

$$(Bf)(P')=-\frac{1}{4\pi}\int_S\frac{\exp(ik_eR)-\exp(ik_iR)}{R}f(P)\,dS.\tag{3.9}$$

For a smooth surface S, Eq. (3.5) was derived by Mitzner.[4]

We now employ the following form of Gauss' integral:

$$\frac{1}{4\pi}\int_S\frac{\partial}{\partial n}\frac{1}{R}dS=\begin{cases}0, & P'\in V_e\\ -\frac{4\pi-\sigma(P')}{4\pi}, & P'\in S\\ -1, & P'\in V_i\end{cases}.\tag{3.10}$$

By multiplying both sides of (3.10) by $-[(\rho_i-\rho_e)/\rho_i]\times u(P')$ and adding to Eq. (3.5), we get the following integral representation

$$\begin{aligned}u^i(P')&+\frac{1}{4\pi}\int_S u^+\frac{\partial}{\partial n}\frac{\exp(ik_eR)-\exp(ik_iR)}{R}\,dS\\&+\frac{\rho_i-\rho_e}{\rho_i}\frac{1}{4\pi}\int_S u^+\frac{\partial}{\partial n}\frac{\exp(ik_iR)-1}{R}\,dS\\&+\frac{\rho_i-\rho_e}{\rho_i}\frac{1}{4\pi}\int_S[u^+(P)-u(P')]\frac{\partial}{\partial n}\frac{1}{R}\,dS\\&-\frac{1}{4\pi}\int_S\frac{\exp(ik_eR)-\exp(ik_iR)}{R}\frac{\partial u^+}{\partial n}\,dS=u(P'),\\&\qquad P'\in\bar V_e.\end{aligned}\tag{3.11}$$

We note that this is a continuous representation of $u(P')$ in $\bar V_e$, it is independent of σ even at conical points, and is "regularized" in the sense of Ahner and Kleinman.[7] In fact, their representation is exactly reproduced by letting $\rho_i\to\infty$.

We can write the last equation, for $P'\in S$, in the form,

$$u^i+Eu^+ +B\frac{\partial u^+}{\partial n}=u^+,\tag{3.12}$$

where

$$(Ef)(P')=\frac{1}{4\pi}\int_S f(P)\frac{\partial}{\partial n}\frac{\exp(ik_eR)-\exp(ik_iR)}{R}\,dS$$

$$+\frac{\rho_i-\rho_e}{\rho_i}\frac{1}{4\pi}\int_S f(P)\frac{\partial}{\partial n}\frac{\exp(ik_iR)-1}{R}\,dS$$
$$+\frac{\rho_i-\rho_e}{\rho_i}\frac{1}{4\pi}\int_S [f(P)-f(P')]\frac{\partial}{\partial n}\frac{1}{R}\,dS \quad (3.13)$$

and B is given by (3.9).

In the case when $k_i=k_e=k$, Eqs. (3.6) and (3.12) no longer involve the normal derivative of the total field and reduce to the boundary integral equations for the total field

$$(I-\lambda A)u^+=\lambda u^i \quad (3.14)$$

and

$$(I-E)u^+=u^i. \quad (3.15)$$

Solutions of these equations suffice to solve Problem A for $k_i=k_e$, and hence Problem B.

B. Integral equations for the normal derivative of the field

When $k_i\neq k_e$, it is necessary to derive an additional equation for the normal derivative of the total field at P', assuming P' to lie on a smooth part of the surface.

When $\rho_i=\rho_e$ and $k_i\neq k_e$, we can obtain this additional equation by direct differentiation of Eqs. (3.5) or (3.11), because the derivative of the integrands are either continuous or integrable as will be shown in Lemmas 4.1 and 4.2 (Ref. 16, p. 217). In either case, we get the regularized integral equation

$$\frac{\partial u^i}{\partial n'}(P')+\frac{1}{4\pi}\int_S u^+\frac{\partial^2}{\partial n'\,\partial n}\frac{\exp(ik_eR)-\exp(ik_iR)}{R}\,dS$$
$$-\frac{1}{4\pi}\int_S\frac{\partial}{\partial n'}\left(\frac{\exp(ik_eR)-\exp(ik_iR)}{R}\right)\frac{\partial u^+}{\partial n}\,dS=\frac{\partial u^+(P')}{\partial n'}. \quad (3.16)$$

When $\rho_i\neq\rho_e$ and $k_i\neq k_e$ direct differentiation of Eq. (3.5) or Eq. (3.11) will involve the term

$$\frac{\partial}{\partial n'}\int_S u^+\frac{\partial}{\partial n}\frac{1}{R}\,dS$$

which exists only under more stringent conditions on u^+ than just requiring that $u^+\in C(S)$.[17] This would place undesirable restrictions on the space of functions in which the solution is sought. However, this disturbing term can be avoided using the method of Mitzner[4] as follows.

Differentiating both sides of Eq. (3.4) in the direction of the normal as P' approaches a smooth portion of S from V_i and using the jump conditions (Ref. 18, p. 685), we get

$$\frac{1}{4\pi}\int_S u^-\frac{\partial^2}{\partial n'\,\partial n}\frac{\exp(ik_iR)-1}{R}\,dS$$
$$+\lim_{P'\in V_i\to P'\in S}\frac{1}{4\pi}\frac{\partial}{\partial n'}\int_S u^-\frac{\partial}{\partial n}\frac{1}{R}\,dS$$
$$-\frac{1}{4\pi}\int_S\frac{\partial}{\partial n'}\left(\frac{\exp(ik_iR)}{R}\right)\frac{\partial u^-}{\partial n}\,dS=-\frac{1}{2}\frac{\partial u^-(P')}{\partial n'},\quad P'\in S. \quad (3.17)$$

Similarly, taking the normal derivative of both sides of Eq. (3.3) as P' approaches S from V_e, yields

$$\frac{\partial u^i(P')}{\partial n'}+\frac{1}{4\pi}\int_S u^+\frac{\partial^2}{\partial n'\,\partial n}\frac{\exp(ik_eR)-1}{R}\,dS$$
$$+\lim_{P'\in V_e\to P'\in S}\frac{1}{4\pi}\frac{\partial}{\partial n'}\int_S u^+\frac{\partial}{\partial n}\frac{1}{R}\,dS$$
$$-\frac{1}{4\pi}\int_S\frac{\partial}{\partial n'}\frac{\exp(ik_eR)}{R}\frac{\partial u^+}{\partial n}\,dS=\frac{1}{2}\frac{\partial u^+(P')}{\partial n'},\quad P'\in S. \quad (3.18)$$

Subtract Eq. (3.17) from ρ_e/ρ_i times Eq. (3.18) and simplify using the boundary conditions and the known result that continuity of f is sufficient to guarantee that

$$\lim_{P'\in V_e\to P'\in S}\frac{\partial}{\partial n'}\int_S f(P)\frac{\partial}{\partial n}\frac{1}{R}\,dS$$
$$-\lim_{P'\in V_i\to P'\in S}\frac{\partial}{\partial n'}\int_S f(P)\frac{\partial}{\partial n}\frac{1}{R}\,dS=0.^{17}$$

This leads to the result

$$\frac{2\rho_e}{\rho_i+\rho_e}\frac{\partial u^i(P')}{\partial n'}+\frac{\rho_e}{\rho_i+\rho_e}\frac{1}{2\pi}\int_S u^+\frac{\partial^2}{\partial n'\,\partial n}\frac{\exp(ik_eR)-\exp(ik_iR)}{R}$$
$$\times\,dS-\frac{1}{\rho_i+\rho_e}\frac{1}{2\pi}\int_S\frac{\partial}{\partial n'}\left(\frac{\rho_e\exp(ik_eR)-\rho_i\exp(ik_iR)}{R}\right)\frac{\partial u^+}{\partial n}\,dS$$
$$=\frac{\partial u^+(P')}{\partial n'},\quad P'\text{ on a smooth part of }S. \quad (3.19)$$

Equation (3.19) is essentially the same weakly singular integral equation derived by Mitzner[4] for the normal derivative of the field. It can be written as

$$\frac{2\rho_e}{\rho_i+\rho_e}\frac{\partial u^i}{\partial n'}+Cu^++D\frac{\partial u^+}{\partial n}=\frac{\partial u^+}{\partial n'}, \quad (3.20)$$

where

$$(Cf)(P')=\frac{\rho_e}{\rho_i+\rho_e}\frac{1}{2\pi}\int_S f(P)\frac{\partial^2}{\partial n'\,\partial n}\frac{\exp(ik_eR)-\exp(ik_iR)}{R}\,dS \quad (3.21)$$

and

$$(Df)(P')$$
$$=-\frac{1}{\rho_i+\rho_e}\frac{1}{2\pi}\int_S\frac{\partial}{\partial n'}\left(\frac{\rho_e\exp(ik_eR)-\rho_i\exp(ik_iR)}{R}\right)f(P)\,dS. \quad (3.22)$$

To achieve a regularized form, we again employ Gauss' integral (3.10) for P' on a smooth part of S and, by multiplying both sides by

$$2\frac{\rho_i-\rho_e}{\rho_i+\rho_e}\frac{\partial u^+}{\partial n'},$$

obtain

$$\frac{\rho_i-\rho_e}{\rho_i+\rho_e}\frac{1}{2\pi}\int_S\frac{\partial}{\partial n}\left(\frac{1}{R}\right)\frac{\partial u^+(P')}{\partial n'}\,dS=-\frac{\rho_i-\rho_e}{\rho_i+\rho_e}\frac{\partial u^+(P')}{\partial n'}. \quad (3.23)$$

Adding this equation to (3.19) and simplifying, we get the desired regularized equation

$$\frac{\partial u^i(P')}{\partial n'}+\frac{1}{4\pi}\int_S u^+\frac{\partial^2}{\partial n'\,\partial n}\frac{\exp(ik_eR)-\exp(ik_iR)}{R}\,dS$$

$$-\frac{1}{4\pi}\int_S \frac{\partial}{\partial n'}\left(\frac{\exp(ik_eR)-\exp(ik_iR)}{R}\right)\frac{\partial u^+}{\partial n}\,dS$$
$$+\frac{\rho_i-\rho_e}{\rho_e}\,\frac{1}{4\pi}\int_S \frac{\partial}{\partial n'}\left(\frac{\exp(ik_iR)-1}{R}\right)\frac{\partial u^+}{\partial n}\,dS$$
$$+\frac{\rho_i-\rho_e}{\rho_e}\,\frac{1}{4\pi}\int_S \left[\frac{\partial}{\partial n'}\left(\frac{1}{R}\right)\frac{\partial u^+(P)}{\partial n}+\frac{\partial}{\partial n}\left(\frac{1}{R}\right)\frac{\partial u^+(P')}{\partial n'}\right]dS$$
$$=\frac{\partial u^+(P')}{\partial n'}\,. \tag{3.24}$$

This is regularized because

$$\frac{\partial}{\partial n'}\left(\frac{1}{R}\right)\frac{\partial u^+(P)}{\partial n}+\frac{\partial}{\partial n}\left(\frac{1}{R}\right)\frac{\partial u^+(P')}{\partial n'}$$
$$=-\nabla\left(\frac{1}{R}\right)\cdot\left(\hat{n}'\frac{\partial u^+(P)}{\partial n}-\hat{n}\frac{\partial u^+(P')}{\partial n'}\right)$$

and the quantity in the square brackets vanishes when $R\to 0$.

We can write this equation in the form

$$\frac{\partial u^i}{\partial n'}+Gu^+ +H\frac{\partial u^+}{\partial n}=\frac{\partial u^+}{\partial n'}\,, \tag{3.25}$$

where

$$(Gf)(P')=\frac{1}{4\pi}\int_S f(P)\frac{\partial^2}{\partial n'\,\partial n}\,\frac{\exp(ik_eR)-\exp(ik_iR)}{R}\,dS \tag{3.26}$$

and

$$(Hf)(P')=-\frac{1}{4\pi}\int_S \frac{\partial}{\partial n'}\left(\frac{\exp(ik_eR)-\exp(ik_iR)}{R}\right)f(P)\,dS$$
$$+\frac{\rho_i-\rho_e}{\rho_e}\,\frac{1}{4\pi}\int_S \frac{\partial}{\partial n'}\left(\frac{\exp(ik_iR)-1}{R}\right)f(P)\,dS$$
$$+\frac{\rho_i-\rho_e}{\rho_e}\,\frac{1}{4\pi}\int_S \left[\frac{\partial}{\partial n'}\left(\frac{1}{R}\right)f(P)+\frac{\partial}{\partial n}\left(\frac{1}{R}\right)f(P')\right]dS. \tag{3.27}$$

Equations (3.6) and (3.20) form a pair of coupled weakly singular surface integral equations (when the field point is on a smooth part of the surface) which can be written in vector form as

$$(I-\mathcal{L})U=U_1^i, \tag{3.28}$$

where

$$U=\begin{pmatrix} u^+ \\ \dfrac{\partial u^+}{\partial n'}\end{pmatrix},\qquad U_1^i=\begin{pmatrix} \lambda u^i \\ \dfrac{2\rho_e}{\rho_i+\rho_e}\,\dfrac{\partial u^i}{\partial n'}\end{pmatrix} \tag{3.29}$$

and $\mathcal{L}$ is the operator matrix

$$\mathcal{L}=\begin{pmatrix} \lambda A & \lambda B \\ C & D\end{pmatrix}. \tag{3.30}$$

Equations (3.12) and (3.25) form a pair of coupled regularized surface integral equations which can be written in vector form as

$$(I-\mathcal{M})U=U^i, \tag{3.31}$$

where U is the same as before,

$$U^i=\begin{pmatrix} u^i \\ \dfrac{\partial u^i}{\partial n'}\end{pmatrix}, \tag{3.32}$$

and $\mathcal{M}$ is the operator matrix

$$\mathcal{M}=\begin{pmatrix} E & B \\ G & H\end{pmatrix}. \tag{3.33}$$

As will emerge in the following sections, both the weakly singular equation (3.28) and the regularized equation (3.31) may be solved iteratively for different but partially overlapping parameter ranges.

4. EXISTENCE OF SOLUTIONS IN THE FORM OF NEUMANN SERIES

We will use the following direct method to solve the equations of Sec. 3. Suppose

$$(I-K)\phi=\psi \tag{4.1}$$

is the integral equation with K being a linear operator. Then we use the iteration scheme

$$\phi^{(N+1)}=K\phi^{(N)}+\psi \quad \phi^{(0)}=\psi. \tag{4.2}$$

This will lead to the Neumann series

$$\sum_{j=0}^{\infty} K^j\psi. \tag{4.3}$$

We can use this series not only to approximate the solution but also to demonstrate the existence of the solution by showing that the series converges in a complete linear vector space V under a suitable norm $\|\ \|$. For this, it is sufficient to prove that K maps V into itself and $\|K\|<1$.

For the convergence of the Neumann series generated by Eqs. (3.14) and (3.15), we use $C(S)$, the space of continuous complex valued functions on S with the norm

$$\|f(P)\|=\sup_{P\in S}|f(P)|. \tag{4.4}$$

For the Neumann series generated by Eqs. (3.28) and (3.31), we consider the space $C(S)\times C(S)$ with the following extension of norm (4.4)

$$\left\|\begin{matrix} f \\ g\end{matrix}\right\|=\max\{\|f\|,\ \|g\|\} \tag{4.5}$$

where the symbol $\|\ \|$ on the right refers to the norm (4.4). For convenience, we will use the same symbol for the norms on $C(S)$ and $C(S)\times C(S)$ and the meaning will be clear from the context.

It can be easily verified that these norms satisfy the required axioms and the corresponding spaces are complete in the norms.

Having defined the appropriate Banach spaces and norms, we next establish that the operators introduced in Sec. 3 map the appropriate spaces into themselves.

Lemma 4.1:

A: $C(S)\to C(S)$ if S is Lyapunoff,

B: $C(S)\to C(S)$, E: $C(S)\to C(S)$ } if S is piecewise Lyapunoff.

Proof: Ahner and Kleinman[7] showed that for S piecewise Lyapunoff, the operator L defined as

$$Lf=\frac{1}{4\pi}\int_S f(P)\frac{\partial}{\partial n}\,\frac{\exp(ikR)}{R}\,dS+f(P')\left(\frac{4\pi-\sigma(P')}{4\pi}\right)$$

mapped $C(S)$ into $C(S)$. Renaming this operator L_{k_e} or L_{k_i} accordingly as k is replaced by k_e or k_i, it is easily verified from the definitions (3.8) and (3.13) that

$$Af = L_{k_e} f - \frac{\rho_e}{\rho_i} L_{k_i} f + \frac{\rho_e - \rho_i}{\rho_i} f(P') \left(\frac{4\pi - \sigma(P')}{4\pi} \right) \tag{4.6}$$

and

$$Ef = L_{k_e} f - \frac{\rho_e}{\rho_i} L_{k_i} f. \tag{4.7}$$

Hence, E maps $C(S)$ into $C(S)$ for S piecewise Lyapunoff, whereas, A maps $C(S)$ into $C(S)$ only if S is Lyapunoff because $\sigma(P')$ is discontinuous at corner points. The operator B (3.9) clearly maps $C(S)$ into $C(S)$ since the kernel is analytic in R.

Lemma 4.2: If S is piecewise Lyapunoff and satisfies Eq. (2.2), the operators C, D, G, H all map $\mathcal{D}(S) \to \mathcal{D}(S)$.

Proof: The operators in question are

$$Cf = \frac{\rho_e}{2\pi(\rho_i + \rho_e)} \int_S f(P) \frac{\partial^2}{\partial n' \, \partial n} \frac{\exp(ik_e R) - \exp(ik_i R)}{R} dS, \tag{4.8}$$

$$\begin{aligned} Df = &- \frac{(\rho_e - \rho_i)}{2\pi(\rho_i + \rho_e)} \int \frac{\partial}{\partial n'} \frac{1}{R} f(P) \, dS \\ &- \frac{1}{2\pi(\rho_i + \rho_e)} \int_S \frac{\partial}{\partial n'} \\ &\times \frac{\rho_e[\exp(ik_e R) - 1] - \rho_i[\exp(ik_i R) - 1]}{R} f(P) \, dS, \end{aligned} \tag{4.9}$$

$$Gf = \frac{1}{4\pi} \int_S f(P) \frac{\partial^2}{\partial n' \, \partial n} \frac{\exp(ik_e R) - \exp(ik_i R)}{R} dS, \tag{4.10}$$

$$\begin{aligned} Hf = &- \frac{1}{4\pi} \int_S \frac{\partial}{\partial n'} \left(\frac{\exp(ik_e R) - \exp(ik_i R)}{R} \right) f(P) \, dS \\ &+ \frac{\rho_i - \rho_e}{4\pi\rho_e} \int_S \frac{\partial}{\partial n'} \left(\frac{\exp(ik_i R) - 1}{R} \right) f(P) \, dS \\ &+ \frac{\rho_i - \rho_e}{4\pi\rho_e} \int_S \left(\frac{\partial}{\partial n'} \frac{1}{R} f(P) + \frac{\partial}{\partial n} \frac{1}{R} f(P') \right) dS, \end{aligned} \tag{4.11}$$

where D has been rewritten slightly for convenience. We see that the above integrals fall into one of the following three types:

(1) Integrals whose kernels have discontinuities of the form $\hat{n} \cdot \hat{R}$ or $\hat{n}' \cdot \hat{R}$ multiplied by analytic functions of R.

(2) $\int f(P) \frac{\partial}{\partial n} \frac{1}{R} dS$ or $\int f(P) \frac{\partial}{\partial n'} \frac{1}{R} dS$.

(3) $\int_S f(P) \frac{\partial^2}{\partial n' \, \partial n} \frac{\exp(ik_e R) - \exp(ik_i R)}{R} dS$.

Type (1) integrals include kernels such as

$$\frac{\partial}{\partial n'} \frac{\exp(ikR) - 1}{R} \quad \text{and} \quad \frac{\partial}{\partial n'} \frac{\exp(ik_e R) - \exp(ik_i R)}{R}$$

and hence are continuous functions of P' on S if $f(P)$ is continuous on the segments S_i and remain bounded as P' approaches a boundary point of S_i. As for Type (2), $\int_S f(P)(\partial/\partial n) R^{-1} dS$ is the potential of a double layer and is continuous on smooth portions of S if f is continuous (Ref. 17, p. 49). Condition (2.1) assures that integrals of this type remain bounded as P' approaches a boundary point of S_i. $\int f(P)(\partial/\partial n') R^{-1} dS$ is the "direct value of the normal derivative of a single layer" and is continuous (Ref. 18, p. 587) on smooth portions of S if f is continuous. It remains bounded as P' approaches boundary points of S_i because of (2.2).

Type (3) may be handled as follows. Observe that

$$\frac{\partial}{\partial n} F(R) = \hat{n} \cdot \nabla F(R) = (\hat{n} \cdot \hat{R}) \frac{d}{dR} F(R) \tag{4.12}$$

and

$$\frac{\partial}{\partial n'} F(R) = \hat{n}' \cdot \nabla' F(R) = - (\hat{n}' \cdot \hat{R}) \frac{d}{dR} F(R), \tag{4.13}$$

where $\hat{R}$ is the unit vector from P' to P. Then the integrand becomes

$$\begin{aligned} &\frac{\partial^2}{\partial n' \, \partial n} \frac{\exp(ik_e R) - \exp(ik_i R)}{R} \\ &= \frac{\partial}{\partial n'} \left[\frac{d}{dR} \left(\frac{\exp(ik_e R) - \exp(ik_i R)}{R} \right) (\hat{n} \cdot \hat{R}) \right] \\ &= - (\hat{n}' \cdot \hat{R})(\hat{n} \cdot \hat{R}) \frac{d^2}{dR^2} \left(\frac{\exp(ik_e R) - \exp(ik_i R)}{R} \right) \\ &\quad + \frac{d}{dR} \left(\frac{\exp(ik_e R) - \exp(ik_i R)}{R} \right) \frac{\partial}{\partial n'} (\hat{n} \cdot \hat{R}). \end{aligned} \tag{4.14}$$

Expanding the exponentials and using Lemma 1 of the Appendix, this can be written as

$$\begin{aligned} &\frac{\partial^2}{\partial n' \, \partial n} \frac{\exp(ik_e R) - \exp(ik_i R)}{R} \\ &= i \sum_{m=1}^{\infty} \frac{(m+1)}{(m+2)!} (k_e^{m+2} - k_i^{m+2}) (iR)^{m-1} \\ &\quad \times [(m-1)(\hat{n} \cdot \hat{R})(\hat{n}' \cdot \hat{R}) + (\hat{n} \cdot \hat{n}')] \\ &\quad - \frac{k_e^2 - k_i^2}{2R} [(\hat{n}' \cdot \hat{n}) - (\hat{n} \cdot \hat{R})(\hat{n}' \cdot \hat{R})]. \end{aligned} \tag{4.15}$$

The first term leads to an integral of type (1), while the second is the potential of a single layer with continuous density on each S_i if f is continuous on S_i, hence the integral is continuous on each S_i (Ref. 18, p. 582). The corollary in the Appendix may be used to show that the potential is bounded at boundary points of S_i. Hence, the lemma.

With these results, it follows that

$$\mathcal{L} = \begin{pmatrix} \lambda A & \lambda B \\ C & D \end{pmatrix} : C(S) \times C(S) \to C(S) \times C(S)$$

if S is Lyapunoff. The fact that A is a mapping of continuous functions into themselves only when S is Lyapunoff imposes the same restriction on $\mathcal{L}$. However, the regularized operators apply to piecewise Lyapunoff surfaces subject to (2.2). Hence,

$$\mathcal{M} = \begin{pmatrix} E & B \\ G & H \end{pmatrix} : C(S) \times \mathcal{D}(S) \to C(S) \times \mathcal{D}(S).$$

Next, we establish a series of bounds on the operators.

Lemma 4.3: If S is Lyapunoff (not piecewise Lyapunoff)

$$\|Af\| \leq \left\{ \frac{3}{16\pi} |k_e^2 - k_i^2| \Sigma + \frac{3}{16\pi} \frac{|\rho_i - \rho_e|}{\rho_i} k_i^2 \Sigma + \frac{|\rho_i - \rho_e|}{4\pi\rho_i} \alpha \right\} \|f\|.$$

Proof:

$$\|Af\| = \sup_{P' \in S} \left| \frac{1}{4\pi\rho_i} \int_S f(P) \frac{\partial}{\partial n} \frac{\rho_i \exp(ik_e R) - \rho_e \exp(ik_i R)}{R} dS \right|$$

$$\leq \sup_{P' \in S} \left\{ \left| \frac{1}{4\pi} \int_S f(P) \frac{\partial}{\partial n} \frac{\exp(ik_e R) - \exp(ik_i R)}{R} dS \right| \right.$$

$$+ \frac{|\rho_i - \rho_e|}{4\pi\rho_i} \left| \int_S f(P) \frac{\partial}{\partial n} \frac{e^{(ik_i R)} - 1}{R} dS \right|$$

$$\left. + \frac{|\rho_i - \rho_e|}{4\pi\rho_i} \left| \int_S f(P) \frac{\partial}{\partial n} \frac{1}{R} dS \right| \right\}.$$

The estimates (A5) and (A7) in the Appendix and the corollary to Lemma 4 lead to the desired result.

Lemma 4.4: If S is piecewise Lyapunoff

$$\|Bf\| \leq (|k_e - k_i|/4\pi) \Sigma \|f\|.$$

Proof:

$$\|Bf\| = \sup_{P' \in S} \left| \frac{1}{4\pi} \int_S \frac{\exp(ik_e R) - \exp(ik_i R)}{R} f(P) dS \right|.$$

The estimate (A4) in the Appendix establishes the result.

A similar use of the estimates in the Appendix serves to establish the following lemmas:

Lemma 4.5: If S is piecewise Lyapunoff

$$\|Cf\| \leq \left| \frac{\rho_e}{\rho_i + \rho_e} \right| \frac{1}{2\pi} \left(\tfrac{7}{12} |k_e + k_i| \Sigma + \tfrac{3}{4}\beta \right) |k_e^2 - k_i^2| \, \|f\|.$$

Lemma 4.6: If S is piecewise Lyapunoff and satisfies (2.2)

$$\|Df\| \leq \frac{\rho_i}{\rho_i + \rho_e} \times \left\{ \frac{3}{8\pi} |k_e^2 - k_i^2| \Sigma + \frac{3}{8\pi} \frac{|\rho_i - \rho_e|}{\rho_i} k_i^2 \Sigma + \frac{|\rho_i - \rho_e|}{2\pi\rho_i} \alpha \right\} \|f\|.$$

Lemma 4.7: If f is piecewise Lyapunoff

$$\|Ef\| \leq \left\{ \frac{3}{16\pi} |k_e^2 - k_i^2| \Sigma + \frac{3}{16\pi} \frac{|\rho_i - \rho_e|}{\rho_i} k_i^2 \Sigma + \frac{|\rho_i - \rho_e|}{2\pi\rho_i} \alpha \right\} \|f\|.$$

Lemma 4.8: If S is piecewise Lyapunoff

$$\|Gf\| \leq \frac{1}{4\pi} \left(\tfrac{7}{12} |k_e + k_i| \Sigma + \tfrac{3}{4}\beta \right) |k_e^2 - k_i^2| \, \|f\|.$$

Lemma 4.9: If S is piecewise Lyapunoff and satisfies (2.2)

$$\|Hf\| \leq \left\{ \frac{3}{16\pi} |k_e^2 - k_i^2| \Sigma + \frac{3}{16\pi} \frac{|\rho_i - \rho_e|}{\rho_e} k_i^2 \Sigma + \frac{|\rho_i - \rho_e|}{2\pi\rho_e} \alpha \right\} \|f\|.$$

These estimates permit us to establish the main results.

Theorem 4.1: If S is Lyapunoff (not piecewise Lyapunoff), then $\|\mathcal{L}\| < 1$ for sufficiently small $|k_i - k_e|$ and $|\rho_i - \rho_e|$.

Proof: Let $\begin{pmatrix} f \\ g \end{pmatrix} \in C(S) \times C(S)$. With the definition (3.30)

$$\mathcal{L}\begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} \lambda Af + \lambda Bg \\ Cf + Dg \end{pmatrix}.$$

Using the norm (4.5), it follows that

$$\left\| \mathcal{L}\begin{pmatrix} f \\ g \end{pmatrix} \right\| = \max\{ \|\lambda Af + \lambda Bg\|, \|Cf + Dg\| \}.$$

Consider first

$$\|\lambda Af + \lambda Bg\| \leq |\lambda| \, \|Af\| + |\lambda| \, \|Bg\|.$$

Since S is Lyapunoff $\lambda = 2\rho_i/(\rho_i + \rho_e)$ [(3.7) with $\sigma = 2\pi$]. Using this together with Lemmas (4.3) and (4.4), we obtain

$$\|\lambda Af + \lambda Bg\| \leq \frac{1}{2\pi} \left\{ \frac{3\rho_i |k_e^2 - k_i^2| \Sigma}{4|\rho_i + \rho_e|} + \frac{3k_i^2 |\rho_i - \rho_e| \Sigma}{4|\rho_i + \rho_e|} \right.$$

$$\left. + \frac{|\rho_i - \rho_e| \alpha}{|\rho_i + \rho_e|} \right\} \|f\| + \frac{\rho_i |k_e - k_i|}{2\pi|\rho_i + \rho_e|} \Sigma \|g\|$$

$$\leq \frac{1}{2\pi|\rho_i + \rho_e|} \left\{ \frac{3\rho_i |k_e^2 - k_i^2| \Sigma}{4} + \rho_i |k_e - k_i| \Sigma \right.$$

$$\left. + \tfrac{3}{4} k_i^2 |\rho_i - \rho_e| \Sigma + |\rho_i - \rho_e| \alpha \right\} \max\{ \|f\|, \|g\| \}.$$

Similarly, with Lemmas (4.5) and (4.6) we find

$$\|Cf + Dg\| \leq \|Cf\| + \|Dg\|$$

$$\leq \left| \frac{\rho_e}{\rho_i + \rho_e} \right| \frac{1}{2\pi} \left(\tfrac{7}{12} |k_e + k_i| \Sigma + \tfrac{3}{4}\beta \right) |k_e^2 - k_i^2| \, \|f\|$$

$$+ \left| \frac{\rho_i}{\rho_i + \rho_e} \right| \frac{1}{2\pi} \left(\tfrac{3}{4} |k_e^2 - k_i^2| \Sigma + \frac{3}{4} \frac{|\rho_i - \rho_e|}{\rho_i} k_i^2 \Sigma \right.$$

$$\left. + \frac{|\rho_i - \rho_e|}{\rho_i} \alpha \right) \|g\|$$

$$\leq \frac{1}{2\pi|\rho_i + \rho_e|} \left\{ \rho_e |k_e^2 - k_i^2| \left(\tfrac{7}{12} |k_e + k_i| \Sigma + \tfrac{3}{4}\beta \right) \right.$$

$$+ \tfrac{3}{4}\rho_i |k_e^2 - k_i^2| \Sigma + \tfrac{3}{4} |\rho_i - \rho_e| k_i^2 \Sigma$$

$$\left. + |\rho_i - \rho_e| \alpha \right\} \max\{ \|f\|, \|g\| \}.$$

Examination of these inequalities shows that they can be written in the form

$$\|\lambda Af + \lambda Bg\| \leq \{ |k_e - k_i| a_1 + |\rho_i - \rho_e| a_2 \} \max\{ \|f\|, \|g\| \},$$

$$\|Cf + Dg\| \leq \{ |k_e - k_i| a_3 + |\rho_i - \rho_e| a_4 \} \max\{ \|f\|, \|g\| \},$$

where the a_i depend on ρ_i, ρ_e, k_i, and k_e but remain bounded as $k_i \to k_e$ and $\rho_i \to \rho_e$. Hence, for any $\delta > 0$, it is always possible to choose $|k_e - k_i|$ and $|\rho_i - \rho_e|$ small enough (e.g., $|k_e - k_i| < \delta/2a_1$, $|\rho_i - \rho_e| < \delta/2a_2$) so that

$$\|\lambda Af + \lambda Bg\| \leq \delta \max\{ \|f\|, \|g\| \}$$

and

$$\|Cf + Dg\| \leq \delta \max\{ \|f\|, \|g\| \},$$

in which case

$$\left\| \mathcal{L}\begin{pmatrix} f \\ g \end{pmatrix} \right\| = \max\{ \|Af + \lambda Bg\|, \|Cf + Dg\| \} \leq \delta \max\{ \|f\|, \|g\| \}$$

$$= \delta \left\| \begin{pmatrix} f \\ g \end{pmatrix} \right\|.$$

Choosing $\delta < 1$, we have

$$\|\mathcal{L}\| = \sup_{C(S) \times C(S)} \frac{\|\mathcal{L}\begin{pmatrix} f \\ g \end{pmatrix}\|}{\|\begin{pmatrix} f \\ g \end{pmatrix}\|} \leq \delta < 1.$$

Theorem 4.2: If S is piecewise and satisfies (2.2), then $\|\mathcal{M}\| < 1$ for sufficiently small $|k_i - k_e|$ and $|\rho_i - \rho_e|$.

Proof: Let $\begin{pmatrix} f \\ g \end{pmatrix} \in C(S) \times \mathcal{D}(S)$. With the definition (3.33)

$$\mathcal{M}\begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} Ef + Bg \\ Gf + Hg \end{pmatrix}.$$

With (4.5), it follows that

$$\left\| \mathcal{M}\begin{pmatrix} f \\ g \end{pmatrix} \right\| = \max\{\|Ef + Bg\|, \|Gf + Hg\|\}.$$

Lemmas 4.4 and 4.7 show that

$$\begin{aligned} \|Ef + Bg\| &\le \|Ef\| + \|Bg\| \\ &\le \frac{1}{2\pi}\left\{\tfrac{3}{8}|k_e^2 - k_i^2|\Sigma + \frac{3}{8}\frac{|\rho_i - \rho_e|}{\rho_i}k_i^2\Sigma \right. \\ &\quad \left. + \frac{|\rho_i - \rho_e|}{\rho_i}\alpha\right\}\|f\| + \frac{|k_e - k_i|}{4\pi}\Sigma\|g\| \\ &\le \frac{1}{2\pi}\left\{\tfrac{3}{8}|k_e^2 - k_i^2|\Sigma + \frac{3}{8}\frac{|\rho_i - \rho_e|}{\rho_i}k_i^2\Sigma \right. \\ &\quad \left. + \frac{|\rho_i - \rho_e|}{\rho_i}\alpha + \frac{|k_e - k_i|}{2}\Sigma\right\}\max\{\|f\|, \|g\|\} \end{aligned}$$

and with Lemmas 4.8 and 4.9, we obtain

$$\begin{aligned} \|Gf + Hg\| \le \frac{1}{4\pi}\Big\{&\left(\tfrac{7}{12}|k_e + k_i|\Sigma + \tfrac{3}{4}\beta\right)|k_e^2 - k_i^2| + \tfrac{3}{4}|k_e^2 - k_i^2|\Sigma \\ &+ \frac{3}{4}\frac{|\rho_i - \rho_e|}{\rho_e}k_i^2\Sigma + 2\frac{|\rho_i - \rho_e|}{\rho_e}\alpha\Big\}\max\{\|f\|, \|g\|\}. \end{aligned}$$

As before, these inequalities may be written more compactly as

$$\|Ef + Bg\| \le \{|k_e - k_i|b_1 + |\rho_i - \rho_e|b_2\}\left\|\begin{pmatrix} f \\ g \end{pmatrix}\right\|$$

and

$$\|Gf + Hg\| \le \{|k_e - k_i|b_3 + |\rho_i - \rho_e|b_4\}\left\|\begin{pmatrix} f \\ g \end{pmatrix}\right\|,$$

where the b_i depend on the parameters ρ_i, ρ_e, k_i, and k_e but remain bounded when $k_i \to k_e$ and $\rho_i \to \rho_e$. Therefore, just as in the previous theorem, for any $\delta > 0$, we can choose $|k_e - k_i|$ and $|\rho_i - \rho_e|$ sufficiently small so that

$$\frac{\|\mathcal{M}\begin{pmatrix} f \\ g \end{pmatrix}\|}{\|\begin{pmatrix} f \\ g \end{pmatrix}\|} \le \delta \quad \text{for all } \begin{pmatrix} f \\ g \end{pmatrix} \in C(S) \times \mathcal{D}(S).$$

Hence, $\|\mathcal{M}\| \le \delta$ and taking $\delta < 1$ establishes the theorem. Theorem 1 guarantees the convergence in the norm (4.5) of the Neumann series solution of Eq. (3.28) while Theorem 2 guarantees convergence of the Neumann series for (3.31).

As observed previously, the problem simplifies considerably if $k_i = k_e$. The terms involving the normal derivative drop out, and it is no longer necessary to work in the product space. The problem is one of finding only the field on S which was formulated in weakly singular and regularized form in (3.14) and (3.15). The Neumann series solution of (3.14) will converge on Lyapunoff surfaces if $\|\lambda A\| < 1$ which with Lemma 4.3 will be assured if

$$\frac{|\rho_i - \rho_e|}{2\pi|\rho_i + \rho_e|}\left(\tfrac{3}{4}k^2\Sigma + \alpha\right) < 1. \tag{4.16}$$

Similarly, Lemma 4.7 shows that the Neumann series solution of (3.15) will converge on piecewise Lyapunoff surfaces even when P' is a conical point if

$$\left|\frac{\rho_i - \rho_e}{2\pi\rho_i}\right|\left(\tfrac{3}{8}k^2\Sigma + \alpha\right) < 1. \tag{4.17}$$

These limitations are different as is more evident in the case when $k = 0$ (Problem B) and S is convex. The first holds for all finite nonzero ρ_i, ρ_e of the same sign and the second holds for all ρ_i, ρ_e such that $0 < \rho_e/\rho_i < 2$. From this it appears that the iterates for the weakly singular operator converge for a wider range of values of ρ_i and ρ_e than the iterates for the regularized operator. However, in the case of the regularized operator, the iterates converge for a larger class of surfaces and the convergence holds good in the limiting case when $\rho_i \to \infty$, although this is not a consequence of the present results. This behavior is demonstrated in the following examples in which case $\alpha = 2\pi$. Then the inequalities (4.16) and (4.17) become

$$\left|\frac{\rho_i - \rho_e}{\rho_i + \rho_e}\right| < 1 \quad \text{and} \quad \left|\frac{\rho_i - \rho_e}{\rho_i}\right| < 1,$$

respectively.

5. SOME EXAMPLES

We demonstrate the iteration technique in the particular case when S is a sphere for three different cases: (1) $k_i = k_e = 0$ (the potential problem, B); (2) $\rho_i \neq \rho_e$, $k_i = k_e = k \neq 0$; (3) $\rho_i = \rho_e$, $k_i \neq k_e$. More detail as well as the general case $\rho_i \neq \rho_e$, $k_i \neq k_e$ is found in (Ref. 19). In these examples, we will always solve for the limiting values of the field and its normal derivative as the field point approaches S from V_e, but for convenience we write u and $\partial u/\partial n$ rather than u^+ and $\partial u^+/\partial n$.

The surface is taken to be a sphere of radius "a" in which case $\sigma(P') = 2\pi$ for all $P' \in S$. A spherical polar coordinate system with origin at the center of the sphere is employed.

A. Potential problem ($k_i = k_e = 0$)

We solve both the weakly singular equation (3.14) and the regularized equation (3.15) assuming

$$u^i(P) = -1/4\pi R(P, P_0) \tag{5.1}$$

where P_0 is a fixed point in V_e. Hence, we are finding the Green's function for this problem.

The iteration scheme for the weakly singular integral equation is, for $P' \in S$,

$$\begin{aligned} u^{(N+1)}(P') = \frac{\rho_i - \rho_e}{\rho_i + \rho_e}\frac{1}{2\pi}\int_S u^{(N)}(P)\frac{\partial}{\partial n}\frac{1}{R(P, P')}\,dS \\ - \frac{\rho_i}{2\pi(\rho_i + \rho_e)R(P', P_0)}, \quad N \ge 1, \end{aligned} \tag{5.2}$$

$$u^{(0)}(P') = -\frac{\rho_i}{2\pi(\rho_i + \rho_e)R(P', P_0)}.$$

The computations are similar to those described in (Ref. 7). We employ the well-known properties of spherical harmonics

$$\frac{1}{R(P_1,P_2)}=\sum_{m=0}^{\infty}\frac{r_1^m}{r_2^{m+1}}P_m(\cos\gamma_{P_1P_2}),\quad r_1<r_2,\tag{5.3}$$

and

$$\int_S P_m(\cos\gamma_{PP_0})P_n(\cos\gamma_{PP'})\,dS=\begin{cases}0, & m\neq n\\ \dfrac{4\pi a^2}{2n+1}P_n(\cos\gamma_{P'P_0}), & m=n\end{cases},\tag{5.4}$$

where P_1 and P_2 have spherical coordinates (r_1,θ_1,ϕ_1) and (r_2,θ_2,ϕ_2), respectively, $\gamma_{PP'}$ is the angle between the position vectors of the points P and P', and P_m are the Legendre polynomials. On the surface, $\partial/\partial n=\partial/\partial r$. It is then a straightforward calculation to determine that

$$u^{(1)}(P')=-\frac{1}{2\pi}\frac{\rho_i}{\rho_i+\rho_e}\sum_{m=0}^{\infty}\frac{a^m}{r_0^{m+1}}P_m(\cos\gamma_{P'P_0})\left(1-\frac{1}{2m+1}\frac{\rho_i-\rho_e}{\rho_i+\rho_e}\right).\tag{5.5}$$

By mathematical induction, we can show that

$$u^{(N)}(P')=-\frac{1}{2\pi}\frac{\rho_i}{\rho_i+\rho_e}\sum_{m=0}^{\infty}\frac{a^m}{r_0^{m+1}}P_m(\cos\gamma_{P'P_0})\times\sum_{n=0}^{N}\left(\frac{\rho_e-\rho_i}{(\rho_e+\rho_i)2m+1}\right)^n.\tag{5.6}$$

Summing the geometric series and letting $N\to\infty$, we get

$$u(P')=-\frac{1}{4\pi}\sum_{m=0}^{\infty}\frac{(2m+1)\rho_i}{(m+1)\rho_i+m\rho_e}\frac{a^m}{r_0^{m+1}}P_m(\cos\gamma_{P'P_0})\tag{5.7}$$

provided

$$\left|\frac{\rho_i-\rho_e}{\rho_i+\rho_e}\frac{1}{2m+1}\right|<1\quad\text{for } m=0,1,2,3,\cdots.$$

This condition holds if

$$\left|\frac{\rho_i-\rho_e}{\rho_i+\rho_e}\right|<1,\tag{5.8}$$

which is true for all finite ρ_i, ρ_e of the same sign. The above expression for $u(P')$ agrees with that computed by the method of separation of variables.

The iteration scheme for the regularized equation (3.15) is

$$u^{(N+1)}(P')=\frac{\rho_i-\rho_e}{4\pi\rho_i}\int_S[u^{(N)}(P)-u^{(N)}(P')]\times\frac{\partial}{\partial n}\frac{1}{R(P,P')}dS-\frac{1}{4\pi R(P',P_0)},\tag{5.9}$$

$$u^{(0)}(P')=-\frac{1}{4\pi R(P',P_0)}.$$

Proceeding as before, we get

$$u^{(1)}(P')=-\frac{1}{4\pi}\sum_{m=0}^{\infty}\frac{a^m}{r_0^{m+1}}P_m(\cos\gamma_{P'P_0})\left(1+\frac{\rho_i-\rho_e}{\rho_i}\frac{m}{2m+1}\right)\tag{5.10}$$

and

$$u^{(N)}(P')=-\frac{1}{4\pi}\sum_{m=0}^{\infty}\frac{a^m}{r_0^{m+1}}P_m(\cos\gamma_{P'P_0})\frac{(2m+1)\rho_i}{(m+1)\rho_i+m\rho_e}\times\left[1-\left(\frac{\rho_i-\rho_e}{\rho_i}\frac{m}{2m+1}\right)^{N+1}\right].\tag{5.11}$$

$\{u^N\}$ converges to the solution provided

$$\left|\frac{\rho_i-\rho_e}{\rho_i}\frac{m}{2m+1}\right|<1,\quad m=0,1,2,3,\cdots,$$

which always holds for $m=0$. For $m\neq0$, the condition is the same as

$$\left|\frac{\rho_i-\rho_e}{\rho_i}\right|<\frac{2m+1}{m},\quad m=1,2,3,\cdots,$$

which is true if $|(\rho_i-\rho_e)/\rho_i|\leq2$ or

$$-1\leq\rho_e/\rho_i\leq3.\tag{5.12}$$

Comparing (5.8) and (5.12), we find that the weakly singular equation yields the solution for a wider range of values of ρ_i and ρ_e than the regularized equation. However, when $\rho_i\to\infty$, only the regularized equation yields the solution.

B. Diffraction by a sphere ($\rho_i\neq\rho_e$, $k_i=k_e=k$)

Assume that the incident field is a plane wave along the z axis given by

$$u^i(P)=\exp(ikz)=\exp(ikr\cos\theta).\tag{5.13}$$

The iteration scheme for the weakly singular integral equation (3.14) is

$$u^{(N+1)}(P')=\frac{\rho_i-\rho_e}{\rho_i+\rho_e}\frac{1}{2\pi}\int_S u^{(N)}(P)\frac{\partial}{\partial n}\frac{\exp(ikR)}{R}dS+\frac{2\rho_i}{\rho_i+\rho_e}\exp(ika\cos\theta')\tag{5.14}$$

$$u^{(0)}(P')=\frac{2\rho_i}{\rho_i+\rho_e}\exp(ika\cos\theta').\tag{5.15}$$

The explicit iteration can again be carried out with the help of the following well-known expansions in spherical harmonics (Ref. 20):

$$\exp(ikr\cos\theta)=\sum_{m=0}^{\infty}i^m(2m+1)j_m(kr)P_m(\cos\theta),\tag{5.16}$$

$$\frac{\exp[ikR(P_1,P_2)}{R(P_1,P_2)}=ik\sum_{m=0}^{\infty}(2m+1)j_m(kr_1)h_m^{(1)}(kr_2)P_m\times(\cos\gamma_{P_1P_2}),\quad r_1<r_2,\tag{5.17}$$

where j_m and $h_m^{(1)}$ are spherical Bessel functions. The first iterate is found to be

$$u^{(1)}(P')=\frac{2\rho_i}{\rho_i+\rho_e}\sum_{m=0}^{\infty}i^m(2m+1)j_m(ka)P_m(\cos\theta')(1+S_m)\tag{5.18}$$

and the Nth term is induced to be

$$u^{(N)}(P')=\frac{2\rho_i}{\rho_i+\rho_e}\sum_{m=0}^{\infty}i^m(2m+1)j_m(ka)P_m(\cos\theta')\frac{1-S_m^{N+1}}{1-S_m}\tag{5.19}$$

where

$$S_m\equiv\frac{\rho_i-\rho_e}{\rho_i+\rho_e}ik^2a^2\left(j_m(ka)\frac{d}{dka}h_m^{(1)}(ka)+h_m^{(1)}(ka)\frac{d}{dka}j_m(ka)\right).\tag{5.20}$$

Letting $N \to \infty$, we get

$$u(P') = \frac{2\rho_i}{\rho_i + \rho_e} \sum_{m=0}^{\infty} i^m (2m+1) \frac{j_m(ka)}{1 - S_m} P_m(\cos\theta') \tag{5.21}$$

provided $|S_m| < 1$. This condition is fulfilled for sufficiently small $|\rho_i - \rho_e|$. A more detailed discussion is given in (Ref. 19).

In a similar fashion, the iterative scheme for solving the regularized Eq. (3.15) is

$$u^{(N+1)}(P') = \frac{\rho_i - \rho_e}{4\pi\rho_i} \int_S \left(u^{(N)}(P) \frac{\partial}{\partial n} \frac{\exp(ikR)}{R} - u^{(N)}(P') \frac{\partial}{\partial n} \frac{1}{R} \right) dS + \exp(ika\cos\theta'), \tag{5.22}$$

$$u^{(0)}(P') = \exp(ika\cos\theta').$$

The Nth iterate is

$$u^{(N)}(P') = \sum_{m=0}^{\infty} i^m (2m+1) P_m(\cos\theta') j_m(ka) \times \frac{1 - [(\rho_i - \rho_e)/2\rho_i + S_m(\rho_i + \rho_e)/2\rho_i]^{N+1}}{1 - S_m} \tag{5.23}$$

which converges as $N \to \infty$ if

$$\left| \frac{\rho_i - \rho_e}{2\rho_i} + \frac{\rho_i + \rho_e}{2\rho_i} S_m \right| < 1. \tag{5.24}$$

Again, this condition is fulfilled for $|\rho_i - \rho_e|$ sufficiently small.

C. Diffraction by a sphere ($\rho_i = \rho_e$, $k_i \neq k_e$)

Again, we choose the incident field to be a plane wave (5.13). The integral equations now involve both u and $\partial u/\partial n$, however, the weakly singular equation (3.30) and the regularized equation (3.33) coincide in this case. The iteration scheme becomes

$$u^{(N+1)}(P') = \frac{1}{4\pi} \int_S \left\{ u^{(N)}(P) \frac{\partial}{\partial n} \frac{\exp(ik_eR) - \exp(ik_iR)}{R} - \frac{\partial u^{(N)}(P)}{\partial n} \frac{\exp(ik_eR) - \exp(ik_iR)}{R} \right\} dS + \exp(ika\cos\theta'), \tag{5.25}$$

$$\frac{\partial u^{(N+1)}(P')}{\partial n'} = \frac{1}{4\pi} \int_S \left\{ u^{(N)}(P) \frac{\partial^2}{\partial n' \partial n} \frac{\exp(ik_eR) - \exp(ik_iR)}{R} - \frac{\partial u^{(N)}(P)}{\partial n} \frac{\partial}{\partial n'} \frac{\exp(ik_eR) - \exp(ik_iR)}{R} \right\} dS + ik\cos\theta' \exp(ika\cos\theta'), \tag{5.26}$$

$$u^{(0)}(P') = \exp(ika\cos\theta'), \tag{5.27}$$

$$\frac{\partial u^{(0)}(P')}{\partial n'} = ik\cos\theta' \exp(ika\cos\theta'). \tag{5.28}$$

Carrying out the iteration in a manner similar to previous cases, we obtain

$$u(P') = \sum_{m=0}^{\infty} \frac{i^m (2m+1) P_m(\cos\theta') j_m(k_i a)}{\Gamma_m} \tag{5.29}$$

$$\frac{\partial u(P')}{\partial n'} = \sum_{m=0}^{\infty} \frac{i^m (2m+1) P_m(\cos\theta') (d/da) j_m(k_i a)}{\Gamma_m}, \tag{5.30}$$

where

$$\Gamma_m = ik_e a^2 \left(h_m^{(1)}(k_e a) \frac{d}{da} j_m(k_i a) - j_m(k_i a) \frac{d}{da} h_m^{(1)}(k_e a) \right) \tag{5.31}$$

provided $|1 - \Gamma_m| < 1$. This holds for $|k_i - k_e|$ sufficiently small.

APPENDIX

Lemma A1:

$$\frac{\partial}{\partial n'} (\hat{R} \cdot \hat{n}) = \frac{-\hat{n}' \cdot \hat{n} + (\hat{R} \cdot \hat{n}')(\hat{R} \cdot \hat{n})}{R}.$$

Proof: Let $\mathbf{R} = \overline{\mathbf{P'P}}$, so that $\hat{R} = \mathbf{R}/R$,

$$\nabla' \cdot (\hat{R} \cdot \hat{n}) = \nabla' \left(\frac{\mathbf{R} \cdot \hat{n}}{R} \right) = \frac{1}{R} \nabla' (\mathbf{R} \cdot \hat{n}) + (\mathbf{R} \cdot \hat{n}) \nabla' \frac{1}{R} = -\frac{1}{R} \hat{n} + \frac{\hat{R} \cdot \hat{n}}{R} \hat{R},$$

$$\frac{\partial}{\partial n'} (\hat{R} \cdot \hat{n}) = \hat{n}' \cdot \nabla' (\hat{R} \cdot \hat{n}) = \frac{-\hat{n}' \cdot \hat{n} + (\hat{R} \cdot \hat{n}')(\hat{R} \cdot \hat{n})}{R}$$

Corollary:

$$\left| \frac{\partial}{\partial n'} (\hat{R} \cdot \hat{n}) \right| \leq \frac{1}{R}.$$

Proof:

$$\frac{\hat{n} \cdot \hat{n}' - (\hat{R} \cdot \hat{n})(\hat{R} \cdot \hat{n}')}{R} = \hat{n}' \cdot \frac{\hat{R} \times (\hat{n} \times \hat{R})}{R} = \frac{(\hat{n}' \times \hat{R}) \cdot (\hat{n} \times \hat{R})}{R}.$$

Hence the result follows.

Lemma A2: If n is a nonnegative integer,

$$\left| \frac{d^n}{d\theta^n} \frac{\sin\theta}{\theta} \right| \leq \frac{1}{n+1}$$

for all θ.

Proof: Let

$$f(\theta) = \frac{d^n}{d\theta^n} \frac{\sin\theta}{\theta}.$$

$f(\theta)$ is a continuous function which can possibly be unbounded as $\theta \to 0$ and $\theta \to \pm\infty$. From the expansion of $\sin\theta/\theta$ in powers of θ, we find

$$\left. \frac{d^n}{d\theta^n} \frac{\sin\theta}{\theta} \right|_{\theta=0}$$

is either zero or $1/(n+1)$ and hence $f(\theta)$ is bounded as $\theta \to 0$.

By mathematical induction, we can show that $f(\theta)$ is a sum of terms of the form $\pm\sin\theta/\theta^m$ and $\pm\cos\theta/\theta^m$ and hence $f(\theta)$ is bounded as $\theta \to \pm\infty$.

Hence $f(\theta)$ is bounded. We can prove by induction that

$$\frac{d^{n+1}}{d\theta^{n+1}} \frac{\sin\theta}{\theta} = \pm \frac{(\cos)\theta}{\theta} - \frac{n+1}{\theta} \left[\frac{d^n}{d\theta^n} \frac{\sin\theta}{\theta} \right],$$

i. e.,

$$f'(\theta) = \pm \frac{\binom{\cos}{\sin}\theta}{\theta} - \frac{n+1}{\theta} f(\theta).$$

Let θ_0 be a critical value of $f(\theta)$, where $\theta_0 \neq 0$

$$\pm \frac{\binom{\cos}{\sin}\theta_0}{\theta_0} - \frac{n+1}{\theta_0} f(\theta_0) = 0,$$

or

$$f(\theta_0) = \pm \frac{1}{n+1} \begin{matrix}\cos \\ \sin\end{matrix} \theta_0.$$

So,

$$|f(\theta_0)| \leq 1/(n+1) \quad \text{if } \theta_0 \neq 0.$$

This is true for each critical value $\theta_0 \neq 0$ and hence, in general,

$$|f(\theta)| \leq 1/(n+1) \quad \text{if } \theta \neq 0.$$

However,

$$|f(0)| \leq 1/(n+1).$$

Hence the Lemma.

Lemma A3:

$$\left| \frac{\exp(ik_eR) - \exp(ik_iR)}{R} \right| \leq |k_e - k_i|.$$

Proof: Appropriate trigonometric identities lead to

$$\frac{\exp(ik_eR) - \exp(ik_iR)}{R}$$

$$= \frac{\cos k_eR - \cos k_iR + i(\sin k_eR - \sin k_iR)}{R}$$

$$= \frac{2\sin(k_e+k_i)\frac{1}{2}R \sin(k_i-k_e)\frac{1}{2}R - 2i\cos(k_e+k_i)\frac{1}{2}R \sin(k_i-k_e)\frac{1}{2}R}{R}$$

$$= i(k_e-k_i)\exp[i(k_e+k_i)\tfrac{1}{2}R]\frac{\sin(k_e-k_i)\frac{1}{2}R}{(k_e-k_i)\frac{1}{2}R}. \tag{A1}$$

Taking the absolute value, recalling that $|\sin\theta/\theta| \leq 1$ (Lemma A2) and $|\exp[i(k_e+k_i)\frac{1}{2}R]| = 1$, yields the desired result.

Lemma A4:

$$\left| \frac{\partial}{\partial n} \frac{\exp(ik_eR) - \exp(ik_iR)}{R} \right| \leq \tfrac{3}{4}|k_e^2 - k_i^2|.$$

Proof: Differentiating (A1) using (4.12) leads to

$$\frac{\partial}{\partial n} \frac{\exp(ik_eR) - \exp(ik_iR)}{R}$$

$$= \hat{n}\cdot\hat{R} \exp[i(k_e+k_i)\tfrac{1}{2}R]\left\{ \frac{-(k_e^2-k_i^2)}{2} \frac{\sin(k_e-k_i)\frac{1}{2}R}{(k_e-k_i)\frac{1}{2}R} + i(k_e-k_i)\frac{d}{dR}\frac{\sin(k_e-k_i)\frac{1}{2}R}{(k_e-k_i)\frac{1}{2}R} \right\}. \tag{A2}$$

Taking the absolute value and using the fact that $|\hat{n}\cdot\hat{R}| \leq 1$, yields

$$\left| \frac{\partial}{\partial n} \frac{\exp(ik_eR) - \exp(ik_iR)}{R} \right|$$

$$\leq \frac{|k_e^2 - k_i^2|}{2} \left| \frac{\sin(k_e-k_i)\frac{1}{2}R}{(k_e-k_i)\frac{1}{2}R} \right| + \frac{|k_e-k_i|^2}{2}$$

$$\times \left| \frac{d}{d\theta}\frac{\sin\theta}{\theta} \right|_{\theta=(k_e-k_i)R/2}.$$

Again using Lemma A2, we obtain

$$\left| \frac{\partial}{\partial n} \frac{\exp(ik_eR) - \exp(ik_iR)}{R} \right| \leq \frac{|k_e^2 - k_i^2|}{2} + \frac{|k_e-k_i|^2}{4}$$

and for real k_i and k_e of the same sign

$$(k_e-k_i)^2 \leq |k_e^2 - k_i^2|.$$

Hence the Lemma. Observe that the estimate is unchanged if the normal derivative is taken at P' rather than P, i.e.,

$$\left| \frac{\partial}{\partial n'} \frac{\exp(ik_eR) - \exp(ik_iR)}{R} \right| \leq \tfrac{3}{4}|k_e^2 - k_i^2|. \tag{A3}$$

Corollary:

$$\left| \frac{\partial}{\partial n} \frac{\exp(ikR) - 1}{R} \right| \leq \tfrac{3}{4}k^2.$$

This follows by setting $k_i = 0$ and dropping the subscript on k_e. This is a sharper bound than that given by Ahner and Kleinman.[7]

Lemma A5:

$$\left| \frac{\partial^2}{\partial n' \partial n} \frac{\exp(ik_eR) - \exp(ik_iR)}{R} \right|$$

$$\leq \frac{3}{4R}|k_e^2 - k_i^2| + \tfrac{7}{12}|k_e + k_i||k_e^2 - k_i^2|.$$

Proof: Differentiating (A1), using (4.13), leads to

$$\frac{\partial^2}{\partial n' \partial n} \frac{\exp(ik_eR) - \exp(ik_iR)}{R}$$

$$= \frac{\partial}{\partial n'}(\hat{n}\cdot\hat{R})\exp[i(k_e+k_i)\tfrac{1}{2}R]\left[-\frac{(k_e^2-k_i^2)}{2}\frac{\sin\theta}{\theta} + i\frac{(k_e-k_i)^2}{2}\frac{d}{d\theta}\left(\frac{\sin\theta}{\theta}\right)\right]_{\theta=(k_e-k_i)R/2}$$

$$+ (\hat{n}\cdot\hat{R})(\hat{n}'\cdot\hat{R})\exp[i(k_e+k_i)\tfrac{1}{2}R] \times \left\{ i\frac{(k_e+k_i)}{4}(k_e^2-k_i^2)\frac{\sin\theta}{\theta} + \frac{(k_e^2-k_i^2)(k_e-k_i)}{2} \times \frac{d}{d\theta}\left(\frac{\sin\theta}{\theta}\right) - i\frac{(k_e-k_i)^3}{4}\frac{d^2}{d\theta^2}\left(\frac{\sin\theta}{\theta}\right)\right\}_{\theta=(k_e-k_i)R/2}.$$

Taking the absolute value, using the estimates in the corollary to Lemma A1 and Lemma A2, we find that

$$\left| \frac{\partial^2}{\partial n' \partial n} \frac{\exp(ik_eR) - \exp(ik_iR)}{R} \right|$$

$$\leq \frac{1}{R}\left[\frac{|k_e^2-k_i^2|}{2} + \frac{(k_e-k_i)^2}{4} \right] + |k_e^2-k_i^2|\frac{|k_e+k_i|}{4} + \frac{|k_e^2-k_i^2||k_e-k_i|}{4} + \frac{|k_e-k_i|^3}{12}.$$

Again, we use the fact that for real k_e and k_i of the same sign we may appropriately replace $|k_e-k_i|$ by $|k_e+k_i|$ which yields the cruder estimate of the statement of the Lemma.

Using the results of Lemmas A3 to A5 enables us to derive the following estimates:

$$\left| \int_S f(P) \frac{\exp(ik_e R) - \exp(ik_i R)}{R} dS \right| \leq |k_e - k_i| \Sigma \|f\|, \tag{A4}$$

$$\left| \int_S f(P) \frac{\partial}{\partial n} \left(\frac{\exp(ik_e R) - \exp(ik_i R)}{R} \right) dS \right| \leq \tfrac{3}{4} |k_e^2 - k_i^2| \Sigma \|f\|, \tag{A5}$$

$$\left| \int_S f(P) \frac{\partial}{\partial n'} \left(\frac{\exp(ik_e R) - \exp(ik_i R)}{R} \right) dS \right| \leq \tfrac{3}{4} |k_e^2 - k_i^2| \Sigma \|f\|, \tag{A6}$$

$$\left| \int_S f(P) \frac{\partial}{\partial n} \frac{1}{R} dS \right| \leq \alpha \|f\|, \tag{A7}$$

$$\left| \int_S f(P) \frac{\partial}{\partial n'} \frac{1}{R} dS \right| \leq \alpha \|f\|, \tag{A8}$$

$$\left| \int_S f(P) \frac{\partial^2}{\partial n' \partial n} \frac{\exp(ik_e R) - \exp(ik_i R)}{R} dS \right| \leq (\tfrac{7}{12} |k_e + k_i| \Sigma + \tfrac{3}{4} \beta) |k_e^2 - k_i^2| \, \|f\|, \tag{A9}$$

where Σ is the surface area of S,

$$\beta = \sup_{P' \in S} \int_S \frac{dS}{R}$$

and

$$\alpha = \max \left\{ \sup_{P' \in S} \int_S \left| \frac{\partial}{\partial n} \frac{1}{R} \right| dS, \ \sup_{P' \in S} \int_S \left| \frac{\partial}{\partial n'} \frac{1}{R} \right| dS \right\}.$$

The estimates involving $\partial/\partial n'$ are valid for P' on smooth parts of S as well as limiting values as P' approaches boundary points of S_i. (A8) requires the additional assumption (2.2).

*Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR 74-2634.

[1]P. M. Morse and K. Uno Ingard, *Linear acoustic theory*, *Handbuch der Physik*, Vol. XI/I (Springer-Verlag, Berlin, 1961).
[2]A. W. Maue, Z. Phys. 126, 601 (1949).
[3]J. Koringa, J. Math. Phys. 6, 1107 (1965).
[4]K. M. Mitzner, J. Math. Phys. 7, 2053 (1966).
[5]V. D. Kupradze, Russ. Math. Surv. 22, 58 (1967).
[6]C. Müller, *Grundprobleme der mathetischen theorie electromagnetischer Schwingungen* (Springer-Verlag, Berlin, 1957).
[7]J. F. Ahner and R. E. Kleinman, Arch. Rat. Mech. Anal. 52, 26 (1973).
[8]J. F. Ahner and G. C. Hsiao, "A Neumann series representation for solutions to boundary-value problems in dynamic elasticity," Quart. Appl. Math. (to be published).
[9]S. G. Mikhlin, *Mathematical Physics: An Advanced Course* (Elsevier, New York, 1970).
[10]S. L. Sobolev, *Partial Differential Equations of Mathematical Physics* (Pergamon, Oxford, 1964).
[11]A. N. Tikhonov and A. A. Samarskii, *Equations of Mathematical Physics* (Pergamon, Oxford, 1963).
[12]O. D. Kellogg, *Foundations of Potential Theory* (Dover, New York, 1953).
[13]S. Bergman and M. Schiffer, *Kernel Functions and Elliptic Differential Equations* (Academic, New York, 1953).
[14]J. Van Bladel, *Electromagnetic Fields* (McGraw-Hill, New York, 1964).
[15]R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Wiley, New York, 1962).
[16]E. McShane, *Integration* (Princeton U. P., Princeton, New Jersey, 1944).
[17]N. M. Günter, *Potential Theory and Its Applications to Basic Problems of Mathematical Physics* (Ungar, New York, 1967).
[18]V. I. Smirnov, *A Course of Higher Mathematics* (Addison Wesley, Reading, Massachusetts, 1964), Vol. IV.
[19]R. Kittappa, "Transition problems for the Helmholtz equation," AFOSR Scientific Report, Grant 69-1794, University of Delaware, October 1973.
[20]W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas on Theorems for Special Functions of Mathematical Physics* (Springer, New York, 1966).

Topology of Higgs fields

J. Arafune*, P. G. O. Freund*, and C. J. Goebel†

*The Enrico Fermi Institute and The Department of Physics, The University of Chicago, Chicago, Illinois 60637

†Department of Physics, University of Wisconsin, Madison, Wisconsin 53706

(Received 9 October 1974)

It is shown that the conserved magnetic charge discovered by 't Hooft in non-Abelian gauge theories with spontaneous symmetry breaking is not associated with the invariance of the action under a symmetry group. Rather, it is a topological characteristic of an isotriplet of Higgs fields in a three-dimensional space: the Brouwer degree of the mapping between a large sphere in configuration space and the unit sphere in field space provided by the normalized Higgs field $\hat{\phi}^a = \phi^a (\phi^b \phi^b)^{-1/2}$. The use of topological methods in determining magnetic charge configurations is outlined. A peculiar interplay between Dirac strings and zeros of the Higgs field under gauge transformations is pointed out. The monopole–antimonopole system is studied.

1. INTRODUCTION

In quantum electrodynamics one can introduce[1] magnetically charged fields at the price of changing Maxwell's equations. The conservation of magnetic charge is related to the appearance of a new (magnetic) $U(1)$ symmetry in the theory.[2]

In non-Abelian gauge theories of the Higgs type, magnetic monopoles can appear without any modification of the field equations as has been demonstrated recently by 't Hooft.[3] Since one is dealing with the unaltered field theory one may wonder as to the origin of this new conservation law of magnetic charge. Is there a "hidden" symmetry in non-Abelian gauge theories that has not hitherto been recognized? We shall answer this question in the negative and will show that the possibility of nonvanishing 't Hooft magnetic charge is a direct consequence of having three scalar (Higgs) fields in a world with three space and one time dimensions. Mathematically expressed, the value of the magnetic charge is determined by the homotopy class of the Higgs field. In this respect our work is closely related to the "kink" and "metricity" concepts introduced earlier by Finkelstein and Misner[4] and Skyrme.[5]

We also explore the role of gauge invariance and derive a connection between the zeros of a Higgs field and the Dirac strings of monopole theory. A particularly simple treatment of the monopole–antimonopole problem then emerges.

2. THE MEANING OF MAGNETIC CHARGE

We start by considering a gauge theory based on the group $SO(3)$. We have a triplet of Yang–Mills fields A^a_μ ($\mu = 0, 1, 2, 3$ = Lorentz index, $a = 1, 2, 3$) and a triplet of Higgs fields ϕ^a. In terms of these fields one can define 't Hooft's "electromagnetic" field tensor

$$F_{\mu\nu} = \hat{\phi}^a G^a_{\mu\nu} - (1/e)\epsilon_{abc}\hat{\phi}^a D_\mu \hat{\phi}^b D_\nu \hat{\phi}^c, \tag{1a}$$

where

$$\phi^a = (\phi^b\phi^b)^{-1/2}\phi^a, \quad D_\mu\hat{\phi}^a = \partial_\mu\hat{\phi}^a + e\epsilon^{abc}A^b_\mu\hat{\phi}^c,$$
$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + e\epsilon^{abc}A^b_\mu A^c_\nu. \tag{1b}$$

The tensor $F_{\mu\nu}$ can be further written in the form

$$F_{\mu\nu} = M_{\mu\nu} + H_{\mu\nu}, \tag{1c}$$

where

$$M_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu,$$
$$B_\mu = \hat{\phi}^a A^a_\mu, \tag{1d}$$
$$H_{\mu\nu} = (1/e)\epsilon_{abc}\hat{\phi}^a\partial_\mu\hat{\phi}^b\partial_\nu\hat{\phi}^c.$$

We now define the magnetic current

$$k_\mu = \tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma}\partial^\nu F^{\rho\sigma} \tag{2}$$

and find from Eqs. (1c, d) and (2) that

$$k_\mu = \tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma}\partial^\nu H^{\rho\sigma} = (1/2e)\epsilon_{\mu\nu\rho\sigma}\epsilon_{abc}\partial^\nu\hat{\phi}^a\partial^\rho\hat{\phi}^b\partial^\sigma\hat{\phi}^c. \tag{3}$$

In deriving Eq. (3) we have used the identity $\epsilon_{\mu\nu\rho\sigma}\partial_\nu M_{\rho\sigma} = 0$ valid in the absence of singularity lines in the gauge field. The remarkable feature is that the magnetic current is completely specified in terms of the scalar triplet of Higgs fields. It is independent of the Yang–Mills fields A^a_μ. The current k_μ being the divergence of an antisymmetric tensor is conserved

$$\partial_\mu k^\mu = 0. \tag{4}$$

In other words the magnetic charge

$$M = (1/4\pi)\int k_0\, d^3x \tag{5}$$

obeys

$$\dot{M} = 0. \tag{6}$$

We now show that M nevertheless does *not* generate a symmetry. The divergence equation (6) is not a dynamical equation. It holds true no matter what action principle determines the dynamics of ϕ^a. It is not a combination of field equations. Noether's theorem (more precisely the reciprocal of what is usually referred to as Noether's theorem) does not apply and we cannot derive a symmetry. In fact M, as can be seen from Eqs. (3) and (5), contains only the fields ϕ^a and their space derivatives and hence commutes with all dynamical variables; thus it generates no symmetry transformations. Then how does the conservation of the magnetic current k_μ come about? To answer this question, we start by writing down the magnetic charge in the form

$$M = \frac{1}{4\pi}\int k_0\, d^3x = \frac{1}{8\pi e}\int \epsilon_{ijk}\epsilon_{abc}\partial_i(\hat{\phi}^a\partial_j\hat{\phi}^b\partial_k\hat{\phi}^c)\, d^3x$$
$$= \lim_{R\to\infty}\frac{1}{8\pi e}\int_{S^2_R}\epsilon_{ijk}\epsilon_{abc}\hat{\phi}^a\partial_j\hat{\phi}^b\partial_k\hat{\phi}^c(d^2\sigma)_i, \tag{7}$$

where the last integral runs over the sphere

$$S_R^2: x^i x^i = R^2 \tag{8}$$

in configuration space. This sphere can be parametrized in terms of two parameters ξ_α ($\alpha = 1, 2$):

$$x^i = x^i(\xi_\alpha). \tag{9}$$

Using

$$d^2\sigma_i = \tfrac{1}{2}\epsilon_{imn}\frac{\partial x^m}{\partial \xi^\alpha}\frac{\partial x^n}{\partial \xi^\beta}\epsilon_{\alpha\beta}d^2\xi,\ \partial_j\hat{\phi}^b = \frac{\partial \xi^\alpha}{\partial x^j}\frac{\partial \hat{\phi}^b}{\partial \xi^\alpha}, \tag{10}$$

we find

$$4\pi e M = \lim_{R\to\infty}\int_{S_R^2}\tfrac{1}{2}\epsilon_{\alpha\beta}\epsilon_{abc}\hat{\phi}^a\partial_\alpha\hat{\phi}^b\partial_\beta\hat{\phi}^c d^2\xi. \tag{11}$$

It is readily checked that the square of the integrand in this expression is precisely

$$g = \det(\partial_\alpha\hat{\phi}^a\partial_\beta\hat{\phi}^a), \tag{12}$$

the determinant of the metric tensor of the unit sphere $\hat{\phi}^a\hat{\phi}^a = 1$. The integrand is then $\pm\sqrt{g}$. While the point (ξ_1, ξ_2) covers the sphere S_R^2 once, the vector $\hat{\phi}$ can cover the unit sphere d_+ times with the positive sign of $\sqrt{g}$ and d_- times with the negative sign. The difference $d = d_+ - d_-$ must be an integer since otherwise the fields $\hat{\phi}^a$ would not be single valued. The integral of $\sqrt{g}\,d^2\xi$ gives the area 4π of the unit sphere with the appropriate sign for each covering. Thus

$$4\pi e M = d4\pi$$

or

$$M = d(1/e), \quad d = \text{integer}. \tag{13}$$

Once we have shown that, in units of $1/e$, M must be integer, we can throw new light on the conservation law of magnetic charge [Eq. (6)]. The continuous time evolution of the scalar fields and their gradients can change the magnetic charge—as defined by the integral (7)—only continuously. This is compatible with the integer spectrum of M only if M is time independent. By using Feynman's path integral quantization this statement is readily carried over to the quantized theory.[4]

We note that the argument that led us from Eq. (11) to Eq. (13) is a special case of a general theorem due to Kronecker.[6] The integer d is called Kronecker's index, and it plays a central role in topology. This is not surprising, since magnetic charge, as was emphasized above, does not originate in dynamics but rather follows from the topological structure of three scalar fields in a three-dimensional space. In the next section we further explore the topological meaning of magnetic charge. We shall give criteria for finding the magnetic charge simply from the field topology. Finally, we note here that the charge density k_0 given by Eq. (3) vanishes everywhere except at the zeros of the Higgs field where it has δ-like singularities that yield the result (13) upon integration. The zeros $\overset{0}{x}$ of the Higgs field are obtained by imposing three conditions $\phi^a(\overset{0}{x}) = 0$, $a = 1, 2, 3$, in a three-dimensional manifold (x^1, x^2, x^3) leaving thus in regular cases a zero-dimensional manifold, i.e., an isolated zero. It is also clear that the arguments of this section can be generalized to $n \neq 3$ space dimensions, provided one considers a theory with n scalar fields. It is remarkable that in all such theories one disposes of an "identically" conserved current of the 't Hooft type. If in n space dimensions one has only $n - p$ scalar fields, these fields will vanish on p-dimensional manifolds, and there are conserved quantities associated also with these manifolds. An example is provided by flux quantization in an "Abelian" superconductor where $n = 3$ and $p = 1$. The quantized and conserved quantity in this case is the magnetic flux.

3. THE TOPOLOGY OF HIGGS FIELDS

In the last section we saw that magnetic charge equals the Kronecker index of the normalized vector field $\hat{\phi}(x)$ defined over a large sphere (S_R^2: $x^i x^i = R^2$) in configuration space. To proceed further with our analysis, let us consider $\hat{\phi}^a(x)$ in more detail. It maps the sphere S_R^2 into the unit sphere in field space $S_{\phi 1}^2$:

$$\hat{\phi}: S_R^2 \longrightarrow S_{\phi 1}^2. \tag{14a}$$

Along with the Kronecker index, such a mapping can be characterized[7] in three further equivalent ways by its: A. Brouwer degree, B. homotopy class, and C. Poincaré—Hopf index. It is worth considering these alternatives as they throw additional light on the topology of Higgs fields.

A. Brouwer degree

Just as in the last section we parametrize the sphere S_R^2 by the two parameters ξ_α [Eq. (9)]. In addition, we also parametrize the sphere $S_{\phi 1}^2$ by two parameters say φ_α ($\alpha = 1, 2$; e.g., polar coordinates). The mapping (14) can then be described by

$$\varphi_\alpha = \varphi_\alpha(\xi_\beta) \tag{14b}$$

and is not necessarily one to one. Let (ψ_1, ψ_2) be the coordinates of a point ψ on $S_{\phi 1}^2$. Let $\varphi^{-1}(\psi)$ denote the set of all points on the sphere S_R^2 that are mapped into the point ψ by the mapping (14b). If the Jacobian of the mapping (14) does not vanish at any of the points of $\varphi^{-1}(\psi)$, then ψ is called a regular point of the mapping (14). Assume now that ψ is a regular point of the smooth mapping (14). The Brouwer degree of the map (14) at the point ψ $d(\hat{\phi};\psi)$, is defined as

$$d(\hat{\phi};\psi) = \sum_{\xi\in\varphi^{-1}(\psi)} \operatorname{sgn}\det\left(\frac{\partial\varphi_\alpha}{\partial\xi^\beta}\right). \tag{15}$$

It is readily shown that d does not depend on the choice of the regular point on $S_{\phi 1}^2$, so that we shall use the notation $d(\hat{\phi})$ for it.

B. The homotopy class

Two smooth mappings

$$f, g: S_R^2 \longrightarrow S_{\phi 1}^2 \tag{16}$$

are *smoothly homotopic* if there exists a smooth mapping

$$F: S_R^2 \otimes [0, 1] \longrightarrow S_{\phi 1}^2 \tag{17}$$

such that

$$F(\xi, 0) = f(\xi), \quad F(\xi, 1) = g(\xi). \tag{18}$$

In other words there exists a continuous class of map-

pings $F(\xi, \lambda)$ labeled by a parameter $\lambda \in [0, 1]$ such that Eqs. (18) are obeyed.

The set of all mappings smoothly homotopic to a given mapping defines its *homotopy class.*

C. The Poincaré-Hopf index

Let $x = \overset{0}{x}$ be an isolated zero of the isovector field $\phi^i(x)$. Consider a small 2-sphere S_ϵ^2 centered at $\overset{0}{x}$. The normalized field $\hat{\phi}^i(x)$ maps S_ϵ^2 onto $S_{\hat{\phi}1}^2$. The Brouwer degree of this mapping is called the Poincarê–Hopf index i of the zero $\overset{0}{x}$.

Now let us state without proof some theorems[7] that show that the Brouwer degree, the homotopy class, the Poincarê–Hopf, and Kronecker indices are equivalent ways of characterizing a smooth mapping. First of all a special case of a general theorem of Hopf states that two smooth mappings (16) are homotopic if and only if they have the same Brouwer degree. Furthermore, it can be shown that if the isotriplet field $\phi^a(x)$ has only nondegenerate zeros, then the Brouwer degree of the mapping (14) equals the sum of Poincarê–Hopf indices of the zeros of $\phi^i(x)$ inside the sphere S_R^2. Finally, the Brouwer degree $d(\hat{\phi})$ of the mapping (14) is equal to its Kronecker index. In short then: Magnetic charge in units of $1/e$ equals Kronecker index of mapping (14) equals Brouwer degree of mapping (14) equals sum of indices of (nondegenerate) zeros of the Higgs field,

$$eM = d = d(\hat{\phi}) = \sum_{\phi^a(x)=0} i(x). \tag{19}$$

All Higgs fields in the same homotopy class have the same magnetic charge. Let us illustrate all this with a few simple and important examples.

Example 1: The constant map

$$\phi^a(x) = C^a, \quad C^a = \text{const}, \quad C^a C^a \neq 0. \tag{20}$$

Then $\hat{\phi}^a = C^a/(C^b C^b)^{1/2}$. The Kronecker index and Brouwer degree obviously vanish. Since the constant field has no zeros, the sum $\sum_{\phi^a(x)=0} i(x)$ is also zero. Moreover, all constant fields are smoothly homotopic. One can even give a simple form for $F^a(x,\lambda)$ that accomplishes the transition from one to another member of the homotopy class: $F^a(x, \lambda) = C^a(1 - \lambda) + D^a\lambda$.

Example 2:

$$\phi^a(x) = x^a f(x^i), \tag{21a}$$

where f is a continuous function with no zeros and a sufficient number of continous derivatives. Then

$$\hat{\phi}^a = x^a/R. \tag{21b}$$

While x^i covers S_R^2, $\hat{\phi}^a$ covers once with positive orientation the sphere $S_{\phi,1}^2$. Hence the Kronecker index is $d = 1$. Choosing both the ξ's and φ's (the intrinsic coordinates on both spheres) to be polar coordinates, it is obvious $\varphi_1 = \xi_1$, $\varphi_2 = \xi_2$, so that the Brouwer degree of the mapping (21) is $d\,(x^i/R) = 1$. The only zero of the Higgs field is at the origin and its index is obviously $i(0) = 1$. We thus see that whichever way we turn it, the magnetic charge of the Higgs field (21) is $+1$ (in units of $1/e$).

Example 3:

$$\begin{aligned} \phi^1(x) &= 2ax^1 f(x^i), \\ \phi^2(x) &= 2ax^2 f(x^i), \\ \phi^3(x) &= (x^i x^i - a^2) f(x^i), \end{aligned} \tag{22}$$

where f has sufficiently many continuous derivatives and no zeros. Here again all four criteria give the total magnetic charge $M = 0$; however, it is obvious that $\phi^a(x)$ has two zeros, one at $x^1 = x^2 = 0$, $x^3 = +a$, the other at $x^1 = x^2 = 0$, $x^3 = -a$. They have indices $+1$ and -1 respectively. The field (22) can thus be interpreted to correspond to one monopole and one antimonopole separated by a distance $2a$.

Example 4: A field with total magnetic charge $M = 2$

$$\begin{aligned} \phi^1(x) &= ax^1 f(x^i), \\ \phi^2(x) &= x^2 x^3 f(x^i), \\ \phi^3(x) &= ((x^3)^2 - a^2) f(x^i). \end{aligned} \tag{23}$$

It has zeros at $x^1 = x^2 = 0$ $x^3 = \pm a$, and the indices are $+1$ for both zeros. One could interpret this as two monopoles separated by the distance $2a$, but the asymmetrical appearance of the field casts doubt on such an interpretation. Of course, here as well as in the previous cases such interpretations are to be validated by searching for solutions of the type (21)–(23) to the field equations. In the context of gauge theories gauge invariance can considerably simplify this task. We now address ourselves to this problem.

4. GAUGE INVARIANCE, DIRAC STRINGS AND HIGGS ZEROS

The discussion of the previous sections concerned ways of determining the magnetic charge from the topology of the Higgs fields. Somehow the Yang–Mills fields A_μ^a played no role. In fact, a conserved magnetic charge can be defined even in the absence of gauge fields. One might therefore believe that gauge fields are irrelevant to the problem of magnetic charge. However, this is not so. To properly understand the effects of gauge fields, let us start by considering the problem of gauge invariance.

From its definition (1)–(2) the magnetic current k_μ is an isosinglet and therefore so is the magnetic charge M [Eq. (5)]. This is true whether gauge fields are present or not. In the presence of gauge fields, $F_{\mu\nu}$ and therefore k_μ and M are invariant also under *local* isospin-gauge transformations (i.e., isospin transformations with space dependent parameters). This comes as no surprise as we expect magnetic charge to be an observable. Though $F_{\mu\nu}$ is gauge invariant its "partition" into $M_{\mu\nu}$ and $H_{\mu\nu}$ [Eqs. (1c, d)] is not. In particular, $M_{\mu\nu}$ is not gauge invariant. We have seen that if the gauge field is not singular, then the magnetic charge depends via $H_{\mu\nu}$ only on the homotopy class of the Higgs field. But this homotopy class can change under gauge transformations while the magnetic charge cannot. This means that the $M_{\mu\nu}$ term in $F_{\mu\nu}$ which in the original gauge did not contribute to the magnetic charge must in the new gauge contribute. In other words, the gauge field must acquire a (stringlike) singularity in the new

gauge. This means that the gauge transformation itself must be singular. Before going on, let us illustrate this on an example.

Let the Higgs and gauge fields for large $|\mathbf{x}|$ be given by the ('t Hooft) expressions[3]:

$$\phi^a(x) = Fx^a/r, \quad \hat{\phi}^a(x) = x^a/r, \qquad A_0^a = 0, \quad A_i^a = (1/e)\epsilon_{aij}x_j/r^2. \tag{24}$$

Now perform the gauge transformation

$$\omega = \exp(-i\varphi I_3)\exp(i\theta I_2)\exp(i\varphi I_3), \tag{25}$$

where $\varphi = \arctan(x^2/x^1)$, $\theta = \arccos(x^3/r)$, $r = (x^i x^i)^{1/2}$ are the standard sperhical coordinates and I_a the isospin generators. This gauge transformation is indeed singular in that $\partial_i \omega \omega^{-1}$ has a singularity line along the positive x^3 semiaxis (such gauge transformations are also considered in Ref. 3). The fields ϕ' and $A_i'^a$ in the new gauge are given by

$$\begin{aligned}\hat{\phi}'^a \tau_a/2 &= \omega\hat{\phi}^a(\tau_a/2)\omega^{-1} = \delta_{a3}\tau_a/2,\\ A_i'^a\tau_a/2 &= \omega A_i^a(\tau_a/2)\omega^{-1} - (i/e)\partial_i\omega\omega^{-1}\\ &= \delta_{a3}(1/e)\epsilon_{i3k}[x^k/r(r-x^3)]\tau_a/2.\end{aligned} \tag{26}$$

The Brouwer degrees of the $S_R^2 \to S_{\hat{\phi}1}^2$ mapping defined by $\hat{\phi}$ and $\hat{\phi}'$ are +1 and 0 respectively as we have seen in Sec. 3. In the new gauge, however, the Yang–Mills field $A_i'^3$ is singular along the whole positive 3 semiaxis. It exhibits the familiar Dirac string of Abelian electrodynamics. In this gauge the magnetic charge is completely given by the $M_{\mu\nu}$ term

$$M = \frac{1}{4\pi}\int_{S_R^2} d^2\sigma_i \epsilon_{ijk} M_{jk}^1 = \frac{1}{4\pi}\int_{S_R^2} d^2\sigma \cdot \mathrm{curl}\mathbf{A}'^3 = \frac{1}{e},$$

where only the regular part of $M_{\mu\nu}$ is integrated [one omits the singular $\sim\delta(x^1)\delta(x^2)$ part[8]]. We thus see that the gauge transformation (25) has "transferred" the magnetic charge from the Higgs to the gauge field.

In the primed gauge (26) the fields have a very simple configuration: The Higgs field is constant over all of space, while of the gauge fields only one (the third) isospace component survives so that one has for all practical purposes an Abelian field. For this reason, we call a gauge in which

$$\hat{\phi}^a = \delta_{a3}, \quad A_i^a = \delta_{a3}A_i^3 \tag{27}$$

an *Abelian gauge*. (Of course, the common isodirection of ϕ and A need not be the 3 direction). In the Abelian gauge the nonlinearity of the Yang–Mills self-couplings is removed and the gauge field becomes additive. We are then led to a natural method for dealing with physically interesting systems of collinear monopoles and antimonopoles. It involves the following steps:

1. Start with an Abelian gauge.

2. In the Abelian gauge let A_i^3 equal the sum of the singular potentials corresponding to the monopoles and antimonopoles of the problem.

3. Make a gauge transformation to a gauge in which all string singularities are removed. The expression obtained in this way for the gauge and Higgs fields in the nonsingular gauge are the proper (boundary) expressions to be used for $|\mathbf{x}| \to \infty$.

The Abelian gauge expressions of the fields are obviously solutions of the dynamical equations in that gauge (except on the string). Since the dynamics is gauge invariant, this proves that the expressions for the fields in the nonsingular gauge will also be solutions of these equations and everywhere at that. One can then impose these solutions asymptotically for $r \to \infty$ and determine the form of the solution for small r by multiplying the solutions by arbitrary functions and minimizing the energy integral.[3] As a trivial example, if we start with the fields (26), apply the inverse ω^{-1} of the gauge transformation (25), then we find the proper asymptotic expressions (29) of 't Hooft.[3] Similarly for one monopole and one antimonopole one starts from

$$\hat{\phi}^a = \delta_{a3}, \quad A_i^a(x^j) = \delta_{a3}\frac{1}{e}\epsilon_{3ab}x^b\left(\frac{1}{r_-(r_- - a - x^3)} - \frac{1}{r_+(r_+ + a - x^3)}\right), \tag{28}$$

where $r_\pm$ is the distance of the point $\mathbf{x} = (x^1, x^2, x^3)$ to the point $\mathbf{A}_\pm = (0, 0 \pm a)$ (at which the monopole/antimonopole is located).

The proper gauge transformation in this case is

$$\omega_{+-} = \exp(-i\varphi I_3)\exp(-i\delta I_2)\exp(i\varphi I_3), \tag{29}$$

where δ is the angle between the vectors $\mathbf{xA}_+$ and $\mathbf{xA}_-$ [$\cos\delta = (r^2 - a^2)/r_+ r_-$]. It is readily checked that this gauge transformation indeed removes the string singularity for A_μ^a and that in the nonsingular gauge the $\phi^a(x)$ take the form (22) which thus acquires physical content. Remarkably enough for a system of two identically charged monopoles no static nonsingular gauge exists. The field (23) is then not physically interesting. We hope to return to these problems and their relation with dual resonance models elsewhere.

To sum up, we have found a remarkable complementarity between Dirac strings and Higgs zeros. Magnetic charge can be carried in the zeros of the Higgs field, but gauge fields can "carry" it only in the presence of Dirac strings. Gauge transformations can change the homotopy class of the Higgs field and thus the magnetic charge they carry. The gauge invariance of magnetic charge then requires gauge transformations that affect the homotopy class of the Higgs field, to remove or supply (as the case may be) Dirac strings in the gauge field.

5. CONCLUSIONS

We have established that the conservation of magnetic charge in non-Abelian gauge theories does not originate in a symmetry of the Lagrangian via the Noether mechanism. It is rather a topological invariant of a mapping between two spheres provided by the normalized Higgs fields. The study of the topology of Higgs fields has led us to a method for easy identification of the magnetic charge assignment to a given field configuration. The inclusion of gauge invariance has further revealed a remarkable "complementarity" between the zeros of Higgs fields and the singularity lines (Dirac strings) of gauge fields.

It may be worthwhile to point out the relevance of all this to hadrodynamics. One would want to identify the monopole—antimonopole system discussed in Sec. 4 with an (approximate) classical picture of a meson. The question is where are the quarks? So far, we have not included fermions, but we can do this in the standard way. Classical solutions of theories of fermions coupled to scalar Higgs and gauge fields have been studied recently.[9] It has been pointed out that the density of fermions is large only in the immediate vicinity of the zeros of the Higgs fields. But we have seen that zeros of the Higgs field localize magnetic charge. We thus see that Fermi fields like to localize around magnetically charged centers. We can thus conclude that the location of magnetic charge determines the location of the quarks even though the quarks do *not* themselves carry magnetic charge.

ACKNOWLEDGMENTS

One of us (P. G. O. F.) would like to acknowledge a most illuminating conversation with Professor Charles Fefferman of the University of Chicago Mathematics Department.

*Work supported in part by the National Science Foundation: NSF GP 32904 X2.
†AEC Grant AT[11-1]800.

[1]P.A.M. Dirac, Phys. Rev. **74**, 817 (1948).
[2]J. Schwinger, Phys. Rev. **144**, 1087 (1966); D. Zwanziger, Phys. Rev. D **3**, 880 (1971).
[3]G. 't Hooft, Preprint TH 1876-CERN.
[4]D. Finkelstein and C.W. Misner, Ann. Phys. (N.Y.) **6**, 230 (1959); D. Finkelstein, J. Math. Phys. **7**, 1218 (1966); D. Finkelstein and J. Rubinstein, J. Math. Phys. **9**, 1762 (1968).
[5]T.H.R. Skyrme, Proc. Roy. Soc. (London) A **247**, 260 (1958); T.H.R. Skyrme, Nucl. Phys. **31**, 556 (1962), this paper contains further references.
[6]J. Hadamard, note in J. Tannery, *Introduction à la théorie des fonctions* (Herman, Paris, 1910), pp. 437—77; C. Allendoerfer, Am. J. Math. **62**, 243 (1940).
[7]See the excellent little book by J.W. Milnor, *Topology from the Differentiable Viewpoint* (Univ. of Virginia Press, Charlottesville, 1965).
[8]G. Wentzel, Progr. Theor. Phys. (Kyoto) Suppl. **37** and **38**, 163 (1966).
[9]See, e.g., R. Dashen, B. Hasslacher, and A. Neveu, IAS Princeton preprint.

A variational principle in statistical mechanics for particle systems with bounded pair interactions

A. Gerardi, N. Ianiro*, C. Maffei*, and C. Marchioro*

Istituto Matematico Università di Roma, Rome, Italy
(Received 13 December 1973)

A gas of particles interacting pairwise via bounded potential is considered in the framework of rigorous statistical mechanics. It is proved that the pressure is the supremum on a class of states of the difference between the entropy and the mean energy.

1. INTRODUCTION

Variational principles are well known in thermodynamics. They have been recently discussed also in the framework of rigorous statistical mechanics. In particular, the principle that defines the pressure as the supremum on a large class of states of the difference between the entropy and the mean energy has been considered by Ruelle,[1] for lattice systems, and by Gallavotti and Miracle-Sole, for hard sphere systems.[2]

In this paper we extend the variational principle to a system of particles interacting via a bounded pair potential. In order to control large fluctuations of the number of particles in bounded regions, we suppose the potential to be superstable.

In Sec. 2 notations and assumptions are defined. In Sec. 3 we discuss the definition of states and of their entropy. In Sec. 4 we define the mean energy and finally in Sec. 5 the variational principle is established.

2. NOTATIONS AND ASSUMPTIONS

We consider a system of identical particles in $\mathbb{R}^\nu$, interacting pairwise via a potential ϕ. We assume the following:

Assumption 1: Let

$$\phi : \mathbb{R}^\nu \to \mathbb{R}$$

be a continuous translational invariant function such that

$$\phi(-x) = \phi(x).$$

(In fact, the Lebesgue measurability is sufficient for the validity of the variational principle).

We define the interparticle configuration energy as

Definition 1:

$$\cup_\psi(x_1, \ldots, x_m) = \underset{1}{\cup}(x_1, \ldots, x_m) + \underset{2}{\cup}_\phi(x_1, \ldots, x_m),$$

where $\cup_1$ is the one-body potential (μ = chemical potential)

$$\underset{1}{\cup}(x_1, \ldots, x_m) = -\mu m,$$

$$\psi \equiv (\mu, \phi)$$

and

$$\underset{2}{\cup}_\phi(x_1, \ldots, x_m) = 0, \quad m = 0, 1.$$

$$\underset{2}{\cup}_\phi(x_1, \ldots, x_m) = \sum_{i<j} \phi(x_i - x_j), \quad m > 1.$$

Assumption 2 (stability): There exists $B \geq 0$, such that for all m, $x_1, \ldots, x_m$

$$\underset{2}{\cup}_\phi(x_1, \ldots, x_m) \geq -mB.$$

The following partition of $\mathbb{R}^\nu$ will be useful in the sequel

Definition 2:

$$\Gamma(r) = \{x \in \mathbb{R}^\nu : r^i - \tfrac{1}{2} \leq x^i < r^i + \tfrac{1}{2}\}$$

where $r \in Z^\nu$.

Assumption 3 (superstability): There exist $A > 0$, $B \geq 0$ such that for every finite set $R \subset Z^\nu$ and $X \equiv (x_1, \ldots, x_m) \subset \cup_{r \in R} \Gamma(r)$

$$\underset{2}{\cup}_\phi(x_1, \ldots, x_m) \geq \sum_{r \in R} [A\, n(X, r)^2 - B\, n(X, r)]$$

where $n(X, r) = \text{Card}\, X \cap \Gamma(r)$.

Assumption 4: There exists a decreasing function

$$\varphi : Z^+ \to \mathbb{R}^+$$

such that

$$\hat{\varphi} = \sum_{r \in Z^\nu} \varphi(|r|) < \infty,$$

$$\sup_{\substack{x \in \Gamma(r) \\ y \in \Gamma(0)}} |\phi(x - y)| \leq \varphi(|r|)$$

where $|r| = \sup r^i$ for $r \in Z^\nu$. Note that Assumption 4 implies weak-tempering and lower-regularity conditions for ϕ.

The set of ψ has a natural structure of a linear space. In this space we introduce the norm of ψ as

$$\|\psi\| = |\mu| + \sum_{r \in Z^\nu} \sup_{\substack{x \in \Gamma(r) \\ y \in \Gamma(0)}} |\phi(x - y)|.$$

We call $\mathcal{N}$ the set of all ψ satisfying Assumptions 1–4 and $\mathcal{H}$ the subset also satisfying Assumption 3.

It can be shown that the following theorem holds:

Theorem 2.1: If $X \equiv (x_1, \ldots, x_m) \subset \cup_{r \in R} \Gamma(r)$ (R finite subset of Z), then

$$\cup_\psi(X) \leq \|\psi\| \sum_{r \in R} n^2(X, r)$$

3. STATES AND ENTROPY

The description of states in classical mechanics as states on an Abelian C^*-algebra, has been considered by several authors (see for instance Refs. 3 and 4). Let

$T \in \mathbb{R}^\nu$ be the configurational space of one particle (as usual we omit the particles momenta), and let

$$\mathcal{T} = \sum_{n \geq 0} T^n$$

be the topological sum of disjoint copies of T^n. Let also k_Λ^n be the space of real continuous functions on T with support in Λ^n ($\Lambda \subset T$ open bounded set) and k_Λ the space of sequences $(f^n)_{n \geq 0}$ where $f^n \in k_\Lambda^n$ and $f^n = 0$, for n large enough. Finally we call k the union of k_Λ. Therefore an element of k may be considered as a function on $\mathcal{T}$.

If $f \equiv (f^n) \in_{n \geq 0} k$ we define a function $Sf \in k$ in such way that its restriction to T^n is

$$Sf(x_1, \ldots, x_n) = \sum_{p \geq 0} \sum_{i_1=1}^{n} \cdots \sum_{i_p=1}^{n} f^p(x_{i_1}, \ldots, x_{i_p}).$$

We now construct an abelian C^*-algebra. For any integer $q \geq 0$, and any bounded continuous complex function φ on $\mathbb{R}^q$, the set $\tilde{\mathcal{A}}$ of the function $\varphi(Sf_1, \ldots, Sf_q)$ on $\mathcal{T}$, with respect to the usual operations, forms a commutative $*$-algebra. Its closure $\mathcal{A}$ with respect to the uniform norm is our Abelian C^*-algebra.

Let E be the space of states on $\mathcal{A}$. We shall consider now a subset of E, related to a family of density distributions. For every bounded open $\Lambda \subset T$ and any integer $n \geq 0$ let $\mu_\Lambda^n \geq 0$ be a measure on Λ^n, symmetric in its n arguments. We say that (μ_Λ^n) is a family of density distributions, if the following conditions are satisfied:

(i) Normalization. For all Λ.

$$\sum_{n=0}^{\infty} \mu_\Lambda^n(\Lambda^n) = 1. \tag{3.1}$$

(ii) Compatibility. Let $\Lambda \subset \Lambda'$ and $\chi_{\Lambda'/\Lambda}$ be the characteristic function of Λ'/Λ where Λ'/Λ is the complement of Λ in Λ'. If $f^n \in k_\Lambda^n$ then

$$\mu_\Lambda^n(f^n) = \sum_{m=0}^{\infty} \frac{(n+m)!}{n!\, m!}\, \mu_{\Lambda'}^{n+m}(f^n \otimes \chi_{\Lambda'/\Lambda}^{\otimes m}), \tag{3.2}$$

where

$$(f^n \otimes \chi_{\Lambda'/\Lambda}^{\otimes m})(x_1, \ldots, x_{n+m}) = f^n(x_1, \ldots, x_n)\chi_{\Lambda'/\Lambda}(x_{n+1}) \times \cdots \chi_{\Lambda'/\Lambda}(x_{n+m}).$$

We consider the state ρ on $\tilde{\mathcal{A}}$ defined by

$$\rho(\varphi(Sf_1, \ldots, Sf_q)) = \sum_{n \geq 0} \int_{\Lambda^n} d\mu_\Lambda^n(x_1, \ldots, x_n)\varphi(Sf_1(x_1, \ldots, x_n), \ldots, Sf_q(x_1, \ldots, x_n))$$

where $f_1, \ldots, f_q \in k_\Lambda$ and we extend it by continuity to a state ρ on $\mathcal{A}$.

We call $\mathcal{J}$ the set of states thus obtained.

If F is a function on $\mathcal{T}$, a translation τ_a by $a \in T$ is defined by

$$\tau_a F(x_1, \ldots, x_n) = F(x_1 - a, \ldots, x_n - a).$$

We call $\mathcal{J}^\perp$ the subset of $\mathcal{J}$, consisting of translation invariant elements that are related to a density distribution satisfying the following requirement:

(iii) Invariance:

$$\mu_\Lambda^n(f^n) = \mu_{\Lambda+a}^n(\tau_a f^n).$$

We require a further condition on the states:

(iv) There exist real numbers γ, δ with $\gamma > 0$, such that for every bounded and open region Δ and for every $f^n \in k_\Delta^n$, we have

$$\int_{\Delta^n} f^n(X)|d\mu_\Delta^n(X) \leq \frac{1}{n!} \int_{\Delta^n} f^n(X) \exp \sum_r [-\gamma n^2(X, r) + \delta n(X, r)]\, dX$$

where $dX \equiv dx_1 \ldots dx_n$ is the Lebesgue measure; we call $\hat{\mathcal{J}}^\perp$ the set of states of $\mathcal{J}^\perp$ satisfying (iv).

This condition is suggested by the fact that it holds for the equilibrium state of a gas of particles interacting via a superstable potential; of course, in this case γ and δ depend on the potential considered. [5]

The following inequality will be useful in the sequel.

Lemma 3.1: Let $\Lambda \subset \mathbb{R}^\nu$ be an open bounded set containing the origin. If (μ_Λ^n) is a family of density distribution satisfying (i)—(iv), then for every $M \in Z^+$

$$V(\Lambda)^{-1} \sum_{n \geq M} \int_{\Lambda'', n} \sum_r n^2(X, r)\, d\mu_\Lambda^n(X) \leq \exp(-\gamma M^2)\epsilon(\Lambda)\, \text{const.}$$

where Λ''^n is the set of $X \subset \Lambda^n$ such that $n(X, 0) = m \geq M$ and $\epsilon(\Lambda) \to 1$ when $\Lambda \to \infty$ in the Van Hove sense.

Proof: Let $\hat{\Gamma}(r) = \Lambda \cap \Gamma(r)$ and let $\mathcal{R} = \{r \in Z^\nu : \hat{\Gamma}(r) \neq \emptyset\}$ be a finite set. We consider separately the contribution of the term $r = 0$ and the other ones. The first one gives

$$\sum_{n \geq M} \int_{\Lambda'', n} n^2(X, 0)\, d\mu_\Lambda^n(X) = \sum_{n \geq M} \sum_{m=M}^{n} m^2 \binom{n}{m} \int_{\hat{\Gamma}(0)^m \times (\Lambda'' \setminus \hat{\Gamma}(0))^{n-m}} \times d\mu_\Lambda^n(X)$$

$$= \sum_{m \geq M} m^2 \sum_{p \geq 0} \frac{(m+p)!}{m!\, p!} \times \mu_\Lambda^{m+p}(\chi_0^m \otimes \chi_{\Lambda/\hat{\Gamma}(0)}^{\otimes p})$$

where χ_0^m is the characteristic function of $\hat{\Gamma}(0)^m$. Using (ii) and (iv) we obtain

$$\sum_{n \geq M} \int_{\Lambda'', n} n^2(X, 0)\, d\mu_\Lambda^n(X) = \sum_{m \geq M} m^2\, \mu_{\hat{\Gamma}(0)}^m(\chi_0^m)$$

$$\leq \sum_{m \geq M} \frac{m^2}{m!} \exp(-\gamma m^2 + \delta m) \leq \text{const} \quad \exp(-\gamma M^2).$$

For the other terms

$$\sum_{n \geq M} \int_{\Lambda'', n} \sum_{r \in \mathcal{R} - \{0\}} n^2(X, r)\, d\mu_\Lambda^n(X)$$

$$= \sum_{r \in \mathcal{R} - \{0\}} \sum_{n \geq M} \sum_{m=M}^{n} \sum_{p=0}^{n-m} p^2 \int_{(\hat{\Gamma}(0)^m \times \hat{\Gamma}(r)^p) \times (\Lambda / [\hat{\Gamma}(0)^m \times \hat{\Gamma}(r)^p])} d\mu_\Lambda^n(X) \binom{n}{m+p}$$

$$\leq \exp(-\gamma M^2)\, \text{const} \quad (\text{card}\, \mathcal{R} - 1)$$

where we made use of (ii) and (iv).

With the aid of the Van Hove limit the lemma is finally proved. QED

The entropy shall now be defined. Let (μ_Λ^n) be the family of density distribution related to a state $\rho \in \mathcal{F}^\perp$. We suppose that for every Λ, n, μ_Λ^n is absolutely continuous with respect to the Lebesgue measure and we put

$$d\mu_\Lambda^n(X) = \frac{1}{n!}\, \sigma_\Lambda^n(X)\, dX.$$

The entropy $S_\rho(\Lambda)$ is defined by:

Definition 4:

$$S_\rho(\Lambda) = -\sum_{n\geq 0} \frac{1}{n!} \int_{\Lambda^n} dX\, \sigma_\Lambda^n(X) \log \sigma_\Lambda^n(X).$$

The following results hold.[4]

Theorem 3.1 (Robinson–Ruelle): The limit

$$s = \lim_{\Lambda\to\infty} \frac{S_\rho(\Lambda)}{V(\Lambda)} = \inf_\Lambda \frac{S_\rho(\Lambda)}{V(\Lambda)} < +\infty$$

exists in the Van Hove sense.

Definition 5 (Mean entropy): $s(\rho)$ is defined on $\mathcal{F}^\perp$ by (i) $s(\rho) = s$ if $\rho \in \mathcal{F}^\perp$ and the measures μ_Λ^n associated with ρ are absolutely continuous with respect to the Lebesgue measure. (ii) $s(\rho) = -\infty$ otherwise.

Theorem 3.2[4]: The functional $s(\rho)$ is affine on $\mathcal{F}^\perp$.

4. MEAN ENERGY

Let $\Lambda \subset \mathbb{R}^\nu$ be an open bounded region, $\rho \in \mathcal{F}^\perp$ and $\psi \in \mathcal{N}$. We define the mean energy as

$$U_{\psi,\Lambda}(\rho) = V(\Lambda)^{-1} \sum_{n\geq 0} \int_{\Lambda^n} U_{\psi,\Lambda}(X)\, d\mu_\Lambda^n(X).$$

Theorem 4.1: Let $\rho \in \hat{\mathcal{F}}^\perp$, then

$$\lim_{\Lambda\to\infty} U_{\psi\Lambda}(\rho) = U_\psi(\rho)$$

when Λ is a net of increasing cubes. The functional $\rho \to U_\psi(\rho)$ is affine.

Proof: We first note that for every Λ not too small, for instance $\Lambda \supset \Gamma(0)$, the following continuity relation holds:

$$|U_{\psi\Lambda}(\rho)| \leq \|\psi\| \text{ const}$$

where we used Theorem 2.1 and Lemma 3.1.

So we can confine ourselves to consider finite range potentials. We put

$$U_{\psi\Lambda}(\rho) = U'_{\psi\Lambda}(\rho) + U''_{\psi\Lambda}(\rho)$$

where $U'_{\psi\Lambda}(\rho)$ is the functional obtained considering only the configurations in which no more than M particles are in each $\Gamma(r)$. By Lemma 3.1 we have

$$U''_{\psi\Lambda}(\rho) \leq \text{const}\ \exp(-\gamma M^2)\, \epsilon(\Lambda)\, V(\Lambda)\, \|\psi\| \tag{4.1}$$

which vanishes in the limit

$$V(\Lambda) \to \infty, \tag{4.2}$$

$$M = \mathcal{J}\left(\log \frac{V(\Lambda)}{\gamma}\right) \quad (\mathcal{J} \equiv \text{integer part}).$$

We prove that in the same limit $U'_{\psi\Lambda}(\rho)$ converges. The proof is similar to that of Ref. 2. For every $X \subset \mathbb{R}^\nu$ let $\theta(X)$ be a function defined as

$$\theta(X) = 0 \qquad \text{if } X \cap \Gamma(0) = \phi,$$

$$\theta(X) = \epsilon(x_1) + \cdots + \epsilon(x_m) \quad \text{where } x_r \in X \cap \Gamma(0), \quad r = 1, \ldots, m.$$

and $\epsilon(\xi)$ is a continuous function with compact support contained in $\Gamma(0)$, such that

$$\int_{\Gamma(0)} \epsilon(\xi)\, d\xi = 1.$$

We define $A_\psi^\theta \in \mathcal{A}$ as

$$A_\psi^\theta(X) = \sum_{T\subset X} \frac{\theta(T)\psi(T)}{N(T)} = S\left(\frac{\theta\psi}{N}\right)(X)$$

with

$$\psi(T) = -\mu \qquad \text{if} \qquad N(T) = 1,$$

$$\psi(T) = \phi(x-y) \quad \text{if} \quad N(T) = 2, \quad T \equiv (x, y),$$

$$\psi(T) = 0 \qquad \text{otherwise.}$$

and

$$\text{supp } A_\psi^\theta \subset \Lambda_0.$$

if Λ_0 is a sufficiently large cube centered at the origin. It is easily seen that, for every $X \equiv (x_1, \ldots, x_n)$ $(x_i \in \mathbb{R}^\nu)$ if $\Lambda(X)$ is a cube centered at the origin containing X and $\Lambda \supset \Lambda(X) + \Gamma(0)$,

$$\int_\Lambda d\xi (\tau_\xi A_\psi^\theta)(X) = \int_\Lambda d\xi\, A_\psi^\theta(X - \xi) = U_\psi(X). \tag{4.3}$$

Now let $\rho \in \hat{\mathcal{F}}^\perp$ and $\{\mu_\Lambda^n\}$ be its family of density distribution. We define

$$\rho_m(A_\psi^\theta) = \sum_{n\geq 0} \int_{\Lambda' n} A_\psi^\theta(X)\, d\mu_\Lambda^n(X) \tag{4.4}$$

where Λ'^n is the set of all configurations $X \subset \Lambda$, $\Lambda \supset \Lambda_0$ with $n(X, r) \leq M$. We note that there exists the limit

$$\lim_{M\to\infty} \rho_m(A_\psi^\theta) = \rho(A_\psi^\theta). \tag{4.5}$$

In fact,

$$|A_\psi^\theta(X)| \leq 2 \max_{\xi\in\Gamma(0)} \epsilon(\xi) \|\psi\| \sum_{s\subset S} n^2(X, s)$$

where

$$S = \{s \in Z^\nu : [\Lambda_0 + \Gamma(0)] \cap \Gamma(s) \neq \phi\}$$

so that, by (4.4) and Lemma 3.1, (4.5) is proved.

The proof then proceeds as in Ref. 2. Using the limit (4.2), we prove that

$$\rho(A_\psi^\theta) = U_\psi(\rho). \qquad \text{QED}$$

5. VARIATIONAL PRINCIPLE

Definition 6: Let Λ a bounded Lebesgue measurable region of $\mathbb{R}^\nu$ and $\psi \in \mathcal{N}$:

$$Z_\Lambda(\psi) = \sum_{n\geq 0} \frac{1}{n!} \int_{\Lambda^n} dX \exp[-\beta U_\psi(X)].$$

Definition 7 (Pressure):

$$P_\Lambda(\psi) = \beta^{-1}\, V(\Lambda)^{-1} \log Z_\Lambda(\psi).$$

For $\Lambda \to \infty$ in the Fisher sense, the thermodynamic limit exists:

$$\lim_{\Lambda\to\infty} P_\Lambda(\psi) = P(\psi).$$

A property of the density distribution associated with an equilibrium state is known:

Lemma 5.1 (Ruelle[5]): Let $\psi \in \mathcal{M}$; the equilibrium state of a system of identical particle in a box Λ is related to a family of density distribution $\bar{\mu}_\Lambda^\eta$ such that

$$d\bar{\mu}_\Lambda^\eta(X) = \frac{\exp[-\beta U_\psi(X)]}{Z_\Lambda(\psi)}\, dX$$

and there exist $\bar{\gamma}$, $\bar{\delta}$ (independent of Λ) so that (iv) of Sec. 3 holds.

Assumption 5: Call $\hat{\mathcal{F}}_\psi^\perp$ the subset of $\hat{\mathcal{F}}^\perp$ consisting of states such that

$$\delta > \bar{\delta},\quad 0 < \gamma < 2^{-2\nu}\bar{\gamma} < \bar{\gamma}$$

where $\bar{\gamma}$ and $\bar{\delta}$ are defined in Lemma 5.1. The assumption $\delta > \bar{\delta}$ and $\gamma < \bar{\gamma}$ assures that $\hat{\mathcal{F}}_\psi^\perp$ contains the equilibrium state. The stronger Assumption 5 is introduced by technical reasons.

We give now the main theorem of this note:

Theorem 5.1 (Variational principle): The functional $\rho \to S(\rho)\beta^{-1} - U_\psi(\rho)$ is affine on $\psi \in \mathcal{M}$ and, for every $\psi \in \mathcal{M}$,

$$P(\psi) = \sup_{\rho \in \hat{\mathcal{F}}_\psi^\perp} [S(\rho)\beta^{-1} - U_\psi(\rho)].$$

The proof runs in the same way as in Ref. 1, also taking into account the technique of Theorem 4.1. Note that the Assumption 5 is introduced to assure that the test measure used in Ref. 1 be in $\hat{\mathcal{F}}_\psi^\perp$.

ACKNOWLEDGMENTS

It is a pleasure to thank G. Gallavotti, E. Presutti, and A. Schiaffino for interesting discussions.

*Research partially supported by the Consiglio Nazionale delle Ricerche (G.N.F.M.).

[1]D. Ruelle, Commun. Math. Phys. 5, 324 (1967).
[2]G. Gallavotti and S. Miracle-Sole, Ann. Inst. Henri Poincaré VIII (3), 287 (1968).
[3]D. Ruelle, J. Math. Phys. 8, 1657 (1967).
[4]D. Robinson and D. Ruelle, Commun. Math. Phys. 5, 288 (1967).
[5]D. Ruelle, Commun. Math. Phys. 18, 127 (1970).
[6]D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969).

Erratum: Multichannel stationary scattering theory in two-Hilbert space formulation [J. Math. Phys. 14, 957 (1973)]

E. Prugovečki

Department of Mathematics, University of Toronto, Toronto, Canada M4P 2A7
(Received 3 October 1974)

The second parts of Theorem 4.1 and 4.2 should state *only* that (3.16) implies (4.7) and (2.5) implies (4.10), respectively. The converses of these two statements are not true. This is an essential observation since, as discussed later, a scattering theory in the Liouville space $\mathcal{B}_2(\mathcal{H})$ based on (4.7) and (4.10) is actually more general than its couterpart on the Hilbert space $\mathcal{H}$.

The errors stem from the incorrect relation in lines 16 and 17 on p. 962. The expression on line 17 is not equal to the one on line 16 but rather to

$$\langle[\underline{\Omega}_\pm^J - \underline{U}(t)J\underline{\tilde{U}}(-t)]\,|\Psi)(\Psi\|[\underline{\Omega}_\pm^J - \underline{U}(t)J\underline{\tilde{U}}(-t)]\,|\Psi)(\Psi|\rangle_2.$$

The *correct* relation is:

$$\langle[\underline{\Omega}_\pm^J - \underline{U}(-t)J\underline{\tilde{U}}(t)]\,|\Psi)(\Psi|\,|\,[\underline{\Omega}_\pm^J - \underline{U}(-t)J\underline{\tilde{U}}(t)]\,|\Psi)(\Psi|\rangle_2$$
$$= ||\,\Omega_\pm^J\Psi\,||^4 + ||\,U(-t)J\tilde{U}(t)\,\Psi\,||^4$$
$$- 2\,|\langle\Omega_\pm^J\Psi\,|\,U(-t)J\tilde{U}(t)\Psi\rangle\,|^2. \tag{1}$$

By using (1) and the same method as in the last part of the proof of Theorem 4.1, we conclude that (3.16) indeed implies (4.7) for any

$$\rho \in \mathcal{B}_1(\mathcal{H}).$$

To see that the converse is not generally true, consider the case of two-body scattering, when $\mathcal{H}=\tilde{\mathcal{H}}$, and $\tilde{H}$ is the free Hamiltonian. Assume that

$$\Omega_\pm = \text{s-}\lim_{t\to\pm\infty} U(-t)\tilde{U}(t) \tag{2}$$

does exist. If we perform an energy shift η of $\tilde{H}$, i.e., consider instead the model with the free Hamiltonian $\tilde{H}' = \tilde{H} + \eta\mathbb{1}$,

$$\Omega_\pm' = \text{s-}\lim_{t\to\pm\infty} U(-t)\tilde{U}'(t), \quad \tilde{U}(t) = \exp(-i\tilde{H}'t)$$

does not exist because of the factor $\exp(-i\eta t)$, which diverges for $t \to \pm\infty$. However, the super-operator limit

$$\underline{\Omega}_\pm'\rho = \text{h-}\lim_{t\to\pm\infty} \underline{U}(-t)\underline{\tilde{U}}'(t)\rho$$

still exists for any $\rho \in \mathcal{B}_1(\mathcal{H})$, and is actually equal to

$$\underline{\Omega}_\pm\rho = \text{h-}\lim_{t\to\pm\infty} \underline{U}(-t)\underline{\tilde{U}}(t)\rho. \tag{3}$$

In general, the existence of the limits $\Psi_\pm = \Omega_\pm\Psi$ in (2) for a given $\Psi \in \mathcal{H}$ requires that

$$\lim_{t\to\pm\infty} ||\,\Psi_\pm - U(-t)\tilde{U}(t)\,\Psi\,|| = 0. \tag{4}$$

On the other hand, according to (1), the existence of the limit $\rho_\pm = |\Psi_\pm\rangle\langle\Psi_\pm|$ given by (3) for the corresponding $\rho = |\Psi\rangle\langle\Psi|$

demands only that

$$\lim_{t\to\pm\infty} |\langle\Psi_\pm\,|\,U(-t)\tilde{U}(t)\Psi\rangle\,| = 1. \tag{5}$$

Evidently, as a condition, (5) is strictly weaker than (4).

